UNIVERSITY OF CAMBRIDGE

The Past, Present & Future in Optimisation for Institutional Investors

S.E. Satchell

University of Cambridge

With special thanks to

J. Knight (UWO) and

C. Koutsoyannis (OMAM)

Presented to:

Past, Present & Future in Investment Management

11 – 12 August 2005, Sydney Harbour Marriott Hotel

UNIVERSITY OF CAMBRIDGE

The Brief

I wish to investigate the relevance of robust optimisation.

2

UNIVERSITY OF CAMBRIDGE

Risk vs. Uncertainty, Bayes vs. Robustness

- **Risk:** when you capture model uncertainty by assigning priors to parameters or structures.
- **Uncertainty:** when you cannot assign priors.

- **Estimation Risk/Uncertainty** (eg of unknown mean, variance)
- **Model Risk/Uncertainty** (eg unknown distribution)

- **Bayes:** assign prior probabilities to scenarios
- **Ambiguity aversion:** does not assign probabilities (uncertainty resolved by multiple priors and some version of maximum expected utility).

3

UNIVERSITY OF CAMBRIDGE

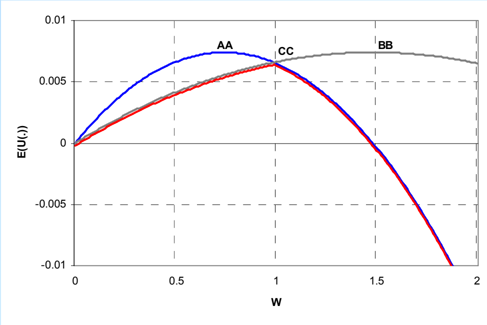
Maximin Expected Utility

- Much of the robustness literature is centred on this concept.
- Origins in Ellsberg's (1961) discovery of ambiguity aversion.
- Has earlier origins in decision theory, problem of not being able to assign probabilities to priors.
- Resolution due to Gilboa and Schmeidler (1989) is to capture ambiguity aversion by taking maximin of expected utility, where the decisions are made over different scenarios.
- Intellectual basis of scenario analysis, more averse you are, more scenarios you consider.
- Lutgens and Schotman (2004) apply this to mean-variance analysis (see diagram)

4

UNIVERSITY OF CAMBRIDGE

Maximin Expected Utility



Notice: CC not necessarily terribly conservative

5

UNIVERSITY OF CAMBRIDGE

Linear Factor Model & Robustness

- Uncertainty/Robustness to implied factor returns: experts with differing views about (unknown) implied factor returns.
- We have added Bayesian updating to both the case of ambiguity aversion and the case of Bayesian priors with probabilities: We combine the overall prior distribution of the implied factor returns with their sample distribution.
- In the case of a simple prior, we recapture the Black-Litterman model (1990).
- In the case of ambiguity aversion, we follow a robust maximum methodology by calculating the expected utility for each expert's separate distribution, and optimising for the minimum of those.

6

UNIVERSITY OF CAMBRIDGE

Linear Factor Model & Robustness: The Details 1

- The weighted least squares estimator of the (unknown) implied factor returns is given by:

$$\hat{V}_t = \left(\beta_t' \hat{D}^{-1} \beta_t\right)^{-1} \left(\beta_t' \hat{D}^{-1} y_t\right) \qquad \hat{V}_t \sim \left(V_t^{\,p}, \left(\beta_t' \hat{D}^{-1} \beta_t\right)^{-1}\right)$$

- The overall prior distribution is a mixture of normals:

$$pdf\left(V_t^{\,p}\right) = \Sigma \pi_j N\left(V_0^{\,j}, \Omega^j\right)$$

- The combined distribution is:

$$pdf\left(V_t^{\,p}\right) = \Sigma \pi_j N\left(V_0^{\,j}, \Omega^j\right) \quad = \Sigma \pi_j N\left(\mu_j, \theta_j\right)$$

7

UNIVERSITY OF CAMBRIDGE

Linear Factor Model & Robustness: The Details 2

- Following a robust maximin methodology on expected utility, the solution will be along the lines of Theorem 1 in Lutgens and Schotman (2004). That is our optimal portfolio will be of the form:

$$w^* = \left(\beta_t \Psi \beta_t' + D\right)^{-1} \left(\Sigma \lambda_j \beta_t \mu_j\right)$$

- Numerical computation and practical implementation of technique?
- Computationally, ambiguity case seems much more difficult to implement.

8

UNIVERSITY OF CAMBRIDGE

Alternative Robust Methodologies

- Robust statistics and mixed distributions (Victoria-Fesser (2004))
- Lower bounds on expected utility (Andersen et al (2000), Chu (2003))
- Convex optimisation (eg Costa and Paiva (2002), huge literature)
- Resampled frontiers (Michaud (1989,1998) – Michaud averages could be interpreted as equal prior probabilities)
- Cavadini Sbuelz Trojani (2001) et al examine estimation risk/uncertainty vs model, in the context on MV analysis. They adjust the coefficient of risk aversion to express uncertainty. No clear intellectual justification.

9

UNIVERSITY OF CAMBRIDGE

Axiomatic basis of Ambiguity Aversion

- Gilboa and Schmeidler (1989) present a set of axioms that guarantee the validity of AA.
- We can demonstrate that one of these axioms will not hold for MV analysis (certainty independence).
- This casts doubts as to the validity of Lutgens and Schotman (2004).

10

Tentative Conclusions

- Robust optimisation is a huge area. Applications to portfolio choice by no means established in the MV context
- Bayes-type models (Black and Litterman) available and implementable.
- Michaud-type models questionable on a substantial number of grounds, more on this later.
- Robust statistics can always be used as an input to optimisation.
- Tentative conclusion is to recommend a BL approach to impied factor returns, with the possible use of robust statistics.
- Data-cleaning methods will have robustified the data in the right direction.
- Winsorisation techniques will also robustify.

11

Robust Optimisation

- The next section deals with robust optimisation based on resampling.

12

UNIVERSITY OF CAMBRIDGE

Abstract

- We revisit the problem of calculating the exact distribution of optimal investments in a mean variance world under multivariate normality. Whilst a number of authors have considered aspects of this problem before, we extend the problem by considering the problem of an investor who wish to maximise quadratic utility defined in terms of alpha and tracking errors. The results derived allow some exact and numerical analysis. Furthermore, they allow us to also solve the more traditional fully invested portfolio problem. The results shed light on when an optimised portfolio simulator can go horribly wrong.

13

UNIVERSITY OF CAMBRIDGE

A Primer in Simulation (Monte Carlo 1)

- Monte Carlo methods consist of the following:
  - I use my computer to randomly generate N values from a given distribution, f(x)
  - The average of these ∞ values will be the mean

$$E(x) = \int_{-\infty}^{\infty} x f(x) dx \ \textit{as} \ N \to \infty.$$

  - If I compute the empirical distribution function (cumulative histogram), $F_N(x)$, it will approach $\int_{-\infty}^{x} f(s) ds$, the true distribution.

14

UNIVERSITY OF CAMBRIDGE

A Primer in Simulation (Monte Carlo 1) contd.

- If I compute some function of the values $(x_1,\ldots,x_N)$ i.e. $g_i = g(x_i)$, then the average $\sum_{i=1}^{N} g_i / N$ will approach $E(g(x))$ as $N \to \infty$

15

UNIVERSITY OF CAMBRIDGE

A Primer in Simulation (Monte Carlo 2)

- If $E(g(x))$ is *unbiased* i.e. $g(x) = \hat{\theta}$ and $E(\hat{\theta}) = \theta$ where θ is now a parameter of interest, then Monte Carlo will lead us to the *true* value assumed in the Monte-Carlo experiment.
- If $E(g(x))$ is biased, then the Monte-Carlo experiment leads to some value different from the true value. Such a value can be quite misleading.

16

Application of Exact Theory to MV Analysis

- The application of exact distribution theory to mean-variance (MV) analysis has been undertaken by a number of authors, see Jobson and Korkie (1989), Jobson (1990), Britten-Jones (1999) and Hillier and Satchell (2003). The usual assumptions are that returns are iid multivariate normal and that there may or may not be a riskless asset. However, in all listed cases the analysis is in terms of absolute, i.e. unbenchmarked, portfolios. This is a limitation since most institutional risk analysis is based on MV analysis using returns relative to a benchmark.
- One purpose of this paper is to consider expected utility in terms of relative returns and compute the exact properties of the optimal alpha, tracking error, and Sharpe ratio

17

The Literature/Issues

- The major motivation for this research has been to try and understand the magnitude of estimation error; that is, the extent to which the outcome of the portfolio decision is influenced by parameter uncertainty. Michaud (1998) has proposed a resampling procedure, the outcome of which can only be really understood by an analysis of the exact properties of optimal portfolios. Michaud's procedure purports to solve some of the problems of portfolio optimisation. Other authors have criticised Michaud's approach, see Harvey et al. (2004) and Scherer (2002).

18

UNIVERSITY OF CAMBRIDGE

The Literature/Issues (contd)

- Our results, whilst being highly simplified, since we do not impose the myriad of constraints that institutional portfolios typically obey, nevertheless exhibit certain key characteristics that shed light on investment issues. Furthermore, we are able to extend the problem to consider the same case with absolute, not relative, weights. This allows us to derive some new results for this problem.

19

UNIVERSITY OF CAMBRIDGE

The Michaud Methodology

- To motivate our analysis we consider how Michaud (1998) carries out his resampling methodology. Quoting from Michaud (op. cit, pages 17, 19 and 37).
- "1. Monte Carlo simulate 18 years of monthly returns based on data in Tables 2.3 and 2.4...
- 2. Compute optimised input parameters from the simulated return data.
- 3. Compute efficient frontier portfolios…
- 4. Repeat steps 1 3500 times…
- 5. ….Observe the variability in the efficient frontier estimation."

20

UNIVERSITY OF CAMBRIDGE

The Michaud Methodology

- The assumption behind the Monte Carlo simulation of returns can vary. It can be based on historical returns and involve resampling, or it may involve using means variance and covariance and simulating via multi-variate normality as Michaud details above, his Tables 2.3 and 2.4 contain first and second sample moments. The key feature of such an analysis is that the mean simulated efficient frontier will differ from the "population" efficient frontier based on the information in step 1 by the degree of finite sample bias. Whilst this may be small for $T = 216$ monthly observations, there are lots of portfolio calculations that will be based on much shorter time-periods due to the usual reasons; regime shifts, institutional change and time-varying parameters. Furthermore, we conjecture, and subsequently show, that it is not T that determines bias alone but T and N (the number of stocks) co-jointly. If N gets large as T goes to infinity, then biases can be very large indeed.

21

UNIVERSITY OF CAMBRIDGE

The Michaud Methodology

- Institutional active portfolios usually have N ranging from 60 to 120. A passive (absolute) portfolio could have N equal to the size of the stock universe.

22

UNIVERSITY OF CAMBRIDGE

Notation

- It is worth noting that the emphasis of the above approach is in terms of MV efficient frontier analysis rather than expected utility. But as we shall show next maximising quadratic utility gives you a solution that is expressed solely in terms of efficient set mathematics; the only additional information is the risk aversion coefficient (λ); as we change λ we move along the MV frontier in any case.
- Consider the active weights ω and the *known* benchmark weights b both (N×1) vectors and both sum to one i.e. $\omega' i = b' i = 1$.

  Let μ and Ω be the (n × 1) mean vector and covariance matrix of the N asset returns where the letter i denotes an (N×1) vector of one.

23

UNIVERSITY OF CAMBRIDGE

Problem

- Our investor chooses to maximise U, by choosing ω where

  $$U = \mu'(\omega - b) - \tfrac{\lambda}{2}(\omega - b)'\Omega(\omega - b);$$

  note that there is also a constraint $(\omega - b)' i = 0$. This framework is widely used in finance, see Grinold and Kahn (1995), Scherer (2002).

24

UNIVERSITY OF CAMBRIDGE

Solution

Our first-order condition is,

$$\frac{\partial U}{\partial(\omega - b)} = \mu - \lambda\Omega(\hat{\omega} - b) + \hat{\theta} i = 0 \quad \text{or}$$

$$\hat{\omega} = b + \tfrac{1}{\lambda}\Omega^{-1}(\mu + \hat{\theta} i).$$

Using $i'(\omega - b) = 0$, we see that $\tfrac{1}{\lambda}(\beta + \hat{\theta}\gamma) = 0$, where

$\beta = i'\Omega^{-1}\mu$ and $\gamma = i'\Omega^{-1}i, \alpha = \mu'\Omega^{-1}\mu.$

25

UNIVERSITY OF CAMBRIDGE

Solution

Thus $\hat{\omega} = b + \tfrac{1}{\lambda}(\Omega^{-1}\mu - \tfrac{\beta}{\gamma}\Omega^{-1} i)$ and hence active returns can be computed as

$$\pi = \mu'(\hat{\omega} - b) = \tfrac{1}{\lambda}\left(\alpha - \frac{\beta^2}{\gamma}\right)$$

$$= \frac{1}{\lambda}\left(\frac{\alpha\gamma - \beta^2}{\gamma}\right) \qquad (1)$$

26

Solution

Other terms of interest can be calculated. For example, we have

$$\hat{\sigma}^2 = \frac{1}{\lambda^2}(\Omega^{-1}\mu - \frac{\beta}{\gamma}\Omega^{-1}\mathrm{i})'\Omega(\Omega^{-1}\mu - \frac{\beta}{\gamma}\Omega^{-1}\mathrm{i})$$

$$= \frac{1}{\lambda^2}(\alpha - \frac{\beta^2}{\gamma} + \frac{\beta^2}{\gamma}) \qquad (2)$$

$$= \frac{1}{\lambda^2}(\alpha - \frac{\beta^2}{\gamma}),$$

27

Solution

we will focus on the $\hat{\sigma}$, tracking error or standard deviation of relative returns. Finally,

$$E(U) = \frac{1}{\lambda}(\frac{\alpha\gamma - \beta^2}{\gamma}) - \frac{\lambda}{2}(\frac{1}{\lambda^2}(\frac{\alpha\gamma - \beta^2}{\gamma})) \qquad (3)$$

$$= \frac{1}{2\lambda}(\frac{\alpha\gamma - \beta^2}{\gamma}).$$

It is straightforward to compute the information ratio defined as $\frac{\alpha}{\hat{\sigma}}$. Notice that in this problem all terms depend essentially on a single term $(\frac{\alpha\gamma - \beta^2}{\gamma})$ or functions of it.

28

An Alternative Problem

A related formulation of the above problem is the following $\min \tfrac{1}{2}(\omega - b)'\Omega(\omega - b)$ subject to $(\omega - b)'i = 0$ and $(\omega - b)'\mu = \pi$. Here the Lagrangean is given by

$$L = \tfrac{1}{2}(\omega - b)'\Omega(\omega - b) - \theta_1(\omega - b)'i - \theta_2((\omega - b)'\mu - \pi)$$

Solving we have $\omega - b = \theta_1\Omega^{-1}i + \theta_2\Omega^{-1}\mu$ with

$$\theta_1 = \frac{\beta\pi}{\alpha\gamma - \beta^2}$$

and

$$\theta_2 = \frac{\gamma\pi}{\alpha\gamma - \beta^2}.$$

29

An Alternative Problem

Thus

$$\hat{\omega} = b + \frac{\pi\gamma}{\alpha\gamma - \beta^2}(\Omega^{-1}\mu - \frac{\beta}{\gamma}\Omega^{-1}i)$$

$$\hat{\omega} = b + \pi\hat{\hat{\omega}}$$

and consequently,

$$\hat{\sigma}^2 = \pi^2\hat{\hat{\omega}}'\Omega\hat{\hat{\omega}}$$

$$\hat{\sigma}^2 = \frac{\pi^2\gamma}{\alpha\gamma - \beta^2}$$

30

UNIVERSITY OF CAMBRIDGE

An Alternative Problem

- Comparing with (2) we see immediately that $\pi = \frac{1}{\lambda}(\alpha - \beta^2/\gamma)$.
- This second problem is simply the computation of the minimum variance frontier. It differs from the earlier version in that it explicitly specifies π, the expected rate of return, rather than λ, the risk aversion coefficient.

31

UNIVERSITY OF CAMBRIDGE

Finite Sample Properties of Estimators of Alpha and Tracking Error

Consider

$$Q = (\mu, \mathrm{i})' \Omega^{-1} (\mu, \mathrm{i})$$

$$= \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix} \qquad (4)$$

It is well known that, under normality $\hat{\mu} \sim \mathrm{N}(\mu, \frac{1}{\mathrm{T}}\Omega)$, and $\hat{\mu}$ and S are independent, where S is the sample covariance matrix. Firstly, by Theorem 3.2.11 of Muirhead (1982), conditional on $\hat{\mu}$.

32

UNIVERSITY OF CAMBRIDGE

Finite Properties of Estimators of Alpha and Tracking Error

$\hat{Q}^{-1} = ((\hat{\mu},i)'\hat{\Omega}^{-1}(\hat{\mu},i))^{-1}$ has a central Wishart: $W_2(T-N+1, \frac{1}{T}\overline{Q}^{-1})$ where $\overline{Q} = (\hat{\mu},i)'\Omega^{-1}(\hat{\mu},i)$. The statistic of interest is given by $\hat{h} = \dfrac{\hat{\gamma}}{\hat{\alpha}\hat{\gamma}-\hat{\beta}^2}$ and is the first principal element of $\hat{Q}^{-1}$. Formally, we have $\hat{h} = (1,0)\hat{Q}^{-1}(1,0)'$ and again from Muirhead (1982) Theorem 3.2.5 we have $\hat{h} \mid \hat{\mu} \sim W_1(T-N+1, \frac{1}{T}(1,0)\overline{Q}^{-1}(1,0)')$ and thus, letting $\varphi = \frac{1}{T}(1,0)\overline{Q}^{-1}(1,0)'$, we have

$$\frac{\hat{h}}{\varphi} \mid \varphi \sim \chi^2_{(\nu)}, \text{ where } \nu = T-N+1$$

and consequently, this result holds unconditionally.

33

UNIVERSITY OF CAMBRIDGE

Finite Properties of Estimators of Alpha and Tracking Error

- Next we examine φ noting immediately that Tφ is the first principal element in $\overline{Q}^{-1}$. That is

$$\varphi = \frac{1}{T}\frac{i'\Omega^{-1}i}{(\hat{\mu}'\Omega^{-1}\hat{\mu})(i'\Omega^{-1}i)-(\hat{\mu}'\Omega^{-1}i)(i'\Omega^{-1}\hat{\mu})}$$

- Now $\hat{\mu} \sim N(\mu, \frac{1}{T}\Omega)$ and thus letting $\omega = \sqrt{T}\Omega^{-1/2}\hat{\mu}$ we have $\omega \sim N(\sqrt{T}\Omega^{-1/2}\mu, I_N)$.

34

Finite Properties of Estimators of Alpha and Tracking Error

- Further, letting $c = \Omega^{-1/2} i$ we have that

$$\varphi = \frac{1}{T}\left(\frac{Tc'c}{\omega'\omega c'c - \omega' cc'\omega}\right)$$

$$= \frac{1}{\omega'(I - c(c'c)^{-1}c')\omega} = \frac{1}{\omega'\overline{P}_c\omega} \qquad (6)$$

- and it follows immediately that $\omega'\overline{P}_c\omega \sim \chi^2_{(N-1,\lambda^*)}$ where $\lambda^* = T\mu'\Omega^{-1/2}\overline{P}_c\Omega^{-1/2}\mu = T/h$ with $h = \left(\frac{\gamma}{\alpha\gamma - \beta^2}\right)$.

35

Finite Properties of Estimators of Alpha and Tracking Error

- Therefore $\varphi^{-1} \sim \chi^2_{(N-1,T/h)}$ and thus the distribution of $\hat{h}$ will be given by the following ratio:

$$\hat{h} \sim \frac{\chi^2_{(\upsilon)}}{\chi^2_{(N-1,T/h)}} \qquad (7)$$

- where the two χ^2 variables in (7) are independent. Thus, $\text{pdf}(\hat{h})$ can be easily found using results related to non-central F distributions.

36

Finite Properties of Estimators of Alpha and Tracking Error

- In this regard we have from Johnson and Kotz (1972 , p. 191) that the $\text{pdf}(1/h = g)$ is, noting $B(\ )$ and ${}_1F_1$ to be Beta and confluent hypergeometric functions respectively,

$$\text{pdf}(g) = \frac{e^{-T/2h} g^{\frac{N-1}{2}-1}}{B\left(\frac{N-1}{2}, \frac{\nu}{2}\right)(1+g)^{\frac{N-1+\nu}{2}}} \, {}_1F_1\left(\frac{N-1+\nu}{2}, \frac{N-1}{2}, \frac{T}{2h}\frac{g}{(1+g)}\right)$$

37

Probability Functions

Using this result and simple transformations one can readily derive the pdf density functions for the quantities of interest, viz, $\hat{\alpha} = 1/\lambda\hat{h}$. Tracking Error $= \overline{TE} = 1/\lambda\sqrt{\hat{h}}$ and the Information Ratio $= \overline{IR} = 1/\sqrt{\hat{h}}$. Thus we have

$$\text{pdf}(\hat{\pi} = \omega) = \frac{\lambda e^{-T/2h} (\lambda\omega)^{\frac{N-1}{2}-1}}{B\left(\frac{N-1}{2}, \frac{\nu}{2}\right)(1+\lambda\omega)^{\frac{N-1+\nu}{2}}} \, {}_1F_1\left(\frac{N-1+\nu}{2}, \frac{N-1}{2}, \frac{T}{2h}\frac{\lambda\omega}{(1+\lambda\omega)}\right), \quad \omega > 0$$

38

UNIVERSITY OF CAMBRIDGE

Probability Functions

$$\text{pdf}(\overline{IR}=x)=\frac{2xe^{-T/2h}(x^2)^{\frac{N-1}{2}-1}}{B(\frac{N-1}{2},\frac{\nu}{2})(1+x^2)^{\frac{N-1+\nu}{2}}}\,{}_1F_1\left(\frac{N-1+\nu}{2},\frac{N-1}{2},\frac{T}{2h}\left(\frac{x^2}{(1+x^2)}\right)\right),\quad x>0$$

$$\text{pdf}(\overline{TE}=y)=\frac{2\lambda^2 ye^{-T/2h}(\lambda^2y^2)^{\frac{N-1}{2}-1}}{B(\frac{N-1}{2},\frac{\nu}{2})(1+\lambda^2y^2)^{\frac{N-1+\nu}{2}}}\,{}_1F_1\left(\frac{N-1+\nu}{2},\frac{N-1}{2},\frac{T}{2h}\left(\frac{\lambda^2y^2}{(1+\lambda^2y^2)}\right)\right),\quad y>0$$

39

UNIVERSITY OF CAMBRIDGE

Moments of our Estimators

From these pdfs or via that of $\hat{h}$ or g, we can easily find moments. That is, since

$$\hat{h}^{-1}=g=\chi^2_{(N-1,T/h)}/\chi^2_{(\nu)}$$

$$E(g^k)=E[(\chi^2_{(N-1,T/h)})^k]E[\chi^2_{(\nu)})^{-k}]$$

and since

$$E[(\chi^2_{(N-1,T/h)})^k]=\frac{e^{-T/2h}2^k}{\Gamma\left(\frac{N-1}{2}\right)}\Gamma\left(\frac{N-1}{2}+k\right){}_1F_1\left(\frac{N-1}{2}+k,\frac{N-1}{2};\tfrac{T}{2h}\right)$$

40

UNIVERSITY OF CAMBRIDGE

Moments of Our Estimators

and

$$E((\chi^2_{(\nu)})^{-k}) = \frac{\Gamma(\frac{\nu}{2}-k)}{2^k\,\Gamma(\nu/2)},\quad \nu > 2k$$

$$E[g^k] = \frac{e^{-T/2h}\,\Gamma\left(\frac{N-1}{2}+k\right)\Gamma\left(\frac{\nu}{2}-k\right)}{\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\nu/2\right)}\;{}_1F_1\left(\frac{N-1}{2}+k,\frac{N-1}{2};T/2h\right)$$

41

UNIVERSITY OF CAMBRIDGE

Moments of our Estimators

$$E[(\hat{\pi})^k] = \lambda^{-k}E(g^k)$$

$$= \frac{e^{-T/2h}\,\Gamma\left(\frac{N-1}{2}+k\right)\Gamma\left(\frac{\nu}{2}-k\right)2^k}{\lambda^k\,\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\nu/2\right)}\;{}_1F_1\left(\frac{N-1}{2}+k,\frac{N-1}{2};T/2h\right) \quad (8)$$

$$E((\overline{TE})^k = \lambda^{-k}E(g^{k/2})$$

$$= \frac{e^{-T/2h}\,\Gamma\left(\frac{N-1}{2}+\frac{k}{2}\right)\Gamma\left(\frac{\nu}{2}-\frac{k}{2}\right)2^k}{\lambda^k\,\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\nu/2\right)}\;{}_1F_1\left(\frac{N-1}{2}+\frac{k}{2},\frac{N-1}{2};T/2h\right) \quad (9)$$

$$E((\overline{IR})^k) = E(g^{k/2}) \quad (10)$$

In particular, if we consider the means of the three quantities we have:

42

UNIVERSITY OF CAMBRIDGE

Unbiased Estimators

$$E(\hat{\pi}) = \frac{\Gamma\left(\frac{N-1}{2}+1\right)\Gamma\left(\frac{\nu-1}{2}-1\right)}{\lambda\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{\nu}{2}\right)} e^{-T/2h}\ {}_1F_1\left(\frac{N-1}{2}+1, \frac{N-1}{2}; T/2h\right)$$

$$= \frac{(N-1)}{\lambda(\nu-2)}\ {}_1F_1(-1, \frac{N-1}{2}; -T/2h)$$

$$= \frac{N-1}{\lambda(\nu-2)}\left[1 + \frac{T}{h(N-1)}\right]$$

$$\underset{\pi}{=} \frac{N-1}{\lambda(\nu-2)} + \frac{T}{\lambda h(\nu-2)}; \qquad \nu = T - N + 1.$$

Since the true $\pi = 1/\lambda h$ we can readily develop an unbiased estimator of π via a simple transformation:

43

UNIVERSITY OF CAMBRIDGE

Unbiased Estimators

$$E\left[\frac{(\nu - 2)}{T}\hat{\pi} - \frac{N-1}{T\lambda}\right] = \pi$$

$$E(\overline{TE}) = \frac{\Gamma\left(\frac{N-1}{2}+\frac{1}{2}\right)\Gamma\left(\frac{\nu}{2}-\frac{1}{2}\right)}{\lambda\Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\nu/2\right)}\ {}_1F_1\left(-\frac{1}{2}, \frac{N-1}{2}; -T/2h\right)$$

and $E(\overline{IR}) = \lambda E(\overline{TE})$.

Also, note that since $E(\hat{\sigma}^2) = E(\overline{TE}^2) = \frac{1}{\lambda}E(\hat{\alpha})$ and an unbiased estimator of σ^2 is easily derived to be

$$\hat{\sigma}^2 = \frac{1}{\lambda}\left[\frac{\nu-2}{T}\hat{\pi} - \frac{(N-1)}{T\lambda}\right].$$

44

UNIVERSITY OF CAMBRIDGE

Unbiased Estimators

While little progress can be made with exact expressions for the expectation of $\overline{\text{TE}}$ and $\overline{\text{IR}}$, we can get more insight by considering approximations.

45

UNIVERSITY OF CAMBRIDGE

Sources of Bias

- We now examine a situation in which both N, the number of stocks or assets and T. The sample size, increase in such a way that the ratio $\frac{N-1}{T}$ remains constant. Thus we now let $\frac{N-1}{T} = n$ so that $N-1 = T.n$. By letting $T \to \infty$ we can readily see the effect on the moments of large N and T. For $\hat{\pi}$ we have from our exact result

$$E(\hat{\pi}) = \frac{Tn}{\lambda(T(1-n)-2)}\left[1 + \frac{1}{hn}\right]$$

46

Sources of Bias

and therefore as $T \to \infty$ we find

$$E(\hat{\pi}) \to \frac{1}{\lambda}\left[\frac{n}{1-n} + \frac{1}{h(1-n)}\right].$$

The corresponding results for $\overline{IR}$ and $\overline{TE}$

$$E(\overline{IR}) \to \left(\frac{n}{1-n} + \frac{1}{h(1-n)}\right)^{1/2}$$

and

$$E(\overline{TE}) \to \frac{1}{\lambda}\left(\frac{n}{1-n} + \frac{1}{h(1-n)}\right).$$

47

Magnitude of Bias

We now illustrate the accuracy of these approximations using two contrasting numerical examples. In both cases we have $\lambda = 12.5$ and $h = 4$, giving true values of $\pi = 0.02$, $\overline{TE} = 0.04$ and $\overline{IR} = 0.5$.

i) T = 180, N = 4, so that n = 1/60 = 0.01667

ii) T = 180, N = 80, so that n = 79/180 = 0.43889.

48

UNIVERSITY OF CAMBRIDGE

Magnitude of Bias

The results are given in the following table.

	$E\hat{\pi}$	$E(\overline{TE})$	$E(\overline{IR})$
The values	0.02	0.04	0.05
Case I (T = 180, N = 4)			
Exact	0.021726	0.041410	0.517621
Approx	0.021695	0.041660	0.520756
Case II (T = 180, N = 80)			
Exact	0.100202	0.089062	1.113272
Approx	0.098218	0.088642	1.108027

49

UNIVERSITY OF CAMBRIDGE

Magnitude of Bias

- What the above results illustrate is the fact that while the estimators $\hat{\pi}$, $\overline{TE}$, $\overline{IR}$ are always biased, the bias is very small when n is small. However, for large n the bias is extremely large being more than four times the true value for π and greater than twice the true value for $\overline{TE}$ and $\overline{IR}$. We also notice that in both cases the approximation is quite accurate.

50

UNIVERSITY OF CAMBRIDGE

Comparing Numerical Results

- Keeping our numerical results consistent with those in Scherer (2002, p. 165) we will consider two cases

i) $$Q = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} = \begin{pmatrix} 0.125 & 6.5 \\ 6.5 & 675.00 \end{pmatrix}$$

$$h = \frac{\gamma}{\alpha\gamma - \beta^2} \approx 16$$

In each case, by choosing different values for λ (risk aversion parameter) we can generate a wide set of values for both active returns $\pi = \mu'\omega = \frac{1}{\lambda h}$ and tracking error $Te = \frac{1}{\lambda\sqrt{h}}$.

51

UNIVERSITY OF CAMBRIDGE

Comparing Numerical Results

The following table highlights this relationship.

Some authors such as Grinold and Kahn (1999) express the units associated with the active return, π and the tracking error, Te, in terms of percent. Others, e.g. Scherer (2002), use the decimal equivalent. However, shifting the units from decimal to percent will alter λ, the risk aversion parameter, by a factor of 100. That is, the λ associated with percent units will be 100^{th} of the value of λ associated with decimal units. Thus, the following constellations of parameter values are consistent:

52

UNIVERSITY OF CAMBRIDGE

Table of Values

h = 16				h = 4			
λ	π	Te	Ir	λ	π	Te	Ir
2	0.03125	0.125	0.25	12.5	0.02	0.04	0.5
0.02	3.125	12.5	0.25	0.125	2	4	0.5

In what follows we choose the decimal representation.

h = 16				h = 4		
λ	π	Te	Ir	π	Te	Ir
2	0.03125	0.125	0.25	0.125	0.25	0.5
4	0.0156	0.0625	0.25	0.0625	0.125	0.5
6	0.0104	0.04167	0.25	0.04167	0.0833	0.5
8	0.0078	0.03125	0.25	0.03125	0.0625	0.5
12.5				0.02	0.04	0.5

53

UNIVERSITY OF CAMBRIDGE

Bias Magnitudes for Realistic Parameter Values

We now examine the tracking error optimization and the performance, i.e. relative bias. of the standard estimators for different portfolio sizes, N = 4 and N = 80 with T = 180 in both cases.

	h = 16, λ = 2			h = 4, λ = 12.5		
θ:	π= 0.03125	Te = 0.125	Ir = 0.25	π = 0.02	Te = 0.04	Ir = 0.5
N=4 $E(\hat{\theta})$	0.0407	0.1378	0.2757	0.0217	0.0414	0.5176
% Rel. Bias	30.24	10.24	10.28	8.50	3.50	3.52
N=80$E(\hat{\theta})$	0.4558	0.4747	0.9494	0.1002	0.0891	1.1133
% Rel. Bias	1358.56	279.76	279.76	401.00	122.75	122.66

54

UNIVERSITY OF CAMBRIDGE

An Aside on Efficient Frontier

We now examine portfolio optimization without a benchmark. Here we maximize $\omega'\mu - \frac{\lambda}{2}\omega'\Omega\omega$ subject to $\omega'\mathrm{i} = 1$. The associated Lagragian is given by:

$$W = \omega'\mu - \frac{\lambda}{2}\omega'\Omega\omega - \theta(\omega'\mathrm{i} - 1)$$

with $\frac{\partial W}{\partial \omega} = \mu - \lambda\Omega\omega - \theta \quad \mathrm{i} = 0$

implying $\omega = \frac{1}{\lambda}(\Omega^{-1}\mu - \theta\Omega^{-1}\mathrm{i})$

since $\mathrm{i}'\omega = 1$ we have immediately that $\theta = (\mathrm{i}'\Omega^{-1}\mu - \lambda)/\mathrm{i}'\Omega^{-1}\mathrm{i}$

i.e. $\theta = (\beta - \lambda)/\gamma$.

55

UNIVERSITY OF CAMBRIDGE

An Aside on Efficient Frontier

Consequently,

$$\hat{\omega} = \frac{1}{\lambda}(\hat{\Omega}^{-1}\hat{\mu} - 1(\frac{\hat{\beta} - \hat{\lambda}}{\hat{\gamma}})\hat{\Omega}^{-1}\mathrm{i})$$

and thus

$$\tilde{\pi} = \hat{\mu}'\hat{\omega} = \frac{1}{\lambda}\hat{\mu}'\hat{\Omega}^{-1}\hat{\mu} - (\frac{\hat{\beta} - \hat{\lambda}}{\hat{\gamma}})\hat{\mu}'\hat{\Omega}^{-1}\mathrm{i}$$

i.e. $\tilde{\pi} = \frac{1}{\lambda\hat{\mathrm{h}}} + \frac{\hat{\beta}}{\hat{\gamma}} = +\tilde{\pi}(\frac{\hat{\beta}}{\hat{\gamma}})$.

Also, $\tilde{\sigma}^2 = \hat{\omega}'\hat{\Omega}\hat{\omega}$

$$\tilde{\sigma}^2 = \frac{1}{\lambda^2\hat{\mathrm{h}}} + \frac{1}{\hat{\gamma}}.$$

56

UNIVERSITY OF CAMBRIDGE

An Aside (II)

Thus we notice immediately that the active return $\tilde{\pi}$ and the $\hat{T}e^2$ are given by our earlier results plus an additional term. Under the normality assumption we again examine some of the statistical properties of these new estimations. We present the results below with the proofs given in the appendix.

57

UNIVERSITY OF CAMBRIDGE

As Aside (II)

Theorem 2

$$E(\tilde{\pi}) = E(\tilde{\pi}) + E(\hat{\beta}/\hat{\gamma})$$

$$= E(\tilde{\pi}) + (\beta/\gamma).$$

$$\mathrm{Var}(\tilde{\pi}) = \mathrm{Var}(\tilde{\pi}) + \mathrm{Var}(\hat{\beta}/\hat{\gamma}), \textit{ since } \tilde{\pi} \textit{ and } \hat{\beta}/\hat{\gamma} \textit{ are independent}$$

$$= \mathrm{var}(\tilde{\pi}) + \frac{1}{T\gamma}(E(\hat{h}^{-1}) + 1)$$

$$E(\tilde{\sigma}^2) = E(\tilde{\sigma}^2) + \frac{T-N+1}{T\gamma}.$$

58

General Linear Restrictions

- The results of the previous section can be readily extended to incorporate general linear restrictions on the relative weights. Here we briefly outline the results giving the full derivation in the Appendix. We now consider the maximization of utility subject to a set of K restrictions: $R(\omega - b) = 0$, where R is a $K \times N$ matrix. The Lagragian and the associated first order conditions are as follows:

59

Solution

$$L = \mu'(\omega - b) - \frac{\lambda}{2}(\omega - b)'\Omega(\omega - b) + \theta' R(\omega - b)$$

$$\frac{\partial L}{\partial \omega} = \mu - \lambda\Omega(\omega - b) + R'\theta = 0$$

$$\frac{\partial L}{\partial \theta} = R(\omega - b) = 0$$

60

UNIVERSITY OF CAMBRIDGE

Solution

- Solving we find $\omega = b + \frac{1}{\lambda}\Omega^{-1}(\mu - R'(R\Omega^{-1}R')R\Omega^{-1}\mu)$

resulting in

$$\pi = \mu'(\omega - b) = \frac{1}{\lambda}\left[\mu'\Omega^{-1}\mu - \mu'\Omega^{-1}R'(R\Omega^{-1}R')^{-1}R\Omega^{-1}\mu\right]$$

$$= \frac{1}{\lambda}(1,0)\hat{Q}_k^{-1}(1,0)'$$

61

UNIVERSITY OF CAMBRIDGE

Solution

- and $\sigma^2 = (\omega - b)'\Omega(\omega - b)$

$$= \frac{1}{\lambda^2}(1,0)\hat{Q}_k^{-1}(1,0)'$$

- where $\hat{Q}_k = (\hat{\mu}, R')'\hat{\Omega}^{-1}(\hat{\mu}, R')$
- with $\hat{Q}_k^{-1} \sim W_{k+1}(T - N + k, \frac{1}{T}\hat{Q}_k^{-1})$.

62

UNIVERSITY OF CAMBRIDGE

Solution 2

- Following earlier results we now define

  $\hat{h}_k = (1,0)\ \hat{Q}_k^{-1}\ (1,0)'$ and we have immediately, corresponding to (7):

$$\hat{h}_k \sim \frac{\chi^2_{(T-N+k)}}{\chi^2_{(N-k,\,T/h_k)}}$$

where $h_k = (1,0)'Q_k^{-1}(1,0)$ with $Q_k = (\mu,R')'\Omega^{-1}(\mu,R')$

63

UNIVERSITY OF CAMBRIDGE

Solution 2

- Thus, by a simple substitution into our earlier results we can readily specify the exact distribution and moments of $\hat{\pi}$, $\overline{\text{IR}}$ and $\overline{\text{TE}}$. That is, we merely replace N – 1 by N – K and ~ by

T – (N – K).

64

UNIVERSITY OF CAMBRIDGE

Bias

- In the previous section with only one restriction we noticed that the fixed T and small N we have very small bias in our estimators. However, for large N relative to T the bias was very significant. In the general, K, restrictions case what we find is that as K increases the bias decreases. This is best illustrated with $\hat{\pi}$.
- Here

$$E(\hat{\pi}) = \frac{((N-k)h_k + T)}{(T-N+k)\lambda h_k}$$

and the following graph shows the effect on the relative bias of increasing k.

65

UNIVERSITY OF CAMBRIDGE

Simulating the Frontiers

- We now consider the standard mean variance optimal portfolio problem and use our earlier exact results to develop an efficient simulation algorithm for the frontier and 95% confidence intervals. Let ω we a (N×1) vector of portfolio weights. Returns are distributed as N (μ, Ω), I is a (N×1) vector of ones and π is a given level of returns.
- Here the problem is to minimize $\omega'\Omega\omega$ subject to:

$\mu'\omega = \pi$ and $i'\omega = 1$, thus the Lagrangean is given by:

$$L = \tfrac{1}{2}\omega'\Omega\omega - \theta_1(\mu'\omega - \pi) - \theta_2(i'\omega - 1)$$

66

UNIVERSITY OF CAMBRIDGE

Simulating the Frontiers

- with

$$\frac{\partial L}{\partial \omega} = \Omega\omega - \theta_1\mu - \theta_2 i = 0$$

$$\frac{\partial L}{\partial \theta_1} = \mu'\omega - \pi = 0$$

$$\frac{\partial L}{\partial \theta_2} = i'\omega - 1 = 0$$

67

UNIVERSITY OF CAMBRIDGE

Simulating 2

- Consequently, $\omega = \Omega^{-1}(\mu, i)\begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix}$
- where $\begin{bmatrix}\mu'\Omega^{-1}\mu & \mu'\Omega^{-1}i \\ i'\Omega^{-1}\mu & i'\Omega^{-1}i\end{bmatrix}\begin{bmatrix}\theta_1 \\ \theta_2\end{bmatrix} = \begin{bmatrix}\pi \\ 1\end{bmatrix}$
- That is, $\hat{\omega} = \hat{\Omega}^{-1}(\hat{\mu}, i)\hat{Q}^{-1}\begin{bmatrix}\pi \\ 1\end{bmatrix}$,
- and therefore $\sigma^2 = \hat{\omega}'\hat{\Omega}\hat{\omega} = (\pi\, 1)'\hat{Q}^{-1}\begin{bmatrix}\pi \\ 1\end{bmatrix}$

$$= p'\hat{Q}^{-1}p, \qquad \text{where } p = \begin{bmatrix}\pi \\ 1\end{bmatrix}.$$

68

UNIVERSITY OF CAMBRIDGE

Simulating 2

- From our earlier results $p'\hat{Q}^{-1}p \mid \hat{\mu} \sim W_1(T-N+1, \frac{1}{T}p\overline{Q}^{-1}p)$

  and therefore where W_1 (.) is a Wishart of dimension 1.
  Where $\overline{Q} = (\hat{\mu}, i)'\Omega^{-1}(\hat{\mu}, i)$, and therefore, $p'\hat{Q}^{-1}p \sim \chi^2_{(T-N+1)} \cdot \Psi$
  where χ^2_m is a chi- squared with m degree of freedom.

- Where $\Psi = \frac{1}{T}p'\overline{Q}^{-1}p.$

69

UNIVERSITY OF CAMBRIDGE

Simulating 3

- Next, letting

$$\overline{Q}^{-1} = \begin{bmatrix} \overline{a} & \overline{b} \\ \overline{b} & \overline{c} \end{bmatrix}$$

we have $\Psi = \frac{1}{T}(\overline{a} - \overline{b}^2/\overline{c})^{-1}(\overline{b}/\overline{c} - \pi)^2 + \frac{1}{T\overline{c}}$

where $TE = \frac{1}{\lambda}$ and via a transformation

$u = \sqrt{T}\Omega^{-\frac{1}{2}}\hat{\mu}$ and $\alpha = \Omega^{-\frac{1}{2}}i$,

70

Simulating 3

- we get $T(\bar{a} - \bar{b}^2/\bar{c}) \sim \chi'^2_{(N-1, T(a-b^2/c))}$
- and $\bar{b}/\bar{c} \sim N(b/c, 1/Tc)$.
- Therefore $\sigma^2 = p'\hat{Q}^{-1}p = k_1\left[(k_2 - \pi)^2/k_3 + \frac{1}{T\bar{c}}\right]$
- where the three random variables

  $k_1 \sim \chi^2_{(T-k+1)}$, $k_2 \sim N(b/c, 1/T\bar{c})$

  and $k_3 \sim \chi'^2_{(N-1, T(a-b^2/c))}$ are mutually independent.

71

Simulating 4

- We now have a straightforward method to generate, via simulation, the mean–variance frontier and confidence intervals.
- For specified $\hat{a}$, $\hat{b}$ and $\hat{c}$ along with T and N, and taking 5000 replications as an example.
  1. Generate 5000 observations on the three independent random variables, k_1, k_2 and k_3 i.e. k_{ij}, k_{2j}, k_{3j}, j = 1,…,5000.
  2. Select 201 values of π centred on $\hat{b}/\hat{c}$ i.e. choose an interval $\hat{b}/\hat{c}$ i.e. choose an interval $\hat{b}/\hat{c} \pm 3(\hat{b}/\hat{c})$ with 100 points above $\hat{b}/\hat{c}$ and 100 below $\hat{b}/\hat{c}$

72

UNIVERSITY OF CAMBRIDGE

Simulating 4

3. Now for each value of π_ℓ, $\ell = 1,\ldots,201$ we can use

$$\sigma_j = \sqrt{k_{1j}\left(\frac{(R_{2j}-\pi_\ell)^2}{R_{3j}} + \frac{1}{T\hat{c}}\right)}_{\sigma_{j\cdot}}$$

to generate 5000 values of σ_j.

73

UNIVERSITY OF CAMBRIDGE

Simulating 5

4. From the 5000 values of σ_j for each π_ℓ we can calculate the mean and the 2.5 and 97.5 percentile i.e. $\bar{\sigma}_\ell, \sigma_\ell^L$ and $\bar{\sigma}_\ell^u$ respectively.
5. Now plot the pairs of points: $(\bar{\sigma}_\ell, \pi_\ell)$, $(\sigma_\ell^L, \pi_\ell)$ and $(\sigma_\ell^u, \pi_\ell)$ for $\ell = 1,\ldots,201$ giving the average-mean variance frontier and the 95% confidence limits.

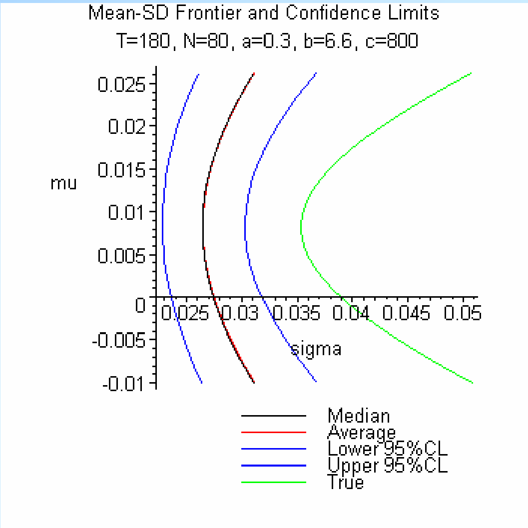
From the discussion in Section 5 of Knight and Satchell it is clear that a similar algorithm to that above can be used to generate the frontier in a situation with general linear restrictions. However, a simple algorithm for inequality constraints based on our approach does not seem easily attainable.

74

Conclusion

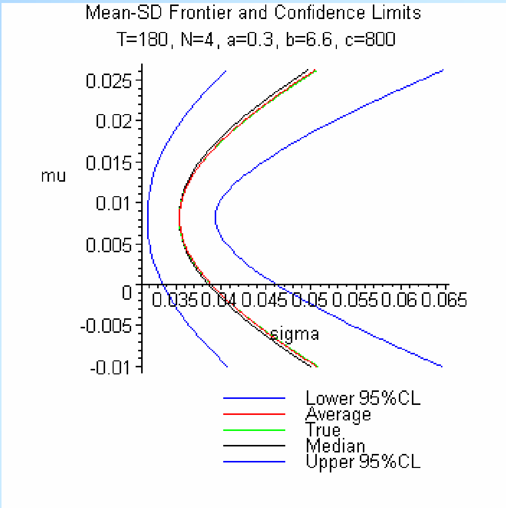
- Using multivariate analysis one can analyse the properties of simulation-based portfolio optimisers.
- These optimisers will inevitably lead to a dispersion, which is useful, about a bias, which is not.
- The bias is especially bad when the number of stocks is near the number of observations.
- Constraining the weights is one solution. If we add k constraints, then effectively we get N – k stocks; this is what practitioners do.
- Theory confirms the wisdom of practitioners in this respect.

75



76

UNIVERSITY OF CAMBRIDGE



77

UNIVERSITY OF CAMBRIDGE

Bibliography

Britten-Jones, M. (1999) The Sampling Error in Estimates of Mean-Variance Efficient Portfolio Weights, *Journal of Finance* 54, 655-671.

Grinold, R.C. and R.N. Kahn (1999) *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Selecting Superior Returns and Controlling Risk*, McGraw Hill Inc.

Harvey, C., J.C. Leichty, M.W. Leichty, P. Muller, (2004) *Portfolio Selection with Higher Moments*

Hillier, G. and S. Satchell (2003) Some Exact Results for Efficient Portfolios with Given Returns, *New Advances in Portfolio Construction and Implementation*, Butterworth-Heinemann, London.

Jobson, J.D. (1990)

Jobson, J.D. and R. Korkie (1989). A performance interpretation of multivariate tests of asset set intersection, spanning, and mean-variance efficiency. *Journal of Financial and Quantitative Analysis* 24, 185-204.

Johnson, N.L. and S.C. Kotz (1972) *Distributions in Statistics: Continuous Multivariate Distributions*, John Wiley and Sons. Inc.

Michaud, R. (1998) *Efficient Asst Management* (Harvard Business School Press).

Muirhead, R.J. (1982) *Aspects of Multivariate Statistical Theory*, John Wiley and Sons Inc.

Scherer, B. (2002) *Portfolio Construction and Risk Budgeting*, Risk Books.

78