Portfolio Selection using Random Matrix Theory and L-Moments

Wardah Ushan

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

November 6, 2015

MPhil in Mathematical Finance,
University of Cape Town.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy at the University of Cape Town. It has not been submitted before for any degree or examination to any other University.

Wardah Ushan

November 6, 2015

Abstract

Markowitz's (1952) seminal work on Modern Portfolio Theory (MPT) describes a methodology to construct an optimal portfolio of risky stocks. The constructed portfolio is based on a trade-off between risk and reward, and will depend on the risk-return preferences of the investor. Implementation of MPT requires estimation of the expected returns and variances of each of the stocks, and the associated covariances between them. Historically, the sample mean vector and variance-covariance matrix have been used for this purpose. However, estimation errors result in the optimised portfolios performing poorly out-of-sample.

This dissertation considers two approaches to obtaining a more robust estimate of the variance-covariance matrix. The first is Random Matrix Theory (RMT), which compares the eigenvalues of an empirical correlation matrix to those generated from a correlation matrix of purely random returns. Eigenvalues of the random correlation matrix follow the Marčenko-Pastur density, and lie within an upper and lower bound. This range is referred to as the "noise band". Eigenvalues of the empirical correlation matrix falling within the "noise band" are considered to provide no useful information. Thus, RMT proposes that they be filtered out to obtain a cleaned, robust estimate of the correlation and covariance matrices.

The second approach uses L-moments, rather than conventional sample moments, to estimate the covariance and correlation matrices. L-moment estimates are more robust to outliers than conventional sample moments, in particular, when sample sizes are small.

We use L-moments in conjunction with Random Matrix Theory to construct the minimum variance portfolio. In particular, we consider four strategies corresponding to the four different estimates of the covariance matrix: the L-moments estimate and sample moments estimate, each with and without the incorporation of RMT. We then analyse the performance of each of these strategies in terms of their risk-return characteristics, their performance and their diversification.

Acknowledgements

I would like to thank my supervisor, Petrus Bosman, and his colleagues Dan Golding and Byran Taljaard, of Avior Capital Markets, for their keen insight and recommendations. I am grateful to my co-supervisor, Prof David Taylor, for not only assisting in the writing of this dissertation, but for facilitating most of my education this past year. Thanks to all my lecturers, tutors and classmates for making 2014 a most memorable year.

Special thanks to all my friends, especially to Bilal Mirza, Rushikesh Alur and the rest of the "Twin Oaks Crew" for their unconditional friendship and support in the most difficult times. I am indebted to Zakir Hossen for his unending mentorship and advice.

Finally, I cannot go without expressing immense gratitude and appreciation to my parents – Ali Akbar Mollah and Anowara Begum – both of whom have dedicated their lives to improving mine. The completion of this dissertation and Masters degree would not have been possible without their endless love and support.

Contents

1. Introduction 1

2. Robust estimation of covariance matrices 6
- 2.1 L-Moments 7
  - 2.1.1 Univariate L-moments 8
  - 2.1.2 Bivariate L-moments 9
  - 2.1.3 L-Moments variance-covariance and correlation matrices . . . 11
- 2.2 Random Matrix Theory 12
  - 2.2.1 Mathematical framework of RMT 12
  - 2.2.2 Filtering the correlation matrix 14
- 2.3 Other approaches to obtaining robust estimates of the covariance matrix 16

3. Data and methodology 17
- 3.1 Data 17
  - 3.1.1 South African Data 17
  - 3.1.2 US Data 19
- 3.2 Methodology for the construction of minimum variance portfolios . . 19

4. Empirical results and analysis 21
- 4.1 Rationale for using RMT in finance 21
- 4.2 Robust Risk Estimation using RMT 29
- 4.3 Global Minimum Variance Portfolio Optimisation 32
  - 4.3.1 Portfolio Performance 32
  - 4.3.2 Comparison of SA results to US Data 43

5. Conclusion 45
- 5.1 Potential further research 46

Bibliography 47

A. Yanou's (2013) US database of stocks 51

B. Additional plots of RMT analysis on efficient frontier 52

List of Figures

2.1 Marčenko-Pastur densitites 14

4.1 Histogram of the empirical eigenvalue spectrum of the sample and L-correlation matrices 22
4.2 Distribution of the eigenvector components using the sample correlation matrix 26
4.3 Distribution of the eigenvector components using the L-correlation matrix 27
4.4 Scatterplot of returns on the hypothetical "market" portfolio against those on the Top40 TR Index 28
4.5 Efficient frontiers displaying in-sample and out-of-sample risk (1) . . 30
4.6 Evolution of the portfolio value of the four different strategies. . . . 37
4.7 Rolling risk of the four different strategies. 38
4.8 Rolling return of the four different strategies. 38
4.9 Diversification of the four different strategies, as measured by the Effective Number of Bets. 39
4.10 Diversification of the four different strategies, as measured by the number of stocks that are invested in, out of those that are available, i.e. Effective Size. 39
4.11 Active Share of three of the strategies, measured relative to the conventional sample moments method. 40
4.12 Active share of the L-moments strategy along with the Effective Size/proportion of stocks that are invested in. 41
4.13 Sorted active share and sorted Effective Size of the L-moments strategy. 41
4.14 Active share of the L-moments strategy along with the Effective Number of Bets. 42
4.15 Sorted active share and sorted Effective Number of Bets of the L-moments strategy. 42

B.1 Efficient frontiers displaying in-sample and out-of-sample risk (2) . . 53
B.2 Efficient frontiers displaying in-sample and out-of-sample risk (3) . . 54

List of Tables

3.1 Data subsets extracted from South African data set 18

4.1 Kolmogorov-Smirnov test on eigenvectors 24
4.2 Portfolio performance on weekly data and 4 week rolling period . . . 33
4.3 Portfolio performance on weekly data and 2 week rolling period . . . 33
4.4 Portfolio performance on weekly data and 1 week rolling period . . . 33
4.5 Portfolio performance on daily data and 4 week rolling period 34
4.6 Portfolio performance on daily data and 2 week rolling period 34
4.7 Portfolio performance on daily data and 1 week rolling period 34
4.8 Portfolio performance on US data 44

Chapter 1

Introduction

Markowitz's (1952) seminal work on Modern Portfolio Theory (MPT) describes a methodology to construct an optimal portfolio of risky stocks for an investor who is concerned only with the means and variances of stock returns. The constructed portfolio is based on a trade-off between risk and reward, and will depend on the risk-return preferences of the investor. Implementation of MPT requires the estimation of the expected returns and variances of each of the stocks, and the associated covariances between them. Historically, the sample mean vector and variance-covariance matrix have been used for this purpose.

However, it is well known that it is more difficult to estimate the expected returns than the variances and covariances between the returns of the stocks (Merton, 1980). Furthermore, estimation errors in the expected returns have a greater impact on portfolio weights than estimation errors in the variance-covariance matrix (Chopra and Ziemba, 1993). These estimation errors result in volatile portfolio weights and poor out-of-sample performance of the optimised portfolio.

As a result, recent research has focused on the Global Minimum Variance (GMV) portfolio. This is the portfolio on the efficient frontier corresponding to the lowest possible variance across all efficient portfolios. In order to construct the optimal weights of a GMV portfolio, one needs to estimate only the variance-covariance matrix; estimation of the expected returns is not required. This technique therefore possesses the advantage of being less prone to estimation errors.

There has been empirical evidence to show that GMV portfolios – and more generally, low-volatility portfolios – outperform both mean-variance portfolios and other strategies with higher volatilities. Haugen and Heins (1975, 1972) find little support for the widespread notion that taking on greater risk (systematic, or otherwise) manifests in higher returns. Their findings indicate that, over the long run, portfolios with lower average volatilities tend to produce greater average returns than portfolios exhibiting higher volatilities. Baker and Haugen (2012) go on to show that this anomaly of out-performance of low risk stocks persists in several equity markets around the world – including both developed and emerging economies. Similarly, Ang *et al.* (2009, 2006) have shown that stocks with high idiosyncratic volatilities have low average returns. They go on to state that this effect persists both within the US and internationally, and has surfaced in periods of expansions, recessions, volatility, and stability, and is consistent with different holding periods. In respect of the minimum variance portfolio itself, Jagannathan and Ma (2003) have shown that this portfolio has lower risk and higher returns than a market capitalisation weighted benchmark. Similarly, de Carvalho *et al.* (2012) have shown that the minimum variance portfolio provides the highest Sharpe ratio when compared to alternative mean-variance or capitalisation weighted strategies.

There are several explanations of the "low-volatility" phenomenon. These explanations are behavioural, following from the actions taken by investors. For example, some investors, particularly those that have leverage or margin constraints, display increased demand for high-beta assets to compensate for the constraints (Frazzini and Pedersen, 2014). Similarly, investors exhibit lottery-type behaviour towards high volatility stocks, being willing to pay a premium for the potential high returns, but at the cost of being exposed to lower returns (Baker *et al.*, 2010). Hsu *et al.* (2012) state that analysts display optimism towards volatile stocks and tend to produce high growth forecasts for such stocks, and consequently display high demand. Furthermore, Baker and Haugen (2012) argue that investment managers themselves are usually compensated by a base salary, with an additional bonus on out-performance of funds, while not being penalised for under-performance. This upside-only exposure creates excess demand for excessively volatile stocks from the investment manager. The authors argue that these activities drive up the prices of high-volatility stocks, hence eroding the returns on them. This results in the out-performance of lower-volatility stocks, explaining the "low-volatility" anomaly.

It is imperative to note that there also exists literature that explains away these empirical anomalies. In particular, Bali and Cakici (2008) have shown that these findings lack robustness when the data frequency is altered, and do not hold when illiquid stocks are screened out of the optimisation process. Furthermore, Amenc *et al.* (2011) demonstrate that the holding period has an influence on the risk-return relationship, and show that the "low-volatility" anomaly no longer exists for holding periods of greater than a year.

Despite the evidence in support of the out-performance of low-volatility portfolios, there is a concern associated with investing in the GMV portfolio. Intuitively, one would expect a portfolio that gives the lowest possible variance to be concentrated in those stocks that correspondingly have the lowest volatilities. It can be argued that investing in the GMV portfolio leads to portfolios that are poorly diversified, as the portfolios tend to be exposed to sector-specific or stock-specific risks. In particular, Clarke *et al.* (2011) note that their minimum variance portfolio includes only 12% of the stocks in their investment universe. Stoyanov and Goltz (2011) explain that this occurs because the portfolio weights for the minimum variance portfolio are selected to ensure the lowest possible variance, while not intelligently exploiting correlation properties to enhance returns. Correlation properties of stocks are exploited only to the extent that they aid the goal of variance reduction.

Nonetheless, by reason of fewer estimation errors and the empirical evidence of out-performance of low-volatility portfolios, the Global Minimum Variance Portfolio has gained popularity in the years following the 2007-2008 global financial crisis. In particular, MSCI created the MSCI Global Minimum Volatility Indices, and in 2011 S&P followed by launching a series of low-volatility indices whose constituents have stock weights that are inversely proportional to their historical volatilities (Hsu and Li, 2013).

Historically, the sample covariance matrix has been used as the empirical estimate of variances and covariances between stock returns. Despite the result that estimation of the covariance matrix results in lower estimation errors than estimation of the mean return of stocks, the sample covariance matrix still has a number of undesirable properties. When the length of time series available is small in comparison to the number of stocks, the sample covariance matrix is usually ill-conditioned and suffers from significant estimation error. This causes small changes in the estimates of the covariance to result in large changes in the portfolio weights. This instability leads to a high portfolio turnover over time (Bai and Shi, 2011). This occurs because implementation of MPT requires inversion of the covariance matrix, and this amplifies the noise in the sample covariance matrix (Michaud, 1989). Furthermore, Bai and Shi (2011) show that the expected value of the inverse is a biased estimator. In the case where the length of the time series is less than the number of stocks, the covariance matrix is not even invertible, as it is not of full rank.

There are several approaches that aim to deal with the problem of obtaining a robust estimate of the covariance matrix. Bai and Shi (2011) provide an overview of some of these methods. They discuss the methods of factor models, shrinkage estimators, and Bayesian estimators of the covariance matrix.

Factor models assume a market model, e.g. the single-index model, and use historical returns data to estimate model parameters, which are then used to simulate returns. The simulated returns are then used to obtain an estimate of the covariance matrix. Shrinkage estimators are a weighted average of the sample covariance matrix and another "target" covariance matrix, which possesses some underlying structure, for example the single-index model covariance matrix as described above. Bayesian estimators incorporate prior information into the covariance matrix estimate in the form of the investor's beliefs, historical experience or economic theory.

Bai and Shi (2011) also discuss Random Matrix Theory and its filtering techniques, which were originally applied in finance by Laloux *et al.* (1999, 2000) and Bouchaud and Potters (2000). This method compares the eigen-structure of the empirical correlation matrix to that generated from a correlation matrix of completely random returns, in order to identify the degree of noise in the empirical correlation matrix. Filtering techniques then remove the noise from the correlation matrix, retaining only those factors that correspond to real information.

Yanou (2013) proposes another approach to obtaining robust estimates of the covariance matrix. This method uses L-moments, rather than conventional sample moments, to estimate the covariance matrix. L-moments were popularised by Hosking *et al.* (1985), and possess the advantage that they are more robust to outliers than sample moments, in particular when sample sizes are small. Yanou (2013) uses L-moments to compute the covariance matrix, and then filters the matrix using Random Matrix Theory. This filtered covariance matrix is then used in portfolio selection. In particular, it is used to construct the minimum variance portfolio.

This dissertation uses the work of Yanou (2013) as the key inspiration to investigate the use of L-moments and Random Matrix Theory in portfolio selection in the South African market. We use these two approaches, both on their own and in conjunction with one another, to construct minimum variance portfolios. In particular, we aim to analyse four strategies for the construction of minimum variance portfolios. The four strategies correspond to the four different methods for estimating the covariance matrix: the conventional sample covariance matrix, the L-moments covariance matrix, the sample moments covariance matrix filtered using the techniques of Random Matrix Theory and the L-moments covariance matrix filtered using the techniques of Random Matrix Theory. The risk-return characteristics and performance of the four portfolios will be analysed. In particular, for each portfolio, we will compute the mean return, the volatility (as a measure of risk) and the risk-adjusted performance, as measured by the Sharpe and Sortino ratios. We will also comment on the diversification and stability of the portfolios, and on the feasibility of practical implementation.

The analysis is performed primarily on the South African market, but results are compared to those of Yanou (2013), which are based on the US market.

The remainder of this dissertation is structured as follows. Chapter 2 begins by discussing the motivation for robust estimation of covariance matrices. It then reviews the concept of L-moments and Random Matrix Theory, and presents the theoretical results which form the basis of this dissertation. Chapter 3 is oriented towards practical application and begins by describing the data that has been used in the analysis. It then details the methodology that has been followed to implement both Random Matrix Theory and L-moments, and construct global minimum variance portfolios. Chapter 4 applies the described methodology to South African and US data, and presents the empirical results, along with a discussion. Finally, Chapter 5 concludes the dissertation with a brief discussion of the key findings, along with a mention of potential avenues of further research.

Chapter 2

Robust estimation of covariance matrices

Empirical correlation and covariance matrices are of considerable importance in both portfolio construction and the risk measurement and management of such portfolios. The risk-return characteristics of portfolios are largely dependent on the correlation of the stocks comprising the portfolio. It is thus imperative to be able to correctly exploit the advantages that these correlations present. This necessitates accurate and robust estimation of both the correlation and covariance matrices.

Modern Portfolio Theory (MPT), due to Markowitz (1952), models stock returns using some elliptical distribution, and assumes that the means, variances and covariances of the returns of all stocks in a universe are known. In practice, these parameters are estimated from finite length time series data. Accurate estimation of the correlation and covariance matrices is difficult due to the finite and limited time series data. For a universe of N assets, computing the empirical correlation matrix requires the estimation of $N(N-1)/2$ unique correlations, while computing the covariance matrix requires estimating N variances and $N(N-1)/2$ covariances. When the length of the time series of returns, T, is not large in comparison to N, the estimates are noisy, i.e. they are, to a large extent, random. The smaller T is, the greater the randomness.

Historically, the sample covariance matrix has been used as the empirical estimate, but, as described in Chapter 1, this has a number of undesirable properties which lead to out-of-sample portfolio risks that are usually much larger than the anticipated risks, resulting in poor out-of-sample performance of optimal portfolios. One should therefore be careful when using this estimate of the covariance matrix, whether to construct optimal portfolios, or estimate their risk-return characteristics. In order to be able to estimate the out-of-sample risk more accurately, and ideally improve the risk-return characteristics of optimised portfolios, robust estimation of the covariance matrix is required.

In this chapter, two methods of improving the empirical estimate of covariance and correlation matrices are discussed. The first relates to the use of L-moments to compute the correlation and covariance matrices. L-moments are more robust to outliers than the sample moments, in particular, when sample sizes are small. The second method discussed is Random Matrix Theory, and its associated filtering techniques. These techniques “de-noise” the correlation matrix by removing spurious correlations, leaving behind a matrix that represents more accurate information. Removal of the noise makes the optimisation process more reliable and stable, leading to a more accurate estimate of the risk associated with the constructed portfolios.

2.1 L-Moments

When analyzing a probability distribution, standard practice is to summarise the shape of the distribution by computing measures of location, dispersion, skewness and peakedness. The most popular approach to compute these quantities is using classical sample moments. Despite their popularity, sample moments are known to suffer from several drawbacks. In particular, they are sensitive to outliers and small sample sizes, resulting in empirical estimates of sample moments being markedly different from their theoretical values. Sample moments also have poor asymptotic efficiency, especially when the underlying probability distributions are fat-tailed (Perez *et al.*, 2003).

An alternative to sample moments are L-moments, originally derived by Sillitto (1969, 1951), without referring to them as such. They were later popularised by Hosking *et al.* (1985), and obtained their name from their construction as expectations of linear combinations of order statistics. Hosking (1990) gathered and extended the works of Sillitto (1969, 1951), Downton (1966), Chan (1967), Konheim (1971), Mallows (1973), and Greenwood *et al.* (1979) to demonstrate that L-moments can be used for statistical analysis of univariate probability distributions. In particular, he showed that L-moments perform competitively against available statistical techniques.

The theoretical distribution of a random variable has a set of population L-moments, analogous to conventional moments. Sample L-moments can be defined and calculated for a sample from the population, and are used as estimators of the population L-moments in order to characterise the shape of the distribution. In addition to computing summary statistics, L-moments are also used to compute parameter estimators and quantiles, and perform hypothesis and goodness-of-fit tests for probability distributions. As such, the theory of L-moments parallels the theory of conventional sample moments (Hosking, 1986, 1989, 1990).

The efficacy of L-moments lies in their similarity to conventional sample moments, while counteracting some of the drawbacks that sample moments possess. The main advantage of L-moments is that they are more robust to outliers than sample moments. When sample sizes are small, parameter estimates obtained via L-moments are more accurate, less subject to bias in estimation and approximate their asymptotic distribution more closely (Hosking, 1986). Furthermore, a finite mean guarantees the existence of higher-order L-moments, even if the corresponding conventional moments may not exist. This allows L-moments to characterise a wider range of distributions (Hosking, 1990).

It is due to these superior sampling properties of L-moments that Hosking and Wallis (1987) advocate that L-moments give a better approximation of the unknown population distribution than conventional sample moments.

2.1.1 Univariate L-moments

Consider a random sample $\{X_1, X_2, ..., X_n\}$ of size n taken from the univariate distribution X. The L-moment of order k for the distribution X is defined as:

$$\lambda_k(X) = k^{-1} \sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} E\left[X_{k-i:k}\right],$$

where $X_{k-i:k}$ is the $(k-i)$-th order statistic from a sample of size k taken from the distribution X. From David and Nagaraja (2003), the expected value of an order statistic $X_{k:n}$ can be written as:

$$E\left[X_{k:n}\right] = \frac{n!}{(k-1)!(n-k)!} \int_0^1 F^{-1}(u) u^{k-1} (1-u)^{n-k} du,$$

where $F^{-1}(\cdot)$ is the quantile distribution function of X.

This results in the following expression for the L-moment of order k for the distribution X:

$$\lambda_k(X) = \int_0^1 F^{-1}(u) P^*_{k-1}(u) du,$$

where,

$$P^*_k(u) = \sum_{i=0}^{k} p^*_{k,i} u^i, \tag{2.1}$$

and,

$$p^*_{k,i} = (-1)^{k-i} \binom{k}{i} \binom{k+1}{i}. \tag{2.2}$$

$P^*_k(u)$ for $0 \leq u \leq 1$ are shifted Legendre polynomials, which are related to the usual Legendre polynomials $P_k(u)$ by $P^*_k(u) = P_k(2u-1)$.

The k-th L-moment may also be represented as (Serfling and Xiao, 2007):

$$\lambda_k(X) = n^{-1} \sum_{i=1}^{n} w_{i:n}^{(k)} E\left[X_{i:n}\right], \tag{2.3}$$

where the weights $w_{i:n}^{(k)}$ are given by:

$$w_{i:n}^{(k)} = \sum_{j=0}^{\min(i-1,k-1)} (-1)^{k-1-j} \binom{k-1}{j} \binom{k-1+j}{j} \binom{n-1}{j}^{-1} \binom{i-1}{j}. \tag{2.4}$$

To estimate the L-moments for a sample, we use the sample version of equation (2.3):

$$\hat{\lambda}_k(X) = n^{-1} \sum_{i=1}^{n} w_{i:n}^{(k)} X_{i:n}. \tag{2.5}$$

In particular, note that

$$w_{i:n}^{(1)} = 1,$$
$$w_{i:n}^{(2)} = -1 + \frac{2(i-1)}{n-1},$$

such that the estimates of the first- and second-order L-moments are given by:

$$\hat{\lambda}_1(X) = \frac{1}{n} \sum_{i=1}^{n} X_{i:n},$$
$$\hat{\lambda}_2(X) = \frac{1}{n} \sum_{i=1}^{n} \left(-1 + \frac{2(i-1)}{n-1}\right) X_{i:n}.$$

Note that the L-moment of order 1 is the same as the conventional sample mean. The L-moment of order 2, however, differs from the sample variance.

2.1.2 Bivariate L-moments

In the context of bivariate random variables, we explore the theory of L-comoments developed by Serfling and Xiao (2007). Consider a bivariate random sample $\{(X_1, Y_1), (X_2, Y_2), ...(X_n, Y_n)\}$ of size n taken from the bivariate distribution $\{X, Y\}$.

Before defining L-comoments, we first consider the concept of concomitants, as defined by Yang (1977). The element from the distribution of X that is paired with the i-th order statistic of Y, $Y_{i:n}$, is called the concomitant of the i-th order statistic of Y, and is denoted by $X_{i:n}^{Y}$. Analogously, the concomitant of the i-th order statistic of X, denoted by $Y^X_{i:n}$ is the element from the distribution of Y that is paired with the i-th order statistic of X, $X_{i:n}$.

The k-th L-comoment of X with respect to Y is then defined as:

$$\lambda_k(X,Y) = k^{-1}\sum_{i=0}^{k-1}(-1)^i\binom{k-1}{i}E\left[X^Y_{k-i:k}\right],$$

where $X^Y_{k-i:k}$ is the $(k-i)$-th concomitant of Y.

Since the expectation of the concomitant X^Y is the expectation of X conditioned on the order statistics of Y, it may be expressed as follows, in a form similar to the univariate L-moment case:

$$E\left[X^Y_{k:n}\right] = \frac{n!}{(k-1)!(n-k)!}\int_0^1 F^{-1}_{X|Y_{k:n}}(u)u^{k-1}(1-u)^{n-k}du,$$

where $F^{-1}_{X|Y_{k:n}}(\cdot)$ is the quantile distribution function of X conditional on the order statistics of Y.

This results in the following expression for the L-comoment of order k of X with respect to Y:

$$\lambda_k(X,Y) = \int_0^1 F^{-1}_{X|Y_{k:n}}(u)P^*_{k-1}(u)du,$$

where $P^*_k(u)$ has been defined in equations (2.1) and (2.2).

The k-th L-comoment of Y with respect to X, $\lambda_k(Y,X)$, can be defined analogously by considering the concomitants $Y^X_{k-i:k}$.

An alternative representation of the k-th L-comoment is given by Serfling and Xiao (2007):

$$\lambda_k(X,Y) = n^{-1}\sum_{i=1}^{n} w^{(k)}_{i:n}E\left[X^Y_{i:n}\right], \tag{2.6}$$

where the weights are the same as in equation (2.4). The sample version of equation (2.6) is given by:

$$\lambda_k(X,Y) = n^{-1}\sum_{i=1}^{n} w^{(k)}_{i:n}X^Y_{i:n}, \tag{2.7}$$

and is used to compute unbiased estimates of the k-th L-comoment of X with respect to Y. Equation (2.7) above is an analogue of equation (2.5) which was used to compute sample univariate L-moments, and uses concomitants in place of order statistics.

Here we are interested in estimating the covariance matrix of stock returns. It is therefore necessary to consider only second order bivariate L-moments, as these will represent the entries of the L-covariance matrix. Consequently, we do not discuss higher-order L-moments. For details on higher-order L-moments, see Serfling and Xiao (2007).

2.1.3 L-Moments variance-covariance and correlation matrices

The covariance and correlation matrices obtained using L-moments are referred to as the L-covariance and L-correlation matrix respectively.

Using the results from Sections 2.1.1 and 2.1.2 above, the L-covariance matrix is defined as:

$$\Omega_{Lcov} = \begin{pmatrix} \hat{\lambda}_2(X) & \hat{\lambda}_2(X,Y) \\ \hat{\lambda}_2(Y,X) & \hat{\lambda}_2(Y) \end{pmatrix}.$$

The L-covariance matrix characterises the concomitance effect between X and Y. The concomitance effect of X with respect to Y is not necesarily the same as that of Y with respect to X. Consequently, the k-th L-comoment of X with respect to Y is not necessarily the same as that of Y with respect to X. As a result, the L-covariance matrix is not symmetric. In contrast, when $k = 2$, estimation of the sample comoment yields the variance, resulting in a symmetric sample variance-covariance matrix. As mentioned by Serfling and Xiao (2007), it is possible to define symmetric L-comoments, although it is preferred to retain the ordered pair of asymmetric comoments, as they carry more information.

In our context, the L-covariance matrix will be used in Markowitz optimisation. This requires the matrix to be symmetric, since a quadratic equation is used to solve for the optimal portfolio. Yanou (2013) proposes the following transformation to convert the L-covariance matrix into a symmetric one:

$$\Omega_{Lcov}^{symm} = \begin{pmatrix} \hat{\lambda}_2(X) & \alpha_1\hat{\lambda}_2(X,Y) + \alpha_2\hat{\lambda}_2(Y,X) \\ \alpha_1\hat{\lambda}_2(X,Y) + \alpha_2\hat{\lambda}_2(Y,X) & \hat{\lambda}_2(Y) \end{pmatrix},$$

where,

$$\alpha_1 = \frac{\hat{\lambda}_2(X,Y)}{\hat{\lambda}_2(X,Y) + \hat{\lambda}_2(Y,X)},$$

$$\alpha_2 = \frac{\hat{\lambda}_2(Y,X)}{\hat{\lambda}_2(X,Y) + \hat{\lambda}_2(Y,X)}.$$

The parameter α_1 allows for the concomitance effect of X with respect to Y, while α_2 takes into account the concomitance effect of Y with respect to X. Yanou (2013) shows that this approach to estimating a symmetric version of the L-covariance matrix peforms better than simple averaging of the L-comoments, as that fails to capture the concomitance phenomenon.

The L-correlation matrix is defined as:

$$\Omega_{Lcorr} = \begin{pmatrix} 1 & \tau_{X,Y} \\ \tau_{Y,X} & 1 \end{pmatrix},$$

where,

$$\tau_{X,Y} = \frac{\hat{\lambda}_2(X,Y)}{\hat{\lambda}_2(X)},$$

and,

$$\tau_{Y,X} = \frac{\hat{\lambda}_2(Y,X)}{\hat{\lambda}_2(Y)}.$$

$\tau_{X,Y}$ represents the L-correlation coefficient of X with respect Y, and $\tau_{Y,X}$ represents the L-correlation coefficient of Y with respect to X.

Like the traditional sample correlation matrix, the entries in the L-correlation matrix lie between ± 1. However, like the L-covariance matrix, the L-correlation matrix is not symmetric. In our case, the correlation matrix will be filtered using the techniques provided by Random Matrix Theory (discussed in Section 2.2). In Chapter 4, we show that Random Matrix Theory can be applied to the L-correlation matrix as defined above. Therefore, it is not necessary to specify a symmetric version of the L-correlation matrix.

2.2 Random Matrix Theory

Random Matrix Theory (RMT), originally developed for use in physics, has been expounded by Dyson and Mehta (1963), Dyson (1971) and Mehta (1991). The theory relies on the inspection of the eigenvalue distribution of a matrix to identify any non-random properties. Close agreement between the eigenvalue distribution of any matrix with those of a matrix of completely random entries would imply a considerable degree of randomness in the original matrix. This is relevant in finance, as the empirical correlation matrix of stock returns can be compared to a correlation matrix generated from completely random returns. This allows us to determine the degree of randomness or noise in the empirical correlation matrix.

2.2.1 Mathematical framework of RMT

Consider a matrix M of size N by T containing independent and identically distributed elements with mean 0 and variance σ^2. This matrix is called a random Wishart matrix. The correlation matrix C is given by:

$$C = \frac{1}{T} M M'.$$

The statistical properties of these random matrices are known from Dyson and Mehta (1963), Marčenko and Pastur (1967), Dyson (1971), Edelman (1988), Mehta (1991), and Sengupta and Mitra (1999). It has been shown that the eigenvalues of the random correlation matrix C are self-averaging, i.e. in the limit as $N, T \to \infty$, with $Q = T/N \geq 1$ remaining fixed, the distribution of the eigenvalues of C is given by the density function:

$$\rho(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda_- - \lambda)}}{\lambda},$$

where σ^2 is the variance of the elements in the original random matrix M, and λ_- and λ_+ are the theoretical minimum and maximum eigenvalues of the random correlation matrix respectively, given by:

$$\lambda_\pm = \sigma^2 \left(1 \pm \sqrt{\frac{1}{Q}}\right)^2.$$

The density $\rho(\lambda)$ is known as the Marčenko-Pastur density.

Note that this result is valid only as $N \to \infty$. For finite N there is a positive probability of finding eigenvalues that may be above or below the theoretical bounds. This probability tends to zero as N tends to infinity. Nonetheless, Gatheral (2008) illustrates that the theoretical distribution approximates empirical densities well, even when N and T are small. This therefore advocates the use of this theoretical density in practical applications.

For a given volatility, the Marčenko-Pastur density depends on the value of $Q = T/N$. Fig. 2.1 shows the density for three different values of Q, and for $\sigma^2 = 1$.

The Marčenko-Pastur density calculated from a completely random matrix can be compared against the eigen-structure of a correlation matrix obtained from financial returns data. The completely random matrix corresponds to the null hypothesis that the set of stocks considered are strictly independent, and that the correlation matrix is the identity matrix. Deviations in the eigen-structure of the empirical correlation matrix from the Marčenko-Pastur distribution suggest the presence of true information, while those eigenvalues falling within the theoretical spectrum are considered to be noise and should be filtered out to obtain a cleaned correlation matrix.

Laloux *et al.* (1999, 2000) demonstrate, using empirical evidence, the validity of using Random Matrix Theory in finance. In particular, they show that the statistical structure of the empirical correlation matrix agrees considerably with the theoretical distribution, indicating that only a few of the eigenvalues and eigenvectors carry relevant information. They suggest that filtering techniques are useful in extracting relevant correlations between stocks. They also demonstrate that using

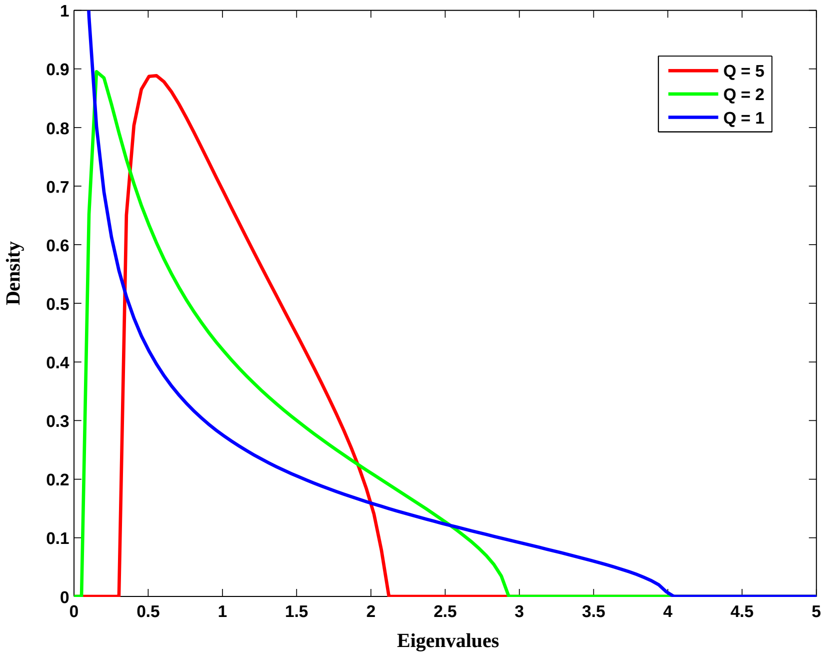


Fig. 2.1: Marčenko-Pastur densitites for different values of $Q = T/M$ and $\sigma^2 = 1$.

RMT significantly reduces the difference between the expected and realised risk of a portfolio, leading to improved risk-return characteristics of optimised portfolios.

More recently, Bouchaud and Potters (2000), Plerou *et al.* (2002), Sharifi *et al.* (2004), Laloux *et al.* (2005), and Conlon *et al.* (2007) have also applied Random Matrix Theory to various financial problems, including portfolio optimisation, and risk measurement and management.

2.2.2 Filtering the correlation matrix

There are several RMT filtering recipes that may be applied to obtain a filtered estimate of the correlation and covariance matrices. A common method, originally suggested and tested by Bouchaud and Potters (2000), referred to as "eigenvalue clipping", first computes the eigen-decomposition of the correlation matrix, C, of N stocks:

$$C = UDU^{-1},$$

where U is an N x N matrix such that each column is an eigenvector of C, and D is a diagonal matrix of size N where the elements on the diagonal are the eigenvalues of C.

The filtering method retains the eigenvalues that lie above the theoretical maximum, λ_+. Eigenvalues below λ_+, falling within the Marčenko-Pastur noise band, are declared as noise, and are replaced by their average. The assumption is that the noisy eigen-components are all of equal significance. Hence they are assigned equal weighting by replacing the eigenvalues with a constant – their average. This ensures that the trace of the correlation matrix is preserved. Note that the trace preservation of the usual eigen-clipping method results in upweighting the eigenvalues below the lower bound, and as such it is only the eigenvalues above the upper bound that are considered to contain useful information. The new set of eigenvalues is used to construct a new diagonal matrix D^*. The filtered correlation matrix is then reconstructed using the new diagonal matrix of eigenvalues, D^*, and the original matrix, U, of eigenvectors:

$$C^* = UD^*U^{-1}.$$

The diagonal elements of C^* are forced to unity to ensure a proper correlation matrix. The filtered covariance matrix, S^*, is then computed as:

$$S^* = V^{\frac{1}{2}}C^*V^{\frac{1}{2}},$$

where V is a diagonal matrix of the original variances.

When computing the filtered L-covariance matrix, one additional step is required. After having calculated the L-correlation matrix directly, using the method described on pg. 12, the L-correlation matrix is filtered, from which a filtered L-covariance matrix is then estimated. By definition, this filtered L-covariance matrix is not symmetric. It must be converted into a symmetric version using the approach described in Section 2.1.3.

This filtered covariance matrix is then applied in portfolio management, including in the construction of optimal portfolios.

Plerou *et al.* (2002) use a different method where the noisy eigenvalues are set to zero, rather than their average. Here, the assumption is that these noisy eigen-components do not possess any relevant information, and hence are assigned of weighting of zero. The non-noisy eigenvalues are reweighted to ensure that the trace of the correlation matrix is preserved. A third method suggested by Sharifi *et al.* (2004) replaces the noisy eigenvalues with components that have the largest separation from one other, while maintaining the trace of the matrix as before. They advocate that this approach is superior than the "eigenvalue clipping" method of Bouchaud and Potters (2000) as it results in a correlation matrix with increased stability.

2.3 Other approaches to obtaining robust estimates of the covariance matrix

There are several other approaches to obtaining robust estimates of the covariance matrix that have not been discussed. Bai and Shi (2011) provide a good overview of several such approaches. In addition to RMT, they discuss the methods of factor models, shrinkage estimators, and Bayesian estimators of the covariance matrix.

The factor model approach first involves the specification of a market model. Possible specifications are the CAPM single index model, due to Sharpe (1964), or multi-factor models such as Arbitrage Pricing Theory, due to Ross (1976). Next, the coefficients of the model are estimated, and these coefficients are used to simulate stock returns, which are then further used to obtain an estimate of the covariance matrix. The implicit assumption in computing this estimate is that the market is fully described in the model specified.

Shrinkage estimators of the covariance matrix involve estimating a target covariance matrix, and then shrinking the sample covariance matrix towards that target. The target matrix will have some inherent structure based on statistical or economic theory. For example, the approach of Ledoit and Wolf (2003) uses a target covariance matrix specified by the single index model. The target covariance matrix is estimated using the same method described in the factor model approach above. The shrinkage estimator is then obtained by considering a linear combination of the target matrix and the sample covariance matrix, where the sample covariance matrix is shrunk by some shrinkage intensity parameter.

Ledoit and Wolf (2004) consider an alternative target. Here, the sample covariance matrix is shrunk towards the identity matrix, weighted by the average variance of all stocks. The assumption here is that the stocks in the market are uncorrelated and have equal variance.

Bayesian methods incorporate prior information in the form of personal beliefs, historical experience or modelling and economic theory in the estimate of the covariance matrix. Bayesian methods may be incorporated along with shrinkage estimators (Bai and Shi, 2011).

The advantage of these methods over the sample covariance matrix is that there is increased structure in the estimate of the covariance matrix. The increased structure is due to the specification of a market model, or incorporation of prior information. The result is that fewer parameters need to be estimated, hence ultimately reducing the estimation errors.

Chapter 3

Data and methodology

This chapter begins by describing the data that we have obtained and used in our empirical analysis. It then details the methodology we have followed to extract relevant data over a specified period, estimate the appropriate covariance matrix, and then construct the minimum variance portfolios. The application of this methodology and corresponding results are discussed in Chapter 4.

3.1 Data

We mainly use South African (SA) data in our analysis, as we are interested in portfolio selection in the South African market. Later we compare our results to that obtained by Yanou (2013) from his analysis on the US market. We aim to replicate his key results, and perform further analysis for comparison to the SA market. We therefore obtain the same US data as Yanou (2013). Each of these data sets is described below, and have been obtained from *Bloomberg*.

3.1.1 South African Data

We consider the constituents of the FTSE/JSE Top40 Index to be our universe of stocks. Since the constituents of the index are adjusted quarterly, we obtain the quoted Total Return (TR) Index (Gross Dividends) of all stocks that have been constituents of the FTSE/JSE Top40 Index at any time over the period 01-Nov-2002 to 30-Jun-2014. There are 74 such stocks in total. At any point in time, however, only 40–42 of these stocks are members of the FTSE/JSE Top40 Index, and therefore the universe is dynamic. We allow for this when estimating the covariance matrix. We sample the TR Index for each stock both daily and weekly over the period 01-Nov-2002 to 30-Jun-2014. We have also obtained the FTSE/JSE Top40 Total Returns (TR) Index, again sampled both daily and weekly over the same period. In the case of daily data, this corresponds to 2912 data points, and in the case of weekly data, 609 data points.

From the full data set, we extract six subsets of data. These subsets of data contain TR Index time series of different lengths, they span different periods, have different sampling frequencies (daily and weekly), and contain different numbers of stocks. The stocks constituting each data subset are those that were members of the FTSE/JSE Top40 Index continuously throughout the relevant period, and for which there were no missing data. For example, data subset A contains the TR Index of 39 stocks, sampled daily, over the period 05-Jan-2004 to 29-Dec-2005. The 39 stocks were members of the FTSE/JSE Top40 Index throughout the period 05-Jan-2004 to 29-Dec-2005, and had no missing data (usually data are missing due to acquisitions or liquidations). These six subsets of data are displayed in Table 3.1 below.

Tab. 3.1: Data subsets extracted from South African data set

Dataset	Period	Sampling Frequency	No. of stocks
A	05-Jan-2004 to 29-Dec-2005	Daily	39
B	02-Jan-2007 to 30-Dec-2009	Daily	31
C	02-Jan-2013 to 27-Jun-2014	Daily	38
D	05-Jan-2004 to 29-Dec-2005	Weekly	39
E	02-Jan-2007 to 30-Dec-2009	Weekly	31
F	02-Jan-2013 to 27-Jun-2014	Weekly	38

The relevance of obtaining subsets of data is that they allow us to repeat our analysis on the different sets of data. This allows us to make more robust conclusions. In particular, we can analyse our results and determine whether there are consistent differences in the results which can be attributed to the characteristics of the data used (e.g. daily vs. weekly data), or whether they are independent of those characteristics.

Although there was no specific reasoning behind the choice of the periods for the subsets of data in Table 3.1 above, the period spanning the years 2007-2009 (corresponding to data sets B and E) was specifically isolated from the complete data set. This period was characterised by a global financial crisis, which although originating in the US, also affected the SA economy. As such, data corresponding to this period was set aside to ensure that should there be any anomalous results specifically due to the recession, these would be contained within the analysis conducted on data sets B and E, and not exhibited when similar analysis is conducted on the remaining data sets. One will therefore note that the periods of the other data sets correspond to times before and after the recession.

It is important to note that choosing stocks for each period using the method described above results in both look-ahead and survivorship bias when any analysis is performed on these data sets. However, since we are assuming perfect information, this bias will not significantly nullify any of the results presented below.

3.1.2 US Data

Yanou (2013) performs global minimum variance optimisation on US market data. He considers two distinct universes. “Database 1” is comprised of 64 stocks from the S&P500 universe, using data sampled on a weekly frequency from 29-May-1981 to 11-Apr-2008. “Database 2” is comprised of 78 stocks from the NYSE universe, again sampled on a weekly frequency over the period 25-May-1981 to 14-Apr-2008. To replicate the Yanou (2013) results, and compare the portfolio performance in the SA market to the US market, we make use of “Database 1”. Refer to Appendix A or Yanou (2013) for specific details on the 64 stocks that comprise the investment universe.

3.2 Methodology for the construction of minimum variance portfolios

The optimisation process involves first estimating a covariance matrix, and then using it to determining a set of portfolio weights that will minimise the portfolio volatility. We include a short sale constraint into the system. This constraint is consistent with those often faced by investors, and is therefore a realistic incorporation. Effectively, we are solving for the minimum variance portfolio subject to a short sale restriction – this portfolio will be different from the unconstrained global minimum variance portfolio.

We aim to solve the following system and estimate w:

$$\min(w'\Sigma w), \tag{3.1}$$

subject to the constraints:

$$w'\mathbf{1} = 1,$$
$$w_i \geq 0, i = 1, .., N,$$

where Σ may be either the sample or L-moments covariance matrix, filtered or non-filtered using RMT, depending on which strategy is being used, and w is a column vector of length N, representing the weights in each of the N stocks in the universe.

When Σ is the L-covariance matrix, it is estimated using the methodology described in Section 2.1.3. When the covariance matrix is required to be filtered using RMT, whether it is on the sample or L-covariance matrix, the filtering is performed using the methodology described in Section 2.2.2.

The procedure for implementing the portfolio optimisation is as follows. We begin by selecting an estimation window, e.g. 12 months. From the full data set, we extract the first 12 months of data. The last date of the estimation window is referred to as the portfolio construction date, since it is the date on which we construct the portfolio using past historical data, i.e. the 12 months of data over the estimation window. To account for survivorship bias, and ensure that the optimisation process is realistic, we filter the extracted data for those stocks that were members of the FTSE/JSE Top40 Index over the estimation window, and for which there was complete data available. This set of stocks is considered to be the investment universe, as at the portfolio construction date. This universe may thus be less than or equal to 42, the maximum number of constituents of the FTSE/JSE Top40 Index at any point in time. Using this data, we calculate the stock returns. The covariance matrix is then estimated, be it using sample moments or L-moments, with or without RMT. The estimated covariance matrix is then used to solve system (3.1) and obtain the portfolio weights.

We invest in the portfolio defined by these weights and hold it for a specified holding period, e.g. one week. We then slide the estimation window forward by the holding period to obtain a new estimation window, and a new portfolio construction date. Data over the new estimation window is extracted as before, again including only stocks that meet the criteria for being part of the investment universe. As before, the covariance matrix is estimated and the new portfolio weights are calculated. We then sell the old portfolio, buy the new one, and hold it for the holding period. This process is continued until we reach the end of the data set.

After performing the optimisation, we observe the evolution of the portfolio value, and determine its risk-return characteristics. In particular, we will calculate the mean return, the volatility – as measured by the standard deviation – and the risk-adjusted performance, as measured by the Sharpe ratio. One may argue that standard deviation is not an appropriate measure of risk, and therefore we compute the Sortino ratio, which considers only downside deviation to be risk. We also compare the composition of the portfolio to some benchmark, using the active share. From a diversification point of view, we compute the proportion of stocks that are invested in out of those available in the universe, as well as the Effective Number of Bets (ENB), proposed by Meucci (2010).

Chapter 4

Empirical results and analysis

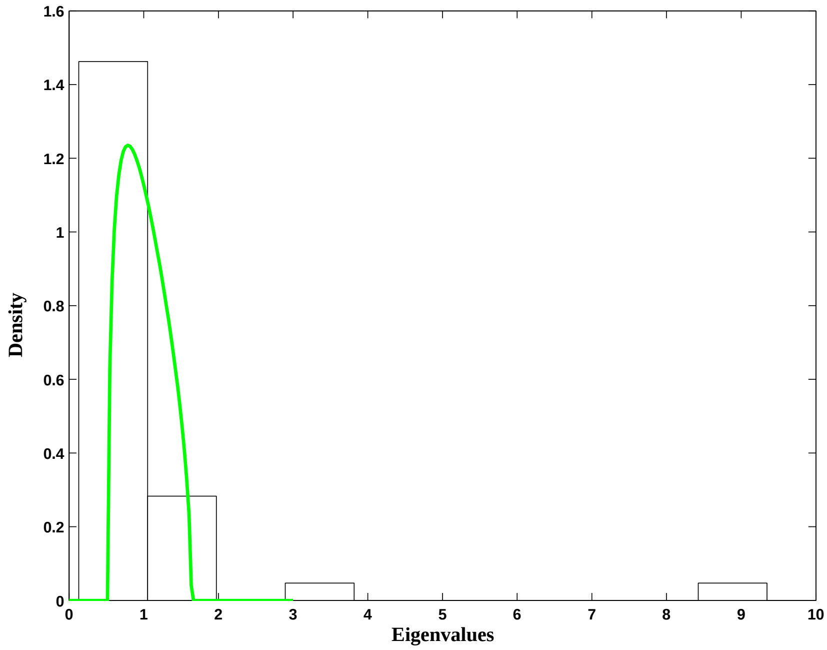
In this chapter, we construct global minimum variance portfolios using four different estimates of the covariance matrix: the conventional sample covariance matrix, the L-covariance matrix, the filtered sample covariance matrix, and the filtered L-covariance matrix. These four approaches will be compared in terms of their performance, diversification and risk-return characteristics.

Before RMT cleaning procedures are incorporated, we first determine whether RMT is consistent with both the sample and L-correlation matrices. Next, we illustrate the advantages that RMT present in the context of the risk-return characteristics of optimised portfolios. Finally we perform the portfolio optimisation, and analyse the different strategies.

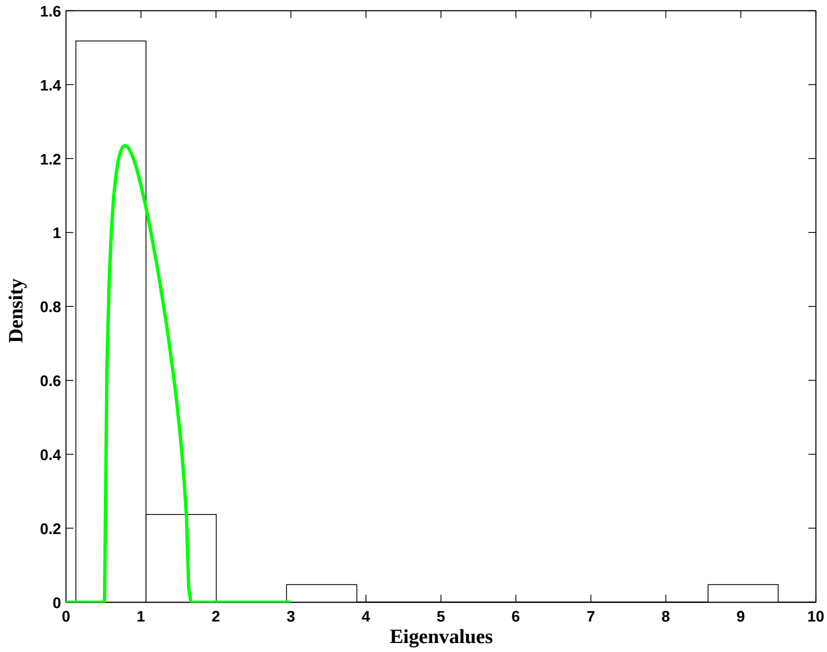
4.1 Rationale for using RMT in finance

To determine whether the use of Random Matrix Theory is valid in finance, we begin by comparing the eigenvalue distribution of the empirical correlation matrix to the Marčenko-Pastur distribution. Using dataset A, we calculate the histogram of eigenvalues of both the sample and L-correlation matrices. We also calculate the Marčenko-Pastur density. The histograms are displayed in Fig. 4.1, along with the Marčenko-Pastur distribution superimposed in green.

From Fig. 4.1, we note that for both the sample and L-correlation matrices, there is some agreement between the eigen-structure of the estimated correlation matrix, and the purely random correlation matrix. Most of the eigenvalues of the empirical correlation matrix lie in the range close to the theoretical bounds, with only a few deviating from the Marčenko-Pastur spectrum. This suggests that a large portion of the correlation matrix is dominated by noise. In particular, we find that for both the sample and L-correlation matrices, only 5.13% of eigenvalues fall outside of the theoretical spectrum, indicating that roughly 95% of the "information" present in the both the sample and L-correlation matrices is noise. Hence, we conclude that the



(a) Empirical eigenvalue spectrum of the sample correlation matrix



(b) Empirical eigenvalue spectrum of the L-correlation matrix

Fig. 4.1: Histogram of the empirical eigenvalue spectrum of the sample and L-correlation matrices, estimated using Dataset A. The theoretical Marčenko-Pastur density has been superimposed in green.

estimated correlation matrix indeed contains spurious and inaccurate correlations, arising due to estimation errors.

Repeating the analysis on the other data sets, B - F, yields similar results. The proportion of eigenvalues falling outside the theoretical spectrum ranges from 5-8%. These results are consistent with those of Laloux *et al.* (1999, 2000), who show that only 6% of eigenvalues exceed the theoretical bounds. Their results, however, consider only the sample correlation matrix. We have shown that this result holds for the L-correlation matrix as well. This implies that although the L-correlation matrix is asymmetric, and has statistical properties that are different to the sample correlation matrix, its eigen-structure is still similar to the sample correlation matrix. Like the sample correlation matrix, the L-correlation matrix is dominated by noise. Hence Random Matrix Theory does indeed find relevance in the context of L-moments being used to compute the covariance and correlation matrices. This result is consistent with Yanou (2013) who shows that the distribution of eigenvalues of the L-correlation matrix does conform, to a large degree, to the Marčenko-Pastur distribution.

In order to be able to conclude that only the eigenvalues falling outside of the theoretical bounds contain relevant information, we must show that those eigen-components do not conform to the properties dictated by Random Matrix Theory, while those falling within the theoretical spectrum do. To show this, we consider the distribution of the components of the eigenvectors. The eigenvectors corresponding to the eigenvalues deviating from the theoretical spectrum should have properties different to that specified by Random Matrix Theory.

We investigate this empirically by comparing the distribution of the components of the eigenvectors to their theoretical distribution. Theoretically, the components of the eigenvectors corresponding to eigenvalues falling within the theoretical bounds are Gaussian. Therefore, to justify agreement between Random Matrix Theory and the sample and L-correlation matrices, we must show that the eigenvectors of the correlation matrices corresponding to eigenvalues falling within the theoretical bounds follow the Gaussian distribution, while eigenvector components corresponding to eigenvalues outside the theoretical spectrum deviate from the Gaussian distribution.

Fig. 4.2 shows histograms of the components of three eigenvectors – eigenvectors 1, 2, and 3 – corresponding to the three largest eigenvalues. It also depicts the histograms of the components of three other eigenvectors – eigenvectors 32, 33, and 34 – corresponding to three eigenvalues falling within the theoretical spectrum of the sample correlation matrix. Fig. 4.3 shows the same, but for the L-correlation matrix. The second set of eigenvectors (i.e. 32, 33 and 34) have been chosen arbitrarily, and bear no significance, other than that their corresponding eigenvalues fall within the theoretical spectrum.

In addition to plotting histograms, we have also conducted Kolmogrov-Smirnov tests on the components of each of the eigenvectors to test for normality. Table 4.1 below shows the results (p-values) of these tests.

Tab. 4.1: p-values from the Kolmogorov-Smirnov normality tests performed on the components of the eigenvectors corresponding to the three largest eigenvalues, and three other (random) eigenvalues that fall within the theoretical upper and lower bound.

	S-correlation matrix	L-correlation matrix
Eigenvector 1	1.3118×10^{-20}	6.9443×10^{-21}
Eigenvector 2	0.1305	0.1468
Eigenvector 3	0.7101	0.5224
Eigenvector 32	0.3069	0.2276
Eigenvector 33	0.8482	0.9750
Eigenvector 34	0.1561	0.7186

At a 5% significance level, in both the sample moments and the L-moments case, the eigenvector corresponding to the largest eigenvalue deviates substantially from the Gaussian distribution. However, the eigenvector components corresponding to the smaller eigenvalues falling within the theoretical spectrum have a distribution that is much closer to the Gaussian distribution. These results demonstrate that there is indeed good agreement between Random Matrix Theory and both the sample and L-correlation matrices.

We proceed to investigate the nature of the "information" represented by the deviating eigenvalues. In particular, we aim to obtain an interpretation of the largest eigenvalue, and its corresponding eigenvector.

From Fig. 4.1, it is clear that in both the sample moments and L-moments case there is one eigenvalue that is considerably larger than the others, falling far outside of the theoretical bounds. It is well known that in the case of the sample correlation matrix, this eigenvalue corresponds to the "market", i.e. that it has roughly equal components in all N stocks (Laloux *et al.*, 1999, 2000; Plerou *et al.*, 2002). We reproduce this result below, and show that it holds for the L-correlation matrix as well.

We construct a hypothetical portfolio using the components of the largest eigenvector as weights. To do this, we have normalised the components of the eigenvector to ensure that they sum to one, allowing them to represent portfolio weights. The hypothetical portfolio defined by these weights is called the "Market Portfolio", and we compare the returns on it to those on the FTSE/JSE Top40 TR Index.

Using dataset A and the sample correlation matrix, we find a 90.27% correlation between the hypothetical portfolio and the Top40 TR Index. Using the L-correlation matrix, we find a correlation of 90.41%. Repeating the analysis on the other data sets with both the sample and L-correlation matrices yields correlations of greater than 90%. We note that due to survivorship bias, this correlation estimate is somewhat inflated. However, as discussed earlier, this does not nullify the result that there is indeed a significant relationship between the "Market Portfolio" and the Top40 TR Index.

Fig. 4.4 shows a scatter plot of the returns on the "Market Portfolio" and the Top40 TR Index, along with a regression line fitted between them. The extent of the correlation between these two returns is measured by the fit of the regression, as quantified by the R^2 of the regression line. From Fig. 4.4, the "Market Portfolio" constructed using both sample and L-moments appears to display identical correlations with the Top40 TR Index. The correlation structures are in fact not identical, evidenced by the fact that the correlation coefficient between the L-moments "Market Portfolio" and Top40 TR Index (90.41%) is different to the correlation coefficient between the sample-moments "Market Portfolio" and the Top40 TR Index (90.27%).

The analysis performed in this section indicates that for both the sample and L-correlation matrices, a large portion of the eigenvalue spectrum falls within the theoretical bounds defined by the Marčenko-Pastur density. Furthermore, using eigenvector components, we have shown that eigenvalues falling outside of the theoretical spectrum deviate from the results dictated by Random Matrix Theory. This validates the proposition that it is only these eigenvalues that contain real information, while others correspond to noise. Hence we conclude that Random Matrix Theory does indeed find relevance in financial applications, irrespective of whether correlation matrices are computed using the L-moments or the conventional sample moments approach. Thus, Random Matrix Theory may be applied to filter both the sample and L-correlation matrices to produce a more robust estimate, containing a truer and more informative correlation structure.

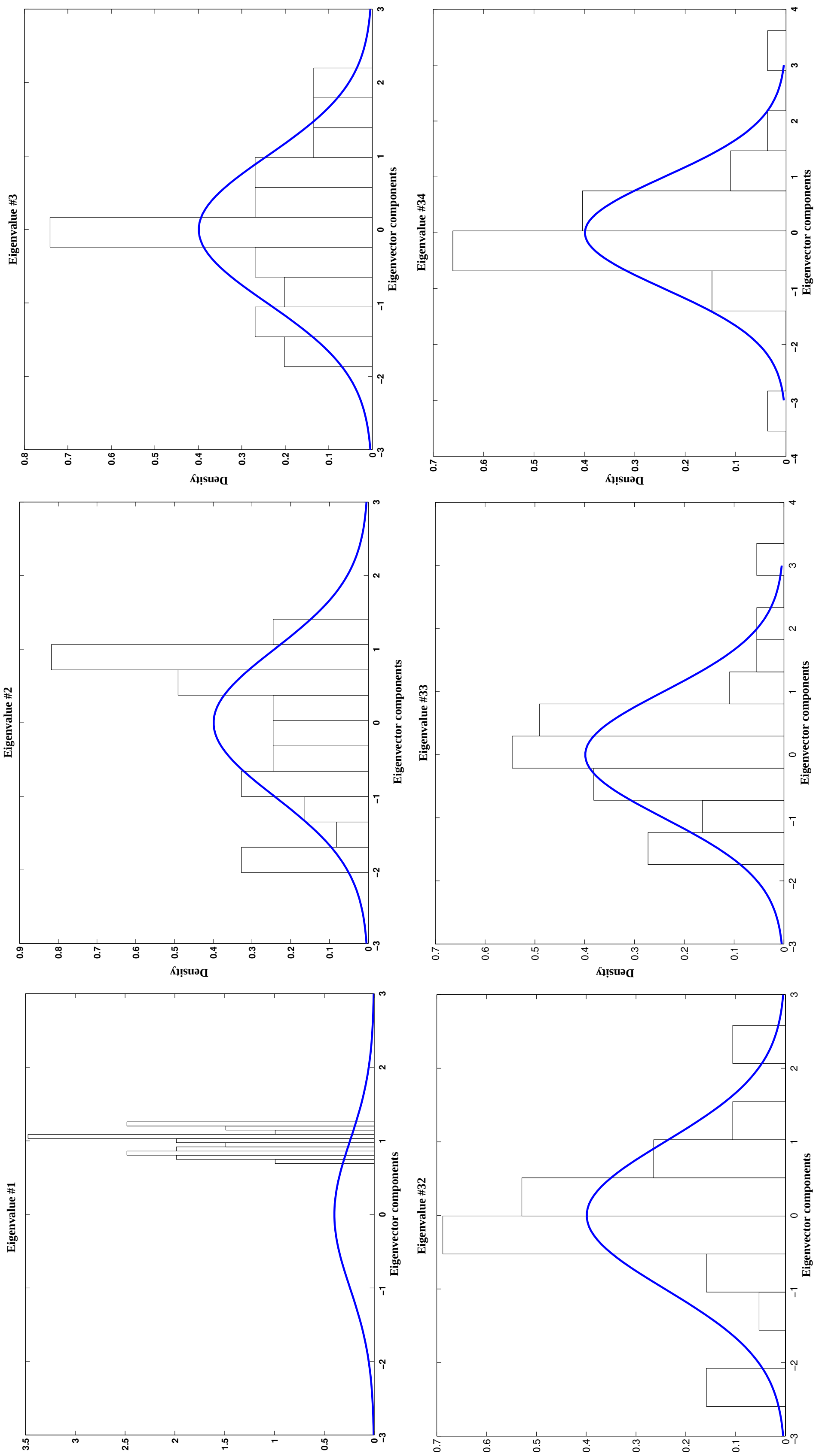


Fig. 4.2: Histograms displaying the distributions of the components of eigenvectors corresponding to the three largest eigenvalues, and three other (random) eigenvalues that fall within the theoretical upper and lower bound. The eigen-spectrum has been taken from the sample correlation matrix. The theoretical density, the standard normal distribution, has been superimposed in blue for comparison.

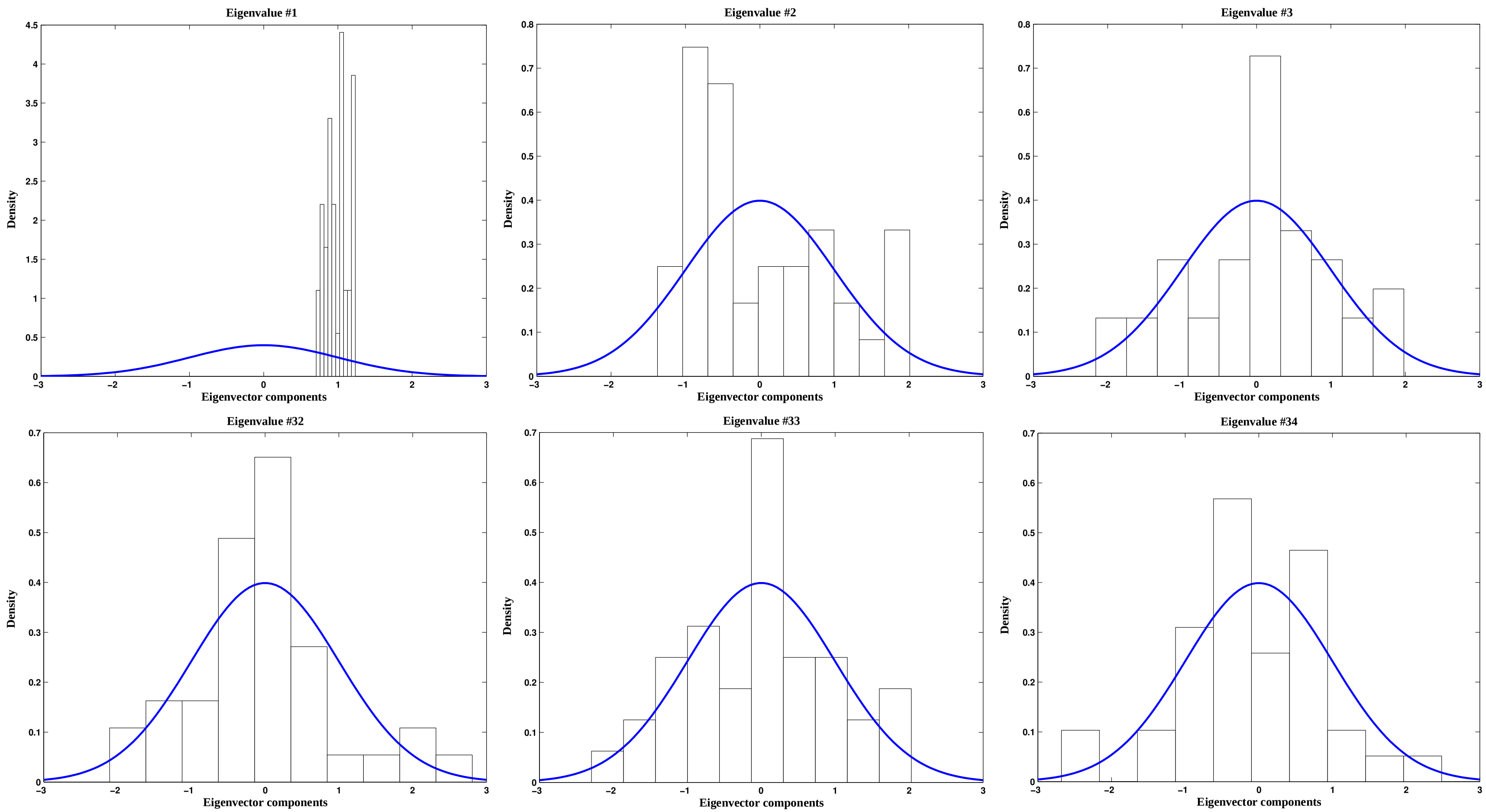
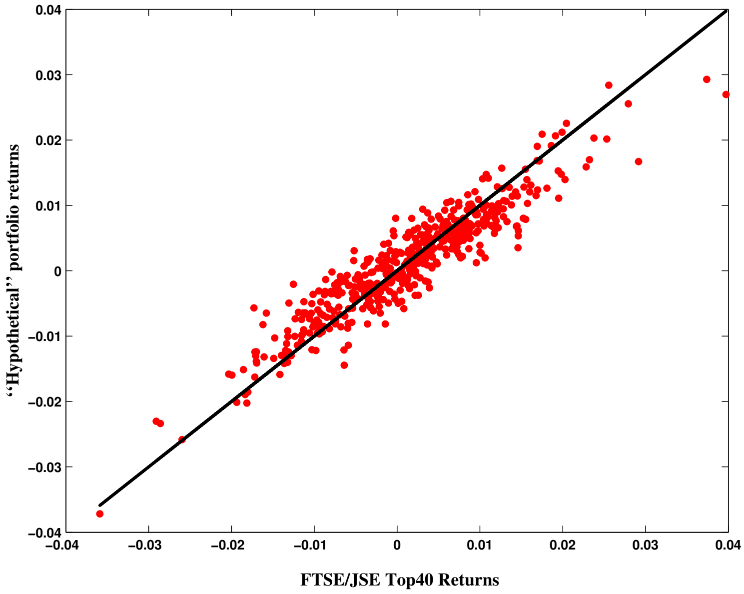
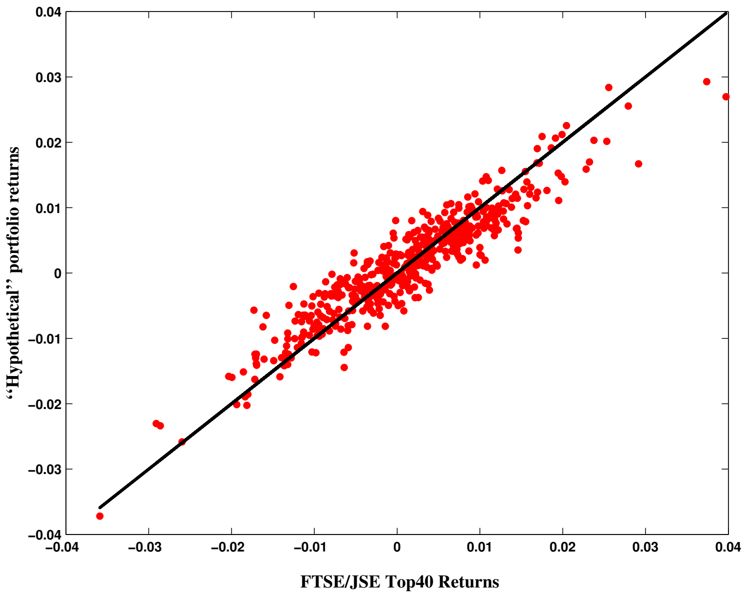


Fig. 4.3: Histograms displaying the distributions of the components of eigenvectors corresponding to the three largest eigenvalues, and three other (random) eigenvalues that fall within the theoretical upper and lower bound. The eigen-spectrum has been taken from the L-correlation matrix. The theoretical density, the standard normal distribution, has been superimposed in blue for comparison.



(a) Using correlation matrix calculated by the method of sample moments



(b) Using correlation matrix calculated by the method of L-moments

Fig. 4.4: Scatterplot of returns on the hypothetical “market” portfolio against those on the Top40 TR Index. The “hypothetical” portfolio has been constructed using the components of the eigenvector corresponding to the largest eigenvalue as weights.

4.2 Robust Risk Estimation using RMT

In this section we demonstrate that the incorporation of RMT allows robust estimation of the correlation matrix, by facilitating an improved risk-return profile of optimal portfolios that have been constructed using mean-variance optimisation.

We begin by using data subset A as described in Table 3.1. Following the approach of Laloux *et al.* (2000) and Conlon *et al.* (2007), the data is divided into two sub-periods of equal length. The first half spans 2004, while the second spans 2005.

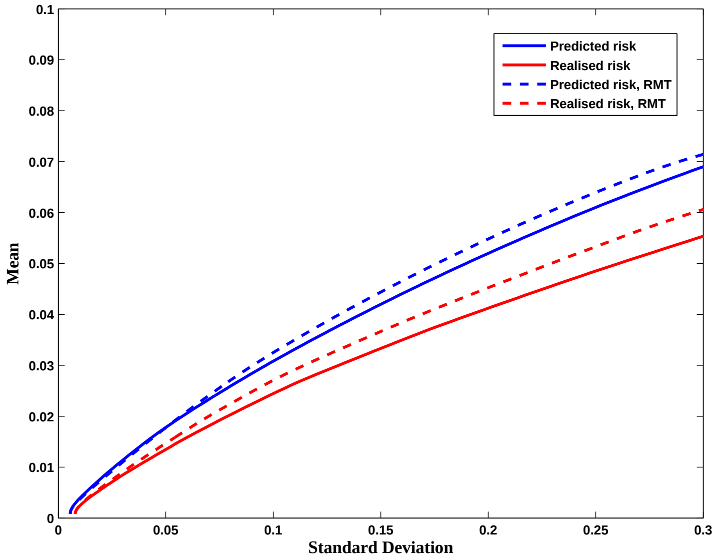
We assume that we are positioned on the date at the end of the first period. On this date we wish to construct optimal portfolios using MPT, and then analyse their-risk return characteristics. MPT requires that we estimate both the variance-covariance matrix and mean vector on this date, using historical data. Thus, we use the data from the first sub-period to compute the covariance and correlation matrices. Since we wish to investigate the effect of RMT on the estimation of the variance-covariance matrix only, we assume that we have perfect foresight on the stock returns, and thus estimate the vector of mean returns using the data from the second sub-period. Using the estimated variances, covariances and means, we calculate the efficient frontier. This efficient frontier represents the predicted (in-sample) risk-return characteristics of optimal portfolios.

The covariance and correlation matrices are re-calculated using data from the second sub-period. Using the weights of the optimal portfolios from the first efficient frontier, along with the mean vector, and the second covariance matrix, we calculate the standard deviation and expected return of each portfolio. This efficient frontier represents the realised (out-of-sample) risk-return profile of the optimal portfolios.

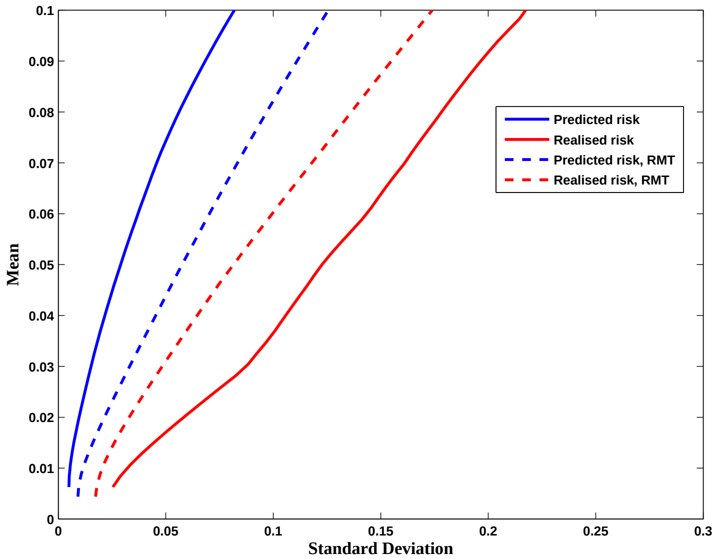
The above procedure is repeated, but now using RMT and the "eigenvalue clipping" method to filter noise from the correlation matrix. Again, the predicted and realised risk-return profile is calculated.

The resulting efficient frontiers are plotted in panel (a) of Fig. 4.5. The blue lines correspond to the predicted (in-sample) risk, and the red lines correspond to the realised (out-of-sample) risk, while the dashed lines indicate that the correlation matrix was filtered using RMT.

Fig. 4.5 shows that the realised risk, for each level of return, is greater than the predicted risk. This result is consistent for both cases where RMT is implemented and where it is not. We note that, in both cases, the incorporation of RMT results in an improved risk-return profile, in that the expected returns for each level of risk are greater. However, we are more interested in the difference between the in-sample and out-of-sample risk, as this gap constitutes the difference between the risk that investors estimated, and that which they experienced. From panel (a) of Fig. 4.5, there is not much difference between the realised and predicted risks in the



(a) Efficient frontiers plotted using data subset A, containing daily data over the period 05-Jan-2004 to 29-Dec-2005.



(b) Efficient frontiers plotted using data subset D, containing weekly data over the period 05-Jan-2004 to 29-Dec-2005.

Fig. 4.5: Efficient frontiers plotted using data subsets A (panel (a)) and D (panel (b)), displaying the in-sample and out-of-sample risk (blue and red, respectively), both when RMT has been incorporated (dashed lines) and when it has not (solid lines).

case where RMT is implemented, compared to the case where RMT has not been implemented. To analyse this further we run the same analysis using data subset D, which is identical to data subset A, but where the data points have been sampled weekly rather than daily. The resulting efficient frontiers are plotted in panel (b) of Fig. 4.5. We note that the realised risks are again larger than the predicted risks, and this is consistent with our previous result using daily data. More importantly, the difference between the realised and predicted risks is much smaller when RMT has been implemented.

While it is not appropriate to directly compare efficient frontiers with different holding periods - daily vs. weekly, in this case - we note that, overall, the incorporation of RMT, in both cases, adds value in reducing the noise in the estimate of the covariance matrix, thus ensuring that the difference between the predicted and realised risks are smaller.

These results are consistent with those obtained by Laloux *et al.* (2000), who find that using the sample correlation matrix, without filtering, leads to an underestimation of the real risk. They find that the risk of the optimised portfolio using the filtered correlation matrix is more stable, again similar to the results presented here using South African data.

Furthermore, we also find that RMT appears to add greater value to robust estimation of the covariance matrix when when the data has been sampled weekly rather than daily, i.e., when data has been sampled less frequently. A possible conclusion might be that weekly sampling of the data, rather than daily, results in greater noise in the estimate of the correlation and covariance matrices. Hence, we find that the incorporation of RMT in the second case adds more value, since its purpose is to de-noise the correlation matrix. This conclusion is, however, inconsistent with the widespread result in financial research that more frequent sampling of data generally results in greater noise in the covariance estimates. Nonetheless, we make note of this result, and also conclude that irrespective of the data sampling frequency, the incorporation of RMT aids in providing a more robust estimate of the covariance and correlation matrices.

We repeat the analysis on the other subsets of data, and present the results in Appendix B. The results from these data sets are consistent with those discussed above. In particular, we find that the realised risk is always larger than the predicted risk, and the gap between the realised and predicted risks is smaller when RMT is implemented. Furthermore, although counter-intuitive, we find that the improvement from incorporating RMT is greater when the data has been sampled weekly rather than daily.

An interesting result is that in panel (a) of Fig. B.1, incorporation of RMT does not add any value to the estimation of the risk-return characteristics of optimal portfolios. This could be because in addition to the data being sampled daily, the data spans the years 2007 – 2009, which incorporates the recent global financial crisis. It is well known that during crisis periods, stocks are highly correlated (Sandoval and Franca, 2012). Therefore, in this case the estimate of the correlation matrix contains high, actual correlations rather than spurious ones. We conclude that it is because of this that RMT did not offer much improvement.

From panel (a) of Fig. B.2 we note that for some levels of return, the predicted risk is actually larger than the realised risk. This result is simply characteristic of the data used, and reflects performance in the market better than what was anticipated by investors.

Our results above suggest that RMT does indeed retain only the relevant information from the correlation and covariance matrices, thus reducing the "noise" and ensuring a more robust estimate. This robustness is illustrated by the fact that the out-of-sample risk is closer to what had been predicted using in-sample data. Furthermore, the efficacy of RMT is more prevalent when data has been sampled less frequently.

4.3 Global Minimum Variance Portfolio Optimisation

In this section, we construct the global minimum variance portfolio using four different estimates of the covariance matrix: the traditional sample covariance matrix, the L-covariance matrix, the filtered sample covariance matrix and the filtered L-covariance matrix.

4.3.1 Portfolio Performance

We analyse the performance of the portfolios and compare them in terms of their risk and return characteristics. In particular, for each portfolio, we compute the annualised mean return (AMR), annualised standard deviation (ASD), and the respective Sharpe and Sortino ratios. For the Sortino ratio, we have set a benchmark of zero, so that any negative return is considered risk.

We have used a 12 month estimation window, various holding periods, and have repeated the analysis on both daily and weekly data. The results of this analysis are displayed in Tables 4.1 to 4.6.

Tab. 4.2: Portfolio performance based on full weekly data, using a 12 month estimation window and 4 week rolling period.

	Top40 Index	S-Moment	S-Moment + RMT	L-Moment	L-Moment + RMT
ASD	0.1929	0.1454	0.1445	0.2023	0.1519
AMR	0.1966	0.2547	0.2477	0.1690	0.2203
Sharpe	1.0189	1.7521	1.7149	0.8354	1.4505
Sortino	1.4167	2.4037	2.3676	1.3620	2.1526

Tab. 4.3: Portfolio performance based on full weekly data, using a 12 month estimation window and 2 week rolling period.

	Top40 Index	S-Moment	S-Moment + RMT	L-Moment	L-Moment + RMT
ASD	0.1841	0.1435	0.1418	0.1858	0.1516
AMR	0.1824	0.2295	0.2242	0.1521	0.1808
Sharpe	0.9907	1.5600	1.5810	0.8187	1.1930
Sortino	1.7107	2.3831	2.2734	1.4921	1.9627

Tab. 4.4: Portfolio performance based on full weekly data, using a 12 month estimation window and 1 week rolling period.

	Top40 Index	S-Moment	S-Moment + RMT	L-Moment	L-Moment + RMT
ASD	0.2053	0.1569	0.1541	0.2226	0.1718
AMR	0.2019	0.2449	0.2477	0.2064	0.1980
Sharpe	0.9834	1.5608	1.6074	0.9272	1.1522
Sortino	1.7671	2.5118	2.6915	1.8068	1.9698

From the results displayed in Tables 4.1 – 4.6, we note that the L-moments approach displays greater risk than the conventional sample moments approach, where risk is measured by annualised standard deviation. In fact, the L-moments approach displays the greatest annualised standard deviation of all four approaches. This greater “risk” however, is not coupled with a greater annualised mean return, as one would expect from the results of Modern Portfolio Theory. The L-moments approach thus results in a Sharpe ratio that is consistently lower than the sample moments approach. One may argue that standard deviation is not an appropriate measure of risk as it considers both upside and downside volatility to be risk.

Tab. 4.5: Portfolio performance based on full daily data, using a 12 month estimation window and 4 week rolling period.

	Top40 Index	S-Moment	S-Moment + RMT	L-Moment	L-Moment + RMT
ASD	0.1697	0.1236	0.1219	0.1428	0.1267
AMR	0.1976	0.2324	0.2250	0.1957	0.2197
Sharpe	1.1639	1.8811	1.8459	1.3707	1.7341
Sortino	2.0836	2.7142	2.6689	2.1853	2.8163

Tab. 4.6: Portfolio performance based on full daily data, using a 12 month estimation window and 2 week rolling period.

	Top40 Index	S-Moment	S-Moment + RMT	L-Moment	L-Moment + RMT
ASD	0.1899	0.1299	0.1294	0.1437	0.1348
AMR	0.1834	0.2141	0.2067	0.1730	0.1867
Sharpe	0.9658	1.6478	1.5976	1.2041	1.3850
Sortino	1.7554	2.6942	2.6298	2.0486	2.3028

Tab. 4.7: Portfolio performance based on full daily data, using a 12 month estimation window and 1 week rolling period.

	Top40 Index	S-Moment	S-Moment + RMT	L-Moment	L-Moment + RMT
ASD	0.2085	0.1526	0.1537	0.1668	0.1631
AMR	0.2008	0.2344	0.2270	0.2151	0.2073
Sharpe	0.9633	1.5358	1.4770	1.2897	1.2707
Sortino	1.6701	2.6868	2.6447	2.2471	2.2861

We therefore evaluate the four strategies on the Sortino ratio, and even then find that the L-moments approach is outperformed by the conventional sample moments approach.

When we incorporate Random Matrix Theory into the sample moments approach, we find that RMT does not offer any advantage in terms of portfolio performance. In fact, in almost all cases, incorporation of RMT into the sample moments approach reduces the performance of the portfolio, as measured by both the Sharpe and Sortino ratios. In a few cases, the Sharpe and Sortino ratio are improved, but the improvement is small. This poorer performance comes about as using the RMT approach results in optimised portfolios with a lower standard deviation, coupled with a lower return. Therefore, inclusion of RMT offers greater stability to the sample moments approach in terms of lower risk, but this does not necessarily translate into better performance of the optimised portfolio.

In the L-moments approach, the opposite is observed. Here, the inclusion of RMT offers substantial improvement to the performance of the portfolio constructed using L-moments. In particular, we see that the originally high standard deviation of the L-moments approach is significantly reduced. Therefore, like the sample moments method, we note that inclusion of RMT results in increased stability in terms of lower risk. However, unlike the sample moments approach, the L-moments approach, coupled with RMT, produces Sharpe and Sortino ratios much higher than the original L-moments approach. Despite this improvement however, the L-moments method, along with Random Matrix Theory, is still not able to outperform the conventional sample moments approach.

The results show that the conventional sample moments approach, irrespective of rolling window and sampling frequency of data, offers the best combination of low risk, higher returns, and greater outperformance measured in terms of the Sharpe and Sortino ratios.

We compare the performance of the four active strategies to the benchmark index, the FTSE/JSE Top40 Total Returns (TR) Index. We find that when using weekly data, three of the four active strategies (the sample moments method, sample moments and RMT method, and the L-moments and RMT method) tend to outperform the benchmark. These three active strategies display lower standard deviations and higher mean returns, resulting in higher Sharpe and Sortino ratios than the benchmark. When using weekly data, the L-moments method performs poorly compared to the benchmark. The standard deviation is higher, and the return is lower, resulting in poorer risk-adjusted performance. However, when daily data is used, we find that the L-moments strategy, like the other three active strategies, outperforms the benchmark. A possible justification for the difference in the performance of the L-moments strategy depending on whether data has been sampled daily or weekly could be that use of more frequently sampled data results in better performance of the strategy. This improved performance is possibly due to the fact that use of more frequently sampled data ensures that there is less noise in the estimate of the covariance matrix. This result extends to the other strategies as well – we note that in all cases, using data sampled more frequently results in improved performance of the optimised portfolios in terms of both the Sharpe and Sortino ratios.

When analysing the impact of Random Matrix Theory, we find that its positive impact on the L-moments strategy is more pronounced when weekly data is used. The improvement in the Sharpe and Sortino ratio when data has been less frequently sampled is greater than when daily data is used. This is consistent with the results of Random Matrix Theory, as the correlation matrix is expected to be dominated by noise when data has been less frequently sampled. It therefore follows that Random Matrix Theory should offer greater advantages when data is less frequently sampled, and this is the result that we observe here. Again, we mention that this result is counter-intuitive to what one would normally expect.

Comparing the sample moments approach based on the sampling frequency of data we noted that when weekly data was used, the addition of Random Matrix Theory to the sample moments method did not offer any improvement in the performance of the portfolios. There is no difference to this result when the data has been sampled daily.

We further illustrate the results discussed here in Figures 4.6 to 4.15. These plots display the performance of the different strategies, their rolling risk and return, as well as their diversification. The plots illustrate the results corresponding to a 12 month estimation window, a four week rolling period and where data has been sampled weekly.

Fig. 4.6 shows the evolution of the portfolio value for the four different active strategies discussed. The figure also displays the evolution of the FTSE/JSE Top40 TR Index for comparison. From Fig. 4.6, we can reiterate the result that the difference between the performance of the portfolio constructed using sample moments and then incorporating Random Matrix Theory is not substantial, as the evolution of the two portfolios is very similar. On the other hand, we can clearly see the significant effect that the inclusion of RMT has on the L-moments approach. While, originally, the L-moments approach performs poorly, in particular, worse than the benchmark, addition of RMT results in the portfolio outperforming the FTSE/JSE Top40 TR Index. However, as mentioned before, we observe that this approach is still not able to outperform the conventional sample moments approach.

Figs. 4.7 and 4.8 show the rolling risk and return of the four different portfolios. The key result that we observe is that the L-moments approach consistently has a much higher risk than the other strategies. It also produces a lower return. This result is consistent with our discussion previously. Furthermore, we note that the inclusion of Random Matrix Theory reduces the risk, and increases the return of the L-moments strategy, making its performance more comparable to the sample moments strategy, both with and without RMT.

Since the risk-return characteristics of the L-moments strategy appear to deviate substantially from the other three strategies, we investigate this further by comput-

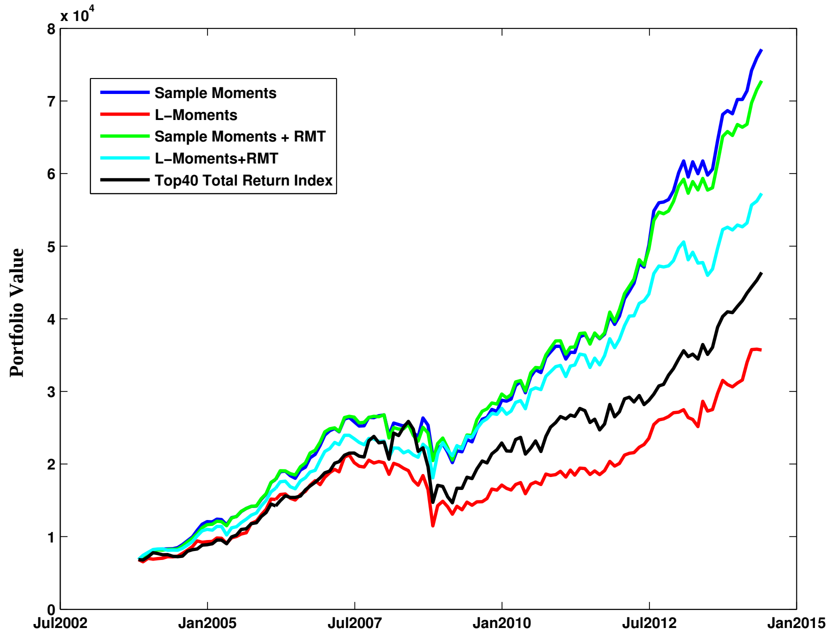


Fig. 4.6: Evolution of the portfolio value of the four different strategies.

ing the active share and diversification of the strategies. We compute diversification using two different measues: Meucci's Effective Number of Bets (ENB) and a simple ratio of the number of stocks that are invested in out of those that are available at any time in the investment universe. We refer to the latter measure as the Effective Size of the portfolio. These two measures of diversification are plotted in Figs. 4.9 and 4.10 respectively. The active share of a portfolio is a measure of how different it is, in terms of its weights and constituents, to some benchmark. We set the benchmark as the conventional sample moments portfolio, and plot the active share of the other three strategies in Fig. 4.11.

The striking result that we observe is that the active share and both diversification measures are much more volatile when using the L-moments strategy compared to the other strategies. In particular, the composition of the portfolio tends to fluctuate significantly and frequently, varying from a composition that is very close to the sample moments portfolio (for example, an active share of around 0.2) to a composition that is entirely different to the benchmark (for example, an active share of close to 1.0). Similarly, the Effective Size of the L-moments strategy varies very frequently. Fig. 4.10 illustrates that the behaviour of the L-moments strategy is due to the largely volatile nature of portfolio weights. On average, at each point in time there are approximately 40 stocks available to invest in, in the investment universe. We see from Fig 4.10 that, very frequently, the proportion of stocks in-

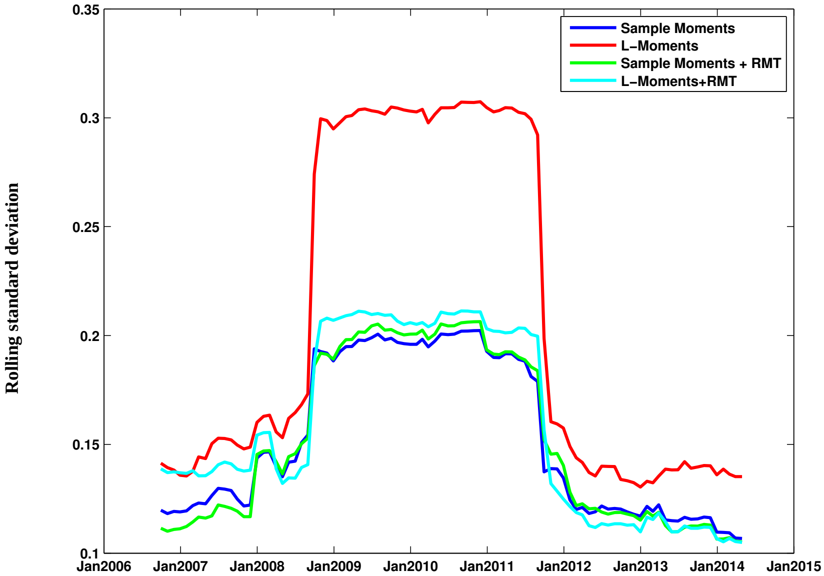


Fig. 4.7: Rolling risk of the four different strategies.

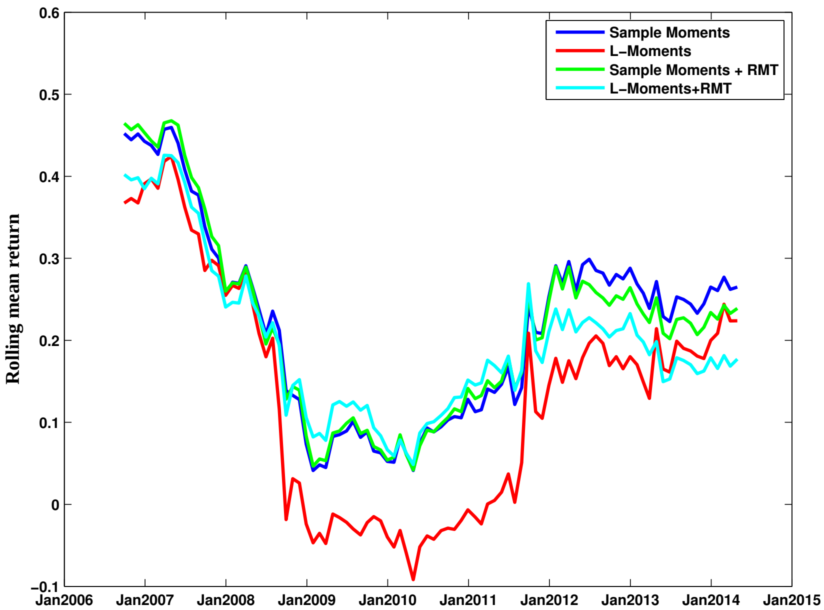


Fig. 4.8: Rolling return of the four different strategies.

vested in drops very close to zero, e.g. investing in only 1 or 2 stocks, and then immediately jumps to a very high number, e.g. investing in 14-16 stocks, only to return to a small number of stocks, e.g. 2-3 stocks again, and this volatile behaviour continues. We also note that it is the L-moments strategy in particular that tends to

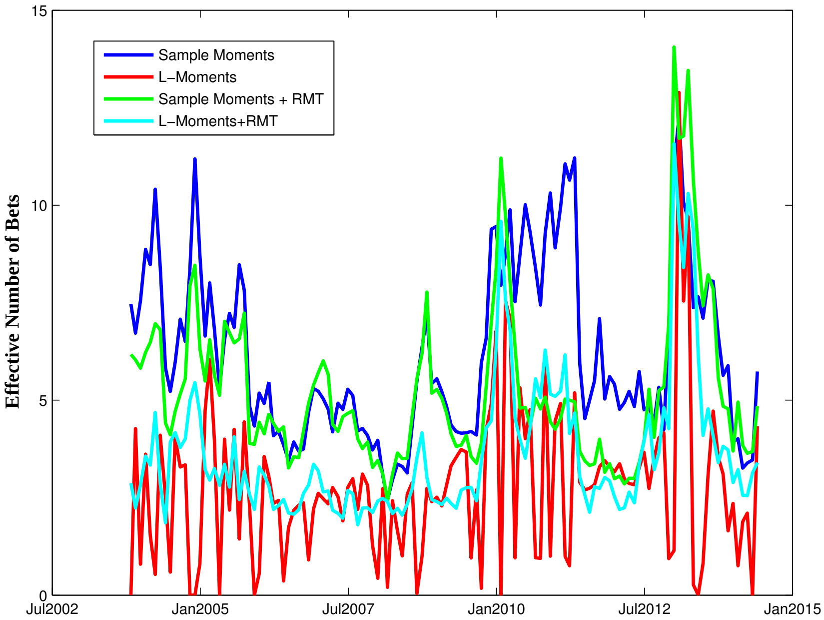


Fig. 4.9: Diversification of the four different strategies, as measured by the Effective Number of Bets.

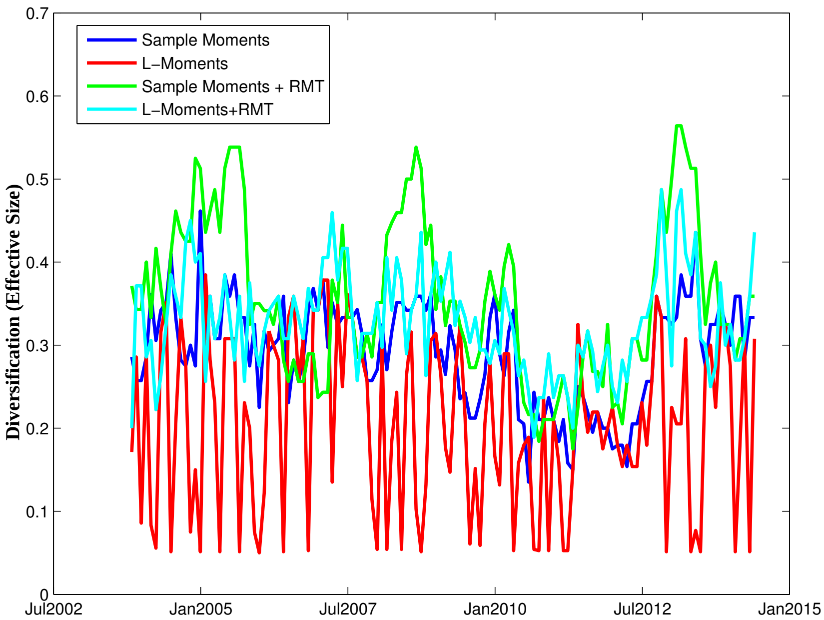


Fig. 4.10: Diversification of the four different strategies, as measured by the number of stocks that are invested in, out of those that are available, i.e. Effective Size.

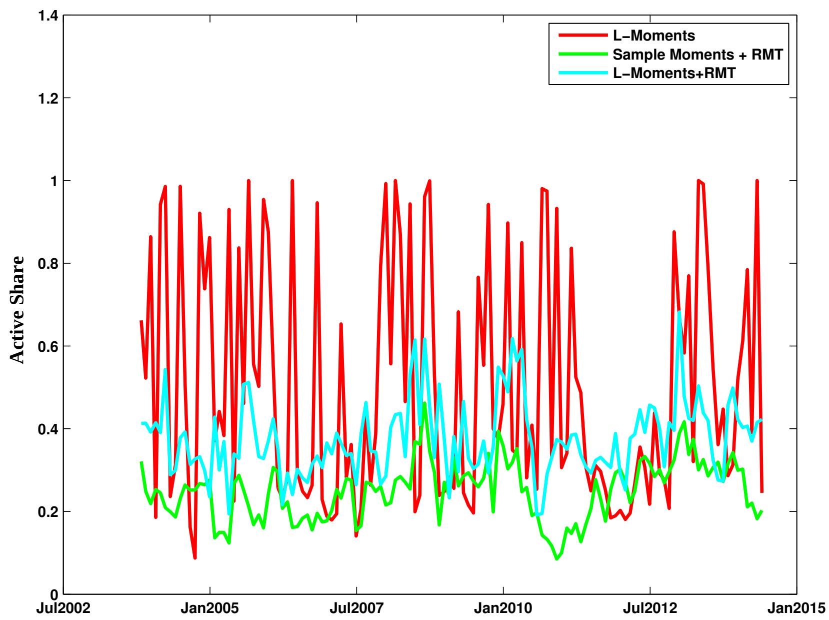


Fig. 4.11: Active Share of three of the strategies, measured relative to the conventional sample moments method.

take very few number of stocks into account compared to the other strategies. We compare this result to the Effective Number of Bets, and again see in this case that the L-moments approach appears to have the poorest diversification. To investigate this further, in Figs. 4.12 and 4.14, we have plotted the active share of the portfolio against the two diversification measures, specifically for the L-moments strategy.

We note that the active share appears to be negatively correlated with both diversification measures. Furthermore, we sort the active share in ascending order, and plot it along with the corresponsing diversification measures. We note that from Figs. 4.13 and 4.15, there is indeed a negative relationship. We note, however, that this negative relationship is stronger between the active share and the Effective Size, than between the active share and the Effective Number of Bets. This is because the strong negative relationship between the active share and Effective Size must hold by definition, whereas the ENB is less strictly related to the active share of the portfolio. Nonetheless, these results indicate that as the L-moments strategy deviates further from the benchmark, it becomes extremely poorly diversified as a result of being concentrated in a small number of stocks. We note that this result is significant in practice, as investors would be unlikely to invest in such concentrated portfolios. We thus comment that the L-moments strategy, on its own, without RMT, appears to be an unfeasible strategy to be used in practice.

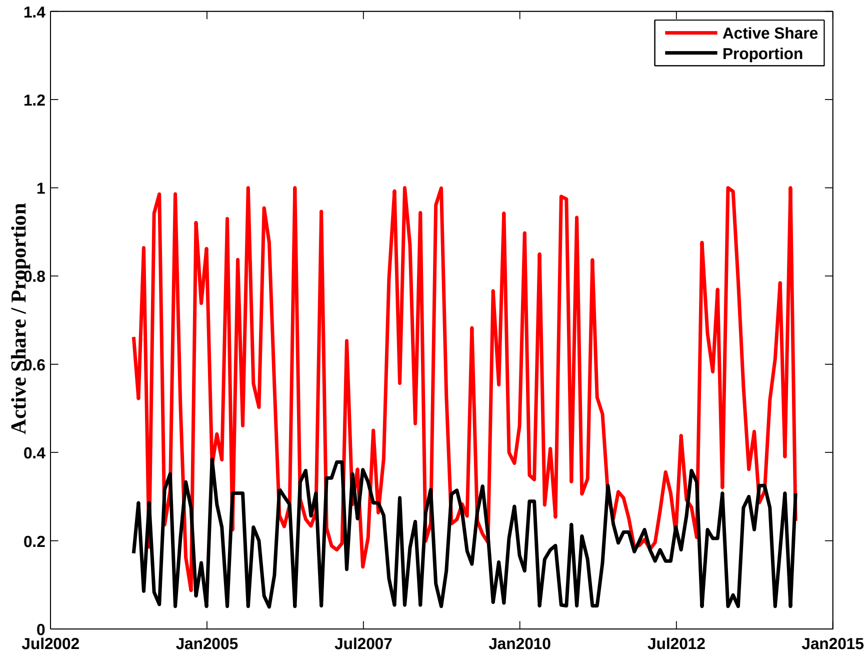


Fig. 4.12: Active share of the L-moments strategy along with the Effective Size/proportion of stocks that are invested in.

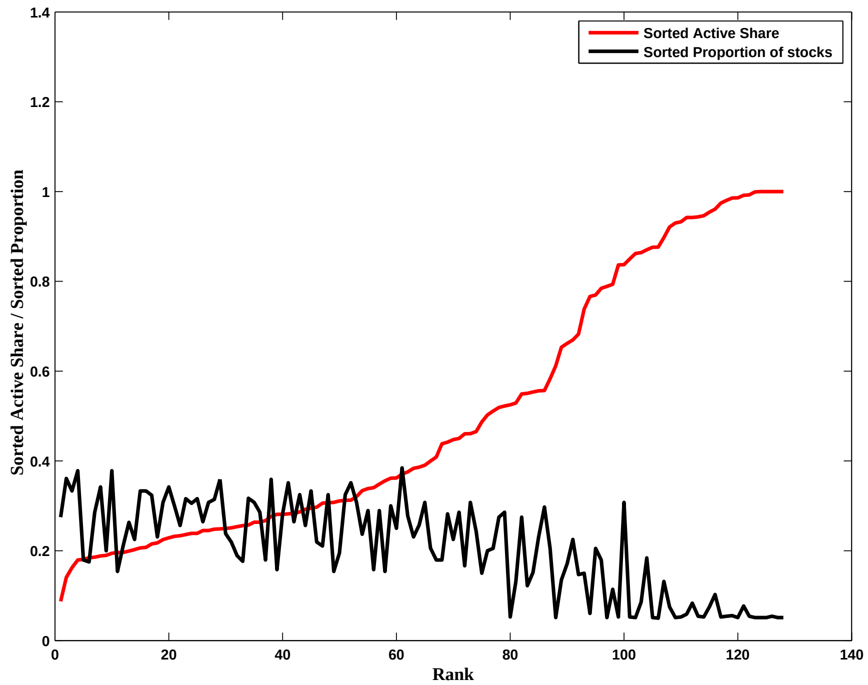


Fig. 4.13: Sorted active share and sorted Effective Size of the L-moments strategy.

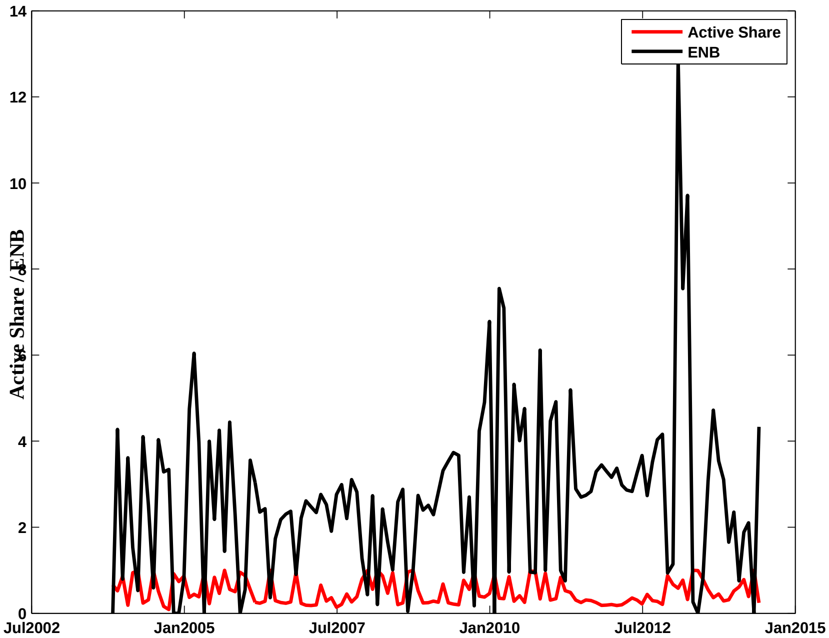


Fig. 4.14: Active share of the L-moments strategy along with the Effective Number of Bets.

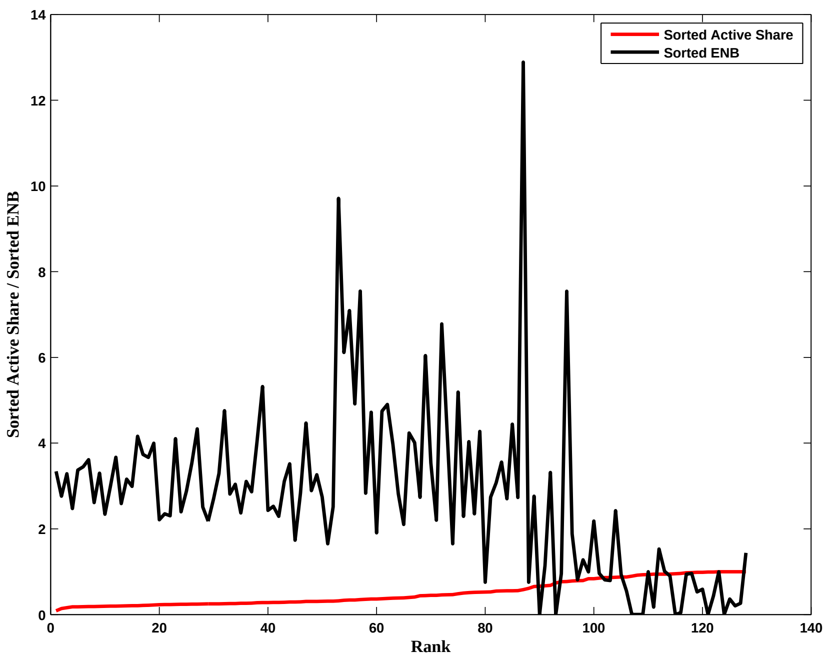


Fig. 4.15: Sorted active share and sorted Effective Number of Bets of the L-moments strategy.

4.3.2 Comparison of SA results to US Data

We compare our results to those of Yanou (2013). Yanou (2013) incorporates Random Matrix Theory into the sample moments and L-moments methods and compares the performance of the global minimum variance portfolio under these two strategies. He does not consider the strategies of sample moments and L-moments on their own, without RMT, as we have done in our analysis.

Yanou (2013) runs the optimisation on US data, in particular on two databases. He runs the out-of-sample optimisation process on 64 stocks from the S&P500 universe, using data sampled on a weekly frequency from 29-May-1981 to 11-Apr-2008. The second database is based on 78 stocks from the NYSE universe, again sampled on a weekly frequency over the period 25-May-1981 to 14-Apr-2008.

Using both of these databases, and a variety of estimation windows and holding periods, Yanou (2013) runs the out-of-sample optimisation process and then characterises the portfolio performance of the two different strategies, based on their performance, measured in terms of the Sharpe ratio. Across both databases, and a variety of estimation windows and holding periods, Yanou (2013) finds that using the sample moments method along with RMT results in portfolios that have a lower volatility than the L-moments method along with RMT. However, despite having a larger volatility, the L-moments strategy with RMT outperforms the sample moments strategy, with RMT, as it has a greater return. This key result is the opposite to the result that we have observed when using SA data.

In order to discuss this further, we perform our own analysis on Yanou's (2013) "Database 1" and aim to replicate his results. In addition to replicating the results, we perform a further analysis in terms of all four strategies (and not just the two that Yanou (2013) considered).

We perform the out-of-sample optimisation process using an estimation window of 510 weeks, and a rolling period of 4 weeks. The results are displayed in Table 4.8 below.

From the table, we observe that we have been able to replicate Yanou's (2013) results. The inclusion of RMT to the L-moments strategy does indeed outperform the sample moments strategy with RMT in terms of both the Sharpe and Sortino ratio. Other than this result, the others appear to be consistent with the results from our SA data.

In particular, we note that the inclusion of RMT to the sample moments strategy actually results in poorer performance, as there is a higher risk and a lower return, resulting in a lower Sharpe ratio. Furthermore we also observe that the L-moments strategy does extremely poorly – it has a much higher return and a lower annualised mean return result in a Sharpe ratio that is much smaller than both

Tab. 4.8: Portfolio performance based on full US weekly data, using a 520 week estimation window and 4 week rolling period.

	S-Moment	S-Moment + RMT	L-Moment	L-Moment + RMT
ASD	0.1040	0.1032	0.1440	0.1212
AMR	0.0722	0.0698	0.0629	0.0911
Sharpe	0.6941	0.6766	0.4369	0.75128
Sortino	1.6576	1.6611	1.0889	1.7311

the sample moments method, as well as the sample moments method with RMT included. However, we then note that the inclusion of RMT to the L-moments strategy enhances the performance of the L-moments strategy. This is similar to the results in the SA market, however the striking result here is that in the US market, the inclusion of RMT to the L-moments approach actually results in a portfolio strategy that outperforms the conventional sample moments method in terms of both the Sharpe and Sortino ratios.

Chapter 5

Conclusion

In this dissertation, we have analysed several different methods of portfolio selection. In particular we have discussed the issues with obtaining robust and reliable estimates of both the covariance matrix of returns and the mean vector of returns. Since the mean vector of returns is more difficult to estimate and has greater estimation errors, we have chosen to focus on the construction of the global minimum variance portfolio, which requires estimation of only the variances and covariances of stock returns. In computing the variance-covariance matrix, we note that due to limited data available, it is often difficult to obtain robust estimates of the covariance matrix. Hence, we consider two approaches to obtaining robust estimates of the covriance matrix - using the L-moments method approach, and using the filtering techniques proposed by Random Matrix Theory.

We thus perform optimisations on four different strategies - the sample moments method, with and without RMT, and the L-moments method, with and without RMT - and we compare the performance of the different strategies, relative to the benchmark, which is the conventional sample moments method.

We run the analysis on SA market data and find that the L-moments strategy appears to be extremely volatile in terms of the portfolio weights and the number of stocks that it invests in. As a result of this volatility, we find that this strategy has a much greater risk than the other strategies and it also has a much lower return. However the incorporation of RMT into this strategy greatly enhances its stability and improves the performance substantially by reducing the risk, and increasing the return, resulting in greater Sharpe and Sortino ratios.

We observe that the inclusion of RMT to the sample moments method does not particularly enhance the portfolio performance, and in many cases actually reduces the Sharpe and Sortino ratios. Furthermore, although the L-moments strategy inclusive of RMT performs well, it is not able to outperform the standard conventional sample moments method, which we have considered to be our benchmark.

This result is however different when we consider data from the US market.

Repeating the analysis on data obtained from the US market, we observe that all the results translate, with the only difference being that the inclusion of RMT to the L-moments strategy enhances the performance of the portfolio to the extent that it does better than the benchmark conventional sample moments method.

We conclude that the L-moments strategy, on its own, results in particularly volatile portfolio holdings. In particular, very often the number of stocks held are quite small, e.g. 2 stocks out of a universe of 40, and very often this is unreasonable as investment managers will, in practice, not be willing to invest in such a poorly diversified manner. However, we note that inclusion of the RMT texhniques adds significant value to the L-moments method, greatly enhacing the stability of the portfolio weights, as well as improving the diversification of the constructed portfolios. Similarly, in the case of the sample moments method, we note that the techniques of RMT do aid the stability, resulting in a portfolio with a lower risk, however this does not necessarily translate into better performance.

We therefore propose that should the L-moments technique be used in portfolio selection, it should be done in conjunction with Random Matrix Theory. In the case of the sample moments method, RMT may or may not be incorporated as it does not offer substantial value. Furthermore, we also note that despite that the fact that RMT does indeed enhance returns on the L-moments strategy, it is not guaranteed to outperform the conventional sample moments strategy. We thus conclude that despite the analysis done on the techniques of RMT and L-moments, the sample moments method performs relatively well and may continue to be used as a reasonable estimate of the covariance matrix, and used in portfolio selection as well as in the measurement of risk-return characteristics of portfolios.

5.1 Potential further research

This dissertation has analysed the use of L-moments to construct the covariance matrix of stock returns, and has also indicated the benefits of incorporating the techniques of Random Matrix Theory from the perspective of stock selection. This research, however, has not delved into the statistical reasoning behind the performance of the L-moments method, and why the RMT techniques appear to add significant value to this method of estimating the covariance matrix. Further potential research may therefore be focused on understanding the mathematical and statistical reasoning behind the results described in this dissertation. Furthermore, it may also be possible to analyse the effect of RMT on other approaches of estimating the covariance matrix - for example, incorporating RMT with estimation of covariance matrices using factor models, shrinkage models or Bayesian estimators.

Bibliography

Amenc, N., Stoyanov, S. and Goltz, F. (2011). A long horizon perspective on the cross-sectional risk-return relationship in equity markets, *EDHEC Risk Institute* .

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns, *The Journal of Finance* **61**(1): 259–300.

Ang, A., Hodrick, R. J., Xing, Y. and Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence, *Journal* **61**(1): 259–300.

Bai, J. and Shi, S. (2011). Estimating High Dimensional Covariance Matrices and its applications, *Annals of Economics and Finance* **12**(2): 119–215.

Baker, M., Bradley, B. and Wurgler, J. (2010). Benchmarks as Limits to Arbitrage: Understanding the Low Volatility Anomaly.

Baker, N. L. and Haugen, R. A. (2012). Low risk stocks outperform within all observable markets of the world, *Manuscript. http://papers.ssrn.com/sol3/papers.cfm* .

Bali, T. and Cakici, N. (2008). Idiosyncratic volatility and the cross-section of expected returns, *Journal of Financial and Quantitative Analysis* **43**: 29–58.

Bouchaud, J.-P. and Potters, M. (2000). Theory of financial risks.

Chan, L. K. (1967). On a characterization of distributions by expected values of extreme order statistics, *American Mathematical Monthly* pp. 950–951.

Chopra, V. K. and Ziemba, W. T. (1993). The effects of errors in means, variances and covariances on optimal portfolio choice, *The Journal of Portfolio Management* **19**: 6–11.

Clarke, R., de Silva, H. and Thorley, S. (2011). Minimum variance portfolio composition, *The Journal of Portfolio Management* **37**(2).

Conlon, T., Ruskin, H. J. and Crane, M. (2007). Random matrix theory and fund of funds portfolio optimisation, *Physica A: Statistical Mechanics and its applications* **382**(2): 565–576.

David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, 3 edn, Wiley: New York.

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

Downton, F. (1966). Linear estimates with polynomial coefficients, *Biometrika* pp. 129–141.

Dyson, F. J. (1971). Distribution of eigenvalues for a class of real symmetric matrices, *Technical report*, Inst. for Advanced Study, Princeton, NJ.

Dyson, F. J. and Mehta, M. L. (1963). Statistical theory of the energy levels of complex systems. IV, *Journal of Mathematical Physics* **4**(5): 701–712.

Edelman, A. (1988). Eigenvalues and condition numbers of random matrices, *SIAM Journal on Matrix Analysis and Applications* **9**(4): 543–560.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta, *Journal of Financial Economics* **111**(1): 1–25.

Gatheral, J. (2008). Random matrix theory and covariance estimation, *NYU Courant Institute Algorithmic Trading Conference*.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C. and Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form, *Water Resources Research* **15**(5): 1049–1054.

Haugen, R. A. and Heins, A. J. (1972). On the evidence supporting the existence of risk premiums in the capital market, *Working paper, University of Wisconsin-Madison. Available at SSRN 1783797* .

Haugen, R. A. and Heins, A. J. (1975). Risk and the Rate of Return on Financial Assets: Some Old Wine in New Bottles, *Journal of Financial and Quantitative Analysis* **10**(5).

Hosking, J. R. M. (1986). The theory of probability weighted moments, *Research Report RC* **12210**: 160.

Hosking, J. R. M. (1989). Some theoretical results concerning L-moments, *Research Report RC* **14492**: 9.

Hosking, J. R. M. (1990). L-moments: Analysis and Estimation of Distributions using Linear Combinations of Order Statistics, *Journal of the Royal Statistical Society, Series B (Methodological)* **52**(1): 105–124.

Hosking, J. R. M. and Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution, *Technometrics* **29**(3): 339–349.

Hosking, J. R. M., Wallis, J. R. and Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments, *Technometrics* **27**(3): 251–261.

Hsu, J., Kudoh, H. and Yamada, T. (2012). When sell-side analysts meet high-volatility stocks: an alternative explanation for the low-volatility puzzle.

Hsu, J. and Li, F. (2013). Low-Volatility Investing, *The Journal of Index Investing* **4**(2): 67–72.

Jagannathan, R. and Ma, T. (2003). Risk reduction in large portfolios: why imposing the wrong constraints helps, *Journal of Finance* **54**: 1651–1683.

Konheim, A. G. (1971). A note on order statistics, *American Mathematical Monthly* pp. 524–524.

Laloux, L., Bouchaud, J.-P. and Potters, M. (2005). Financial applications of random matrix theory: Old laces and new pieces, *arXiv preprint physics/0507111* .

Laloux, L., Cizeau, P., Bouchaud, J.-P. and Potters, M. (1999). Noise dressing of financial correlation matrices, *Physical Review Letters* **83**(7): 1467–1470.

Laloux, L., Cizeau, P., Potters, M. and Bouchad, J.-P. (2000). Random matrix theory and financial correlations, *International Journal of Theoretical and Applied Finance* **3**(3): 391–397.

Ledoit, O. and Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *Journal of Empirical Finance* **10**: 603–621.

Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* **88**(2): 365–411.

Mallows, C. L. (1973). Bounds on distribution functions in terms of expectations of order-statistics, *The Annals of Probability* pp. 297–303.

Marčenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices, *Sbornik: Mathematics* **1**(4): 457–483.

Markowitz, H. (1952). Portfolio Selection, *The Journal of Finance* **7**(1): 77–91.

Mehta, M. L. (1991). *Random Matrices*, Academic Press, New York.

Merton, R. C. (1980). On estimating the expected return on the market: an epxloratory investigation, *Journal of Financial Economics* **8**: 323–361.

Meucci, A. (2010). Managing diversification.

Michaud, R. O. (1989). The Markowitx optimization enigma: Is 'optimised' optimal?, *Financial Analysts Journal* **45**(1): 31–42.

Perez, N. H., Menendez, S. C. and Seco, L. (2003). A theoretical comparison between moments and L-moments, *Unpublished paper retrieved from http://risklab.erin.utoronto.ca/members.htm* .

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., Guhr, T. and Stanley, H. E. (2002). Random matrix approach to cross correlations in financial data, *Physical Review E* **65**(6).

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing, *Journal of Economic Theory* **13**: 341–360.

Sandoval, L. and Franca, I. D. P. (2012). Correlation of financial markets in times of crisis, *Physica A: Statistical Mechanics and its Applications* **391**(1): 187–208.

Sengupta, A. M. and Mitra, P. P. (1999). Distributions of singular values for some random matrices, *Physical Review E* **60**(3): 3389–3392.

Serfling, R. and Xiao, P. (2007). A Contribution to Multivariate L-Moments: L-Comoment Matrices, *Journal of Multivariate Analysis* **9**(3): 1765–1781.

Sharifi, S., Crane, M., Shamaie, A. and Ruskin, H. J. (2004). Random matrix theory for portfolio optimization: a stability approach, *Physica A: Statistical Mechanics and its Applications* **335**(3): 629–643.

Sharpe, W. (1964). Capital Asset Prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* **19**(3): 425–442.

Sillitto, G. P. (1951). Interrelations between certain linear systematic statistics of samples from any continuous population, *Biometrika* pp. 377–382.

Sillitto, G. P. (1969). Derivation of approximants to the inverse distribution function of a continuous univariate population from the order statistics of a sample, *Biometrika* **56**(3): 641–650.

Stoyanov, S. and Goltz, F. (2011). Advantages and shortcomings of minimum variance portfolios, *EDHEC Risk Institute* .

Yang, S. (1977). General disribution theory of the concomitants of order statistics, *The Annals of Statistics* **5**(5): 996–1002.

Yanou, G. (2013). Extension of the random matrix theory to the L-moments for robust portfolio selection, *Quantitative Finance* **13**(10): 1653–1673.

Appendix A

Yanou's (2013) US database of stocks

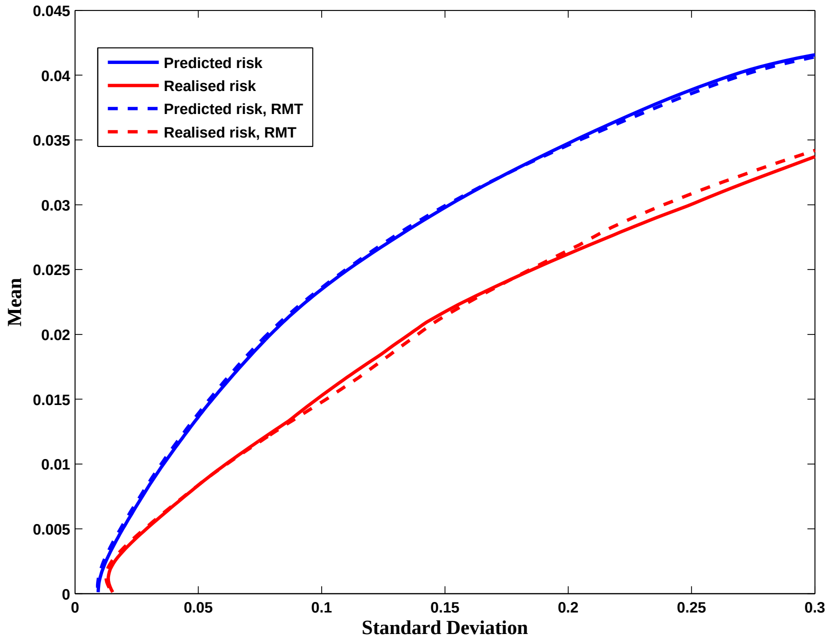
The following is the list of 64 stocks that Yanou (2013) takes into account in his universe called "Database 1". These 64 stocks are taken from the S&P500 universe.

Constell Energy
Integrys Energy Group
Xcel Energy
Duke Energy
Publ Serv Enter
Souther
Progress Energy
FPL Group
American Elec Power
Smith intl
Rowan Companies
Haliburton
Questar
Apache
Noble Energy
Conocophillips
Murphy Oil
Schlumberger
HEss
Occidental
Exxon Mobile
Chevron
Procter & Gamble
Bristol Myers
HJ Heinz
Eli Lily
Merck & Co
Abbott Labs
Pfizer
Johnson & Johnson
American Express
Adv Micro Dev
IBM
Corning
Molex
JP Morgan Chase
Hewlett Packard
Teradyne
Natl Semiconduct
Merrill Lynch
Motorola
Analog Devices
Texas Instruments
Marsh & Mclenann
Lennar Class A
AON
Lincoln Natl
Torchmark
Centex
Mellon Bank of NY
Wells Fargo
BOA
General Mills
Donnelley Sons
New York Times
Washington Post
Gannett
Masco
Campbell Soup
Conagra Foods
Wendy's Intl
Pulte Homes
Varian Medical
Hershey Co.

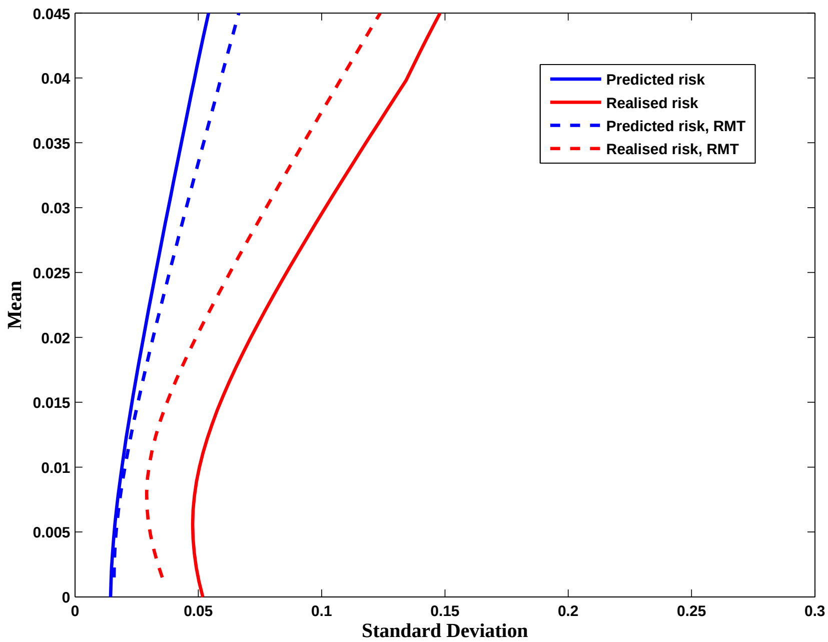
Appendix B

Additional plots of RMT analysis on efficient frontier

The plots on the following page display the efficient frontiers illustrating the in-sample and out-of-sample risk (blue and red, respectively), both when RMT has been incorporated (dashed lines) and when it has not (solid lines). Panels (a) and (b) in each of the plots contain the same data, except that the data used in panel (a) has been sampled daily, while that for panel (b) has been sampled weekly. The results of these plots are discussed in Section 4.2.

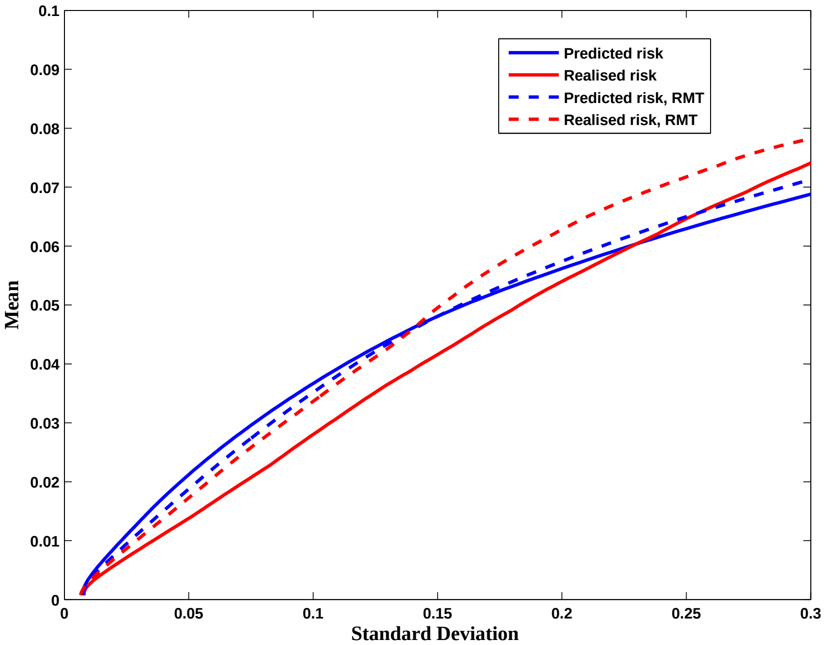


(a) Efficient frontiers plotted using data subset B, containing daily data over the period 02-Jan-2007 to 30-Dec-2009.

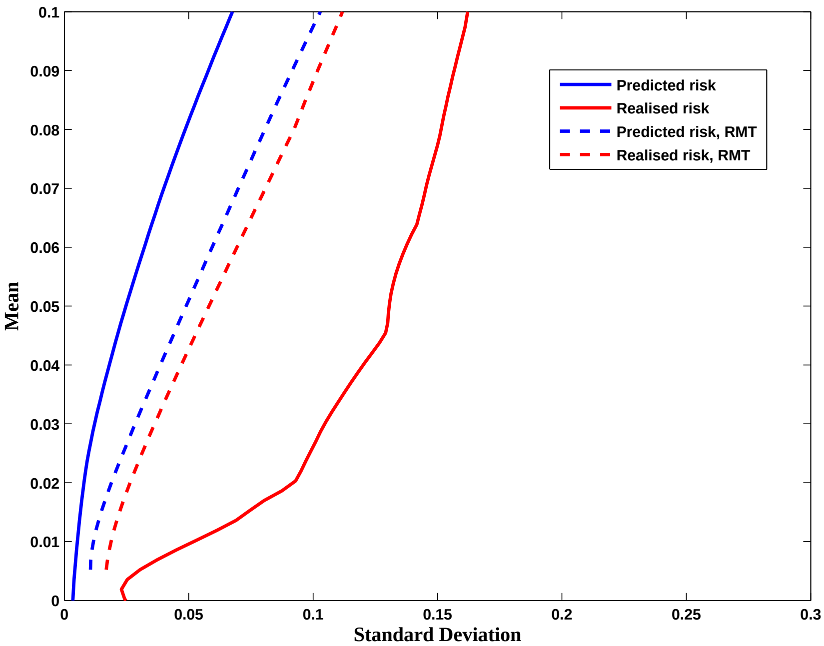


(b) Efficient frontiers plotted using data subset E, containing weekly data over the period 02-Jan-2007 to 30-Dec-2009.

Fig. B.1: Efficient frontiers plotted using data subsets B (panel (a)) and E (panel (b)), displaying the in-sample and out-of-sample risk, with and without RMT.



(a) Efficient frontiers plotted using data subset C, containing daily data over the period 02-Jan-2013 to 27-Jun -2014.



(b) Efficient frontiers plotted using data subset F, containing weekly data over the period 02-Jan-2013 to 27-Jun -2014.

Fig. B.2: Efficient frontiers plotted using data subsets C (panel (a)) and F (panel (b)), displaying the in-sample and out-of-sample risk, with and without RMT.