

UNIVERSITY OF CAPE TOWN

MASTERS THESIS

Robust portfolio construction: Using Resampled Efficiency in Combination with Covariance Shrinkage

Author:
James COMBRINK

Supervisor:
Professor David
BRADFIELD



*A thesis submitted in partial fulfilment of the requirements
for the degree of Masters in Advanced Analytics and Data Sciences*

in the

Department of Statistical Sciences



August 16, 2017

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

Declaration of Authorship

I, James COMBRINK, declare that this thesis titled, “Robust portfolio construction: Using Resampled Efficiency in Combination with Covariance Shrinkage” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

UNIVERSITY OF CAPE TOWN

Abstract

Science Faculty
Department of Statistical Sciences

Masters in Advanced Analytics and Data Sciences

Robust portfolio construction: Using Resampled Efficiency in Combination with Covariance Shrinkage

by James COMBRINK

The thesis considers the general area of robust portfolio construction. In particular the thesis considers two techniques in this area that aim to improve portfolio construction, and consequently portfolio performance. The first technique focusses on estimation error in the sample covariance (one of portfolio optimisation inputs). In particular shrinkage techniques applied to the sample covariance matrix are considered and the merits thereof are assessed. The second technique considered in the thesis focusses on the portfolio construction/optimisation process itself. Here the thesis adopted the 'resampled efficiency' proposal of Michaud (1989) which utilises Monte Carlo simulation from the sampled distribution to generate a range of resampled efficient frontiers. Thereafter the thesis assesses the merits of combining these two techniques in the portfolio construction process.

Portfolios are constructed using a quadratic programming algorithm requiring two inputs: (i) expected returns; and (ii) cross-sectional behaviour and individual risk (the covariance matrix). The output is a set of 'optimal' investment weights, one per each share whose returns were fed into the algorithm. This thesis looks at identifying and removing avoidable risk through a statistical robustification of the algorithms and attempting to improve upon the 'optimal' weights provided by the algorithms. The assessment of performance is done by comparing the out-of-period results with standard optimisation results, which are highly sensitive and prone to sampling-error and extreme weightings.

The methodology looks at applying various shrinkage techniques onto the historical covariance matrix; and then taking a resampling portfolio optimisation approach using the shrunken matrix. We use Monte-Carlo simulation techniques to replicate sets of statistically equivalent portfolios, find optimal weightings for each; and then through aggregation of these reduce the sensitivity to the historical time-series anomalies. We also consider the trade-off between sampling-error and specification-error of models.

Acknowledgements

SUPERVISOR Professor David Bradfield

INSTITUTION University Of Cape Town

SPONSOR An NRF funded bursary, through Statistics South Africa



Contents

Declaration of Authorship	i
Abstract	ii
Acknowledgements	iii
Financial Background	xiv
0.1 Risk and Return	xiv
0.2 The Efficiency Frontier	xiv
0.3 Sharpe Ratio	xv
0.4 Quadratic Programming Problem	xv
0.4.1 Standard form Quadratic Programming (QP) Problem	xv
0.4.2 Short-sale Constrained Optimisation	xvi
0.4.3 L1-Norm Weight Constrained Optimisation	xvi
0.4.4 L2-Norm Weight Constrained Optimisation	xvii
1 Introduction	1
1.1 Background	1
1.2 Overview	1
1.2.1 Problem Statement	1
1.2.2 Outline of Dissertation	2
1.3 The Efficient Frontier	2
1.3.1 An Instance of the Efficient Frontier	2
1.3.2 Portfolio Performance	3
1.3.3 Sampling Error	3
1.4 Data	4
1.5 Resampled Efficiency	5
1.5.1 Synopsis	5
1.6 Covariance Shrinkage	5
1.6.1 Synopsis	5
2 Literature Review	7
2.1 Modern Portfolio Theory	7
2.2 Sensitivity of MV Portfolios	7
2.3 Diversification	8
2.4 Approaches to Sampling Error	8
2.4.1 Minimum-Variance Portfolio	9
2.4.2 Mean Variance Portfolio	9
2.4.3 Resampled Efficiency Frontier	9
2.4.4 Shrinkage	10
2.5 Model Specification Error	10
2.5.1 Sample Error - Specification Error Trade Off	10
2.5.2 Equivalent Regression Procedures	11
2.5.3 Direct Extensions of the use of Shrinkage	12

2.6	Evaluation of Portfolios	12
2.7	Practicality of Theoretical Research	13
3	Theory and Methodology	14
3.1	The Efficient Frontier without Augmentation	14
3.2	Parameter Estimation	15
3.2.1	Estimation Uncertainty	15
3.2.2	The Bias-Variance Trade Off	15
3.2.3	Backtesting	16
3.3	Reduction of Sample Volatility	16
3.3.1	Resampled Efficiency	16
3.3.2	Resampling Methodology	16
3.3.3	Theoretical Usage	17
3.3.4	Shrinkage	17
3.3.5	Simple Sample Estimators	18
	i: Historical Sample Covariance (SC)	18
	ii: Diagonal Covariance Model (DC)	19
3.3.6	Shrinkage Estimators	19
	i: Default (MED / SS)	19
	ii: Average Correlation Model (ACM)	19
	iii: Constant Correlation Model (CCM)	20
	iv: Principal Component Analysis (PCA)	20
	v: Single Index Model (SIM)	20
3.3.7	Dimension Retaining Techniques	21
3.3.8	Direct Dimension Reduction Techniques	22
3.3.9	Blended Models	22
	Blend i: SC; DC; CCM	22
	Blend ii: SC; DC; PCA	22
	Blend iii: SC; DC; SIM	22
	Blend iv: SC; CCM; PCA	22
3.3.10	Parameter Estimation from Shrinkage Loss Function	23
3.4	Overview of Resampling and Shrinkage Applications	23
3.4.1	Perfect Information Coefficient	23
3.4.2	Construction of the Resampled Frontier	26
	Optimality out-of-Sample	30
3.4.3	Vector Algebra: Generating correlated Series	31
4	Pre-Shrinkage Results	33
4.1	Overview of Chapter	33
4.2	Initial Returns and Risk Comparative	33
4.2.1	Longitudinal Risk Behaviour of Resampled Portfolios Total Risk of Portfolios	36
4.2.2	Reduction of the IC	37
4.2.3	Practical Application	37
5	Shrinkage of Covariance Matrix	44
5.1	Application of Shrinkage	44
5.1.1	Shrinkage Factor	44
5.1.2	Historical Optimal Shrinkage Factor	46
5.2	Application of Shrinkage Techniques	47
5.2.1	Shrinkage onto Mean-Variance Portfolios	47

5.2.2	Cumulative Returns under Shrinkage	48
5.2.3	Annualised Risk under Shrinkage	53
5.2.4	Number of Constituents under Shrinkage	58
5.2.5	Shrinkage onto Minimum Variance Portfolios	61
5.2.6	Cumulative Returns under Shrinkage	62
5.2.7	Annualised Risk under Shrinkage	67
5.2.8	Number of Constituents under Shrinkage	73
6	Sensistivity Analysis: Assessment of IC Level	76
6.1	Return Comparative	77
6.2	Risk Comparative	82
6.3	Overall Assessment of Possible IC Confounding	85
7	Conclusion	87
7.1	Summary of Shrinkage Performances	87
7.2	Findings	91
7.3	Shortcomings and Further Research	92
7.3.1	IC level	92
7.3.2	IC Dependence Modelling	92
7.3.3	Transaction Costs	92
7.3.4	Shrinkage Techniques: SIM and Blends	93
7.3.5	Shrinkage Techniques: DC	93
7.3.6	Pseudo Relationships	93
A	Data and External Inputs	98
A.1	Software	98
A.2	Data	98
B	Resampling Paper Replication in SA context	99
B.1	Identifying and dealing with survivor bias from <i>Datastream</i>	99

List of Figures

1.1	Markovitz Efficiency Frontier	3
1.2	Data required for efficiency frontier	4
3.1	Markovitz Efficiency Frontier	14
3.2	Efficiency Frontier with Perfect Information	24
3.3	Efficiency Frontier without Perfect Information	25
3.4	Resampled Efficiency Frontier I iteration	27
3.5	Resampled Efficiency Frontier V iterations	28
3.6	Resampled Efficiency Frontier D iterations	29
3.7	Resampled Efficiency Frontier D iterations 3	30
3.8	Vector Space of Correlated Returns	31
3.9	Vector Algebra of Correlated Returns	32
4.1	Comparison of Resampled Performance i	34
4.2	Comparison of Resampled Performance ii	35
4.3	Longitudinal Comparison of Total Risk	36
4.4	Longitudinal Comparison of Total Risk	37
4.5	Risk under Reduced IC	38
4.6	Return under Reduced IC	39
4.7	Constituents under Reduced IC	40
5.1	Eigenvalues of PCA	45
5.2	Longitudinal Shrinkage α	46
5.3	Longitudinal Shrinkage α	47
5.4	Returns Baseline	48
5.5	Returns under ACM	49
5.6	Returns under CCM	50
5.7	Returns under PCA	51
5.8	Annualised Risk Baseline	53
5.9	Annualised Risk under ACM	54
5.10	Annualised Risk under CCM	55
5.11	Annualised Risk under PCA	56
5.12	Number of Constituents Baseline	58
5.13	Number of Stocks under ACM	59
5.14	Number of Stocks under CCM	60
5.15	Number of Stocks under PCA	61
5.16	Returns Baseline	63
5.17	Returns under ACM	64
5.18	Returns under CCM	65
5.19	Returns under PCA	66
5.20	Annualised Risk Baseline	68
5.21	Annualised Risk under ACM	69
5.22	Annualised Risk under CCM	70
5.23	Annualised Risk under PCA	71

5.24	Number of Stocks under ACM	73
5.25	Number of Stocks under CCM	74
5.26	Number of Stocks under PCA	75
6.1	Risk under Reduced IC	77
6.2	Risk under Reduced IC	78
6.3	Risk under Reduced IC	79
6.4	Risk under Reduced IC	82
6.5	Risk under Reduced IC	83
B.1	Longitudinal Cumulative Performance, Under Historical Returns	100

List of Tables

4.1	Annualised Risks June 2006 : March 2016	39
4.2	Cumulative Returns June 2006 : March 2016	40
4.3	Number of Stocks June 2006 : March 2016	41
4.4	Performance Analytics: Benchmark Performance	42
4.5	Performance Analytics: Benchmark Performance	43
5.1	Average Shrinkage to Covariance	47
5.2	Performance Rankings by Returns	48
5.3	ACM vs Median Shrinkage Returns with IC=0.05	49
5.4	CCM vs Median Shrinkage Returns with IC=0.05	50
5.5	PCA vs Median Shrinkage Returns with IC=0.05	51
5.6	Performance Analytics 1 with IC=0.05	52
5.7	Portfolio Rankings by Risk	53
5.8	ACM vs Median Shrinkage Risk with IC=0.05	54
5.9	CCM vs Median Shrinkage Risk with IC=0.05	55
5.10	PCA vs Median Shrinkage Risk with IC=0.05	56
5.11	Performance Analytics 2 with IC=0.05	57
5.12	ACM vs Median Shrinkage Constituents with IC=0.05	59
5.13	CCM vs Median Shrinkage Constituents with IC=0.05	60
5.14	PCA vs Median Shrinkage Constituents with IC=0.05	62
5.15	Portfolio Rankings by Returns	62
5.16	ACM vs Median Shrinkage Returns	64
5.17	CCM vs Median Shrinkage Returns	65
5.18	PCA vs Median Shrinkage Returns	66
5.19	Performance Analytics 3	67
5.20	Portfolio Rankings by Risk	68
5.21	ACM vs Median Shrinkage Risk	69
5.22	CCM vs Median Shrinkage Risk	70
5.23	CCM vs Median Shrinkage Risk	71
5.24	Performance Analytics 4	72
5.25	ACM vs Median Shrinkage Constituents	73
5.26	CCM vs Median Shrinkage Constituents	74
5.27	PCA vs Median Shrinkage Constituents	75
6.1	Portfolio Comparative of IC by Return	80
6.2	Portfolio Analytics 5	81
6.3	Portfolio Comparative of IC by Risk	84
6.4	Portfolio Analytics 6	85
7.1	MV Shrinkage Returns Narrative	87
7.2	Min Var Shrinkage Returns Narrative	88
7.3	MV Shrinkage Risks Narrative	89
7.4	Min Var Shrinkage Risks Narrative	89
7.5	MV Shrinkage Constituents Narrative	90

7.6 Min Var Shrinkage Constituents Narrative 90

List of Abbreviations

MV	Mean Variance
SR	Sharpe Ratio
ARIMA	Auto-Regressive Integrated Moving Average
ARCH	Auto-Regressive Conditional Heteroscedasticity
GARCH	Generalized Auto-Regressive Conditional Heteroscedasticity
IR	Information Ratio
IC	Information Coefficient
DC	Diagonal Covariance Matrix
SC	Historical Sample Covariance Matrix
F	Structured Covariance Matrix
PCA	Principal Component Analysis
ACM	Average Correlation Model
CCM	Constant Correlation Model
Var(*)	Variance of *
E(*)	Expected value (return) of *
d.p.	decimal place
min var	minimum variance
R	Resampling
S	Shrinkage

List of Symbols

r	risk free interest rate	
W	weights	$\in \mathcal{R}^n$
w_i	the i^{th} entry in W	
$\boldsymbol{\mu}$	return vector	$\in \mathcal{R}^n$
μ_i	the i^{th} entry in $\boldsymbol{\mu}$	
$\hat{\Sigma}$	sample covariance matrix	$\in (\mathcal{R}^n)^2$
$\hat{\sigma}_{i,j}$	sample covariance between shares i and j	
$\hat{\sigma}_i^2$	sample variance of share i	
X_t	share price at time t	
D_t	dividend paid over period: $[t:t+1]$	
R_t	return achieved over period: $[t:t+1]$	
\mathbf{r}_i	return vector achieved by share i	
$r_{i,t}$	return achieved by share i , over period: $[t:t+1]$	

Dedicated to those who never left my side, and never lost faith... My father Richard, who unconditionally encourages, supports, and pushes for brilliance. My mother Debbie, who will compliment and support me for a task as menial as putting a suit on; and who would sooner criticise the stove for daring to burn me, rather than criticise me for touching it the 5th time in a month.

My girlfriend, Thembelihle, who made the period bearable despite all pitfalls. My brother, Joshua; and my incredible friends Jonathan and Jason for the respect they have shown me and for their persistent lack of any doubt in me. My step mother Robyn, who's love is consistent and unfailing. My colleagues, Tim; Qobo; Raymond; Gciniwe; Thembani; Lipika who fast became my friends who fostered a genuine, engaging and productive environment. My baby sister Julia, because she is learning to read and loves reading her own name.

My eternal God, the Holy Trinity.

Financial Background

0.1 Risk and Return

Returns are calculated under the normal methodology:

$$R_t = \frac{X_t - X_{t-1}}{X_{t-1}} \quad (1)$$

where $X_i = \text{Price of Asset } X \text{ at time } i$

We use the historical returns only to forecast future returns, and the historical returns as a time-series to estimate the volatility of shares, as well as the correlation and covariances between two shares.

Note that the following properties apply to a portfolio constructed out of an aggregation of multiple shares, assuming independence:

$$E[R_p] = \sum_{i=1}^n E[R_i] \frac{w_i}{\sum_{j=1}^n w_j} \quad (2)$$

$$\sigma_p = \sum_{i=1}^n \sigma_i \left(\frac{w_i}{\sum_{j=1}^n w_j} \right)^2 \quad (3)$$

When the possibility for covariances between shares is allowed, (3) is extended to include the addition of all covariances as follows (for simplicity, we substitute in the definition that $\sum_{i=1}^n w_i = 1$), and note that the covariance of share x and itself is $\sigma_{x,x} = \sigma_x \sigma_x \rho_{x,x} = \sigma_x^2$:

$$\sigma_p^2 = \sum_{i=1}^n \sum_{j=1}^n \rho_{i,j} \sigma_i \sigma_j w_i w_j \quad (4)$$

0.2 The Efficiency Frontier

An efficient portfolio is a portfolio which offers the highest expected return for a defined level of risk and conversely; the lowest risk for a defined level of expected return. For instance, if one were investing within the 10 parent-sectors (such that $\sum_{i=1}^{10} w_i = 1$; and $w_i \geq 0$ for all i), each set of weightings would allow for an entirely new portfolio. For a portfolio j which holds risk level γ , if there exists no portfolio i with equivalent risk γ and a higher expected return, then portfolio j would be efficient. We refer to the set of 'optimal' portfolios as all portfolios for which as the same amount of risk is compensated more by the rewards; otherwise known as the set of Mean-Variance efficient portfolios, which are precisely the set of portfolios which constitute the efficiency frontier.

The efficient frontier, as well as all relevant coding, has been written by the author in *R_v_3.1.1*.

0.3 Sharpe Ratio

The Sharpe ratio is defined by Sharpe (1970) as the expected reward per unit of risk taken on in a portfolio. It is calculated as:

$$SR = \frac{\text{excess return above benchmark}}{\text{excess risk above benchmark}}$$

$$SR = \frac{E(R_p) - E(R_b)}{\sigma_p} \quad (5)$$

The higher the Sharpe ratio of a portfolio; the higher the return per unit risk, assuming the phenomenon experienced within the sample are reflective of the global behaviour.

0.4 Quadratic Programming Problem

There are multiple formulations of the objective function within the risk-return space. The four of interest are mentioned below for ease of reference. The philosophy of the quadratic programming problem is simply outlined as using the estimated returns and estimated covariance matrix to solve for the associated optimal investment as set of weight vectors, one vector per each level of λ , where λ is a penalising multiplicative constant of the risk.

We gain a continuous curve of optimal portfolios. These portfolios have a risk as the vector product of the weights (from the output of the quadratic programming problem) with the expected return: $\sum_{i=1}^n E[R_i]w_i$, and a risk as the sum of squared product of the weights (the output of the quadratic programming problem) with the associated shares' volatilities: $\sum_{i=1}^n \sum_{j=1}^n \rho_{i,j} \sigma_i \sigma_j w_i w_j$. Following Markowitz (1952), Sharpe (1970) and Best & Grauer (1991) we formulate the quadratic problem under various constraints. We gain a continuous convex curve of all such portfolios, ranging from the minimum-variance portfolio, through to the MV portfolio (the portfolio with the maximum Sharpe-ratio), up to the portfolio of a single constituent of only the share which has the highest expected return.

0.4.1 Standard form Quadratic Programming (QP) Problem

This is the basic portfolio selection strategy to solve for weight vector W such that the portfolio, whilst achieving a pre-specified expected return, is achieving the lowest variance of all portfolios which hold this expected return.

We solve the quadratic programming problem defined below, over a range of values for λ , each providing a unique point upon the efficiency frontier.

$$\max_W z = W' \mu - \lambda W' \hat{\Sigma} W \quad (6)$$

$$\text{subject to } W'e = 1$$

where

$$\begin{aligned} e &\in \mathcal{R}^n \\ W &\in \mathcal{R}^n && (\text{weights}) \\ \mu &\in \mathcal{R}^n && (\text{return vector}) \\ \hat{\Sigma} &\in (\mathcal{R}^n)^2 && (\text{covariance matrix}) \\ \lambda &\in \mathcal{R} \cap [0, \infty) \end{aligned}$$

Alternative versions of the quadratic problem are referred to within the project, thus we define the following alternatives to the traditional minimum variance problem.

0.4.2 Short-sale Constrained Optimisation

This form is adapted from the above, by forcing all investments to be positive, i.e. no short sales. This is used by Ledoit & Wolf's (2003) paper and is a specific case of the QP problem formulated by Best & Grauer (1991).

$$\max_W z = W'\mu - \lambda W'\hat{\Sigma}W \quad (7)$$

$$\begin{aligned} \text{subject to } W'1 &= 1 \\ w_i &\geq 0 && \forall i \end{aligned}$$

where

$$\begin{aligned} W &\in \mathcal{R}^n && (\text{weights}) \\ \mu &\in \mathcal{R}^n && (\text{return vector}) \\ \hat{\Sigma} &\in (\mathcal{R}^n)^2 && (\text{covariance matrix}) \\ \lambda &\in \mathcal{R} \cap [0, \infty) \end{aligned}$$

0.4.3 L1-Norm Weight Constrained Optimisation

The constrained portfolios are formulated by DeMiguel *et al* (2009), based off of the theoretical work of Ledoit & Wolf (2004).

$$\max_W z = W'\mu - \lambda W'\hat{\Sigma}W \quad (8)$$

$$\begin{aligned} \text{subject to} \quad & \|W\|_1 \leq \delta \\ & w_i \geq 0 \qquad \qquad \qquad \forall i \end{aligned}$$

where

$$\begin{aligned} W &\in \mathcal{R}^n && \text{(weights)} \\ \mu &\in \mathcal{R}^n && \text{(return vector)} \\ \hat{\Sigma} &\in (\mathcal{R}^n)^2 && \text{(covariance matrix)} \\ \lambda &\in \mathcal{R} \cap [0, \infty) \end{aligned}$$

The x -norm $\|w\|_x$ is defined as the x^{th} - root of the sum of each individual element of the vector w raised to the x . Thus: $\|w\|_x = (\sum_{i=1}^N |w_i|^x)^{\frac{1}{x}}$.

0.4.4 L2-Norm Weight Constrained Optimisation

$$\max_W z = W'\mu - \lambda W'\hat{\Sigma}W \tag{9}$$

$$\begin{aligned} \text{subject to} \quad & \|W\|_2 \leq \delta \\ & w_i \geq 0 \qquad \qquad \qquad \forall i \end{aligned}$$

where

$$\begin{aligned} W &\in \mathcal{R}^n && \text{(weights)} \\ \mu &\in \mathcal{R}^n && \text{(return vector)} \\ \hat{\Sigma} &\in (\mathcal{R}^n)^2 && \text{(covariance matrix)} \\ \lambda &\in \mathcal{R} \cap [0, \infty] \end{aligned}$$

The x -norm $\|w\|_x$ is defined as the x^{th} - root of the sum of each individual element of the vector w raised to the x . Thus: $\|w\|_x = (\sum_{i=1}^N |w_i|^x)^{\frac{1}{x}}$.

Chapter 1

Introduction

1.1 Background

Modern portfolio theory is founded upon the research of Markowitz (Markowitz, 1952), wherein a 'feasible subset' of portfolios is established from the infinite set of portfolios which exist in the specific market of interest. Litner (1965) and Mossin (1966) develop this framework further and introduce further theory of a premium for risk. An efficient portfolio is defined, within this paper, as any portfolio which is expected to find the highest return relative to all other portfolios which hold the same level of risk. Conversely; to find a portfolio which has the lowest risk, out of the subset of all portfolios with the same return. The efficiency frontier is the set of all efficient portfolios, presented in the risk-return space.

The formulation of the two inputs into the optimization algorithm requires the estimation of a large number of parameters, which induces the potential for problematic sampling error. The covariance matrix is clearly of dimensions $N \times N$, thus $\frac{N(N+1)}{2}$ unique parameters require estimation in the covariance matrix alone. The sample error arising and the impracticality of applying the theoretical results empirically is well documented by Jobson & Korkie (1980). The most extreme values frequently arise as a result of extreme error, and it is the shares associated (to the highest predictive inaccuracy) which are assigned the largest weights. Michaud (1989) describes the impractical phenomenon as 'error maximisation'.

1.2 Overview

1.2.1 Problem Statement

Portfolio theory is the study of investment practices to maximise expected return, and to minimise the associated risk. This dissertation introduces the reader to both shrinkage techniques as well as resampling procedures. These two procedures are introduced and implemented across the Johannesburg Stock Exchange (JSE), singularly and then in conjunction with one another. With the understanding from Michaud of the error maximisation tendencies arising in quadratic programming, this dissertation looks at the above two methods to mitigate the 'extreme weightings', and thus improve out-of-period performance of portfolios. To replicate the investor ability, we

allow for an information coefficient (IC) which reflects an investor's portfolio construction ability (Grinold & Khan, 1999).

1.2.2 Outline of Dissertation

Chapter 3 introduces the methodology of the dissertation including the implementation of IC reduction, resampling and shrinkage. This is followed in Chapter 4 by demonstrating the effects of resampling, both with and without a reduced IC. Chapter 5 introduces, on an IC level select of 0.05, the effect of shrinkage compared in effectiveness with resampling and without resampling. Finally, the last new insight is brought forward in Chapter 6, a sensitivity analysis to confirm that the effects are not merely a facet of the specific IC chosen. The conclusions and shortcomings of this dissertation are brought forward in Chapter 7.

1.3 The Efficient Frontier

The efficient frontier is a curve of solutions within a 2-dimensional risk-return space, satisfying the optimisation problem of max-return/min-risk. There is a unique solution per each level of risk-tolerance, thus the frontier is comprised of an infinite set of unique portfolios. Literature uses quadratic programming to find this set of solutions; using inputs of 'return' and 'risk'. The return is understood to be poorly predicted out-of-sample (via the first principal of efficient markets); however this does not apply to covariance patterns, where one expects to see repeated patterns.

By example, if share X has a strong tenancy to move with share Y, and Z has been seen in-sample to be uncorrelated to X and Y we can expect X and Y to move together out-of-sample, and we can expect behaviour of X and Y to be uninfluenced by behaviour in Z and visa versa. However, the historical return values of R_X , R_Y and R_Z would give no indication of future return values.

1.3.1 An Instance of the Efficient Frontier

The efficiency frontier is the subset of feasible portfolios (refer to Financial Background chapter for more information). With reference to 3.1, one can see (refer to equations below) how a portfolio's composition allows the portfolio to gain risk-reduction through diversification.

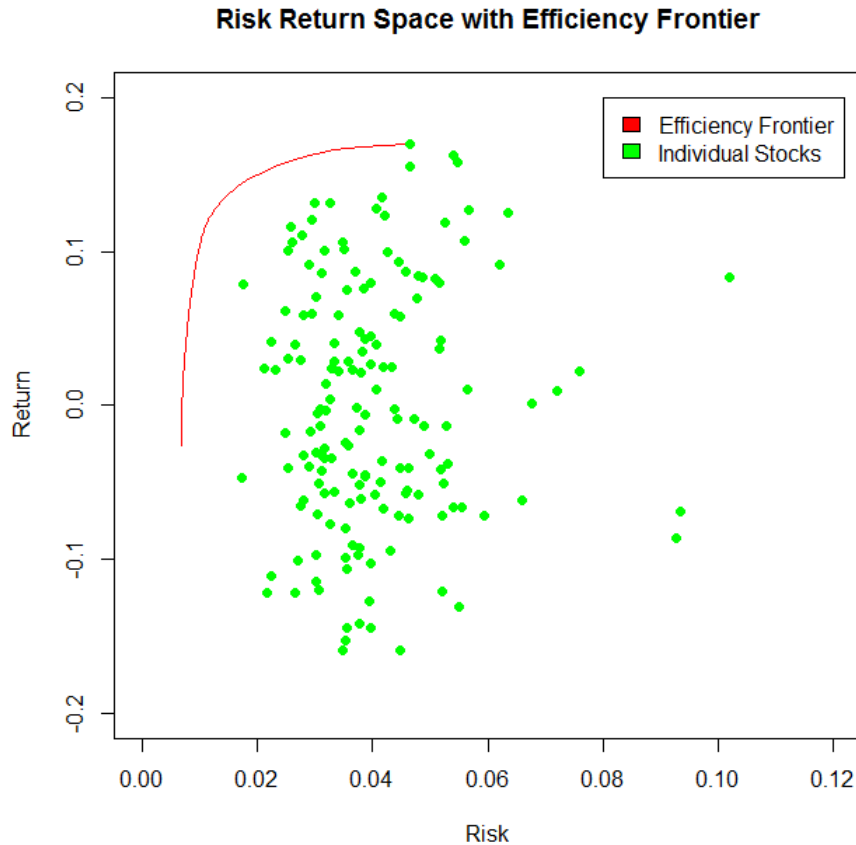


FIGURE 1.1: Efficiency Frontier for June 2006, Σ calculated over period March 2003 to May 2006

1.3.2 Portfolio Performance

As insinuated, there are multiple sources from which risk arises: the literature review analyses in detail the developments within the spectrum of risk sampling error; specific error; and specification error associated with a portfolio. The theoretical reduction of sampling error is introduced by Jorion (1986) in shrinkage of the covariance matrix, whereby PCA and similar approaches mitigate extreme sampling error. Michaud (1989) introduces the first effective procedure of resampling which allows insight into the distribution of the cross-sectional returns, rather than a point-estimate. By dropping a portion of the sampling error, the risk associated to the portfolio is directly reduced, thus constructing a more efficient portfolio.

1.3.3 Sampling Error

The errors arising from sampling error of the Multi-variable time-series returns are assumed to arise randomly, according to some distribution. It is not possible to differentiate between the true underlying series, and the errors which have occurred due to the complexity. Within this dissertation no value is gained from the imposition of an ARCH or GARCH error modelling. Similarly although ARIMA methodology or factor-models may allow for a slight increase in ability to model the return-series structure, this will give rise to specification error and potential spurious relationships, whilst providing little-to-no insight into direct portfolio optimisation.

Refer to equations (1) and (2) in the financial background. Following Green & Hollifield (1992) the ease to reduce residual/share-specific risk through diversification is demonstrated through weighting. One can diversify away the majority of all specific risk of individual shares by assigning weights of $\frac{1}{N}$ to each share. This is, however, impractical in practice as this destroys all purposes of active portfolio management.

1.4 Data

The data used over the duration of this dissertation are the share returns for the JSE shares, of the registered stocks, from 1 March 2003 to 31 March 2016. This is just over 13 years of data in total, and by using weekly returns this translates to 157 months of data, or otherwise 680 individual weekly observations of the returns. The efficiency frontier is calculated monthly, using the previous 38 months worth of data.

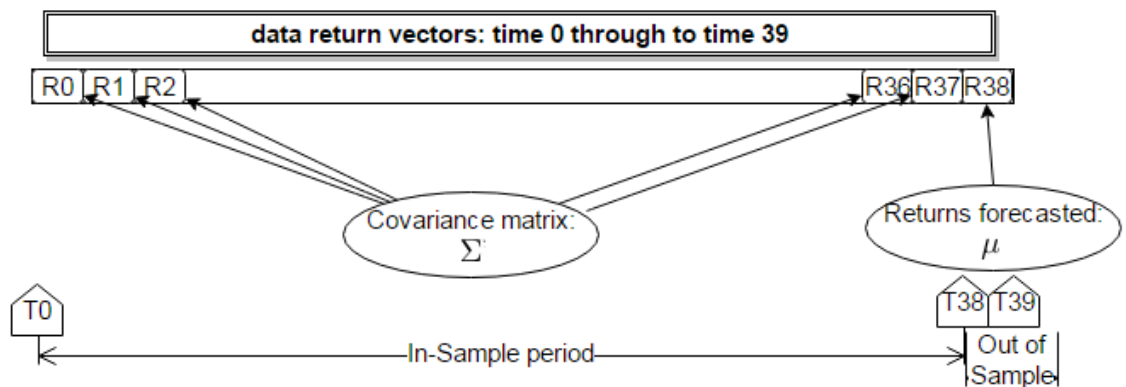


FIGURE 1.2: Data required for efficiency frontier

In 1.2 we demonstrate the source of the data required for the efficiency frontier construction. In particular, the covariance matrix Σ is constructed using the historical 38 months worth of weekly data. We allow for the first 38 months (from time 0 to time 38) to be used to estimate the covariance matrix, we refer to this as the in-sample period, the returns are indicated in 1.2 as $R_0; R_1; \dots; R_{37}$. The historical covariance matrix and the returns for the following period μ are clearly demonstrated to be independent of one another. By assumption, we begin the assessment with an assumption of 100% forecast accuracy ($IC=1$). This allows for an estimation of $\hat{\mu} = R_{38}$ to be used in the optimisation algorithm. This assumption is a good starting point to gain rudimentary understanding into the nature of the efficiency frontier, and later the assumption is dropped to reflect a more realised understanding of market performance.

Once we have solved the weightings for a portfolio, we then solve the realised returns by finding the return of the portfolio out of period (refer to 1.2). This requires cumulatively adding the weekly returns of the portfolio over the out-of-period month, per each week within the month (this is orchestrated such that there is no overlap between the data used to find the

out-of-period return, and the data used to calculate the covariance matrix).

This process has thus finds the 'forecasted' return of the 38th period. The weights are then recalibrated (from time 1 to time 39) as the in-sample data to construct the covariance matrix, and assessed on the month between time 39 to time 40. Then the weights are recalibrated again... This process is continued up until the most recent data (the 157th month) is used to find the portfolio returns. The cumulative returns over the 10 year period, from June-2006 to March-2016, is of interest in comparing portfolio performance. This period has included both a period of economic growth and economic depression, a requirement for robustness of the analysis performed by Jagannathan & Ma (2002).

1.5 Resampled Efficiency

1.5.1 Synopsis

The efficiency frontier is a convex curve, constructed according to iterative solving of one of the minimum-variance problems outlined in the *Financial Background* of this paper. Each value of λ within the maximisation of $z = W'\mu - \lambda W'\Sigma W$ (subject to relevant constraints) provides a different unique solution along the frontier. The quadratic programming problem operates directly from the inputs; intrinsically assuming the parameter inputs are known with certainty. Refer to Best & Grauer (1991) for a detailed analysis of the problems which arise based on the sampling error arising.

Resampling techniques are shown by Munro & Bradfield (2016) to allow for insight into the distribution of the point estimates used by the optimiser. The dispersion of the efficiency frontier can be imposed within the optimisation problem, solving the quadratic problem with a *sampled* set of returns and variances within the risk-return space. Each sample is constructed such that it can be mapping onto the original shares in the actual risk-return space. Thus the frontier can be transformed back to the original space and hence assessed on its performance. The process is repeated multiple times under Monte Carlo simulation. In essence, an empirically constructed confidence band can be established using the set of statistically equivalent efficiency frontiers.

1.6 Covariance Shrinkage

1.6.1 Synopsis

Through a shrinkage applied to the sample-covariance matrix, values are pulled towards various standardised matrices. The algorithm intrinsically targets the shares which are most affected by sampling error, and by the majority of adjustment occurring at the more problematic entries a large portion of error is eroded. The Shrinkage reduced tracking error relative to the benchmark; increases a portfolio manager's ability to bet on shares (an improved IR); and effectively allows for greater risk control.

It is important to try and avoid destroying true patterns within the covariance matrix as far as possible whilst eroding the sampling error; thus we take a weighted sum of a simple covariance matrix; and a shrinkage target.

Chapter 2

Literature Review

2.1 Modern Portfolio Theory

Modern portfolio theory has expanded into several frameworks which work to address the sampling-error issues and inaccurate forecasting. Maximization of the growth of a portfolio consisting of correlated Brownian-motion assets extends beyond simple mean-variance optimization (Chellathurai & Draviam, 2008). Extension of the original framework has been explored under both modern portfolio theory- the focus of this dissertation, and the stochastic analysis of the derivatives in three generic frameworks: i. stochastic optimal control theory under a martingale approach (Merton 1969); ii. Capital growth theory (Fernholz & Shay, 1982); iii. Stochastic programming (Dempster & Thompson, 2002). These frameworks were introduced subsequently to the mean-variance principle; and allow for analysis beyond the scope of simply using the first two moments to attempt to characterize complex share behavior.

It has been well established in the literature that sample estimates perform poorly out of sample (Michaud, 1989; Best & Grauer, 1991).

2.2 Sensitivity of MV Portfolios

Michaud (1989) describes how the weighting returned from maximization algorithms tilt in the direction of sampling error. Potential solutions to the problem are suggested within imposing structure on the covariance matrix; using shrinkage estimators in place of sample estimates and using higher frequency data, striving for a reduction in sampling volatility due to the law of large numbers. Green & Hollifield (1992) introduce the problem of constraining the portfolio weights, being the specification error arising as a result of the coerced reduction of sampling error. Jagannathan & Ma (2002) extend the research to identify the nature of the increase in error which arises from any share upon whose weighting a constraint has been imposed: both the variance of a stock and its respective covariances increase. Of further note by the authors, the upper bound of shares which are actively constrained are more frequently shares which hold low covariates. These shares typically have a 'downwards biased estimation error' (an error working in the opposite direction of error introduced through constraints).

2.3 Diversification

The financial Capital Asset Pricing Model (CAPM) was developed by Sharpe (1964), Lintner (1965), and Black (1972) offered an intuitive relationship between risk and return. Under the CAPM mean-variance framework dictates that all investors will hold some fraction of the market portfolio (Sharpe 1964) which in short is derived from the principal of diversification when future performance is entirely unknown. Treynor & Black (1973) discuss the advantages to a simultaneous 'active' strategy stemming from CAPM pricing, alongside a 'passive' hold for diversification purposes. Even managers who have bullish (or bearish) views on a specific stock tend to still only buy (or sell) excess by a fraction of the benchmark holding (Best & Grauer, 1991).

The traditional expectation of reward (higher returns) for investing in high-beta shares was critiqued by Fama & French (1992); who found that the relationship between β and average return did not hold towards the end of the 20th Century. Instead, we turn to the efficient frontier and select 'optimal' MV portfolios, which are subject to a different set of problems. MV portfolios (solved through quadratic programming) are known to be highly sensitive to the inputs of the expected return vector and the covariance matrix. This allows even low sampling error in an empirical study to cause drastic swings in the weights vector (Michaud, 1989). A small change in the expected returns; which has virtually no bearing on the expected return and volatility of a portfolio, can unintuitively have a drastic affect on constituents of the portfolio. Best & Grauer (1991) demonstrate empirically how a fractional change of the expected return of a single asset by under 2% ceteris paribus can cause an MV-efficient portfolio to have over half of its constituents to a weight of 0.

The principle of diversification shifts the risk of a portfolio to be majority apportioned to idiosyncratic risk, and away from market (systematic) risk, (Carke, *et al.* 2010). Even under a situation of a volatile market, where the longitudinal cross-section of covariances is constantly changing (i.e. there is a low correlation between past and future covariance matrices) Chan (1999) demonstrates a reduction in portfolio volatility through passive diversification.

2.4 Approaches to Sampling Error

Two approaches to attempt to address the covariance matrix estimation are relevant to this dissertation; and they are:

- i. A coerced factor-structure (Chan *et al.* 1999);
- ii. Covariance shrinkage methods (Ledoit & Wolf. 2003).

Jagannathan & Ma (2003) found significant results from a study of American minimum variance portfolios, which outperform all other feasible portfolios. The research involved high-frequency (daily) returns being used in analysis; an abnormality from the typical use of monthly or weekly returns. A fourth method, is short-sale constraints, which is beyond the scope of this dissertation as short-sales are not explored. Of particular interest, it has

been shown that Shrinkage techniques are able to outperform constrained cases (Jagannathan & Ma. 2003).

2.4.1 Minimum-Variance Portfolio

Covariance patterns tend to provide insight into true patterns between shares, however historical returns are regarded as poor to useless in providing forecasting insight. It is for this reason, that academia frequently uses the minimum-variance portfolio as a benchmark comparison in practice as the minimum variance portfolio does not require returns for estimation (DeMiguel *et al.* 2009).

Jagannathan & Ma's (2002) performance analysis of a mean-variance portfolio is extended from Chan *et al* (1999) out-of-sample comparisons to the minimum variance portfolio, a direct understanding that the minimum variance portfolios are expected to have as large an out-of-period Sharpe Ratio than other MV portfolios. Monte Carlo simulation is used by Ledoit & Wolf (2003) which allows for global minimum variance portfolios to be expected to perform as well as a portfolio constructed using the sample-covariance matrix (DeMiguel *et al*, 2009).

The empirical realization of a negative correlation between market beta and idiosyncratic risk allows for a natural variance-decrease due to the low-beta shares (the majority) of the minimum-variance portfolio out-of-period (Clarke *et al*, 2010).

2.4.2 Mean Variance Portfolio

The 'mean-variance' portfolio is defined by Sharpe for the first time in 1964 as the portfolio offering the highest expected return per unit of risk (refer to equation 5 in the Financial Background section). Thus note that although the Sharpe ration can be calculated for any portfolio, the Mean-Variance portfolio is the portfolio a such that the Sharpe ratio is at its maximum. We can thus define the MV portfolio a in relation to the Sharpe ratio mathematically, as the argument which maximizes the equation:

$$\text{Mean Variance Portfolio} = \max_a SR = \frac{E[R_a - R_b]}{\sigma_a} \quad (2.1)$$

Thus the MV portfolio is the portfolio which provides the highest excess expected return above the risk free rate, per unit of risk held.

2.4.3 Resampled Efficiency Frontier

Michaud's (1989) identification of error-tilting-estimators follows directly from pseudo-random observed inputs being set as the underlying inputs, with no account for the sampling error. Resampling methodology allows one to find a set of statistically equivalent portfolios using Monte Carlo simulation techniques; thus mitigating a portion of sampling error. Each iteration of the algorithm in which a new portfolio is 'sampled' involves a draw from a constructed multivariate normal distribution; optimised over the sampled parameters; and then plotted upon the original risk-return space.

The resampled efficiency frontier is an aggregation of the Monte-Carlo portfolios, plotted within the original risk-return space.

Through the resampling process, one is able to gain insight into the distribution of inputs. The South African equivalent replication of Michaud's work has been performed by Munro & Bradfield (2016), using the 10 Economic parent-sectors of the individual shares of the JSE. It was found by both Munro & Bradfield and Michaud that a more robust estimate of the efficiency frontier can be achieved. Out-of-period performance was found to be enhanced at varying levels of an active manager's IC.

2.4.4 Shrinkage

The limitations explored by Jobson & Korkie (1980) were apparent in early attempts to reduce sampling error shrinkage techniques by Frost & Savarino (1986) and Jorion (1986). Shrinkage was only feasible under condition that the number of historical returns is larger than the number of shares. Ledoit & Wolf (2003*) proposed the use of an altered sample-covariance matrix: specifically the form used, makes use of a convex mixture of a sample-covariance matrix and a shrinkage 'target' matrix $\hat{\Sigma}_{Target}$. Ledoit & Wolf construct a linear combination of a positive definite shrinkage target; and a positive semi-definite sample covariance matrix; thus gaining a shrunk covariance matrix which is necessarily positive definite. Several different estimates are suggested within the literature for the target, and slight adjustments on the sample-covariance matrix are similarly explored. The following equation describes the shrinkage of the $\hat{\Sigma}_{LW}$ towards $\hat{\Sigma}_{Target}$, the shrunk matrix (DeMiguel *et al.* 2009).

$$\hat{\Sigma}_{LW} = \frac{1}{1 + \alpha} \hat{\Sigma} + \frac{\alpha}{1 + \alpha} \hat{\Sigma}_{Target} \quad (2.2)$$

The limitations discussed of the sample-covariance matrix (Jobson & Korkie 1980) have been solved in a proposal (Ledoit & Wolf 2004) to use distribution-free estimates, which are proven to be invertible and well-conditioned. Finding a structured estimator involves coercing a matrix to be 'well-conditioned' and introduces bias, making ML estimates unreliable. General asymptotics framework used by Ledoit & Wolf (2004) revealed a superior estimator, as a weighted average of the original sample covariance matrix, and the structured estimator. Optimal linear shrinkage requires a shrinkage parameter to be solved for (to minimize expected quadratic loss of the L1-norm difference between the true covariance matrix, and the weighted estimate), and empirically verified through Monte Carlo methods. Ledoit & Wolf (2003*) provide empirical evidence for convincing out-performance of shrinkage techniques over the lack-thereof.

2.5 Model Specification Error

2.5.1 Sample Error - Specification Error Trade Off

It is well documented that a single dominant factor, even if caused by estimation error, can lead to obtusely weighted investments (Jagannathan & Ma. 2002; Michaud, 1989; Green and Hollifield, 1992). Jagannathan & Ma

(2002) demonstrate the benefit of using a positive-constraint on the weights (refer to 'shortselling constraint'). The observation extends from the rationale that coercing positive weights is equivalent to reducing specific entries within the sample-covariance matrix. Further, this process can further reduce sampling error as higher covariance matrix entries are the cause for negative weightings, thus contra-positively by coercing the weights to be positive offsets one side of the potential harm of sampling error.

Disatnik & Benninga (2007) identify the balance between the minimization of sampling error and specification error. The shrinkage methods, under PCA, PLS, and shrinkage to the diagonal will be compared against standard simple models as well as the true covariance matrix. This method has been shown internationally and within a South African context (Munro & Bradfield 2016) to reduce the sensitivity of the algorithms to noise. Lakhoo (2017) explores the tilting tendency of portfolio weights due to active distance, concentration and volume, as well as the use of shrinkage to reduce portfolio turnover.

Jagannathan & Ma (2002) compare the specification-error and sampling-error trade-off, in which work by Frost & Savarino (1988) demonstrates the likelihood for harm arising from the use of constraints within a shrinkage methodology. Munro & Bradfield (2016) and others have recently avoiding a compounding of the error effect by simply avoiding an upper-bound weight constraint.

2.5.2 Equivalent Regression Procedures

DeMiguel *et al.* (2009) discusses the analogy between various weight constraints to standard regression shrinkage-techniques. In particular, shrinkage via constraining the L1-norm is identical to Lasso-regression (Tibshirani, 2011); and shrinkage via the use of constraint of the squared weights (the L2-norm) is the equivalent to ridge regression. The identification of established regression techniques within portfolio-weight shrinkage was made use of by Ledoit & Wolf (2003) for portfolio theory in shrinking the covariance matrix.

There is a one-to-one mapping of the solutions for Ledoit & Wolf's shrinkage portfolio and an equation described as the 'A-norm-constrained portfolio', which adheres to the traditional min-variance problem with the additional constraint that $W'AW \leq \delta$ if matrix A equals the 1-factor covariance matrix $\hat{\Sigma}$ (DeMiguel *et al.* 2009), provided that the problem is well defined. The constraint is mathematically equivalent to constraining the portfolio variance. An extension of interpretation into a Bayesian framework follows use of the double exponential distribution as the prior distribution for portfolio weights; from DeMiguel's application of Tibshirani's (2011) lasso in identification of the equivalence to the L2-norm.

Parameter selection within the literature is performed using cross-validation techniques (DeMiguel *et al.*, 2009). In particular, the optimal level under the aforementioned L1-norm constraint uses cross-validation in selection of the parameter δ . It should be noted that the short-sale constrained problem is

identical to the L1-norm constraint with parameter δ set equal to 1. The conditions intrinsic to a short-sale constrained optimization thus do not necessitate the need of cross-validation in the basic problem. Although the constraints typically require positive definiteness, non-positive-definite covariance matrices have been shown by Matoti (2009) to be usable in portfolio theory through conversion into a positive-definite partnering form.

2.5.3 Direct Extensions of the use of Shrinkage

Jagannathan & Ma (2002) demonstrate the implicit use of shrinkage methodology, in which specification error is unlikely to be a problem as shrinkage naturally mitigates this error. Ledoit & Wolf's work (2003) then looks empirically towards finding the optimal shrinkage level. DeMiguel *et al* (2009) concludes there is insight within a direct shrinkage application onto the weights, as opposed to the covariance matrix, namely that the norm of a portfolio weight is smaller than a selected threshold.

2.6 Evaluation of Portfolios

Evaluation of performance out of sample is suggested within the literature as the out-of-sample variance; out-of-sample Sharpe ratio, and turnover (DeMiguel *et al*, 2009). Benchmark portfolios are used (often the index and equally weighted portfolios, assigning a weight of $\frac{1}{N}$ to each share) for comparative evaluation. Bootstrapping methods can be used to assess the statistical significance of a consistent difference in behaviour between two portfolios (Ledoit & Wolf, 2008). Turnover, the change in shares required within the portfolio between two consecutive periods, is a buying and selling process required for an active manager and is a comparative indicator of the transaction costs which will arise; as a higher turnover implies a higher total transaction cost.

$$Turnover = \frac{1}{T - \tau - 1} \sum_{t=\tau}^{T-1} \sum_{j=1}^N (\|w_{j,t+1}^i - w_{j,t}^i\|) \quad (2.3)$$

Jagannathan & Ma (2002) find that stocks are unlikely to have a stable, time invariant structure. Literature (including this paper) thus is required to demonstrate having analyzed data pertaining to an economic depression period, as well as an economic growth period in order to be complete in its findings.

Also pertaining to stability: shrinkage allows for less volatile performance, as the more extreme values in the system are pulled down in excess, relative to more neutral values. The benefits of this methodology allow for a higher probability of active managers maintaining a low tracking error; as well as enhancing the information ratio (Ledoit & Wolf 2003).

2.7 Practicality of Theoretical Research

The primary disadvantage of portfolio theory is the assumption of perfect forecasting information; thus the aforementioned 'error-maximisation' induced whilst solving the quadratic problem. As it stands, the volatility of shares can easily be under-represented, and the returns forecasts are likely to be highly inaccurate. These concerns have been discussed above, we now address the investment side assuming that an investor has a realised information coefficient above 0.

The investment in all available stocks is not a common practise, and advised (Grinold & Kahn, 1999) due to difficulties arising in active management. This is beyond the scope of this paper which is founded in a theoretical stance, thus no additional constraints are imposed upon the number of shares which can be invested in. In the same vein, the mathematical models of unconstrained optimization is well known and documented, and whilst the analysis of this paper is extendable into this case, it has been decided to follow the proceedings of Munro & Bradfield, Clarke *et al*, Ledoit & Wolf and others within the constrained case without risk of having interpretable analytical results which are not far displaced from empirical behaviour.

Some practical implications of the current markets do need to be considered at all times: It should be noted that the lowest feasible return values obtained along the Markowitz efficiency frontier should be recorded at $\lambda = r$, the risk free rate (Chellathurai & Draviam, 2008). Also with recent dips into financial crisis, lower risk investments have received increased support from private and public entities (Clarke *et al*. 2010); thus risk-averse decisions should be preferred.

It has been demonstrated in international studies that the risk associated to a portfolio, including the 'optimal' Sharpe portfolio can be reduced even further. This dissertation looks at the application to a South African context of advanced statistical techniques including simulation (resampling efficiency) as well as shrinkage techniques, following from the work of Beaven and Gebbie who introduced Michaud's resampling work into South Africa within industry.

Chapter 3

Theory and Methodology

3.1 The Efficient Frontier without Augmentation

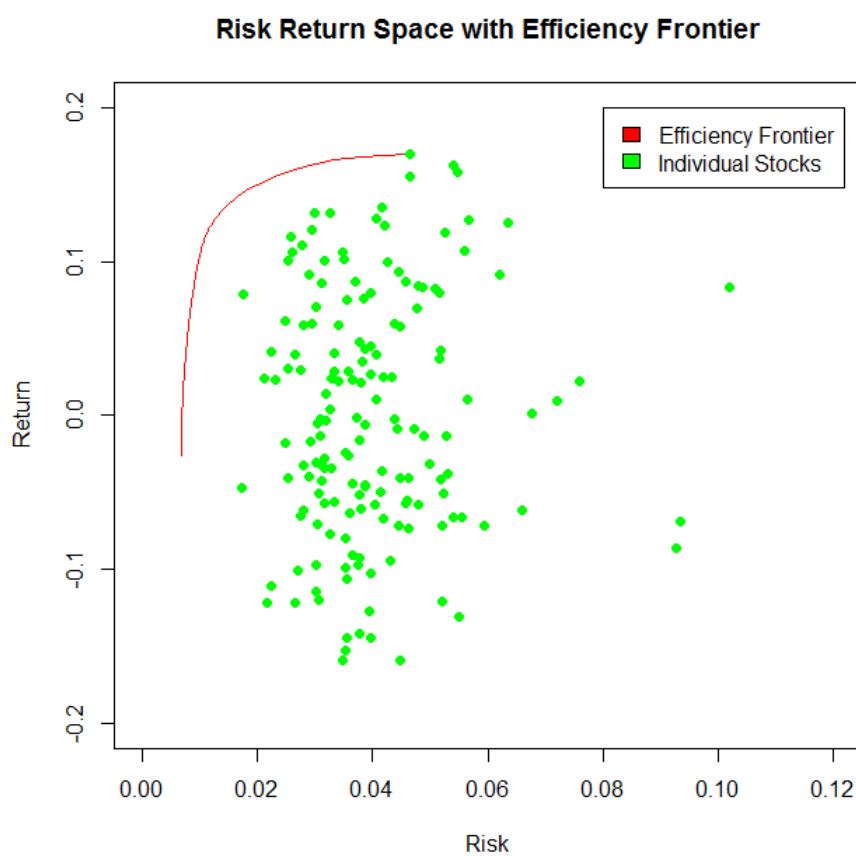


FIGURE 3.1: Efficiency Frontier for June 2006, Σ calculated over period March 2003 to May 2006

The efficiency frontier is the subset of feasible portfolios (refer to Financial Background chapter for more information). With reference to 3.1, one can see (refer to equations below) how a portfolio's composition allows the portfolio to gain risk-reduction through diversification.

Refer to equations (1) and (2) in the financial background. Following Green & Hollifield (1992) the ease to reduce residual/share-specific risk through diversification is demonstrated through weighting. One can diversify away the majority of all specific risk of individual shares by assigning

weights of $\frac{1}{N}$ to each share. This is, however, impractical in practice as this destroys all purposes of active portfolio management.

3.2 Parameter Estimation

3.2.1 Estimation Uncertainty

An issue arising within this scope has been identified in that quadratic programming algorithms take inputs (parameter estimates), as certain. Thus, the following statistical abomination is used, assuming the true parameter α with estimate $\hat{\alpha}$: $\hat{\alpha} = \alpha$, as opposed to the correct interpretation that $E[\hat{\alpha}] = \alpha$. Effectively the prior equation asserts the following, that there is no sampling error; or equivalently that $R_i \sim N(\mu_i, \sigma^2)$, where $\sigma^2 \leq \sigma_i^2$, the true variance; drastically reducing what should be used as an estimate of the variance.

3.2.2 The Bias-Variance Trade Off

We briefly consider the expected error considering the algorithm being used to calculate, as defined within parameter estimation of a bias-variance trade-off (Hastie *et al.*, 2009), which will be used later in the paper again. We define the decomposition of variance fully, and then simplify to eliminate zero-terms.¹

$$E(y - f(\hat{x}))^2 = bias[f(\hat{x})]^2 + Var[f(\hat{x})] + \sigma^2 \quad (3.1)$$

where

$$\begin{aligned} bias[f(\hat{x})] &= E(f(\hat{x})) - f(x) \\ Var[f(\hat{x})] &= E(f(\hat{x}) - E[f(\hat{x})])^2 \end{aligned}$$

Note that there is no adjustment of the input as of yet, and thus the functional form of the inputs are unbiased: $E(f(\hat{x})) = f(x)$. The equation simplifies to

$$E(y - f(\hat{x}))^2 = Var[f(\hat{x})] + \sigma^2 \quad (3.2)$$

¹For those unfamiliar with machine learning bias-variance trade off notation, the following definitions are used (notation on left, with its definition within an application within portfolio-theory on right):

$$\begin{aligned} y &= \text{output (covariance matrix } \Sigma) \\ x &= \text{data (historical returns)} \\ f(x) &= \text{underlying best estimation of } y \text{ using } x \\ f(\hat{x}) &= \text{the estimation of } y \text{ used} \\ \sigma^2 &= \text{random error matrix} \end{aligned}$$

Thus we formally define the problem of underestimation of the volatility of the variability of our estimation, by the fact that $E(y - f(\hat{x}))^2$ is estimated by σ^2 due to $Var[f(\hat{x})]$ being erroneously set equal to 0.

3.2.3 Backtesting

This dissertation makes use of *ex-post* analysis, identifying key performance indicators to compare portfolio construction methodology based on empirical portfolio results. This is understood as sufficient within portfolio theory, as these are the tools used in practise, as opposed to more rigorous tools such as cross-validation. The process of cross-validation is optimal within machine learning to find the best estimate for shrinkage parameters, and although this can be used here, the machine learning approach is not adopted within this dissertation. The key distinction removing the applicability of cross-validation is that the assumption of independence and identical-distribution is in juxtaposition to portfolio theory which uses individual (and joint) historical performances as statistics within the optimisation algorithms.

3.3 Reduction of Sample Volatility

3.3.1 Resampled Efficiency

Resampling from the efficiency frontier, using the original estimates of the parameters i.e. sampling returns from a $N(\mu, \hat{\Sigma})$; allows for the imposition of randomness within the parameters. The 'resampling' process is repeated multiple times, allowing for the returns to be assessed with consideration for sampling error (by finding a pseudo-distribution for each share) rather than being assessed as point-estimates. This adjustment is interesting, as it directly compensates for Michaud's identification that weights are selected by an algorithm which over-weights the shares which are most affected by sampling error.

3.3.2 Resampling Methodology

Michaud's resampling algorithm is adhered to within this dissertation, making adjustments for the frequency of the data as well as the number of constituents within the index (benchmark) considered for investment.

1. Use historical data to estimate covariance matrix Σ , as well as an estimation for the returns μ .
2. Simulate 43'333² observations from a multivariate normal distribution with parameters Σ and μ .
3. Sample with replacement 480 observations from the observations stored after [2].

²Michaud (1989) uses monthly data, applied to each sectors upon which resampling is performed, and a sample of size 10'000 is used. For a comparative study, note that weekly data is used thus the frequency has increased by a magnitude of $\frac{52}{12} = 4.3333$

4. Calculate a new expected return μ^* and covariance matrix Σ^* from the sample in [3].
5. Solve the short-sale constrained quadratic optimisation problem (refer to equation (7) in the financial background), for a range of 50 values of λ to allow for equally spaced points across the frontier.
6. Map the 50 portfolios from the efficiency frontier optimised over μ^* and Σ^* , onto portfolios of the original shares. Thus, using the original parameters Σ and μ ; we solve for the mean and variance of each of the portfolios: R_{p_i} and σ_{p_i} respectively.

Repeat steps 3 to 6; 500 times

This constitutes the algorithm; which results in a total of 500 sets of 50 weight-vectors, each of dimension $1 \times N$.

3.3.3 Theoretical Usage

Resampling is applied directly onto the estimated sample covariance matrix and return vector. 500 portfolios are generated; and the aggregated performance of these generated portfolios then forms the resampled efficiency frontier. The resampled portfolio in this dissertation will be constructed using 100% forecast accuracy in the first few chapters and thereafter be adjusted downwards. This follows from literature written on shrinkage techniques, and for congruence between this paper's resampling and shrinkage methodology. The sample covariance matrix is constructed using 3.25 years of historical returns, using weekly data. This is an extension of literature, which has used monthly returns, however has used far fewer stocks. To mimic the requirement for number of stocks being lower than number of point-observations (a requirement for a positive definite covariance matrix) this paper makes use of higher-frequency data than the literature. We ensure a positive definite matrix by holding more return periods (169) than the number of primary South African stocks listed on the JSE (varying between 155-165 in the data pertinent to this thesis).

This point has been mentioned prior, and will be mentioned later on when explicitly used: estimation will be performed initially with 100% forecast accuracy. This will subsequently be lowered to a more reasonable 5% to represent industry implementation.

3.3.4 Shrinkage

Shrinkage is a technique which can be applied directly onto the covariance matrix. In an alternative attempt to redress the sampling error intrinsic to the historical data, the data is 'shrunk' using principal components, diagonalisation, and the like. This process has been shown to shrink error 'faster' than true patterns. Thus, although it is not possible to eliminate error, nor to identify where it is, it can be marginalised. It is within the use of shrinkage that a bias-variance trade-off must be considered i.e. $E(f(\hat{x})) \neq f(x)$. The more powerful the shrinkage technique the more error is eroded from the output simultaneously with an increasing bias imposed on the model.

3

We select the parameter α such as to allow an optimal trade-off

$$\hat{\Sigma}_S = \alpha F + (1 - \alpha)SC \quad (3.3)$$

with : $SC = \text{Sample covariance matrix}$

$F = \text{Structured covariance matrix}$

$\alpha = \text{Shrinkage constant} \in [0, 1]$

3.3.5 Simple Sample Estimators

i: Historical Sample Covariance (SC)

This is the standard matrix, which has been discussed under dispute from the likes of Michaud (1989); Best & Grauer (1991); Jagannathan & Ma (2002) for its use in quadratic optimisation algorithms. Thus this matrix, calculated from the data, is the starting point from which we hope to eliminate a portion of sampling error.

$$SC = [sc_{jk}] \quad (3.4)$$

$$sc_{jk} = \frac{1}{T-1} \sum_{t=1}^T (r_{i,t} - \bar{r}_j)(r_{i,t} - \bar{r}_k)$$

where

$T = \text{Number of time periods}$

$\bar{r}_h = \text{Average of stock } h \text{ returns}$

$r_{h,t} = \text{Return of stock } i \text{ at time } t$

³We have opted for a simpler notation than used by DeMiguel *et al* (1989): $\hat{\Sigma}_{LW} = \frac{1}{1+\alpha} \hat{\Sigma} + \frac{\alpha}{1+\alpha} \hat{\Sigma}_{Target}$

ii: Diagonal Covariance Model (DC)

This matrix has off-diagonal entries of 0; and diagonal entries equal to the sample covariance, i.e. it retains the variance information but eliminates all covariances.

$$DC = [dc_{jk}] \quad (3.5)$$

$$dc_{jk} = \begin{cases} sc_{jk} & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}$$

where : $sc_{j,j}$ is the j^{th} diagonal element of SC

This estimator adds a large bias, regardless of what matrix it is added to. Statistically the error imposed by this model is a large increase in the specification error; thus aside for a few 'blends' this will not be considered practical, aside for comparative purposes.

3.3.6 Shrinkage Estimators**i: Default (MED / SS)**

Standard comparison will be done using **R** default shrinkage; in accordance to Schafer & Strimmer (2005)'s shrinkage towards median. This is the default shrinkage technique used in *R*. This technique tends to incur a large shrinkage factor recommendation through Ledoit & Wolf's (2003*) derivation. The covariance shrinkage uses the shrunken variances ($s_{i,i}^*$) as the diagonal elements; and calculates the shrunken covariances with respect to the shrunken correlations ($r_{i,j}^*$).

$$f_{i,j}^* = \begin{cases} s_{ii}^* & \text{if } i = j \\ r_{ij}^* \sqrt{s_{ii}^* s_{jj}^*} & \text{if } i \neq j \end{cases} \quad (3.6)$$

$$s_{i,i}^* = \lambda^\times s_{median} + (1 - \lambda^\times) s_{i,i}$$

$$r_{i,j}^* = \begin{cases} 1 & \text{if } i = j \\ r_{ij} \times \min(1, \max(0, 1 - \hat{\lambda}^*)) & \text{if } i \neq j \end{cases}$$

$$s_{i,j} = i - j^{\text{th}} \text{ covariance}$$

$$r_{i,j} = i - j^{\text{th}} \text{ correlation}$$

$$\hat{\lambda}^* = \frac{\sum_{j \neq i} \text{Var}(s_{ij})}{\sum_{j \neq i} r_{ij}^2}$$

$$\hat{\lambda}^\times = \frac{\sum_{k=1}^p \text{Var}(s_{kk})}{\sum_{k=1}^p (s_{kk} - \text{median}(\text{diag}(SC)))^2}$$

ii: Average Correlation Model (ACM)

The ACM uses very little parameter estimation as it is a remarkably simple model. This is useful should data be feared to be difficult to distinguish

sampling error from true patterns, but poor in its ability to use known differences between shares.

$$f_{ij} = \begin{cases} \bar{s}_{ii} & \text{if } i = j \\ \bar{s}_{ij} & \text{if } i \neq j \end{cases} \quad (3.7)$$

with $f_{ij} = i, j^{\text{th}}$ element of structured covariance matrix F

\bar{s}_{ii} = average sample variance

\bar{s}_{ij} = average sample covariance

iii: Constant Correlation Model (CCM)

The CCM distinguishes between variances and covariances, with a marginal level of complexity increase from ACM, specifically with the additional parameter estimation of the average correlation between shares.

$$f_{ij} = \begin{cases} \bar{s}_{ii} & \text{if } i = j \\ \bar{\rho}\sqrt{s_{ii}s_{jj}} & \text{if } i \neq j \end{cases} \quad (3.8)$$

with $f_{ij} = i, j^{\text{th}}$ element of structured covariance matrix F

$\bar{\rho}$ = average correlation

s_{ii} = i^{th} sample variance

s_{ij} = i, j^{th} sample covariance

iv: Principal Component Analysis (PCA)

PCA, in general, allows for the majority of variation within a data structure to be captured in far fewer dimensions. One primary advantage, aside from the ease with which data can be viewed, understood and analysed; is that PCA identifies and keeps the stronger patterns within the data, which are typically less prone to sampling error. PCA has a natural aversion to sampling error, often leaving it to the last few components which are excluded from analysis anyway. Following the literature, we shall use the first η principal components, such that at least 80% of the variation is captured.

v: Single Index Model (SIM)

The SIM shrinkage makes use of a CAPM model, first solving for a stock's beta using a regression of share returns against market returns; then constructing the structured covariance matrix with the $\beta\beta'_{N \times N}$ matrix plus an error matrix:

$$\mathbf{r}_i = \alpha_i + \beta_i \mathbf{r}_m + \mathbf{e}_i \quad (3.9)$$

with \mathbf{r}_m = vector of market returns
 β_i = beta of share i
 \mathbf{e}_i = error vector
 \mathbf{r}_i = vector of share i returns
 α_i = intercept

$$F = s_m^2 \boldsymbol{\beta} \boldsymbol{\beta}' + D$$

with $\boldsymbol{\beta}$ = vector of β_i 's
 β_i = beta of share i
 s_m = sample standard error of r_m
 D = diagonal matrix of sample variances of the residuals

The SIM shrinkage is included for comparative reference, but it was chosen to not be included in the dissertation as the recommended associated shrinkage values found through Ledoit & Wolf's (2003*) equations were consistently below 1%, and an exploration into improving upon these equations is beyond the scope of this dissertation.

3.3.7 Dimension Retaining Techniques

The CCM, ACM, SIM and SS/MED shrinkages are applied without allowance for a mapping of the problem into a smaller space. The advantage of these methods is that the outcome is directly interpretable and easily comparable. The disadvantage of these methods is eminent as they inherit specification error.

CCM and ACM are not sensitive to outliers, which allows for a reduction in sampling error. SIM, which is based on an individual analogy of each share, is slightly more susceptible to sampling error; but this simultaneously allows for a more accurate capturing of the patterns of the data. Granted that there is an alternative error introduced which otherwise would have not been present, these methods do not introduce bias in any manner, and they drastically reduce the sampling variance thus they are in many ways still powerful. SIM has been identified as a convoluted method whilst using as many data points as this dissertation does, and is thus excluded from the study.⁴

⁴SIM shrinkage method, under the Ledoit and Wolf methodology recommended an average shrinkage factor of $\alpha = 0.9\%$; thus suggesting a $\hat{\Sigma}_S$ which was too close to SC to make a significant impact.

$$\hat{\Sigma}_S = 0.009F + (0.991)SC$$

3.3.8 Direct Dimension Reduction Techniques

PCA is the method we use for dimension reduction. The primary advantage of this is the removal of specification error. PCA is implemented through an effective dimension reduction of the data. This generally removes sampling error observed across individual data points; however with a repeated sampling error this can be confounded and appear within the PCA decomposition spuriously.

Using all N-dimensions of a PCA decomposition will exactly replicate the original data set (after respective scaling and rotation). This will reproduce, by definition, $y = f(x) + \sigma^2$ and it will be impossible to extract the true relationship. PCA has been well researched in its uses (Bengtsson & Holst. 2002) within portfolio theory, and to achieve a shrinkage with a reasonable expectancy to eliminate some sampling error but retain the majority of the true patterns, we choose the first k principal components, such that 80% of the variation within the data is captured.

3.3.9 Blended Models

Blend i: SC; DC; CCM

This is an equally-weighted combination of the outlined methods: $\frac{1}{3}(SC + DC + CCM)$. This set of estimators was established by Ledoit & Wolf (2003), and the performance shall be compared in this paper to the other methods.

Blend ii: SC; DC; PCA

This is an equally-weighted combination of the outlined methods: $\frac{1}{3}(SC + DC + PCA)$. This set of estimators was established by Bengtsson & Holst (2002), and the performance shall be compared in this paper to the other methods.

Blend iii: SC; DC; SIM

This is an equally-weighted combination of the outlined methods: $\frac{1}{3}(SC + DC + SIM)$. This set of estimators was established by Jagannathan & Ma (2002), and the performance shall be compared in this paper to the other methods.

Blend iv: SC; CCM; PCA

This is an equally-weighted combination of the outlined methods: $\frac{1}{3}(SC + DC + SIM)$. The weighting selection is, from the author's opinion, preferable as it seems to introduce a lower specification error, as opposed to **Blend ii**, but still maintaining the use of shrinkage towards PCA.

The South African market is an emerging market and is known to provide a higher expected return on investment than developed markets.⁵ In

⁵In addition, the investment in South African stock markets can be used as a 'gateway' into Africa, demonstrated by foreign investment within SA being overweighted relative to other foreign investment in other African countries compared as a proportion of their GDP.

the author's opinion, the usefulness of underestimating covariances relative to variances is risk-mitigating behaviour which is unnecessary. The introduction of specification error (shrinking off-diagonal elements towards zero) could rather be used by active portfolio managers in a higher-weighted portfolio, achieving a similar level of risk; but based more on historical patterns than on specification error. Thus, bearing in mind the market performs well on average, the bets an active manager places should be used to increase his IR, rather than further diversify. Thus, the decision of shrinkage towards a CCM rather than a DC. The use of blended models has been seen by Munro & Bradfield (2016) to outperform the other methods and, although it is beyond the scope of this thesis, this is an area for further research which has a strong potential for performance improvement.

3.3.10 Parameter Estimation from Shrinkage Loss Function

From the trade off equation between the sample covariance matrix and structured covariance matrix (2.3), it is pertinent that the shrinkage factor is not arbitrarily selected. Ledoit & Wolf (2003) impose the Frobenius norm of the difference between the the shrinkage estimator and the true covariance matrix. The Frobenius norm is defined by:

$$\|M\| = \sum_{i=1}^N \sum_{j=1}^N m_{i,j}^2 \quad (3.10)$$

This yields the following quadratic loss function of α :

$$L(\alpha) = \|\alpha F + (1 - \alpha)SC - \Sigma_S\|^2 \quad (3.11)$$

The minimisation of the equation solves for the optimal estimate of the shrinkage factor α below:

$$\hat{\alpha} = \frac{\hat{K}}{T}$$

with : $\hat{K} = \frac{\hat{\pi} - \hat{\rho}}{\hat{\gamma}}$

$\hat{\pi}$ = sum of the asymptotic variances in SC
 $\hat{\rho}$ = asymptotic covariances between entries in SC and F
 $\hat{\gamma}$ = squared difference between terms of SC and F

3.4 Overview of Resampling and Shrinkage Applications

3.4.1 Perfect Information Coefficient

With an information coefficient of 1 (a perfect IC), we refer to this anomaly as allowing for 'perfect foresight'. Effectively this allows for perfect knowledge into the future returns of the shares. The machine learning equivalent is to use the same training and test data set; resulting is severe (in this case, absolute) over-fitting. This allows for a theoretic understanding to be

gained from the image, and for testing purposes it is critical that the algorithms applied with perfect foresight return theoretic sound behaviour.

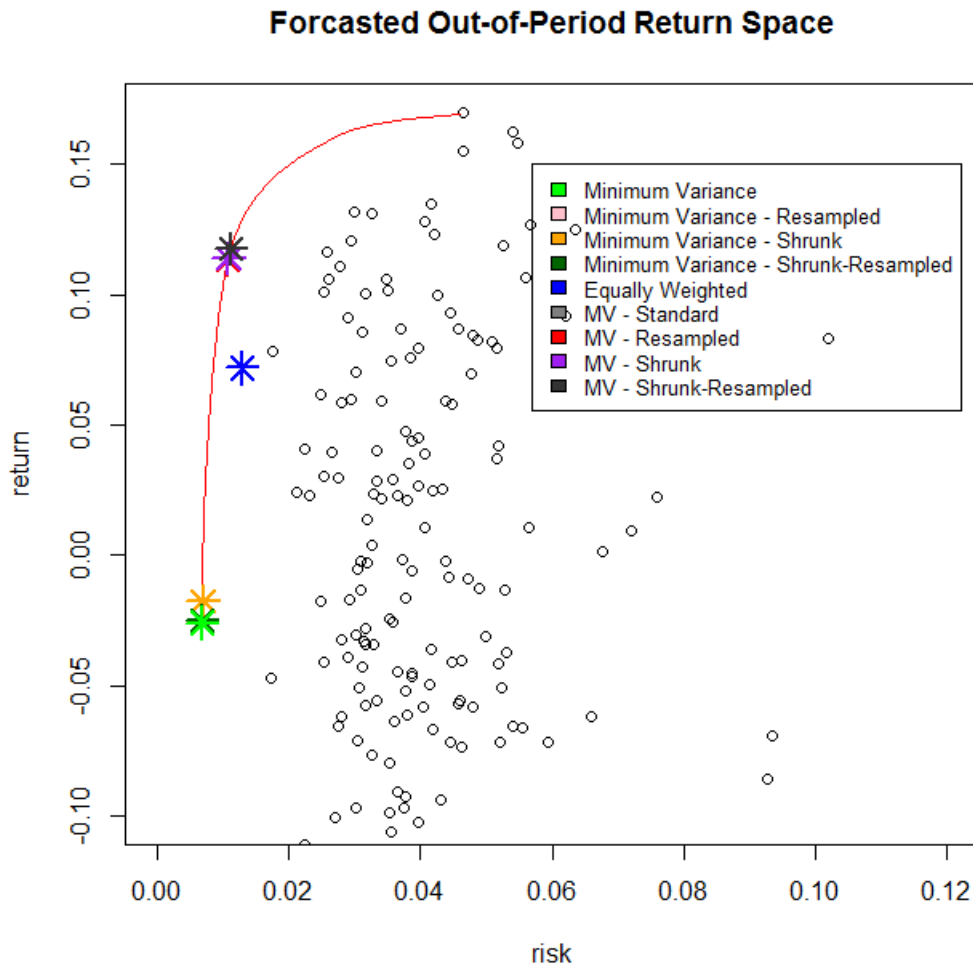


FIGURE 3.2: Efficiency Frontier with Perfect Information

In 3.2 one can identify the **predicted** performance of each portfolio in the risk-return space. This image reflects the actual share returns and portfolio returns over the month of June 2006; the covariance matrix where relevant was constructed using data from March 2003 to May 2006. The black circles represent the risk-return coordinates for each individual share; and the various portfolios are represented by the stars, refer to the legend, and finally the red line displays the efficiency frontier itself.

Both the minimum-variance portfolio as well as the mean-variance portfolio lie precisely on the efficiency frontier as per their respective definitions in the previous chapter. The equally-weighted portfolio will provide be the arithmetic average of the returns of the individual shares, and the variance being the average of each share's individual variance. Lastly, we see the shrinkage and resampling effects. Theoretically these portfolios almost necessarily will not lie on the efficiency frontier; however they will retain very similar behaviour to their original counterpart.

Note the cluster of light green (minimum-variance); pink (resampled minimum-variance); orange (shrunk minimum-variance) and dark green (resampled-shrunk minimum-variance). We observe that the minimum-variance portfolio's return can be outperformed by simply using these techniques. Within this cluster we observe close patterns for both the risk and the return. Similarly for the variance shrinkage and resampling applications to the mean-variance portfolio (grey) similar results are logically obtained. However, we also see that the MV return or MV risk can be outperformed (although not both simultaneously, by definition). These patterns are desired to be seen when the returns are not known with perfect foresight as well.

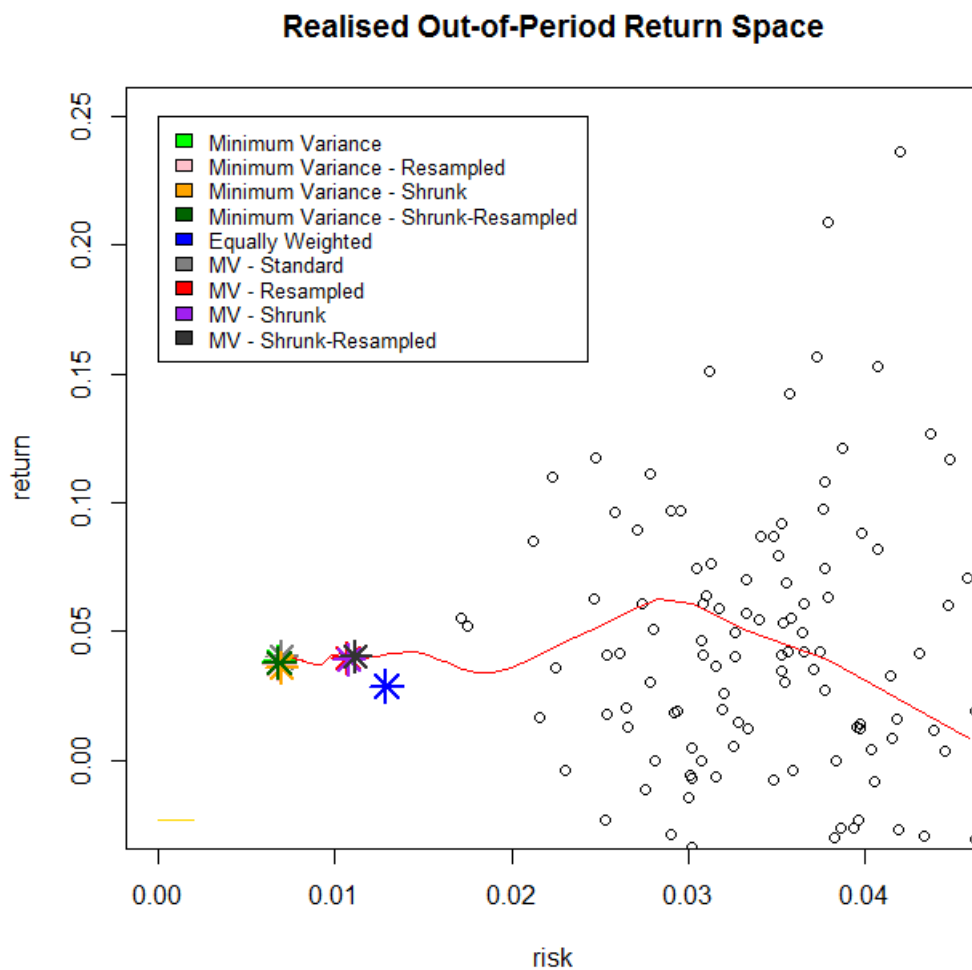


FIGURE 3.3: Efficiency Frontier without Perfect Information

In 3.3 one can identify the **realised** performance of each portfolio in the risk-return space. This image reflects the *realised* share returns over the month of June 2006 if the investor does *not have perfect information*. Thus, if the returns are not forecasted with an $IC=1$, our estimated efficiency frontier does not take the theoretical convex curve, however this is realistic of what an individual may predict (which is far from reality). As before, this corresponds to having calculated the covariance matrix using data from March 2003 to May 2006. As with the prior image, the black circles represent the risk-return coordinates for each individual share; and the various

portfolios are represented by the stars, refer to the legend, and finally the red line displays the efficiency frontier itself.

The portfolios correspond to the original image 3.2; with the difference that returns are not assumed to be known. Under the assumption of a "good" investor skill (Grinold and Kahn, 1999), we adjust the returns to be fed into the quadratic optimisation problem as "predicted returns". The predicted returns will be correlated to a varying degree with the actual returns, to be representative of an investor's ability. We move away from directly using the actual returns since no one has the ability to forecast with 100% accuracy. As expected through the literature, one of the worst performing portfolios is the equally-weighted portfolio; with the minimum-variance portfolio set maintaining comparable returns to the mean-variance portfolio group's returns. Note that the risk is not adjusted out of period, thus the levels of risk are identical between the two graphs. This dissertation continues through the use of both the ALSI and the equally weighted portfolios as benchmarks; and investigates the ability to use shrinkage and resampling techniques to improve the relative performances of both minimum-variance; and mean-variance portfolios. In Chapter 4 a detailed overview is provided into the practicality of perturbing returns (refer to 3.3) as we delve into industry-applicability of this research.

Both the minimum-variance portfolio as well as the mean-variance portfolio lie precisely on the efficiency frontier as per their respective definitions in the previous chapter. The equally-weighted portfolio will provide the arithmetic average of the returns of the individual shares as its return, and the variance being the average of each share's individual variance. Lastly, we see the shrinkage and resampling effects. Theoretically these portfolios almost necessarily will not lie on the efficiency frontier; however they will retain very similar behaviour to their original counterpart.

3.4.2 Construction of the Resampled Frontier

Within this section, the efficiency frontier is resampled under the methodology described in the previous section. The images below demonstrate the iterative process through which a resampled frontier is constructed. There are periodically between 155-165 active shares on the JSE; thus at minimum 165 data points are required to ensure $p > n$. The algorithm has been automated to construct the covariance matrix over the preceding 38 months (ranging from 166-169 weekly returns) and ergo ensure properties of positive definiteness. The demonstration below portrays the efficiency frontier calculated for the weeks of June 2006; thus making use of data between March 2003 and May 2006. The returns used for this section are assumed forecast with 100% accuracy, an assumption which is revised later in this dissertation.

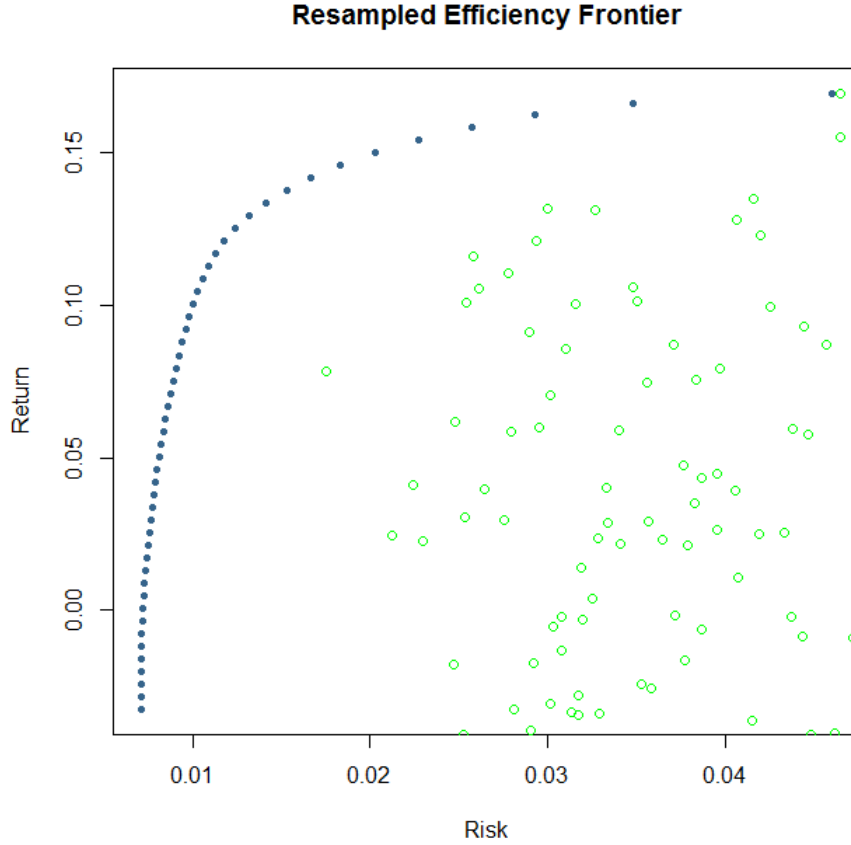


FIGURE 3.4: Resampled Efficiency Frontier (one iteration)

In 3.4 the reader can see the output of the algorithm⁶, whereby through each adjustment of λ in the maximisation problem, one solves for a new, unique portfolio (corresponding to the solution vector W). Each green circle corresponds to the risk-return coordinates of the individual shares, and the solid blue dots correspond to the output per various values of λ . These portfolios have the following standard portfolio properties i.e. with Risk (x-axis) $W_P' \Sigma^* W_P$ and Return (y-axis) $W_P' \mu^*$.

⁶We solve the quadratic programming problem defined below, over a range of values for λ , each providing a unique point upon the efficiency frontier.

$$\begin{aligned} \max_W z &= W' \mu - \lambda W' \hat{\Sigma} W \\ \text{subject to} \quad & W' e = 1 \end{aligned}$$

where

$$e \in \mathcal{R}^n$$

$$W \in \mathcal{R}^n$$

$$\mu \in \mathcal{R}^n$$

$$\hat{\Sigma} \in (\mathcal{R}^n)^2$$

$$\lambda \in \mathcal{R} \cap [0, \infty]$$

(weights)

(return vector)

(covariance matrix)

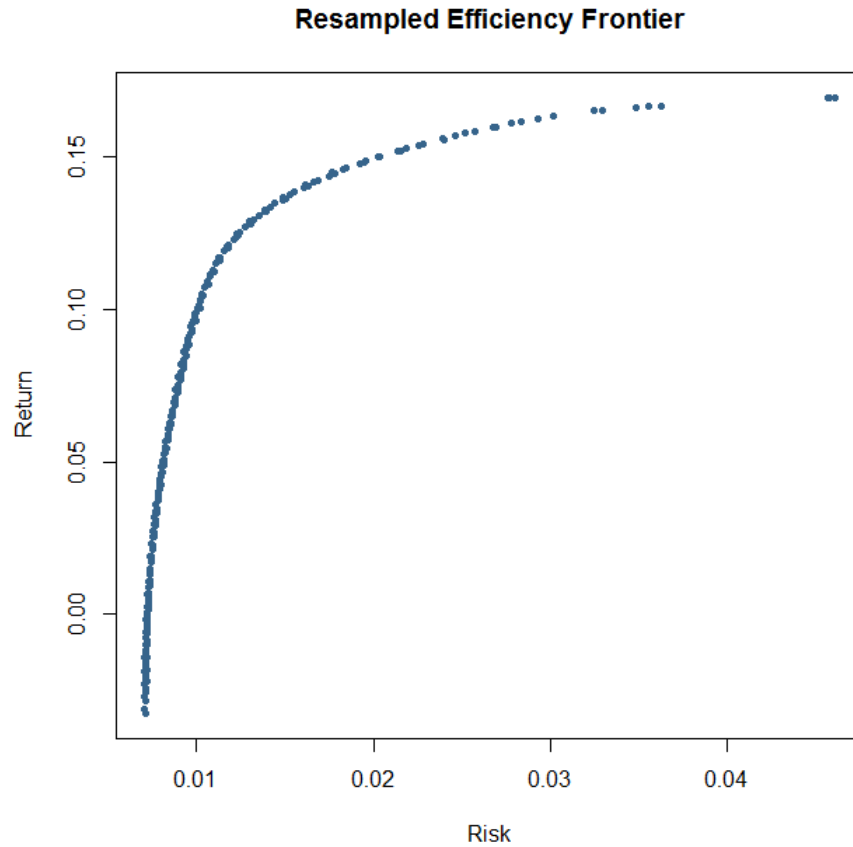


FIGURE 3.5: Resampled Efficiency Frontier (After Five Iterations)

In 3.5 it is evident that there is little volatility between multiple draws from the $MVN(\hat{\mu}, \hat{\Sigma})$ -distribution. the lack of variability is confirmed via an assessment of the variances, covariances, and mean return elements within the multivariate normal.

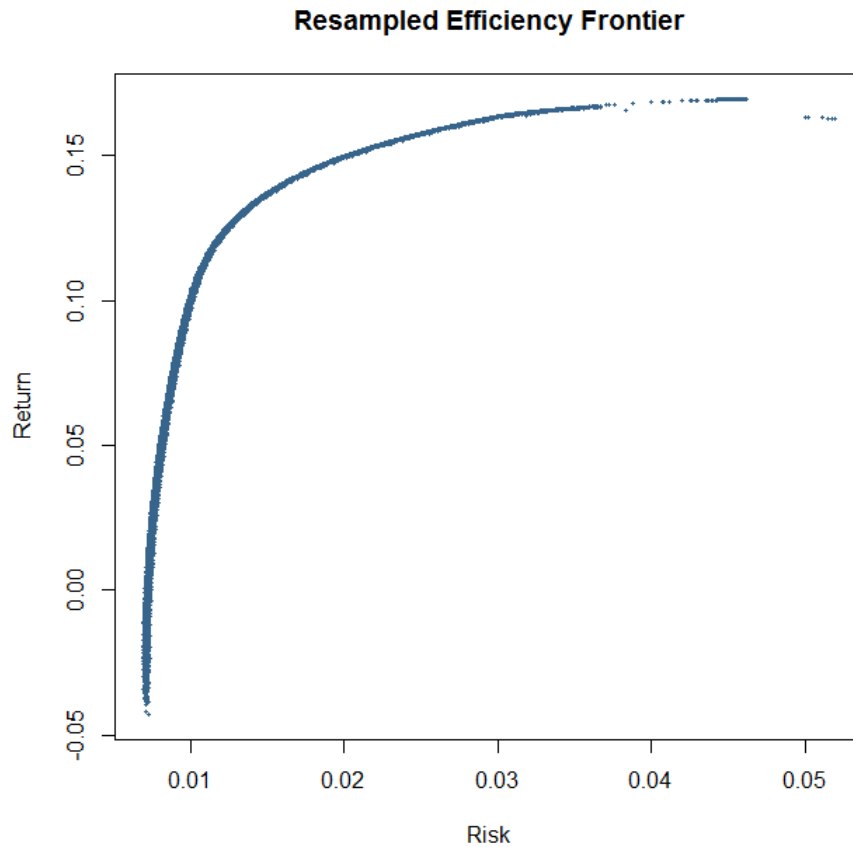


FIGURE 3.6: Resampled Efficiency Frontier (All Five Hundred Iterations)

Pertaining to 3.6; the individual 25'000 portfolios are indicated by smaller dots to allow partial-distinguishability. As expected from the patterns in 3.5, there is visible volatility within the resampled portfolios. The variability between the portfolios is dictated through the parameters used in the resampling process. The use of a 160×160 matrix and the relatively small entries in the matrix does not allow for large variability to be seen in the efficiency frontier during the resampling process, as there are many similarly performing stocks which are replaceable on the frontier. The distinct advantage; is seen when the weights between the various stocks are shrunk naturally upon aggregation of the resampled efficiency frontier. This is discussed further in the assessment of portfolio analytics in Chapters 4-6.

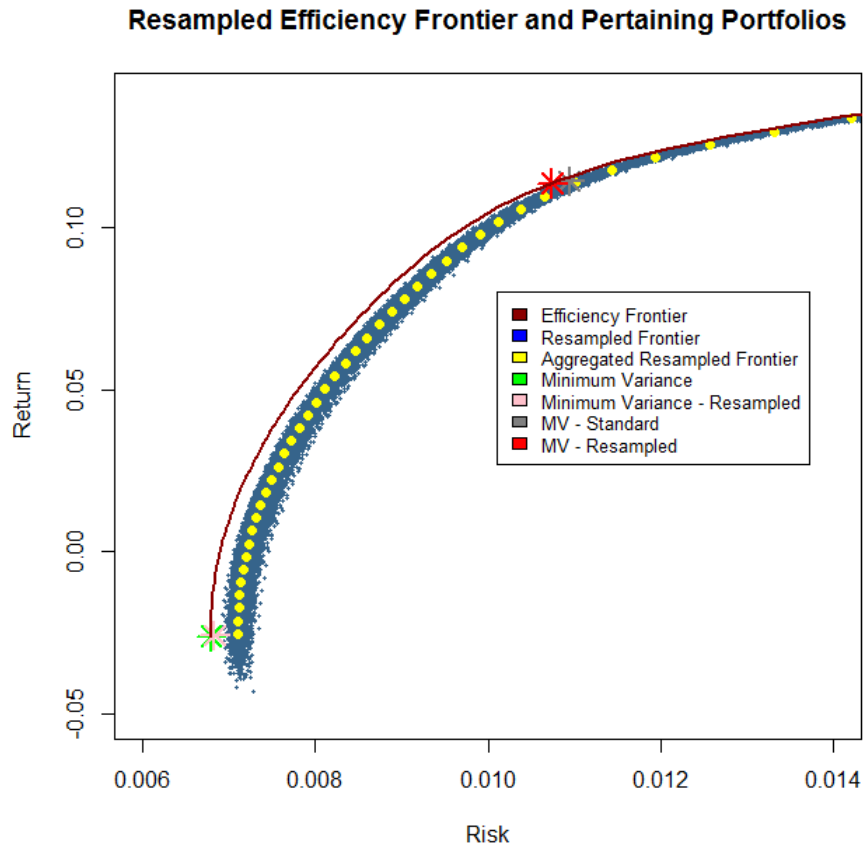


FIGURE 3.7: Close-up of above frontier, with relevant portfolios indicated

We have chosen to view this last image of the process in close-up to distinguish crucial details. It is statistically impossible for any portfolio to out-perform the efficiency-frontier (refer to the definition): thus the true efficiency frontier necessarily has to lie on top of all others, seen by the red line in 3.7. We see the realisation of what will henceforth be referred to as the *resampled efficiency frontier*- the yellow dots indicating the aggregation of the 500 iterations. Further we distinguish the position of the MV portfolio (grey) towards the apex of the curve- this is the point at which the Sharpe Ratio is maximised, and we note as well the minimum-variance portfolio (green). The resampling procedure visibly alters the performance of the respective portfolio through an adjustment of the weights. In later chapters the effect this has on performance will be assessed.

The use of the resampled frontier reduces the extreme weightings, and much like the performance of a random forest model will improve the expected performance from allowing parameters to be less deterministic, the allowance for a lower dependence within the efficiency frontier.

Optimality out-of-Sample

We have no strict definition of the 'best' portfolio when we are in a period out-of-sample using perfect forecast. Thus it is impossible to assess a portfolio without comparison to other portfolios as benchmarks; it is for this

reason that we need to empirically measure performance under non-perfect foresight. Through analysis and comparison of portfolios out-of-sample, we can assess the portfolio-construction methodologies which are likely to be preferable. A perfect foresight is clearly impractical (if the reader does not understand this, consider the arbitrage implications), thus after building up the theoretical understanding of the methodologies, the perfect foresight assumption will be dropped. However, even under the assumption of perfect foresight the associated risk a portfolio holds is regarded as reliable, and portfolio risks are comparable to one another.

3.4.3 Vector Algebra: Generating correlated Series

Knill (2011) describes the relationship which shall be exploited in this section: Given two random variables X and Y both with zero-mean, we have

$$\text{Cov}(X, Y) = E(X \times Y) - E(X) \times E(Y) = E(X \times Y)$$

thus the covariance is the dot product of the vectors. The standard deviation is the length of the vector. Thus should both vectors be of length 1:

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \times \sigma_Y} = \text{Cov}(X, Y)$$

Through substitution between the above two equations it is clear that the correlation is the dot product between the unit vectors X and Y ; which is equal to the cosine of the angle between the vectors. The following steps demonstrate the use of this property:

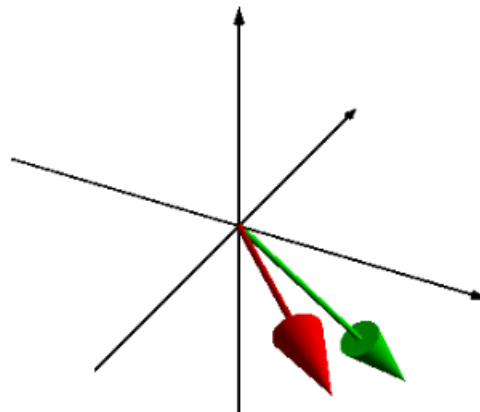


FIGURE 3.8: Vector Space of Correlated Returns

Begin with return series R_t and desire to generate a series S_t such that $Cor(R_t, S_t) = r$.

1. Generate a random normal series Q_t with dimension equal to that of R_t
2. Center $Q_t \rightarrow \bar{Q}_t$
Center $R_t \rightarrow \bar{R}_t$
3. Orthogonally project \bar{Q}_t onto U_t
 $\rightarrow U_t \perp R_t$
4. Scale \bar{R}_t to length 1 $\rightarrow x1$
Scale U_t to length 1 $\rightarrow x2$

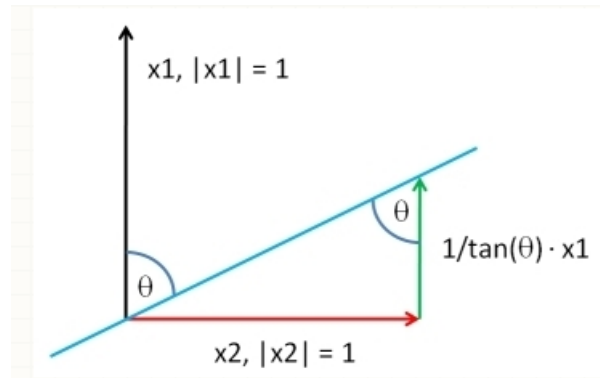


FIGURE 3.9: Vector Algebra of Correlated Returns

5. Using $\theta = \arccos(r)$ calculate the vertical (green) component
6. Add the vectors : $S_t = x1 + \frac{1}{\tan(\theta)}x2$
 $\rightarrow Cor(x1, S_t) = r$
 $\rightarrow Cor(R_t, S_t) = r$

To clarify the steps above, the projection (step 3) makes use of either Gram-Schmidt orthonormalization; or otherwise a QR-decomposition. Once an orthogonal vector to R_t has been generated, with trigonometry the vector parallel to R_t is calculated and added to the horizontal component. The sum of these vectors is correlated with the scaled and centred returns and as correlation is unchanged during linear transformation.

Chapter 4

Pre-Shrinkage Results

4.1 Overview of Chapter

In this chapter we introduce the models and analysis prior to shrinkage and IC reduction. Note that IC reduction is a prerequisite to applying shrinkage in order for the results to be coherent. The results present the returns, the number of constituents within the portfolio and the risk, assuming an IC of 1. After this risks and returns under reduced IC (at 0.1; 0.05 and 0.01) are provided. Other metrics are also provided, including R^2 , realised beta, correlation to the index (ALSI) and the tracking error (of the ALSI). Beyond this chapter, an IC of 0.05 is concluded to be used for the application of shrinkage, this will be justified within this chapter as representative of an experienced investor's ability.

4.2 Initial Returns and Risk Comparative

In this section we compare performance out-of-sample for two standard portfolios: being the minimum-variance (min-var); and the mean-variance (MV) portfolios, to the performance of the resampled, equivalent frontiers. The JSE all shares index (ALSI) is used as a benchmark. Recall, within this section a 100% forecast accuracy is assumed- this intuitively will cause the minimum variance portfolio to consistently under-perform and the MV portfolio to consistently over-perform by implication of 'perfect foresight'.¹

¹With perfect foresight, the end return figure is virtually useless, and not able to be used as a comparative measure. One is using the exact 'out-of-period' returns as the 'expected returns' within the quadratic algorithm. This is useful for comparing performances to resampled and to shrunken portfolios; but not within a comparison of different types of portfolios. As an illustration as to why the returns cannot be used comparatively, consider the highest-return portfolio, which is a 100% investment in the single share which is known (with perfect foresight).

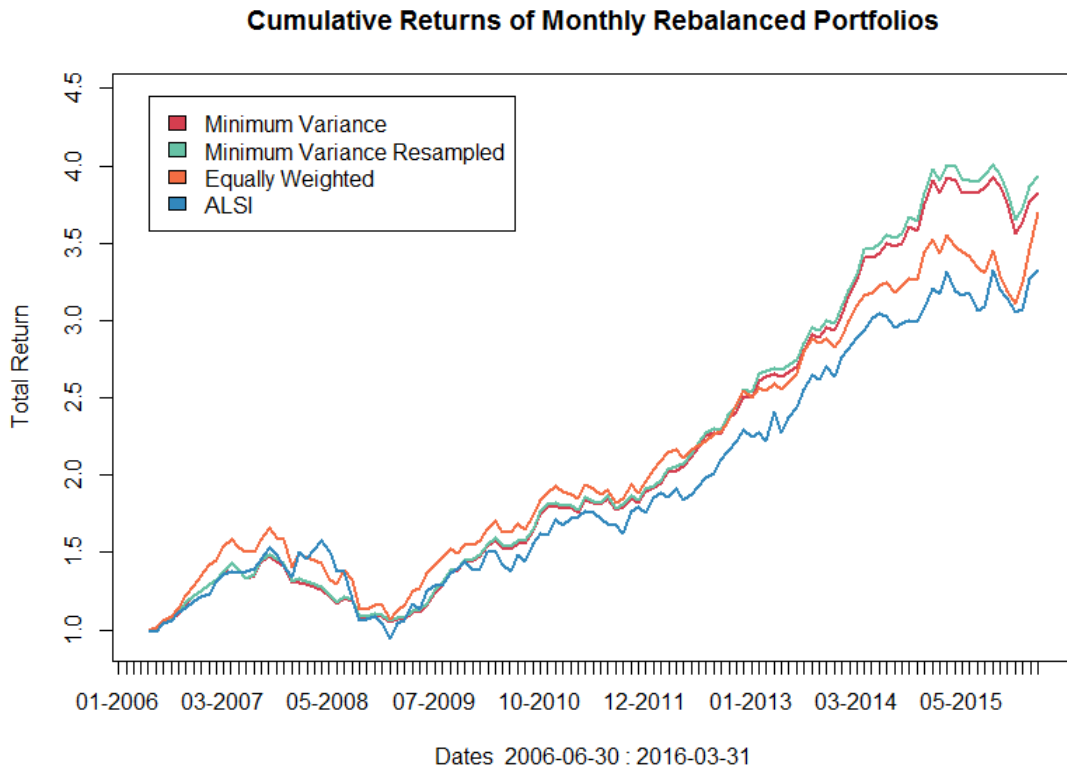


FIGURE 4.1: Cumulative returns 1: Minimum variance; re-sampling and indexes

In 4.2 the resampled minimum variance portfolio is seen to have comparative behaviour to the actual minimum variance portfolio. The longitudinal analysis demonstrates the resampled minimum variance portfolio to consistently hold a higher cumulative return. This is of particular interest, as this holds no assumption about forecast ability of future returns; and allows for an industry-accepted methodology to be improved with very little effort.

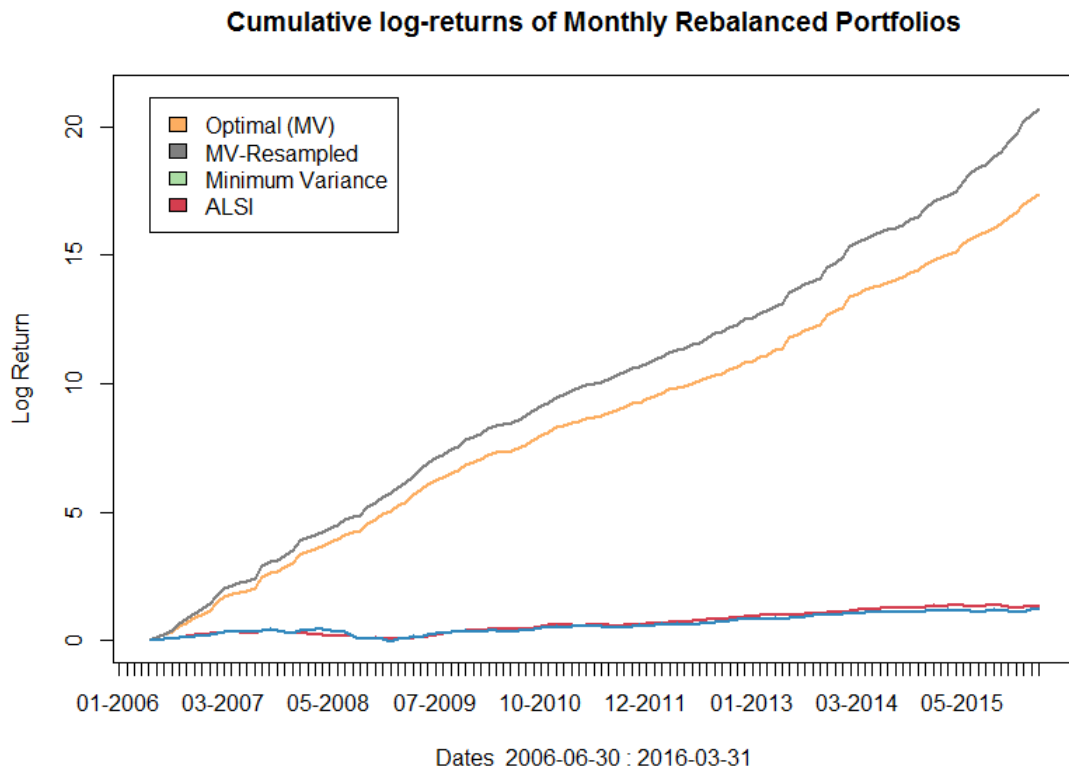


FIGURE 4.2: Cumulative returns 2: log of MV; resampling and indexes

Similarly, 4.3 shows closely linked behaviour between the resampled MV and the actual MV portfolios. There is suggestion of the benefit towards furthering the analysis of the literature with the resampling procedure. The resampled portfolio is seen to attain a far higher cumulative return; by a factor of order $\exp(4)$.

4.2.1 Longitudinal Risk Behaviour of Resampled Portfolios

Total Risk of Portfolios

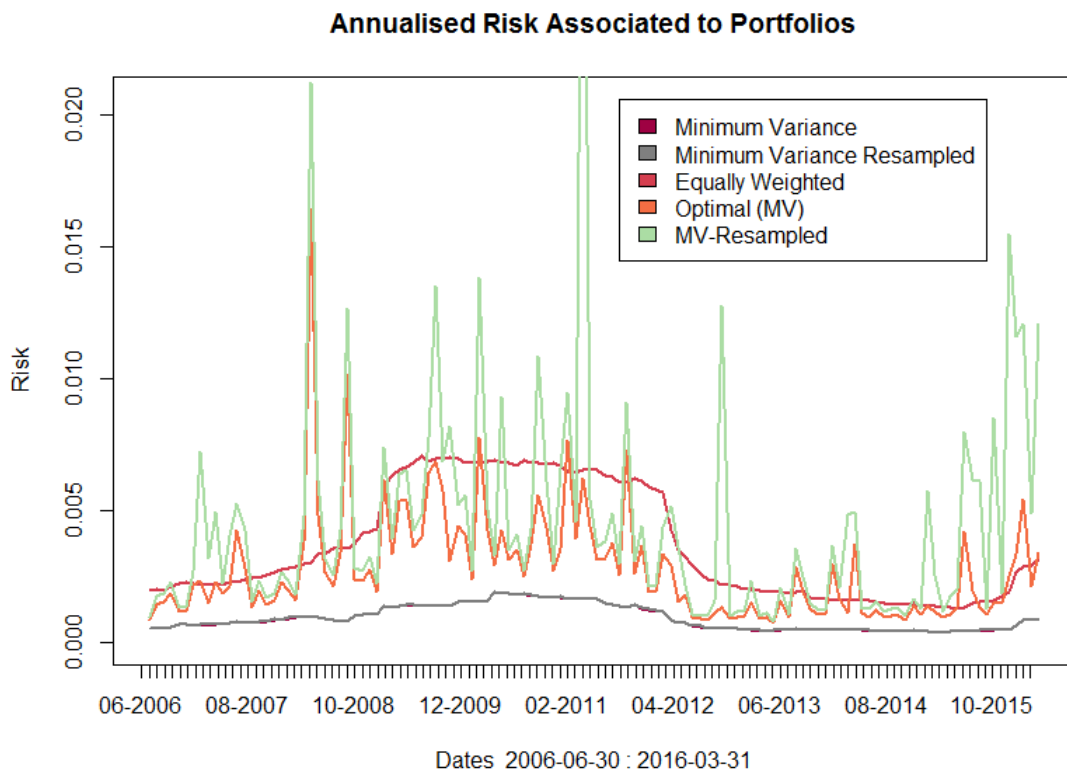


FIGURE 4.3: Longitudinal Comparison of Total Risk exhibited by various portfolios

4.3 portrays a virtually identical longitudinal risk between a portfolio and its corresponding resampled version. The underlying patterns are easier to distinguish through aggregation, and observing a rolling average (of the previous year) as in 4.4. The similarity in behaviour is necessary whilst on the premise of an aggregation of 'statistically equivalent' portfolios to the original. Risk for the portfolios, including the SWIX has been done using the portfolio variance metric. Another necessity is to bear in mind the tracking error, as this is one of the primary risk-metrics used by active portfolio managers.

Clearly demonstrated; the minimum variance portfolio (dark red) holds the least risk; and the resampled minimum variance portfolio is superimposed almost directly on top, with a very small increase in risk between the two. In contrast, the act of resampling upon the MV portfolio increased risk drastically (compare orange to green in 4.4). Thus the effect on the minimum variance portfolio appears all round a better performer with a larger return and similar risks; however the MV portfolio shows no immediate improvement from resampling under a perfect forecast. In the next chapter, the applications of resampling are superimposed into the methodology of covariance-shrinkage.

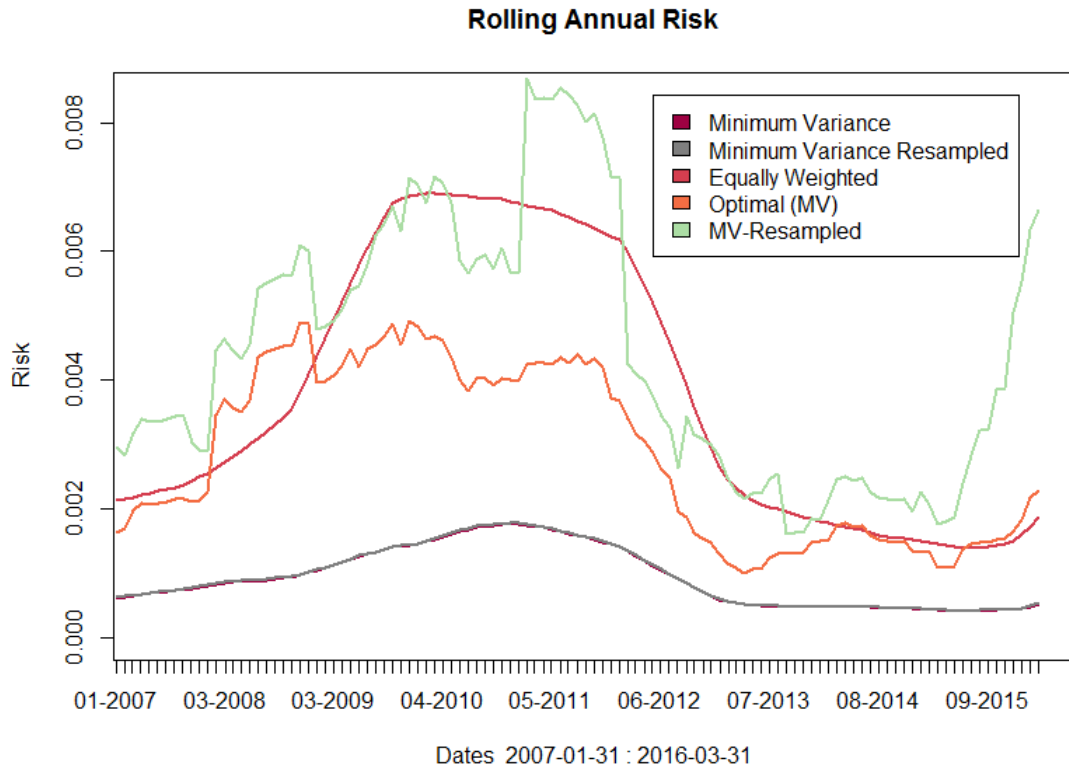


FIGURE 4.4: Rolling annualised risk exhibited by various portfolios

4.2.2 Reduction of the IC

The information coefficient (IC) is a measure of the ability an analyst has to forecast future returns. Up to this point, the MV portfolio has been constructed under the assumption of perfect foresight, using the actual returns to proxy the investor's forecasted returns (i.e. assuming 100% accuracy in forecasting returns). This has allowed for theoretic research, but clearly is not applicable in practice. To allow for practical implications of the research one can coerce a set of returns representative of an analyst's skill level. The suggested skill level for practical research is found within the literature: Grinold and Kahn (1999) describe a great IC at 0.1; and a good IC at 0.05. The IC is exactly the correlation between the forward looking returns and the estimated returns for the corresponding period; thus multivariate tools (vector mathematics) can be used to artificially construct an appropriate set of forecast returns.

Original methodology of Ye (2008) could be used however this paper prefers to justify the use of artificially generated returns through a vector-algebra proof.

4.2.3 Practical Application

Through the generation of a correlated series to the return R_t at any time t we are altering the information coefficient from 1 down to a more realistic

representation of an analyst's ability to forecast. This paper uses an "excellent analyst ability" (0.1); and a "good analyst ability" (0.05) (Grinold and Kahn, 1999) level IC. Further a low IC, allowing for a correlation of only 1% between forecasts and the true returns is used for assessment of the preference of investment strategies when a marginal forecast ability is held. It should be noted that the generation of the correlated returns S_t does not consider the covariance matrix during the generation. This was allowed as a precaution to avoid compounding sample sensitivity within the covariance matrix. There are other processes available beyond the scope of this paper, as it would add unnecessary complication to an algorithm with already acts under non-perfect information.

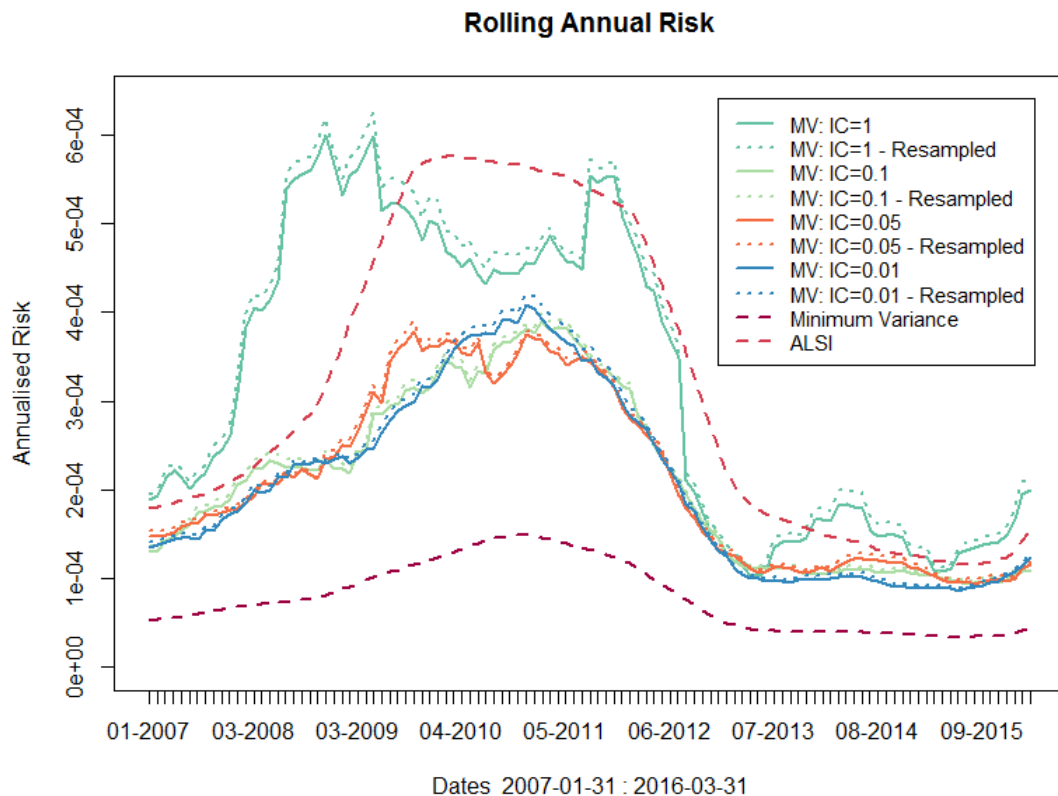


FIGURE 4.5: Resampling Risk under Reduced IC

Demonstrated in 4.5 we find the required performance of minimum variance risk lower than all others (by definition), taking an average value of 7.8×10^{-5} , tabulated in 4.2. Other performances comparable, notably we see resampled performances typically taking on a higher volatility than their non-resampled correspondent. This behaviour is not allowing for any other confounding influences thus we see at various IC levels, a resampling applied to it, increases risk of the MV portfolio. Note that an equally-weighted portfolio in fact holds what appears to be the highest risk, even in comparison to the MV portfolio under various non-perfect foresight ($IC < 1$).

TABLE 4.1: Annualised Risks June 2006 : March 2016

Portfolio	IC level	Average Risk	Risk at 3/2016
MV	1	0.000323	0.000199
MV Resampled	1	0.000337	0.000215
MV	0.1	0.000209	0.000110
MV Resampled	0.1	0.000215	0.000114
MV	0.05	0.000211	0.000116
MV Resampled	0.05	0.000216	0.000119
MV	0.01	0.000205	0.000124
MV Resampled	0.01	0.000211	0.000128
Mininum Variance		0.000078	0.000044
Equally Weighted		0.000315	0.000154

Cumulative Returns of Monthly Rebalanced Portfolios

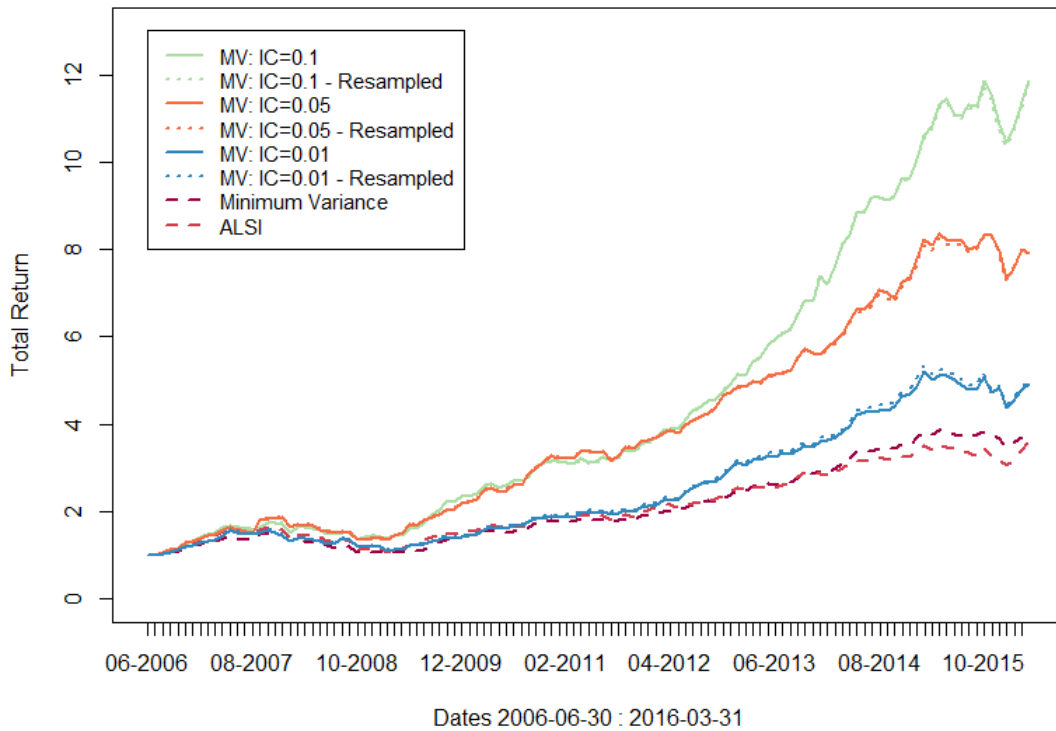


FIGURE 4.6: Resampling Return under Reduced IC

In 4.6 we observe behaviour described in the introduction to IC reduction - the lower the IC the less correlated the returns forecasted (input to the quadratic optimisation) are to the actual returns. Ergo, we observe, as expected a positive relationship between the returns and the IC-level. Resampling in this instance appears to not provide a significant contribution;

in an IC of 0.01 we see resampled MV portfolio (cumulative return 5.00) lying slightly above the equivalent portfolio without resampling (cumulative return 4.93), but the IC 0.1 performance contrasting at a resampled portfolio (cumulative 11.78) below the equivalent non-resampled (cumulative 11.86). Observed from the tabulated graph results in 4.2.

TABLE 4.2: Cumulative Returns June 2006 : March 2016

Portfolio	IC level	50 th Percentile	Total at 3/2016
MV	1	11116.112	390232657
MV Resampled	1	13492.171	609503740
MV	0.1	2.800076	11.86838
MV Resampled	0.1	2.788884	11.78734
MV	0.05	2.853835	7.907233
MV Resampled	0.05	2.874510	7.919742
MV	0.01	1.779587	4.928043
MV Resampled	0.01	1.808845	4.995545
Mininum Variance		1.739086	3.745821
Equally Weighted		1.576561	3.648825

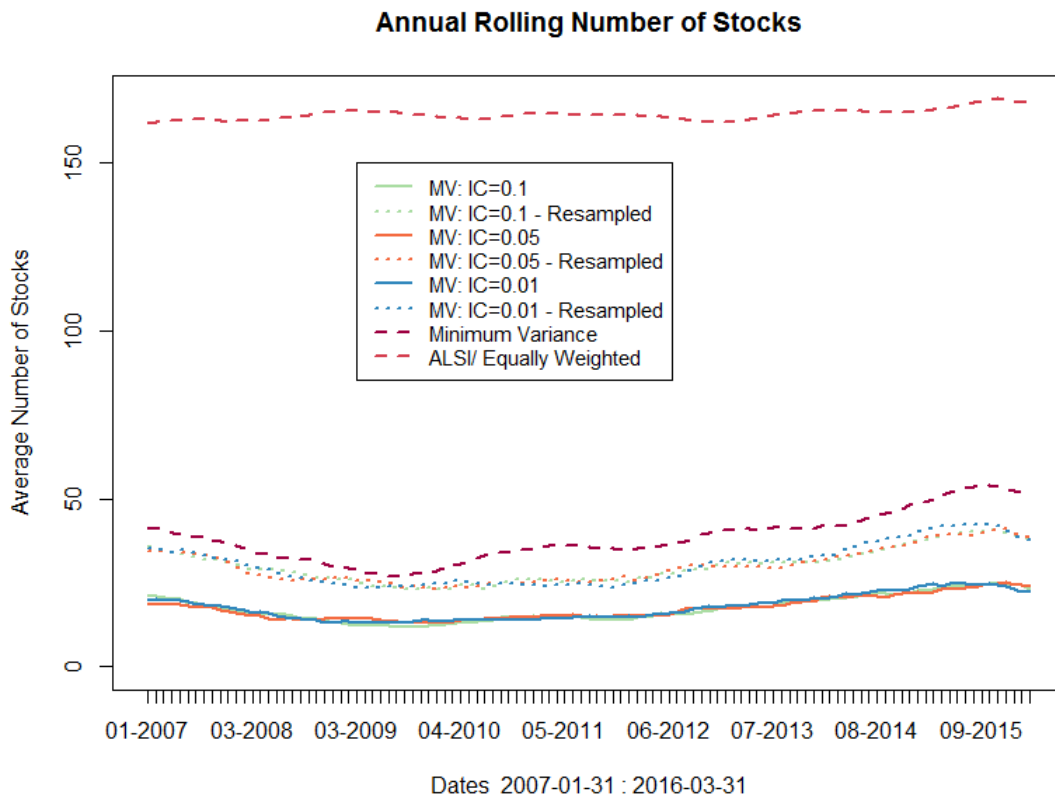


FIGURE 4.7: Constituents under Reduced IC

4.7 displays the number of stocks (constituents) per portfolio. We can observe the number of stocks active through the ALSI/Equally Weighted portfolio count. By definition, they contain all stocks; and in contrast we observe a low number in especially the MV portfolios. Although there does not appear much difference between each MV portfolio at IC of 0.1 (ave 17.23); 0.05 (ave 17.32); or 0.01 (ave 17.58) we do see a consistent (marginal) increase as IC decreases, coupled with a large increase upon the use of resampling procedures, seeing the respective portfolios above holding 29.92; 29.90; and 30.40 shares respectively on average. ²

To delve into the intuition: the resampling will aggregate several optimal portfolios based on multivariate generation. As such, the randomness allows for the subset of shares selected to differ per each iteration and over the set of all iterations therefore multiple shares have been included over and above in a single iteration (i.e. non-resampled).

TABLE 4.3: Number of Stocks June 2006 : March 2016

Portfolio	IC level	Mean	Constituents 3/15-3/16
MV	1	38.207207	50.83333
MV Resampled	1	73.525526	86.83333
MV	0.1	17.231231	23.16667
MV Resampled	0.1	29.922673	37.83333
MV	0.05	17.32432	23.833333
MV Resampled	0.05	29.904651	38.66667
MV	0.01	17.584835	22.41667
MV Resampled	0.01	30.398649	37.75
Minumum Variance		38.33183	51.5
Equally Weighted		164.34685	167.5

4

We assess each portfolio according to various metrics. 4.4 displays the results of the various tools used which allow us to compare the performances beyond the simple risk and return metrics used up to here. The beta realised, corresponding to the correlation between share movement vs. market (ALSI) movement, moves between -1 and 1. We observe a comparative beta decreasing under lower IC-levels; logically the lower one's ability to predict shares movements; they will be less able to foresee the market movements which is simply a weighted sum of individual shares. Resampling is indicative of neither large, nor significant differences as opposed to no resampling.

The R^2 statistic is a measure of the volatility of the market, which can be explained through the portfolio. We note the largest R^2 , aside from the

²Numeric details in paragraph extracted from 4.3.

⁴IC values, should they differ from 1, indicated as subscript due to spacial constraints

TABLE 4.4: Performance Analytics: Benchmark Performance

Portfolio ³	Realized Beta	R^2	Correlation	Tracking Error
MV	0.6346169	0.05197001	0.2450976	0.20899323
MV Resampled	0.6701703	0.05079465	0.2427087	0.21710218
$MV_{0.1}$	0.5472420	0.39764239	0.6346580	0.03705628
$MV_{Resampled_{0.1}}$	0.5475329	0.39799005	0.6349295	0.03702799
$MV_{0.05}$	0.5856234	0.38940234	0.6281888	0.03749579
$MV_{Resampled_{0.05}}$	0.5874496	0.39495824	0.6325579	0.03723122
$MV_{0.01}$	0.4940229	0.36618982	0.6095958	0.03612627
$MV_{Resampled_{0.01}}$	0.4964976	0.36471185	0.6083927	0.03625161
Minimum Variance	0.393193	0.356711	0.601859	0.035231
Equally Weighted	0.641353	0.569523	0.757101	0.029035

equally-weighted portfolio at 0.57, corresponds to the $MV_{0.1}$ at 0.40. The R^2 decreases with either a larger IC-level (as the "perfect" portfolio does not follow market patterns) and similarly as the IC-level is decreased, as a more random set of returns is naturally not very useful in describing market movement. The resampled portfolios share the same number correct to 3 decimal places, indicating no large impact.

Correlation of a portfolio to the index is calculated as the covariance divided by the standard deviation of each: market and the portfolio. We note a large correlation between the $MV_{0.1}$ at 0.63, again being second only to the equally weighted portfolio with a correlation of 0.76. The same logic as above applies: correlation decreases with either a larger IC-level (as the "perfect" portfolio does not follow market patterns) and similarly as the IC-level is decreased, as a more random set of returns is naturally not very useful in describing market movement.

Tracking Error describes the ability of a portfolio to follow the index returns, with a quadratic penalisation for deviations. The higher the tracking error (TE); the larger the deviations from the index. Tracking error does not discriminate for positive or negative deviation; and thus we find a lowest tracking error associated to the portfolios which obtain a typically lower return. Thus, we find a lowest tracking error of 0.029 from an equal weighting, followed by minimum variance at 0.035. Of the MV portfolios, note the highest TE at the resampled MV with IC=1; i.e. the best performer, and a lowest TE of 0.036 associated to the MV non-resampled portfolio with the lower IC=0.01.

TABLE 4.5: Performance Analytics: Benchmark Performance

Portfolio ⁵	s.d. σ	Treynor	Return/Risk Ratio
MV	0.11397299	10.09100813	0.177518183
MV Resampled	0.12154278	10.07319248	0.181995172
MV _{0.1}	0.03795500	0.39024192	0.016186374
MV Resampled _{0.1}	0.03795894	0.38839848	0.016127080
MV _{0.05}	0.04103530	0.27581998	0.012677980
MV Resampled _{0.05}	0.04087894	0.27530023	0.012691612
MV _{0.01}	0.03567258	0.20969168	0.008608218
MV Resampled _{0.01}	0.03592217	0.21192599	0.008725086
ALSI	0.04401791	0.05740875	0.005230046
Minimum Variance	0.02951679	0.17963933	0.006254734
Equally Weighted	0.03769604	0.10528170	0.006029923

In accordance with 4.5 we can look towards other direct risk assessments. The standard deviation of the portfolio returns is calculated by the returns of the portfolio only, with no reference to the covariance matrix. We observe a large volatility from the MV portfolios with IC=1, of 0.12, associated to the large jumps the returns can provide. As predicted minimum variance allows for mitigation of risk, and thus holds the most consistent set of realised returns. It is useful to note that the MV portfolios hold a lower volatility in returns (MV_{0.1} at 0.038 and MV_{0.05} at 0.041) than even the ALSI at 0.044. The equally weighted portfolio does, however hold a lower s.d. at 0.0376 of all but the min var portfolio and the MV with IC=0.01.

The Treynor measure is a measure amalgamated from several factors: it is the excess return per unit of risk, defining risk as the beta of the share. This holds the MV portfolios as attaining the highest level, with IC=1. At lower IC levels recall there is no large change in beta, and thus a consistent pattern of the Treynor measure decreasing directly with the IC decreasing. Even under reduced IC-levels, the MV portfolios maintain a higher return per risk unit (MV_{0.01} with 0.21) than minimum variance at 0.18; equally weighted at 0.11 and the index at 0.06.

The return/risk ratio is identical to the Treynor measure above, but using the traditional definition for the risk associated with a share movement. This does however follow the identical ranking of return per risk unit of the portfolios as above. At the lowest IC-level of 0.01, the MV portfolios maintain a higher return per risk unit (MV_{0.01} with 8.6×10^3) than minimum variance at 6.2×10^3 ; equally weighted at 6.0×10^3 and the index at 5.2×10^3 .

Chapter 5

Shrinkage of Covariance Matrix

5.1 Application of Shrinkage

This section is executed under a constructed IC of 0.05 to represent the abilities of a competent forecaster whilst assessing the performance of shrinkage and resampling applied to a reasonable mean-variance portfolio. The arbitrary selection of 0.05 will be evaluated in chapter 6, where we examine the applicability of the alternative IC-levels.

In addition, this chapter uses Median-shrinkage as a benchmark shrinkage to which other shrinkage methodologies (ACM; CCM and PCA) are compared across various metrics. Thus as a template for the diagrams to follow, expect to see the performance metrics of six portfolios in comparison: i.e. the cross-section of resampled vs non-resampled; and no-shrinkage vs median vs another shrinkage.

5.1.1 Shrinkage Factor

Recall equation 3.3: the definition of the shrunk covariance $\hat{\Sigma}_S$:

$$\hat{\Sigma}_S = \alpha F + (1 - \alpha)SC$$

The shrinkage factor α is optimised via Ledoit & Wolf's (2003*) mathematical derivation outlined in equation 3.11 This paper finds the optimal longitudinal shrinkage factors relevant for the various structured matrices (F).

Using PCA, a direct dimension reduction technique described in Chapter 3.3.7 we find a shrinkage target using eigenvalues which explain the majority of the variation within the dataset.

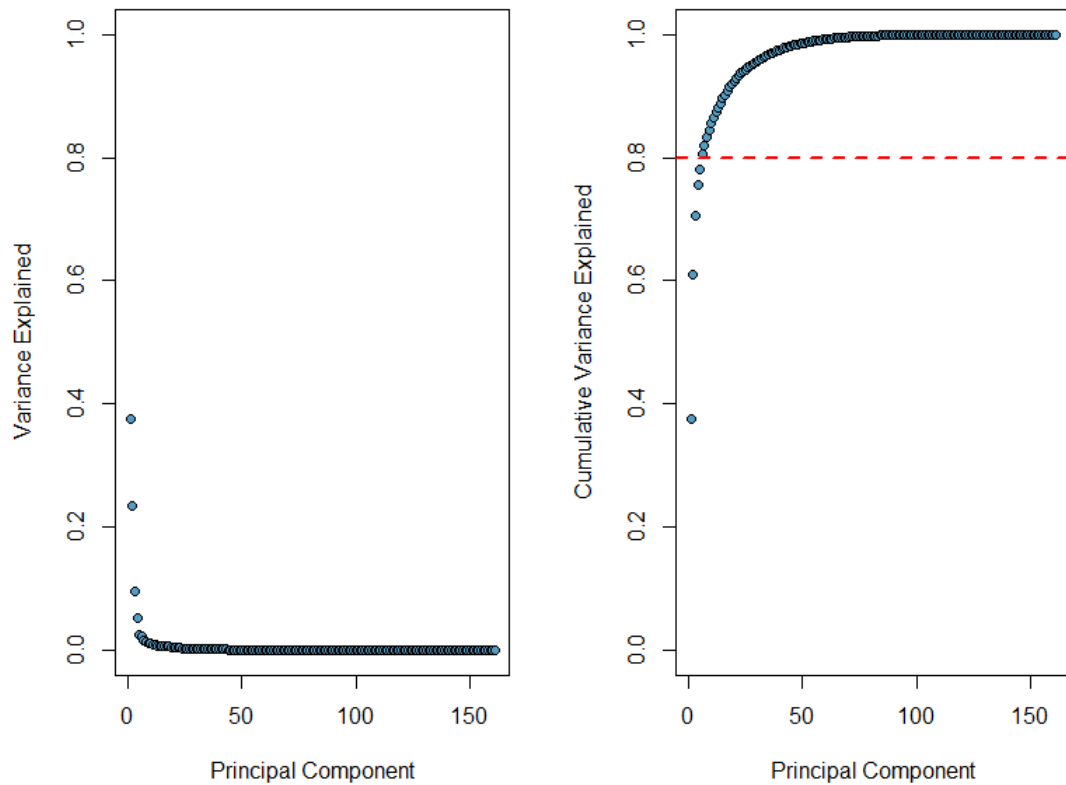


FIGURE 5.1: Variation captured by principal components, order respective in decreasing order of corresponding eigenvalues of each PC over period: June-2003 : May-2006

The exponential decrease in the variation associated to the eigenvectors (the eigenvalues) allows one to explain the majority of the information with a much lower specification error. In 5.1 it is clear that within 6 principal components, over 80% of all the variation within the 162 stocks of interest has been captured. For the first period of historical data, (June-2003 : May-2006), used for analysis over June 2006.

5.1.2 Historical Optimal Shrinkage Factor

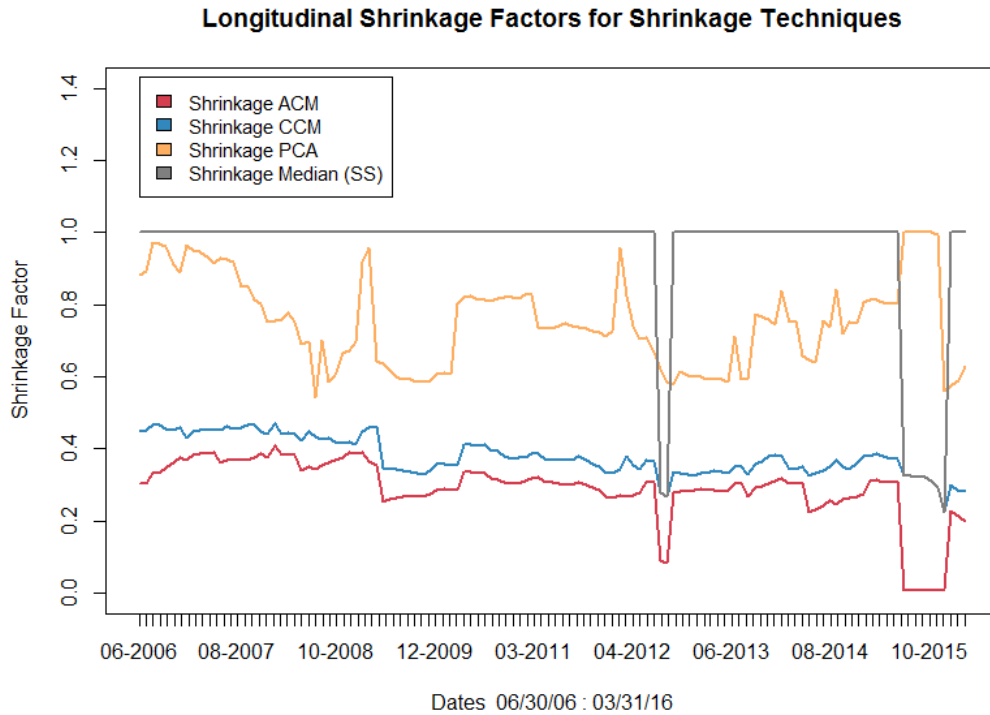


FIGURE 5.2: Shrinkage factor over period, allowing a comparative view between PCA; ACM; CCM and Median

The shrinkage factor α in 5.2 has been calculated in accordance with 3.11. We find a strong coherence in movement between ACM and CCM naturally due to the similar composition of each. On the other side however, we find that the SIM shrinkage is consistently the smallest, and the PCA shrinkage consistently the largest.

We can see the 2008 market crash in South Africa having a large impact on the ACM and the CCM; suggesting that in times of instability the benefits of shrinkage becomes less effective during these times is lower than during times that the economy is more predictable. The optimal shrinkage using these aggregated methods is thus lower, i.e. the shrinkage factor is reduced. This notion is inherent to the type of risk a portfolio is exposed to, where during a crash the sampling error becomes a much smaller component of the total portfolio variance thus strategies of hedging risk require resources dedicated elsewhere.

We calculate the averages, displayed in 5.1 confirming large differences in the average behaviour, where we see the optimal shrinkage factor for ACM α around 0.29 larger than the SIM α ; CCM α almost 0.09 above that at 0.38, and finally PCA α is seldom below 1. The aggregate of the PCA α is at 0.999 due to a single point (that of Dec-2015) of 0.933, with all other shrinkage factors at precisely 1. This period corresponds to the exact same period in which ACM α increased to 0.208 (Dec-2015) from a six month slump of 0.008 and below (June-2015: Nov-2015).

TABLE 5.1: Average Shrinkage to Covariance

Shrinkage to Covariance Estimator	Average Shrinkage Factor
Constant Correlation Model (CCM)	0.37793353
Average Correlation Model (ACM)	0.29209042
Principal Component Analysis (PCA)	0.76819502
Median (SS)	0.94476496

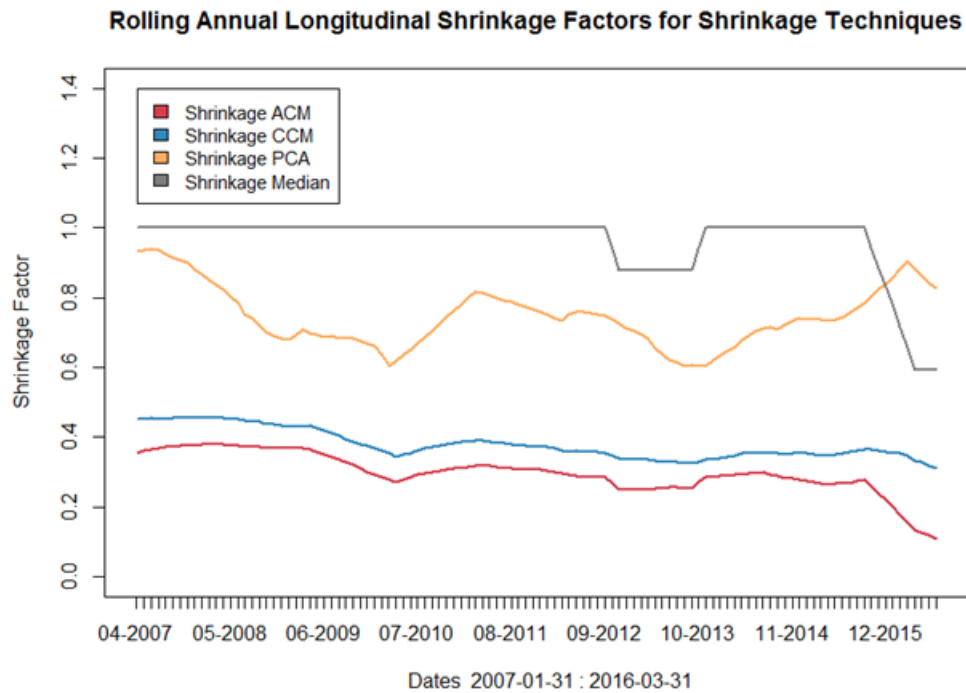


FIGURE 5.3: Shrinkage rolling average, allowing a comparative view between PCA; ACM; CCM and Median

5.2 Application of Shrinkage Techniques

The graphs that follow use **R** default shrinkage; in accordance to Schafer & Strimmer (2005)'s shrinkage towards median:

$$\lambda_{var}^* = \left(\sum_{k=1}^p Var(s_{kk}) \right) / \sum_{k=1}^p (s_{kk} - medians)^2$$

The graphs following use the comparison to the above covariance resulting from the shrinkage applied above.

5.2.1 Shrinkage onto Mean-Variance Portfolios

For comparison's sake, this chapter is performed under consistent assumption of an IC at 0.05. Additionally to avoid confounding effects, the same set of randomly generated correlated returns was repeatedly used for each shrinkage method (separately generated and imported prior to each analysis). Thus the comparisons are purely looking to assess the differences

between risk; returns; and other performance metrics as phenomena on the shrinkage methods.

5.2.2 Cumulative Returns under Shrinkage

In 5.4 observe the baselines of minimum variance and equally-weighted portfolios. These are outperformed, by far, under the assumption of an IC of 0.05 by all portfolios versions of mean-variance portfolios. In this section of the chapter we observe the following behaviour in comparing the various shrinkage techniques:

TABLE 5.2: Performance Rankings by Returns

Shrinkage-Type	Highest Return	Lowest Return
Median	MV	MV-Resampled Shrinkage
ACM	MV	MV-Resampled Shrinkage
CCM	MV-Resampled Shrinkage	MV-Resampled
PCA	MV-Resampled Shrinkage	MV-Resampled

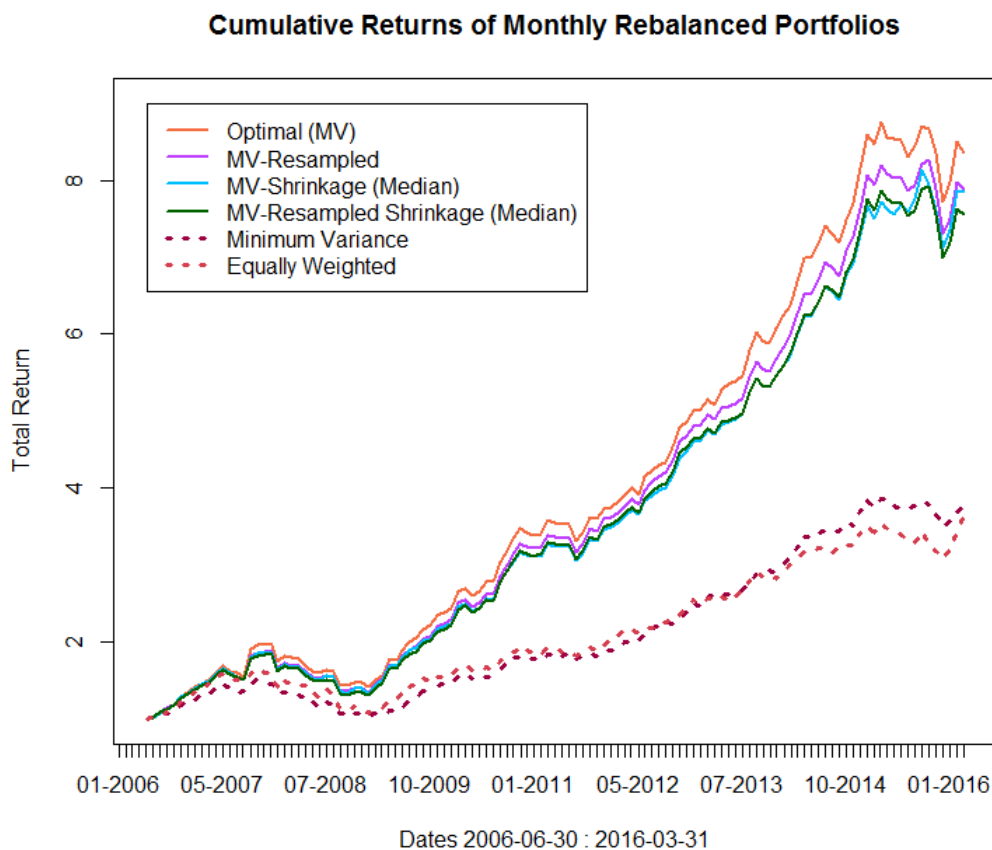


FIGURE 5.4: Returns Baseline with IC=0.05

The remainder of this section will look exclusively at resampling and various shrinkage methods applied upon the mean-variance portfolio formulation. Reference to equally-weighted and minimum-variance portfolio

performance will be excluded from the remainder of this section as the focus is on a comparison of mean-variance performance under the shrinkage and resampling techniques.

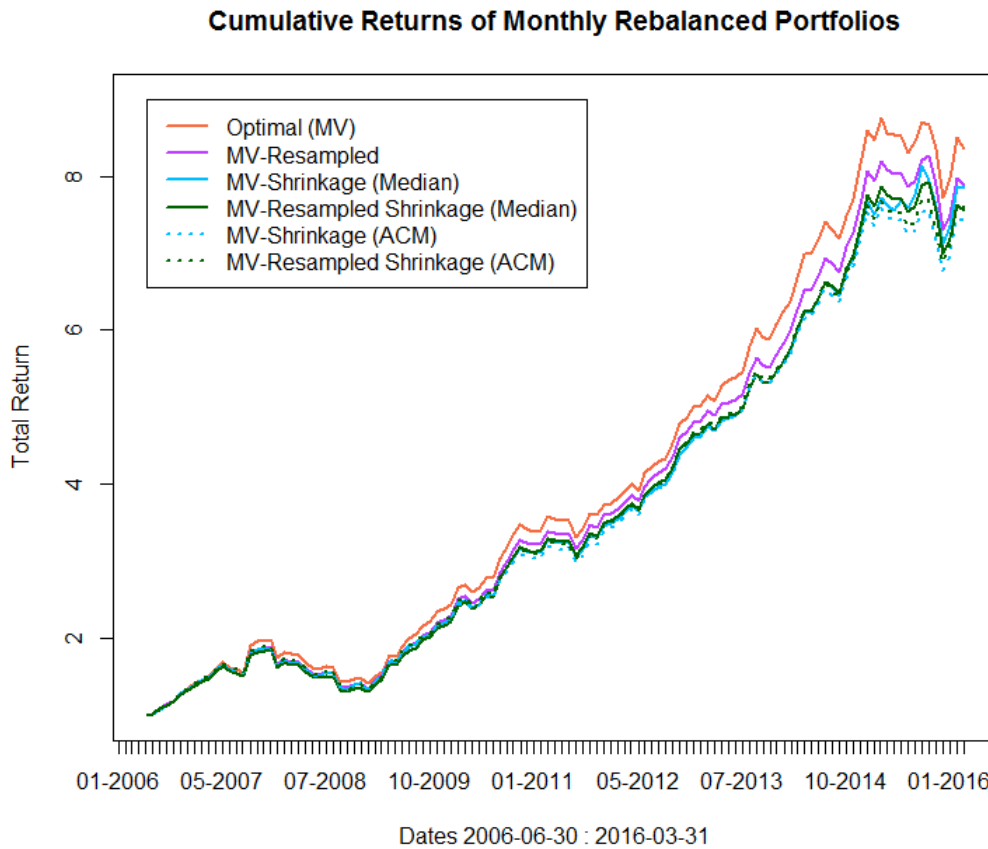


FIGURE 5.5: Returns under ACM with $IC=0.05$

5.5 is portraying ACM cumulative returns uses the Median shrinkage methodology as a benchmark, as described at the beginning of this chapter. We observe the clear best performer of the mean-variance portfolio, under no adjustments. The MV portfolio boasts a cumulative return of 8.35 by the end of March 2016; outstripping the second best (MV resampled) at 7.88 and thereafter (MV-shrinkage median) at 7.84 (at the same date).

TABLE 5.3: ACM vs Median Shrinkage Returns with $IC=0.05$

	Resampling	50 th Percentile	Total at 3/2016
MV		3.42207	8.35124
MV	Yes	3.23689	7.88684
MV - Shrinkage Median		3.12813	7.84105
MV - Shrinkage Median	Yes	3.14799	7.54975
MV - Shrinkage ACM		3.06319	7.43824
MV - Shrinkage ACM	Yes	3.12594	7.59807

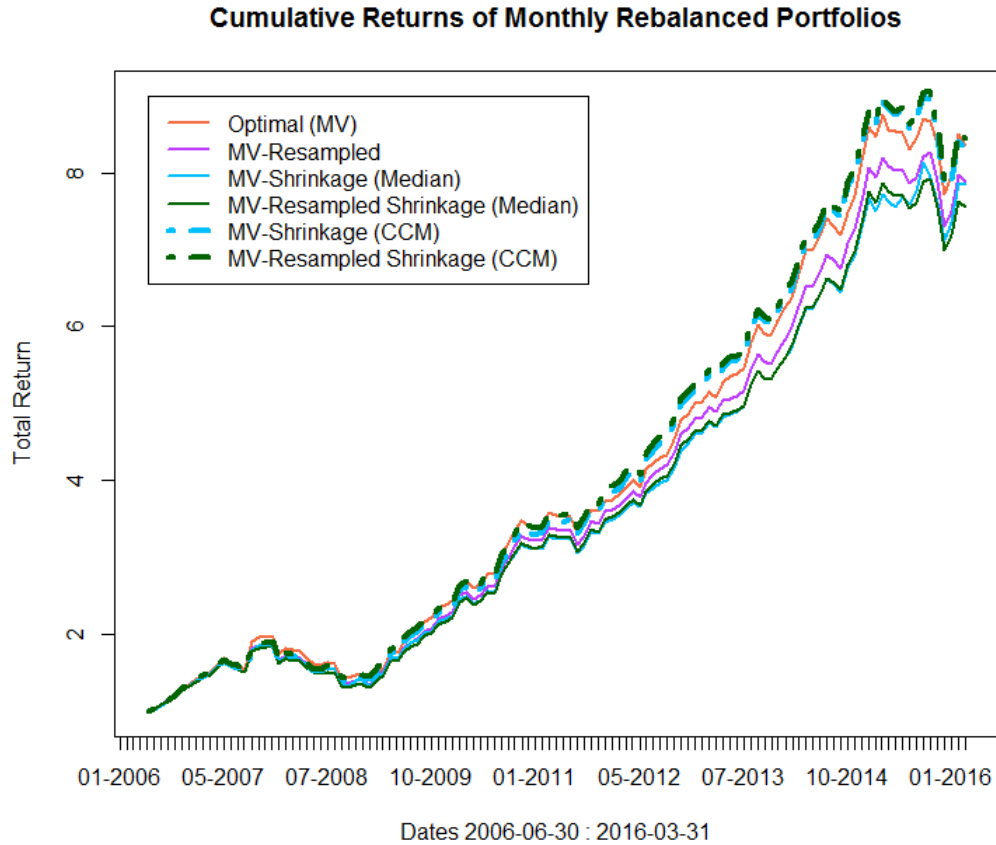


FIGURE 5.6: Returns under CCM with IC=0.05

A contrasting behaviour is seen between the shrinkage methods in 5.6; in which the resampled CCM provides the best cumulative return performance at 8.44 yet the resampled median shrinkage performance at 7.55 is worse off than without either any techniques applied (MV at 8.35). One can also note the contrasting interaction between resampling applied to CCM shrinkage, within which an improvement in performance is seen by 0.12; whereas the application to median shrinkage resulted in a decrease by almost 0.3. These contrasts alone have demonstrated that there is no simple description of the interaction between resampling and shrinkage.

TABLE 5.4: CCM vs Median Shrinkage Returns with IC=0.05

	Resampling	50 th Percentile	Total at 3/2016
MV		3.42207	8.35124
MV	Yes	3.23689	7.88684
MV - Shrinkage Median		3.12813	7.84105
MV - Shrinkage Median	Yes	3.14799	7.54975
MV - Shrinkage CCM		3.32796	8.32054
MV - Shrinkage CCM	Yes	3.41884	8.44076

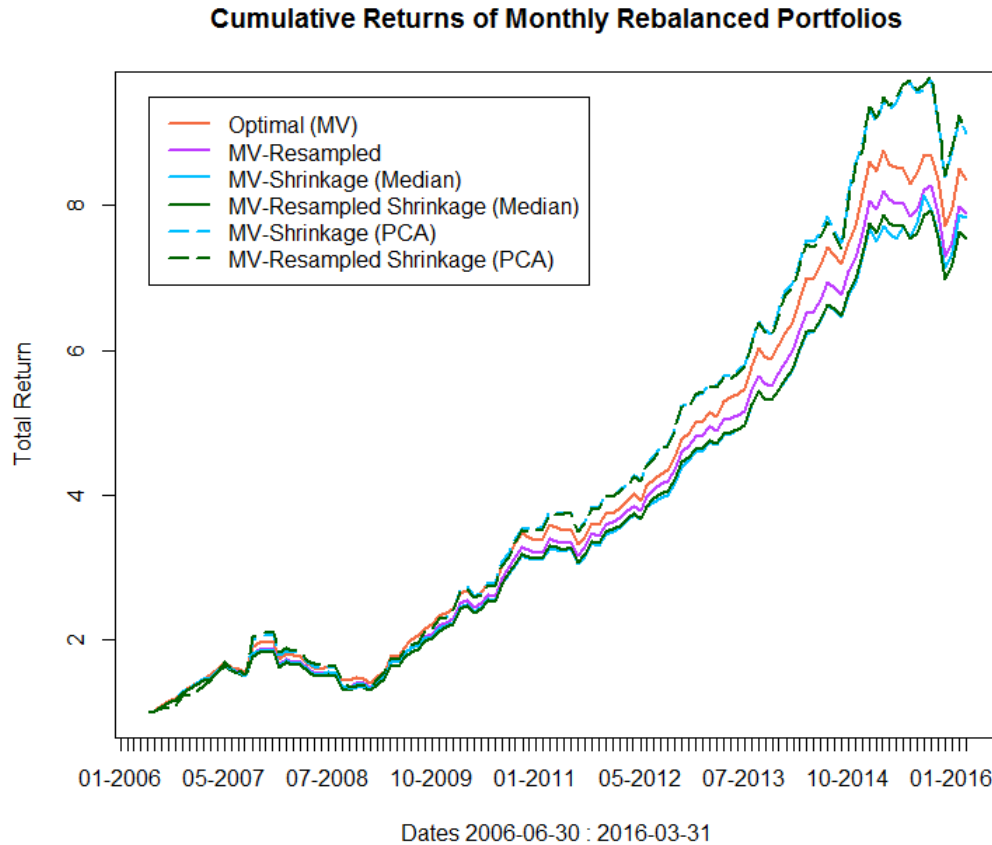


FIGURE 5.7: Returns under PCA with IC=0.05

5.7 demonstrates a follow-in-suit of the behaviour observed under CCM; whereby both shrinkage alone, and further, the conjunction in respect of resampling, provides an increase in performance. The graph demonstrates a clear out-performance of PCA shrinkage in relation to the absence thereof. The MV return without shrinkage (8.35) is outperformed by 0.65 (9.00) through shrinkage and by an additional 0.03 (at 9.03) through resampling on top of the PCA shrinkage. By assessment through other risk we will assess whether the returns performance operates in alignment or in contrast to the risk performance; as shrinkage and resampling are implemented.

TABLE 5.5: PCA vs Median Shrinkage Returns with IC=0.05

	Resampling	50 th Percentile	Total at 3/2016
MV		3.42207	8.35124
MV	Yes	3.23689	7.88684
MV - Shrinkage Median		3.12813	7.84105
MV - Shrinkage Median	Yes	3.14799	7.54975
MV - Shrinkage PCA		3.54494	9.00193
MV - Shrinkage PCA	Yes	3.51519	9.03557

Table 5.5 provides the tabulation of the graphic (5.7); we now progress onto a detailed analytical breakdown of additional information in the subsequent table for further insights. 5.6 displays several metrics of performance

TABLE 5.6: Performance Analytics 1 with IC=0.05

	Beta	R^2	Correlation	Tracking Error
MV	0.60752	0.38732	0.62654	0.03830
MV - R	0.58679	0.39423	0.63198	0.03725
MV - S (Median)	0.60065	0.41226	0.64597	0.03658
MV - R/S (Median)	0.58989	0.40051	0.63690	0.03690
MV - S (ACM)	0.59774	0.40928	0.64368	0.03659
MV - R/S (ACM)	0.60364	0.40755	0.64235	0.03686
MV - S (CCM)	0.54025	0.35728	0.60230	0.03824
MV - R/S (CCM)	0.55578	0.37302	0.61512	0.03778
MV - S (PCA)	0.61685	0.26437	0.52025	0.04848
MV - R/S (PCA)	0.62050	0.26048	0.51653	0.04909
Equally Weighted	0.64770	0.56834	0.75633	0.02914
ALSI	1.00000	1.00000	1.00000	0.00000

S Shrinkage
R Resampled
S/R Both applied

previously referenced- and similarly these are used here to assess the performances and similarities between the various shrinkage and resampling techniques.

We find the largest realized beta from the MV-set of portfolios as the PCA-shrinkage resampled at almost 0.65; with the non-resampled close behind (0.62). This indicates an ability to leverage the index performance; but a lack of necessity of the index movement to perform well. The nature of PCA allows the strongest patterns (components) to be extracted, which appear to coincide with the make-up of the movements of the index, thus allowing more leverage than other shrinkage techniques. The lowest beta has been the CCM, both resampled and non lie at under 0.56; in comparison to the other performances this is lower, but still relatively close to benchmark (equally-weighted at 0.65). Considering R^2 , a similar metric to beta, yet we find completely contrasting behaviour: the largest value sits at 0.41 for the median-shrinkage and similar scores for the ACM shrinkage. The lowest R^2 in this case falls to 0.26 for both PCA non-resampled and resampled. The correlation to the ALSI is observed as being the highest for the median-shrinkage at 0.65; and lowest for the PCA, at 0.52 (the only correlation below 0.6). Lastly in the table we note the largest tracking error is associated to the PCA both resampled and non-resampled at 0.040 and 0.048 respectively; lowest tracking error is associated to the median-shrinkage and ACM shrinkage which are near-identical with and without shrinkage, all rounded at 3 d.p.¹ to 0.037.

The use of resampling does not appear to have a consistent effect upon the different shrinkage techniques. The shrunken portfolios (non-resampled) do not appear to either gain a higher or lower value consistently under any

¹decimal place

metric through the addition of shrinkage. This will be assessed further to establish whether there is consistency in the minimum-variance portfolio, once the IC effect has been removed.

5.2.3 Annualised Risk under Shrinkage

In 5.8 observe the baselines of minimum-variance, which by definition has a lower volatility than all other portfolios, and equally-weighted portfolio which is always larger than at least one of the MV-group. This comparison is provided here to solidify the reader's understanding of the positions of the typical mean-variance portfolios, relative to various benchmarks. In this section of the chapter we observe the following behaviour in comparisons of the various shrinkage techniques:

TABLE 5.7: Portfolio Rankings by Risk

Shrinkage-Type	Highest Risk	Lowest Risk
Median	MV Shrinkage	MV-Resampled
ACM	MV-Resampled Shrinkage	MV-Resampled
CCM	MV-Resampled Shrinkage	MV-Resampled
PCA	MV-Resampled Shrinkage / MV-Shrinkage	MV-Resampled

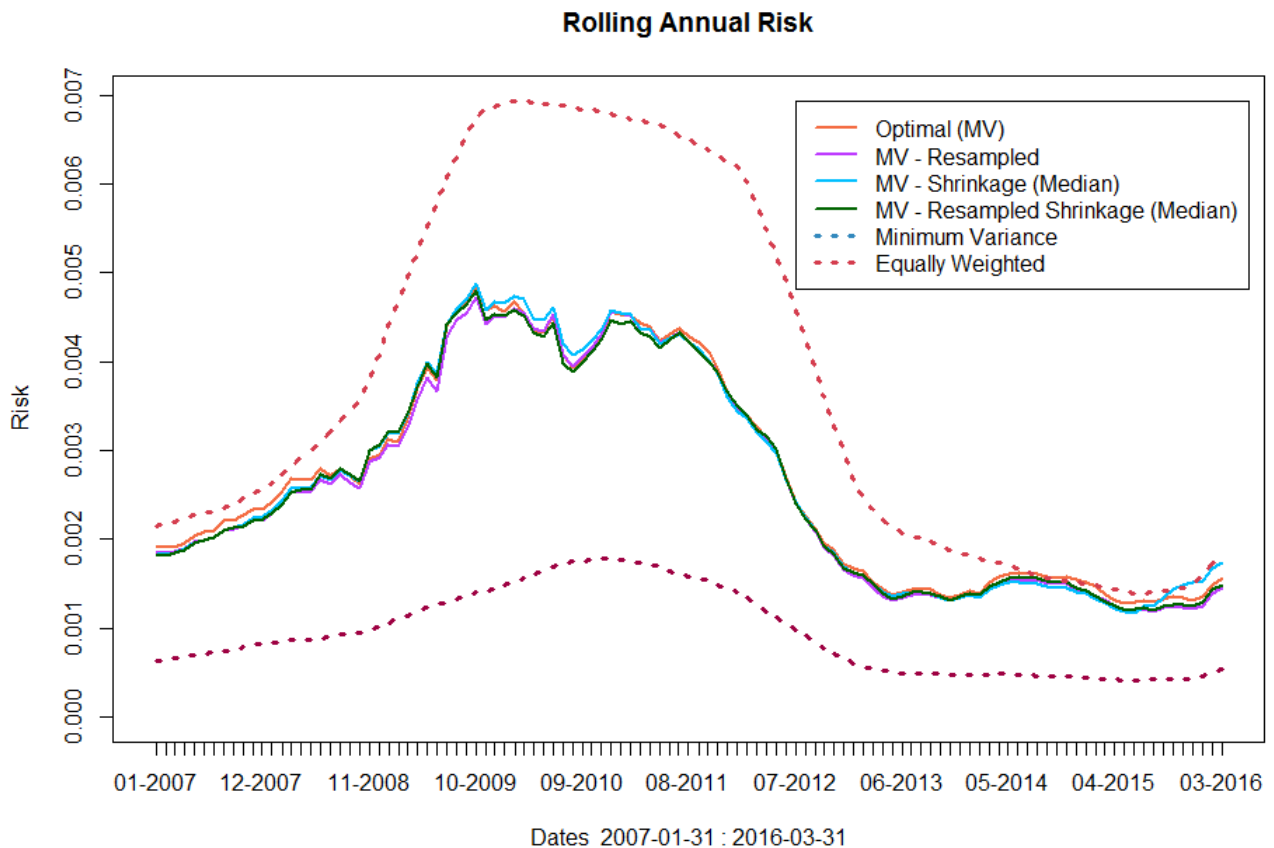


FIGURE 5.8: Annualised Risk Baseline with IC=0.05

The remainder of this section will look exclusively upon resampling and various shrinkage methods upon the mean-variance portfolio. Due to the nature of comparison of identification of the optimal shrinkage method, reference to equally-weighted and the minimum variance portfolio performance will be excluded from the remainder of this section.

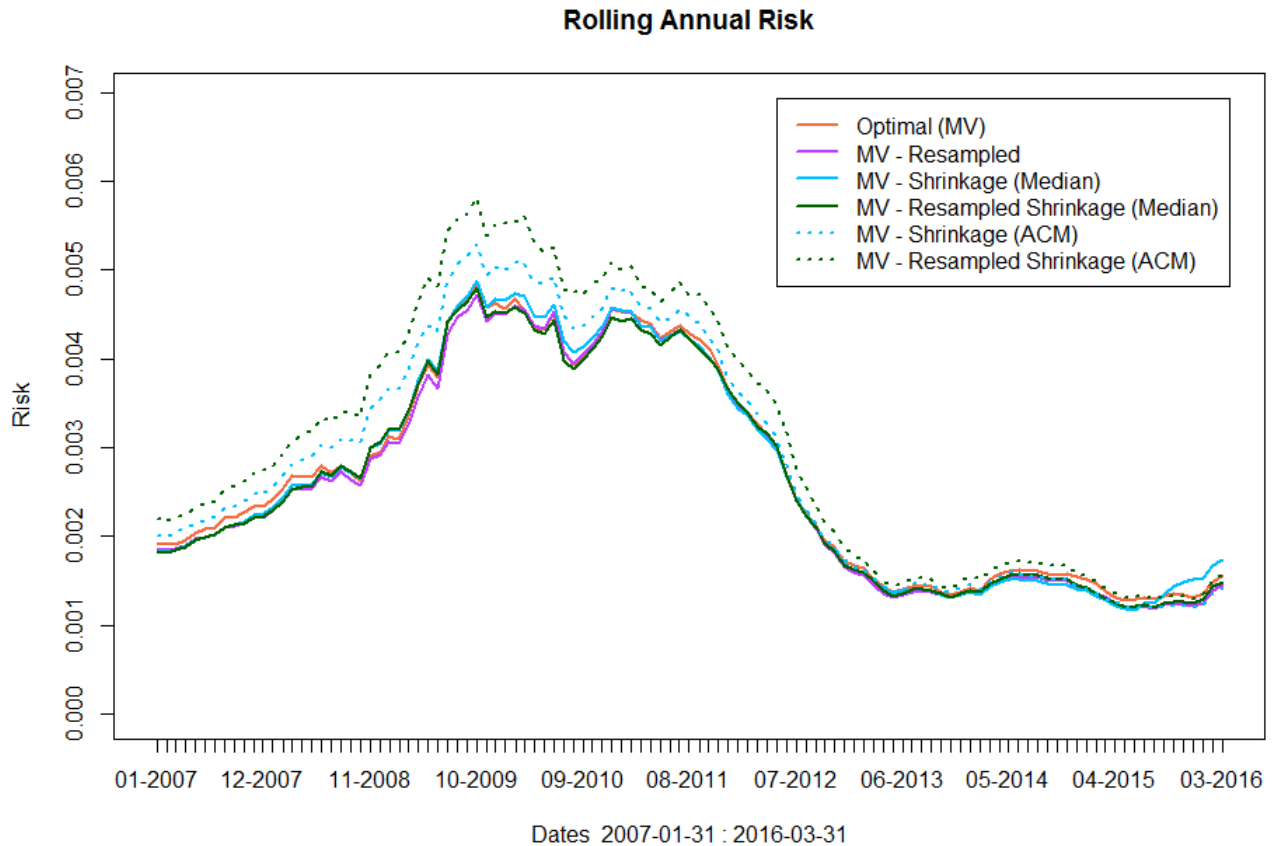


FIGURE 5.9: Risk under ACM with $IC=0.05$

5.9 portrays the first comparative plot of the associated risks under various shrinkage methods. The riskier, between the median and ACM shrinkage, is seen to be ACM at an average of 0.0028. There is a brief period of anomaly however: for the year from March 2015- March 2016 the median shrinkage risk overtakes ACM, ending at 0.0017 in March (0.0003 higher than ACM non-resampled shrinkage). We note that resampling of ACM portfolios increases annualised risk; by an average increase from 0.0028 to 0.0031.

TABLE 5.8: ACM vs Median Shrinkage Risk with $IC=0.05$

	Resampling	Average Risk	Risk at 3/2016
MV		0.00267856	0.00156463
MV	Yes	0.00261238	0.00145287
MV - Shrinkage Median		0.00266116	0.00173307
MV - Shrinkage Median	Yes	0.00262945	0.00149138
MV - Shrinkage ACM		0.0028297	0.00142125
MV - Shrinkage ACM	Yes	0.00307599	0.00158292

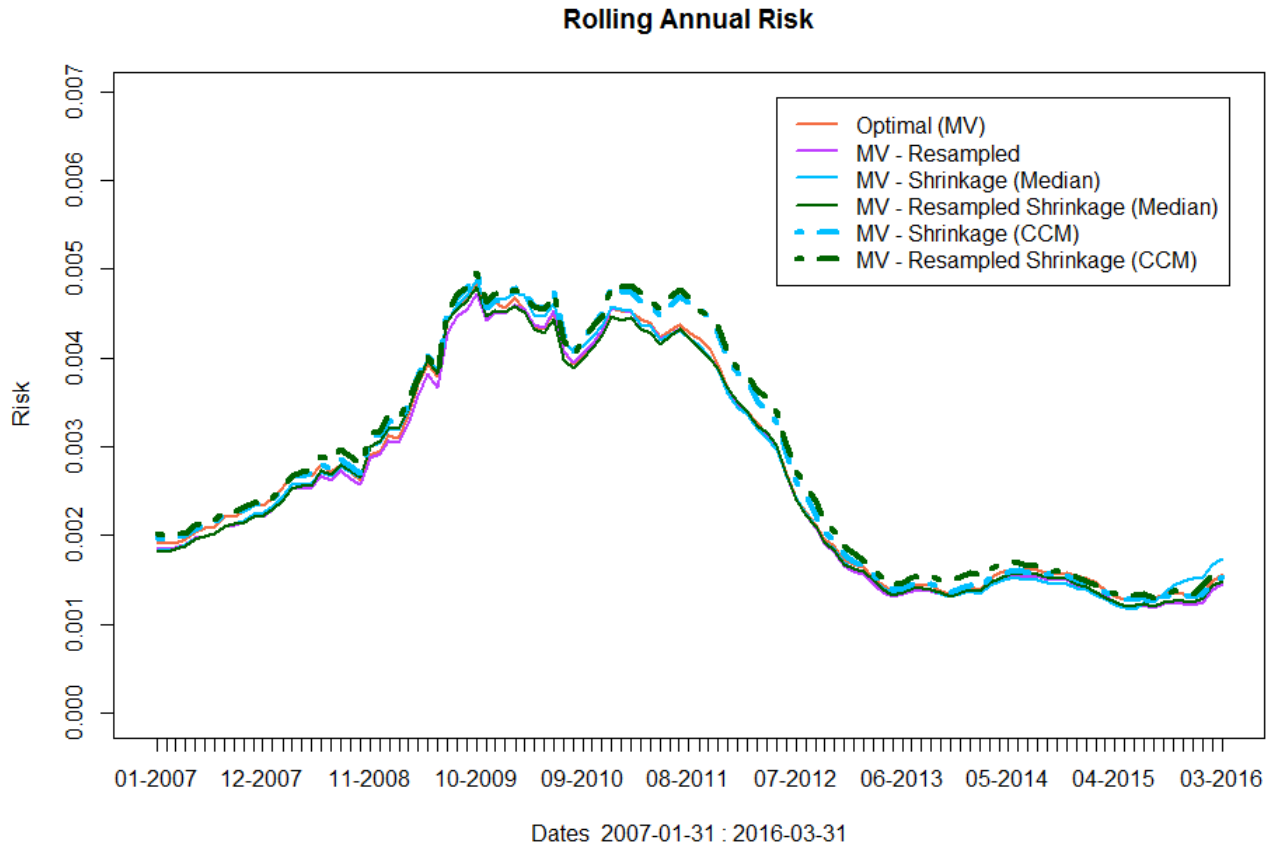


FIGURE 5.10: Annualised Risk under CCM
with $IC=0.05$

5.10 shows the ACM riskiness; in which again we see a higher riskiness of the ACM as opposed to both median-shrinkage portfolios and those without shrinkage. The difference is roughly 0.0002 higher under ACM at 0.0028 as opposed to 0.0027 under median. Resampling allows a marginal increase in riskiness of the portfolio under ACM shrinkage but a decrease in riskiness under median shrinkage.

TABLE 5.9: CCM vs Median Shrinkage Risk with $IC=0.05$

	Resampling	Average Risk	Risk at 3/2016
MV		0.00267856	0.00156463
MV	Yes	0.00261238	0.00145287
MV - Shrinkage Median		0.00266116	0.00173307
MV - Shrinkage Median	Yes	0.00262945	0.00149138
MV - Shrinkage CCM		0.00276957	0.00153515
MV - Shrinkage CCM	Yes	0.00282171	0.00162843

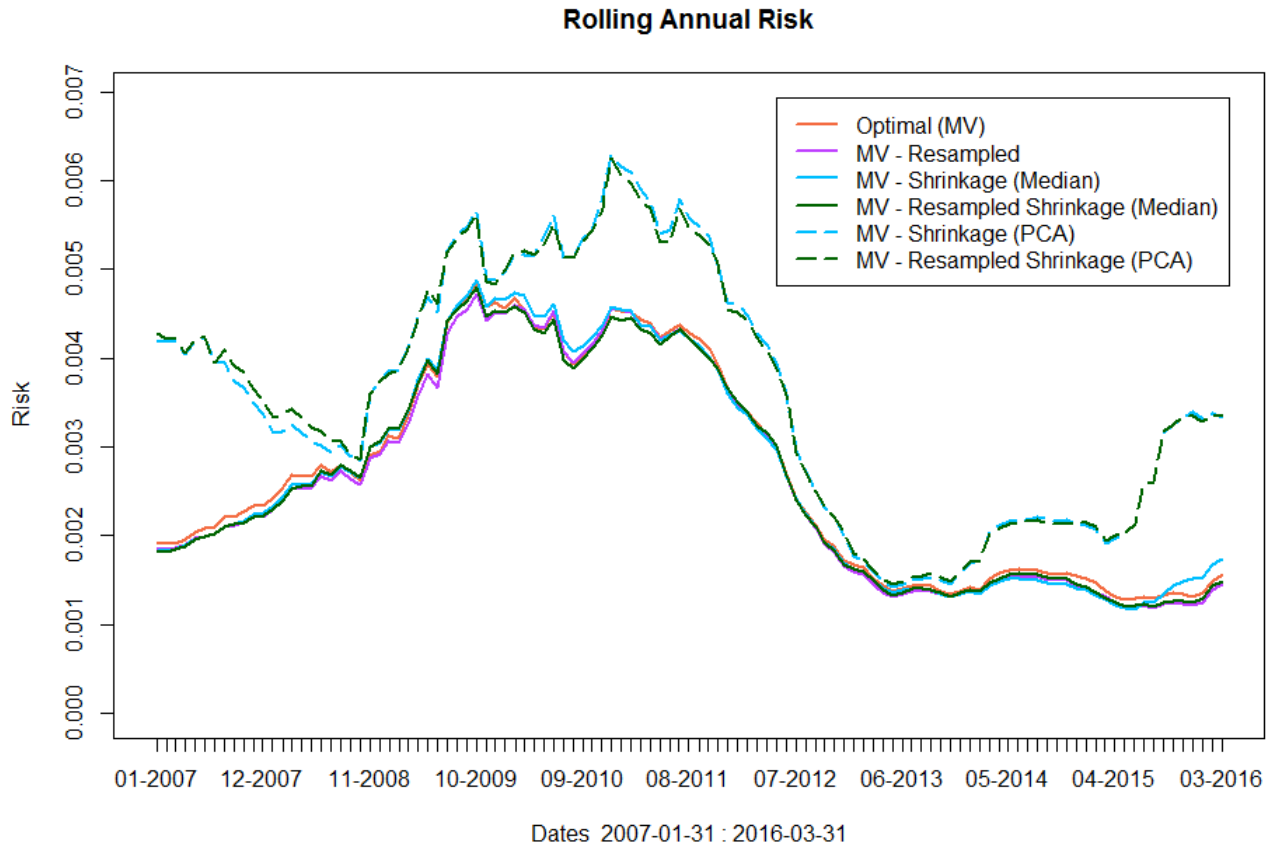


FIGURE 5.11: Annualised Risk under PCA with IC=0.05

5.11 again portrays the comparison wherein median-shrinkage is less risky than the alternative. In this case we see a large and consistently higher risk; with both the resampled and non-resampled being strictly larger than any alternatives at all times. On average, the PCA-shrinkage portfolio held a risk of 0.0036; which was almost 0.0010 higher than the average riskiness of the median shrinkage portfolio at 0.0027. Resampling makes little difference to the PCA-shrinkage portfolio's associated risk as it remains identical to the 4th d.p.

TABLE 5.10: PCA vs Median Shrinkage Risk with IC=0.05

	Resampling	Average Risk	Risk at 3/2016
MV		0.00267856	0.00156463
MV	Yes	0.00261238	0.00145287
MV - Shrinkage Median		0.00266116	0.00173307
MV - Shrinkage Median	Yes	0.00262945	0.00149138
MV - Shrinkage PCA		0.00357669	0.00333663
MV - Shrinkage PCA	Yes	0.00358092	0.00334132

Alternative risk metrics are used, beyond merely the historical risk level of the portfolio, under $w'\Sigma w$. We assess and report in 5.11 the observed s.d.² between the return stream; in which is largest for the PCA-shrinkage, demonstrating a large fluctuation in returns. The standard deviation for

²standard deviation

TABLE 5.11: Performance Analytics 2 with IC=0.05

	s.d. (σ)	Treynor	Return/Risk Ratio
MV	0.04268	0.16857	0.01315
MV - R	0.04087	0.16637	0.01266
MV - S (Median)	0.04093	0.16420	0.01261
MV - R/S (Median)	0.04077	0.16071	0.01228
MV - S (ACM)	0.04088	0.15621	0.01215
MV - R/S (ACM)	0.04137	0.15653	0.01233
MV - S (CCM)	0.03948	0.17993	0.01312
MV - R/S (CCM)	0.03977	0.18029	0.01324
MV - S (PCA)	0.05219	0.15575	0.01380
MV - R/S (PCA)	0.05288	0.15487	0.01383
EQ	0.03770	0.08845	0.00603
ALSI	0.04402	0.08399	0.00523

S Shrinkage
R Resampled
S/R Both applied

PCA sits at 0.052; and slightly larger at 0.053 for the resampled PCA portfolio. In contrast the portfolio holding the lowest variance is CCM at 0.0039; and 0.0040 for the resampled CCM portfolio.

In a similar manner we assess the Treynor measure, observing a large observation at 0.180 for the CCM resampled portfolio, and a lowest measure of only 0.0155 for both PCA non-resampled and resampled. This stands in stark contrast to the s.d. where the portfolio here holds the lowest risk from the respective metric. Lastly we assess the risk per unit of return. This metric will be largest for the portfolio closest to the optimal MV out-of-sample portfolio and under an IC of 1; the MV itself with no adjustments would return the highest ratio. In this case we observe a slightly larger return per unit risk from the PCA at 0.0138 and resampled PCA, identical to the 4th d.p. The lowest return per unit risk is seen from the median resampled portfolio, at 0.0121. This demonstrates an under-performance of the resampled median portfolio in the joint risk-return metric; thus even if this performance had not been comparably poor under either risk or return; jointly this portfolio is arguably one of the worst.

5.2.4 Number of Constituents under Shrinkage

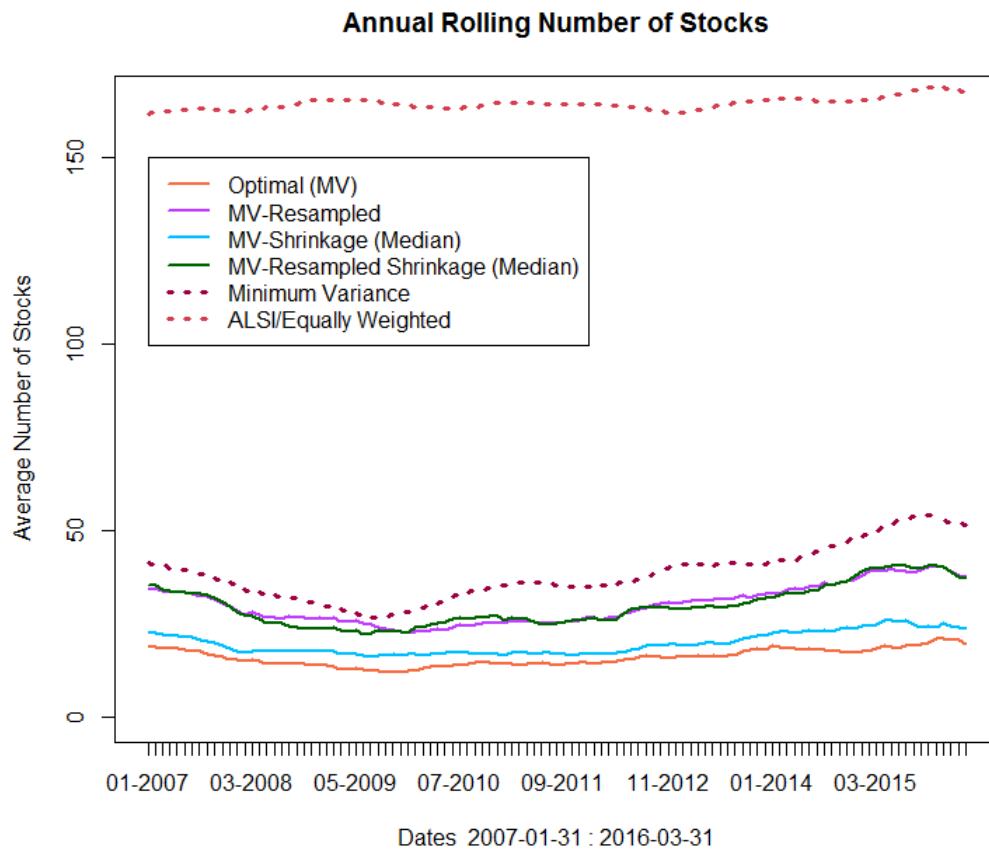


FIGURE 5.12: Number of Constituents Baseline with $IC=0.05$

In a comparison of the number of constituents, we observe that the ALSI/Equally-weighted portfolio contains every live stock on the JSE, which is consistently between 155-165 over the duration. We note the comparison to the baselines, as done prior with relation to the risk levels and cumulative returns: the median shrinkage contains a lower number of stocks than either baseline i.e. ALSI (obviously) and minimum-variance. The shrinkage techniques however allow for a higher number of stocks consistently than the MV portfolio without shrinkage. Offhand, this appears positive in terms of an element of diversification naturally being included by allowing shrinkage.

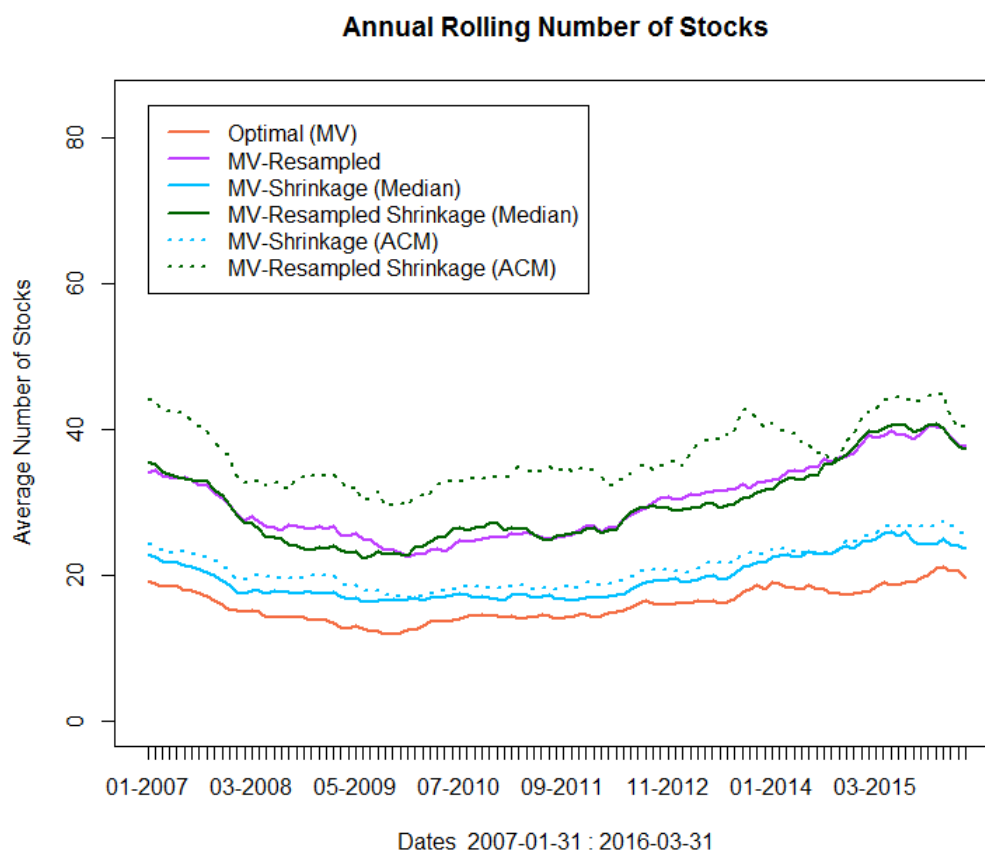


FIGURE 5.13: Number of Stocks under ACM with IC=0.05

In 5.13 we compare median shrinkage to ACM. Both clearly allow for the diversification element aforementioned; with the mean of the portfolio under ACM shrinkage at 21.18 stocks, and that under median at 19.69 stocks. This is a difference of under 6 stocks in both cases away from the MV portfolio at 15.98. A larger factor discrepancy arises when resampling is applied where we observe the three resampled portfolios containing far more stocks than the non-resampled portfolios. Largest number of stocks is seen consistently in the ACM-resampled shrinkage at 36.60 and ending at 40.59 over the final period. The resampling effect by its nature includes a larger number of stocks (as described in Chapter 3), and it appears further exacerbated under the ACM shrinkage effect.

TABLE 5.12: ACM vs Median Shrinkage Constituents with IC=0.05

	Resampling	Mean	Constituents 3/15-3/16
MV		15.9737	19.5833
MV	Yes	30.0848	37.6667
MV - Shrinkage Median		19.6884	23.6667
MV - Shrinkage Median	Yes	29.7508	37.3333
MV - Shrinkage ACM		21.1817	25.6667
MV - Shrinkage ACM	Yes	36.5916	40.5833

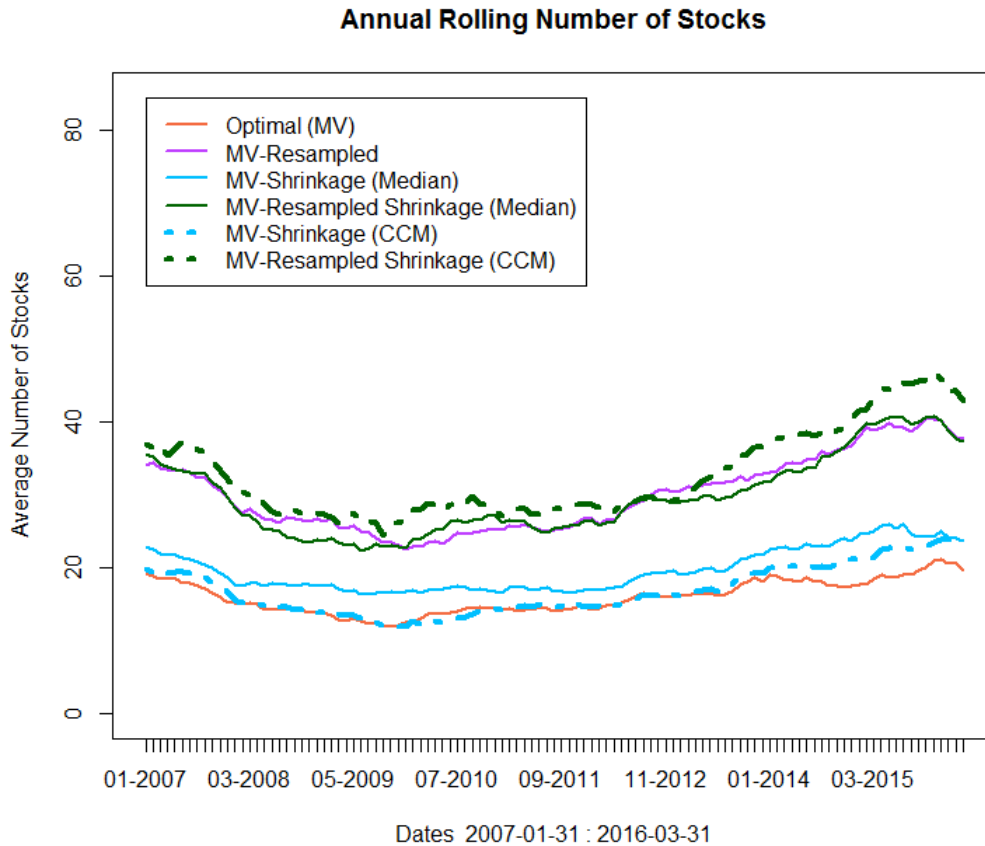


FIGURE 5.14: Number of Stocks under CCM with IC=0.05

5.14 paints a comparison of the CCM methodology versus median shrinkage. Similarly with the ACM shrinkage, we see resampling being the differentiating factor; and a mostly consistent higher number of stocks in the resampled CCM methodology versus all others under resampling. Note that in mid-2012 the resampled MV did contain a higher number of stocks however the average number remains at 32.80 (resampled CCM) over 2 stock units above the average of 30.08 for MV. The lowest number is rivalled between the CCM shrunken MV (16.88) and the regular MV portfolio (15.98); although the CCM shrunken portfolio does out-stock the non-shrunken portfolio from 2012-2016.

TABLE 5.13: CCM vs Median Shrinkage Constituents with IC=0.05

	Resampling	Mean	Constituents 3/15-3/16
MV		15.9737	19.5833
MV	Yes	30.0848	37.6667
MV - Shrinkage Median		19.6884	23.6667
MV - Shrinkage Median	Yes	29.7508	37.3333
MV - Shrinkage CCM		16.8844	23.25
MV - Shrinkage CCM	Yes	32.8003	42.8333

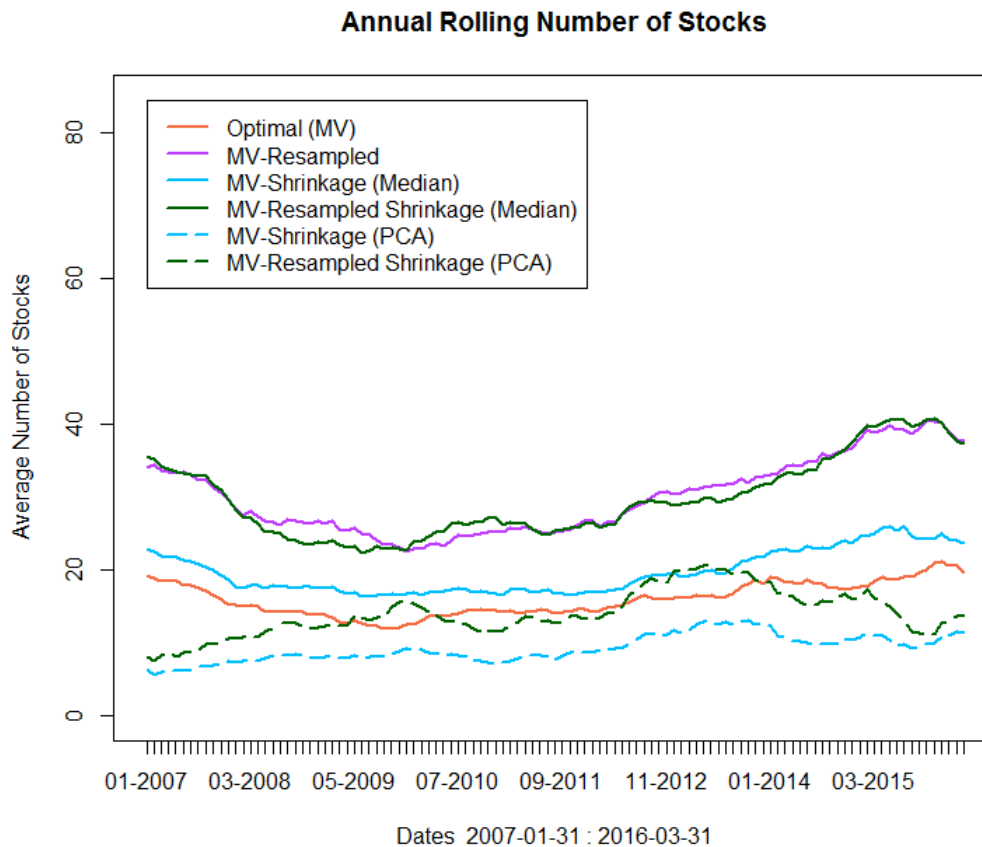


FIGURE 5.15: Number of Stocks under PCA with $IC=0.05$

5.15 Portrays our final stock-count graphic for the MV portfolios. A contrasting behaviour is seen in this graph, whereby PCA-shrinkage applied to the MV portfolio reduced the number of stocks invested in. Note how the PCA-shrunk MV portfolio contains an average of 9.20, over 6 stocks lower than the MV at 15.98. Under resampling, however, we do see that the PCA portfolio picks up additional stocks as expected, averaging 14.04. This is a fascinating and perhaps counter intuitive result; however as seen in analysis of the PCA risks we observe that natural patterns are extracted from the data. In this graphic we observe that the number of stocks contributing to the largest patterns are minimal, and indeed this indicates that perhaps PCA shrinkage is the most volatile technique. Of interest will be to observe whether these patterns are maintained under zero-forecasting information; as we apply shrinkage and resampling to the minimum variance portfolios. In the absence of foresight, the minimum variance portfolio is the obvious choice of investment (Clarke, *et al*, 2010).

5.2.5 Shrinkage onto Minimum Variance Portfolios

This section of the chapter looks exclusively at the minimum variance portfolio set. Note again that generated returns are not a component of either minimum variance construction (the only input is the covariance matrix) thus there are no confounding effects. Thus the comparisons are purely

TABLE 5.14: PCA vs Median Shrinkage Constituents with IC=0.05

	Resampling	Mean	Constituents 3/15-3/16
MV		15.9737	19.5833
MV	Yes	30.0848	37.6667
MV - Shrinkage Median		19.6884	23.6667
MV - Shrinkage Median	Yes	29.7508	37.3333
MV - Shrinkage PCA		9.2042	11.4167
MV - Shrinkage PCA	Yes	14.0435	13.6667

looking to assess the differences between risk; returns; and other performance metrics as phenomena on the shrinkage methods and resampling methods applied onto the minimum variance portfolios.

5.2.6 Cumulative Returns under Shrinkage

In 5.16 observe the baselines of the minimum-variance and equally-weighted portfolios; which are far outperformed (under the assumption of an IC of 0.05) by all mean-variance portfolios. In this section of the chapter we observe the following behaviour in comparisons of the various shrinkage techniques:

TABLE 5.15: Portfolio Rankings by Returns

Shrinkage-Type	Highest Return	Lowest Return
Median	Min Var	Min Var-Resampled Shrinkage
ACM	Min Var	Min Var-Resampled Shrinkage
CCM	Min Var-Resampled Shrinkage	Min Var-Resampled
PCA	Min Var-Resampled Shrinkage	Min Var-Resampled

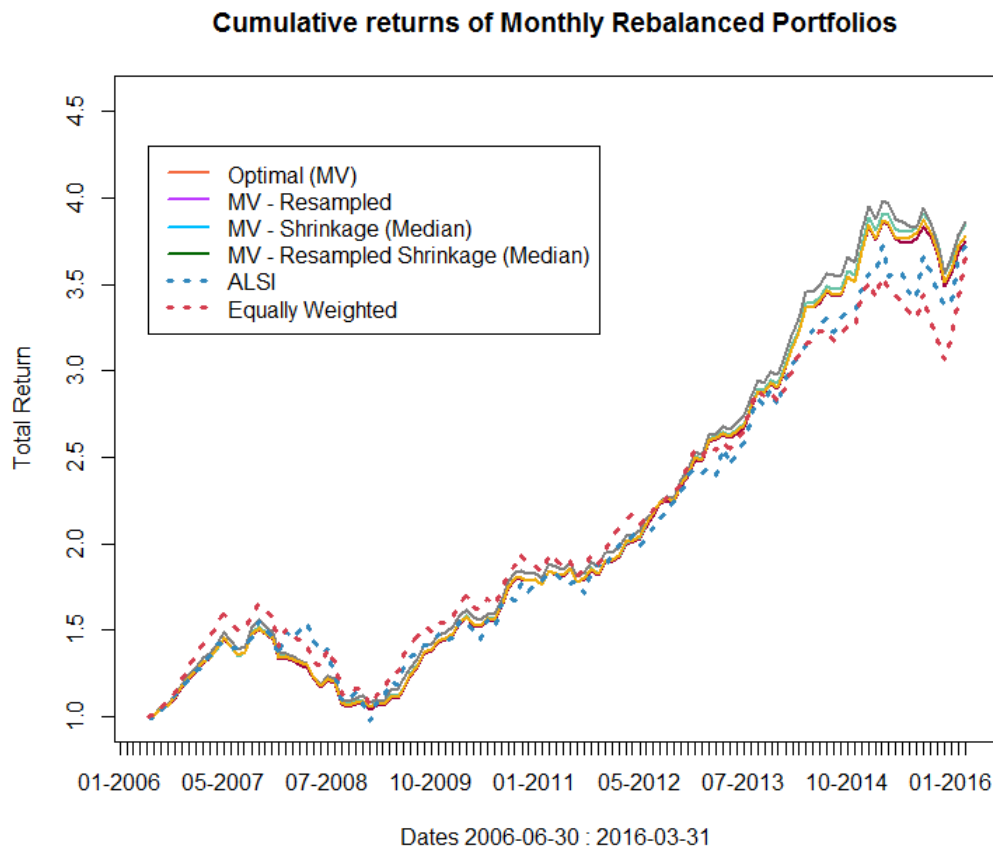


FIGURE 5.16: Returns Baseline with $IC=0.05$ for MV Portfolios

The remainder of this section will look exclusively at resampling and various shrinkage methods applied to the mean-variance portfolio construction. Unlike the MV portfolios which require an IC, the minimum variance portfolios do not (however, the $IC=0.05$ MV portfolios are displayed above for a baseline comparison). Due to the nature of comparison of identification of the optimal shrinkage method, reference to equally-weighted and the minimum variance portfolio performance will be excluded from the remainder of this section. Offhand, one can observe that there is little improvement through simple median-shrinkage however, there is a persevering hope that more analytic shrinkage methodologies will allow for a greater improvement.

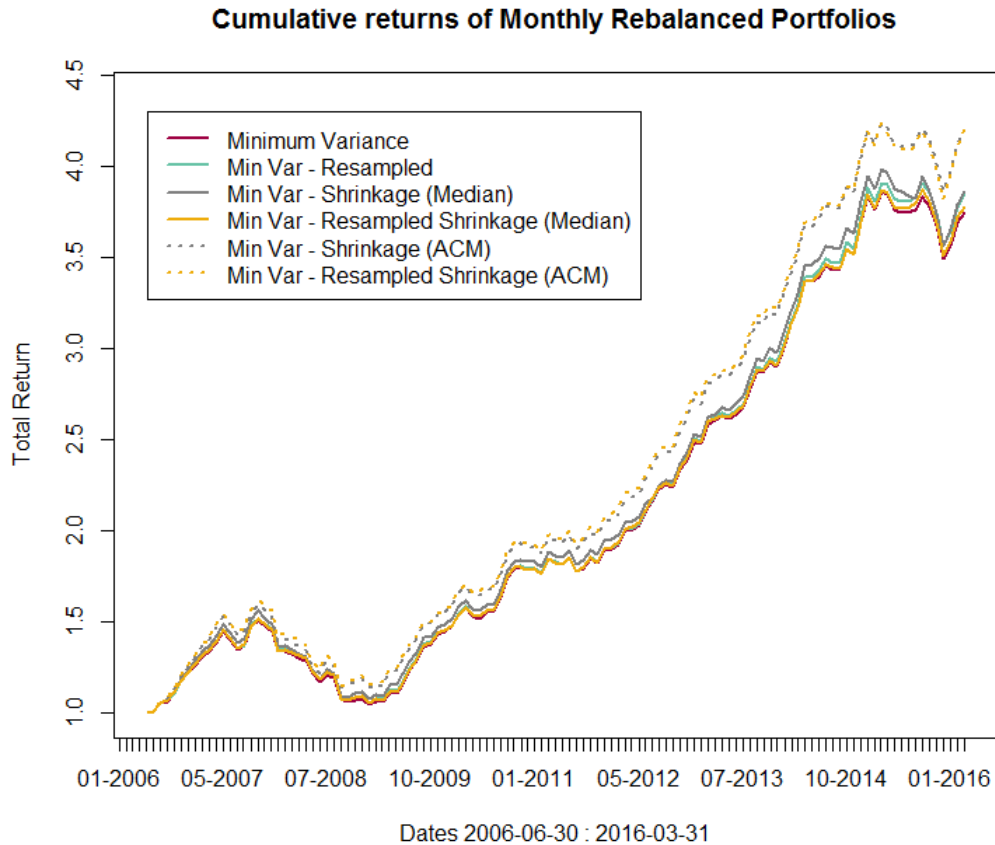


FIGURE 5.17: Returns under ACM

5.17 displays our first comparison between shrinkage methodology under zero forecast assumptions. We note immediately that there is a clear out-performance of everything else by the ACM shrinkage (at a cumulative of 4.20 under resampling and 4.19 without). This retains a full 33% of the initial investment above the next best performer at 3.86 of median shrinkage.

TABLE 5.16: ACM vs Median Shrinkage Returns

	Resampling	50 th Percentile	Total at 3/2016
Minimum Variance		1.79288	3.74607
Minimum Variance	Yes	1.80125	3.84405
Min Var - Shrinkage Median		1.83385	3.86391
Min Var - Shrinkage Median	Yes	1.80045	3.77927
Min Var - Shrinkage ACM		1.92122	4.19365
Min Var - Shrinkage ACM	Yes	1.93863	4.20055

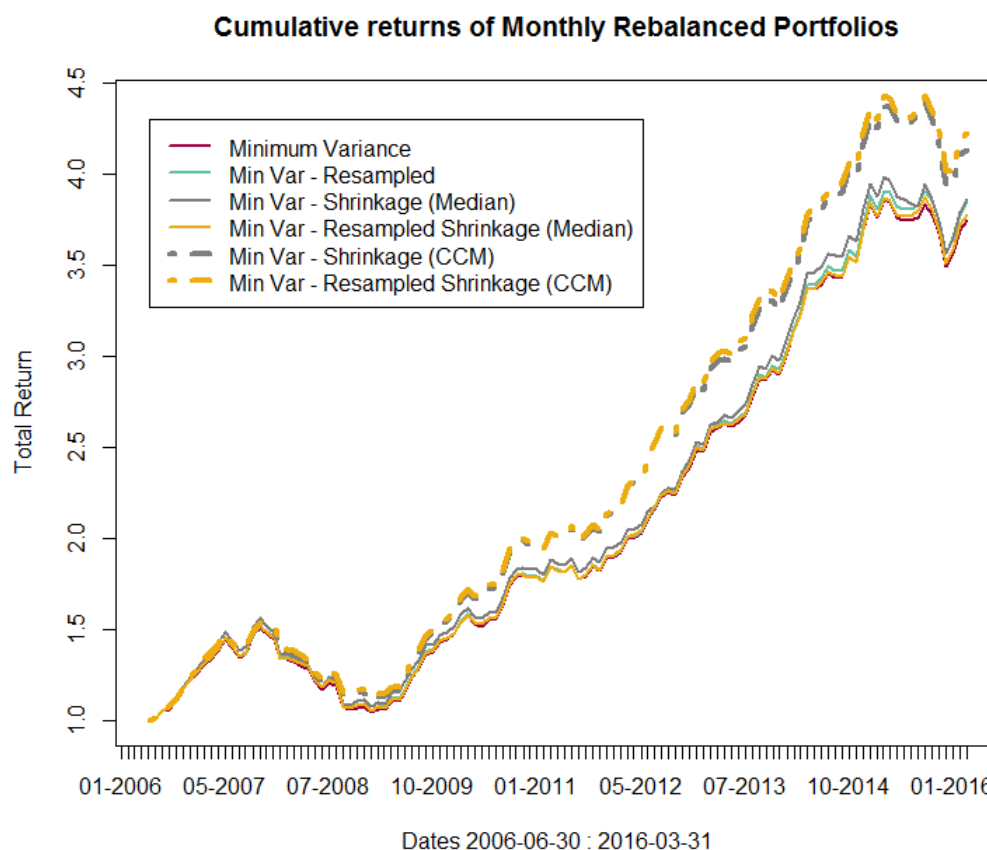


FIGURE 5.18: Returns under CCM

Next we observe the CCM performance, again with the standard use of median shrinkage and non-shrinkage as baselines. 5.18 allows us to observe the performance in which, as with the ACM above, we see the analytical shrinkage methods far outperforming the others. The cumulative return for the CCM shrinkage outperforms the minimum variance portfolio (at 3.75) by almost 40% of the initial investment, at 4.13. Additionally the resampled CCM allows an additional 9% above, at 4.23. The median shrinkage is comparable in ability to the absence of shrinkage.

TABLE 5.17: CCM vs Median Shrinkage Returns

	Resampling	50 th Percentile	Total at 3/2016
Minimum Variance		1.79288	3.74607
Minimum Variance	Yes	1.80125	3.84405
Min Var - Shrinkage Median		1.83385	3.86391
Min Var - Shrinkage Median	Yes	1.80045	3.77927
Min Var - Shrinkage CCM		1.98734	4.13359
Min Var - Shrinkage CCM	Yes	2.00143	4.22691

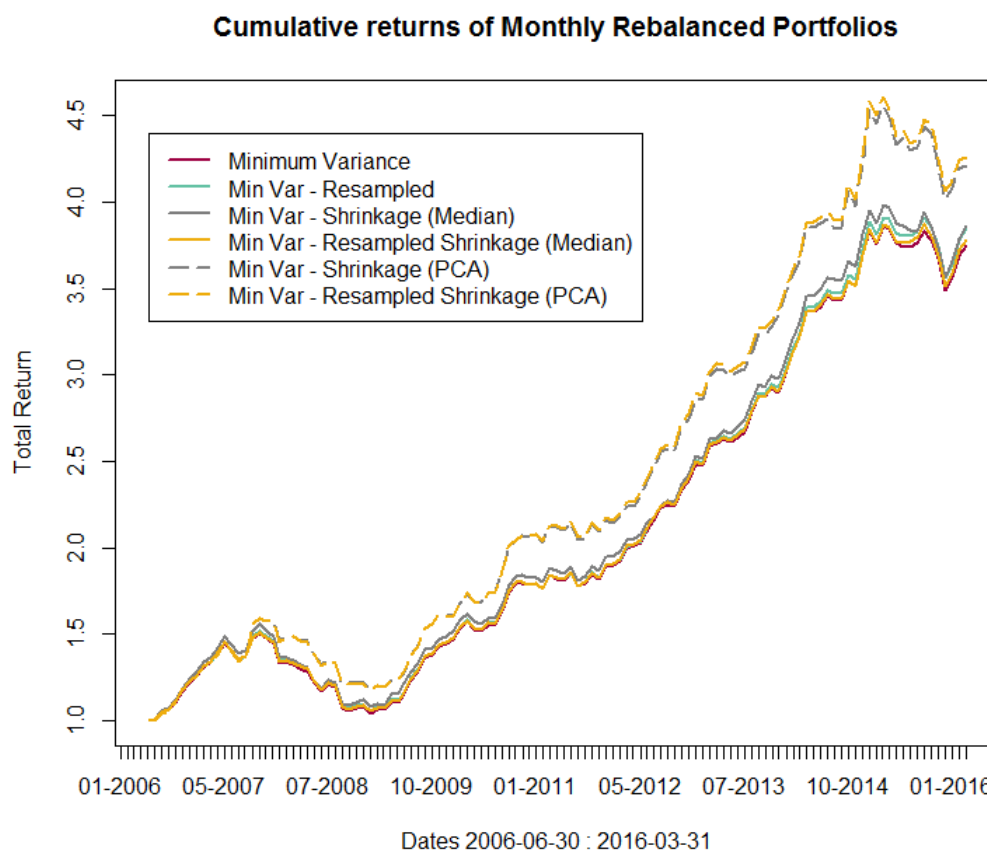


FIGURE 5.19: Returns under PCA

5.19 gives our final cumulative returns comparison: PCA vs median shrinkage; in which we see the clear out-performer again as the non-median approach. PCA finds a cumulative return of 4.21, and even larger at 4.26 with resampling. This stands at over a 50% initial-investment out-performance of the minimum variance without shrinkage (and almost 40% out-performance of median shrinkage). Note how the median shrinkage is the only technique to have dis-benefited through resampling, all other techniques gained a larger net return.

TABLE 5.18: PCA vs Median Shrinkage Returns

	Resampling	50 th Percentile	Total at 3/2016
Minimum Variance		1.79288	3.74607
Minimum Variance	Yes	1.80125	3.84405
Min Var - Shrinkage Median		1.83385	3.86391
Min Var - Shrinkage Median	Yes	1.80045	3.77927
Min Var - Shrinkage PCA		2.06365	4.21375
Min Var - Shrinkage PCA	Yes	2.06788	4.2613

5.19 displays several metrics of performance previously referenced- and similarly these are used here to assess the performances and similarities between the various shrinkage and resampling techniques. We find the largest realized beta from the minimum-variance-set of portfolios as the

TABLE 5.19: Performance Analytics 3

	Beta	R^2	Correlation	Tracking Error
Minimum Variance	0.39667	0.34426	0.59149	0.03567
Min Var - R	0.40006	0.35237	0.59826	0.03542
Min Var - S (Median)	0.42243	0.37427	0.61614	0.03481
Min Var - R/S (Median)	0.39909	0.34977	0.59609	0.03549
Min Var - S (ACM)	0.42447	0.37766	0.61885	0.03473
Min Var - R/S (ACM)	0.44483	0.39317	0.63116	0.03433
Min Var - S (CCM)	0.35290	0.29290	0.54675	0.03713
Min Var - R/S (CCM)	0.36086	0.30533	0.55792	0.03677
Min Var - S (PCA)	0.42545	0.32118	0.57183	0.03693
Min Var - R/S (PCA)	0.42725	0.32631	0.57626	0.03674
Equally Weighted	0.64770	0.56834	0.75633	0.02914
ALSI	1.00000	1.00000	1.00000	0.00000

S Shrinkage
R Resampled
S/R Both applied

ACM-shrinkage resampled at 0.65; with PCA (both resampled and non-resampled) at a close 0.63 in second. The lowest beta has been the CCM, both resampled and non lie at under 0.56. Considering R^2 ; a similar metric to beta we find the matching winner with the largest value at 0.39 for the ACM-shrinkage resampled. Lowest R^2 falls to CCM non-resampled at 0.29 and resampled at only 0.30. The correlation to the ALSI is observed as highest for the ACM-shrinkage resampled; and lowest for the CCM non-resampled at 0.55. Lastly in the table we note the largest tracking error is associated to the PCA both resampled and non-resampled at 0.037; identical to the third decimal place with resampled CCM. Lowest tracking error is associated to ACM resampled; at 0.034.

Outside of median-shrinkage; we consistently observe an identical directional change in each shrinkage method through the addition of resampling. When including resampling (as opposed to not) we find an increase in R^2 ; an increase in beta; an increase in correlation and a decrease in tracking error. We note that median-shrinkage is the differentiating factor from allowing this phenomenon to hold in all instances; and investigation as to the nature causing this may provide interesting insights.

5.2.7 Annualised Risk under Shrinkage

In 5.20 we provide the identical image providing the comparative between the mean-variance portfolio set and the minimum-variance portfolio [Note this is the same image as 5.8]. The minimum variance portfolios by definition holds a lower volatility than all other portfolios; and cannot be improved upon with any techniques- however we wish to describe how divergent the varying risks are away from the absolute minimum. In this section of the chapter we observe the following behaviour in comparisons of the various shrinkage techniques upon the minimum variance portfolio:

TABLE 5.20: Portfolio Rankings by Risk

Shrinkage-Type	Highest Risk	Lowest Risk
Median	Min Var Shrinkage	Min Var-Resampled
ACM	Min Var-Resampled Shrinkage	Min Var-Resampled
CCM	Min Var-Resampled Shrinkage	Min Var-Resampled
PCA	Min Var-Shrinkage	Min Var-Resampled

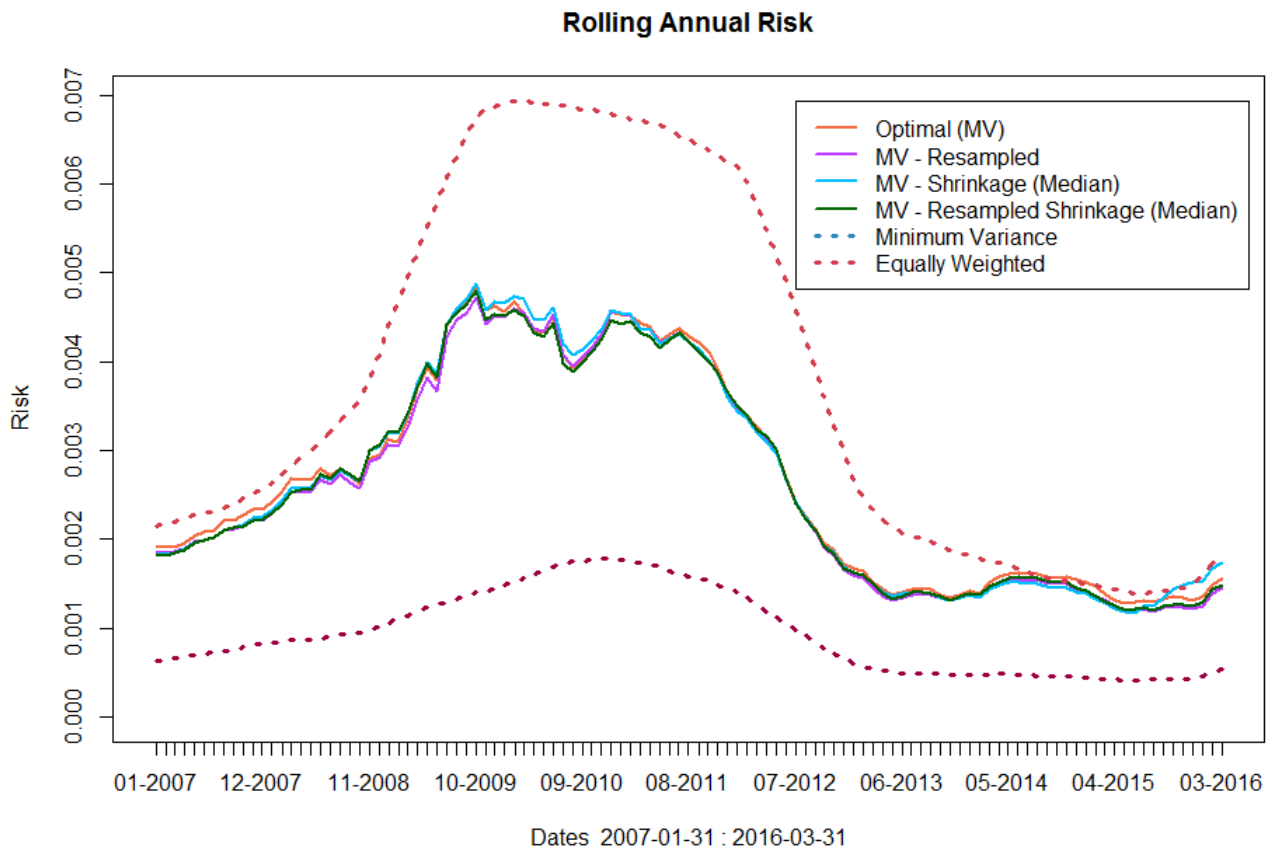


FIGURE 5.20: Annualised Risk Baseline

The remainder of this section will look exclusively upon resampling and various shrinkage methods upon the minimum variance portfolio. Due to the nature of comparison of identification of the optimal shrinkage method, reference to equally-weighted and the mean variance portfolios performance will be excluded from the remainder of this section.

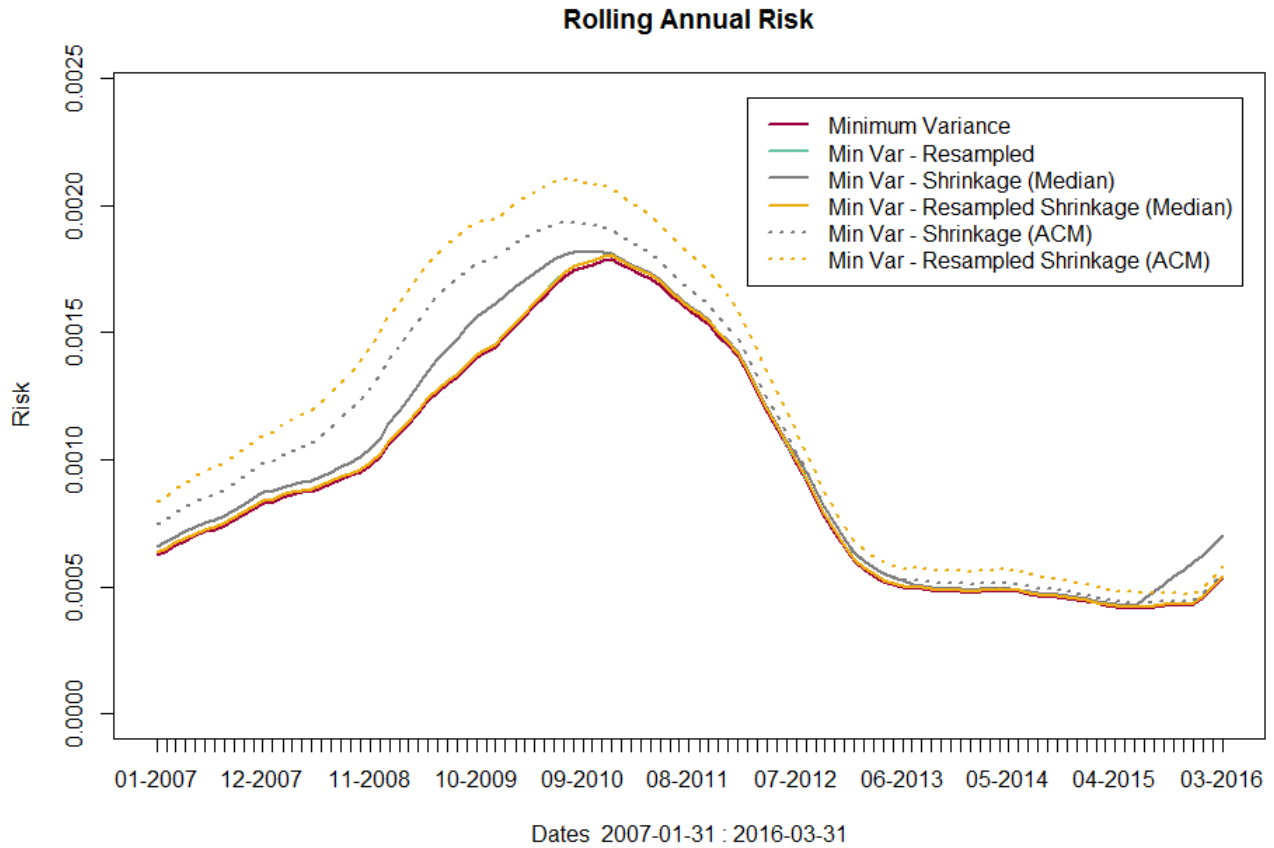


FIGURE 5.21: Risk under ACM

We assess in comparison to median shrinkage, as before. 5.21 distinguishes the risks of ACM and median shrinkage, and the application of resampling over them. Bear in mind the infeasibility of having a risk level lower than the minimum-variance portfolio; however we can assess how close the various shrinkage techniques get to minimum. We observe that the median shrinkage technique, especially under resampling, is exceptionally close to minimum variance. The median shrinkage with resampling has empirically been only 1.02%³ riskier than the minimum; at 0.000956 compared to the minimum at 0.000946. In comparison, ACM shrinkage has been on average 13.5% riskier at 0.00107; and 24.19% riskier at 0.001175.

TABLE 5.21: ACM vs Median Shrinkage Risk

	Resampling	Average Risk	Risk at 3/2016
Minimum Variance		0.000946228	0.000536703
Minimum Variance	Yes	0.000955895	0.000543652
Min Var - Shrinkage Median		0.000998637	0.000702242
Min Var - Shrinkage Median	Yes	0.000955912	0.000543829
Min Var - Shrinkage ACM		0.00107419	0.000550768
Min Var - Shrinkage ACM	Yes	0.00117515	0.000578637

$$^3 \frac{averisk_{portfolio} f_{interest}}{averisk_{minvar}} = \frac{0.000955912}{0.000946228} = 1.010234$$

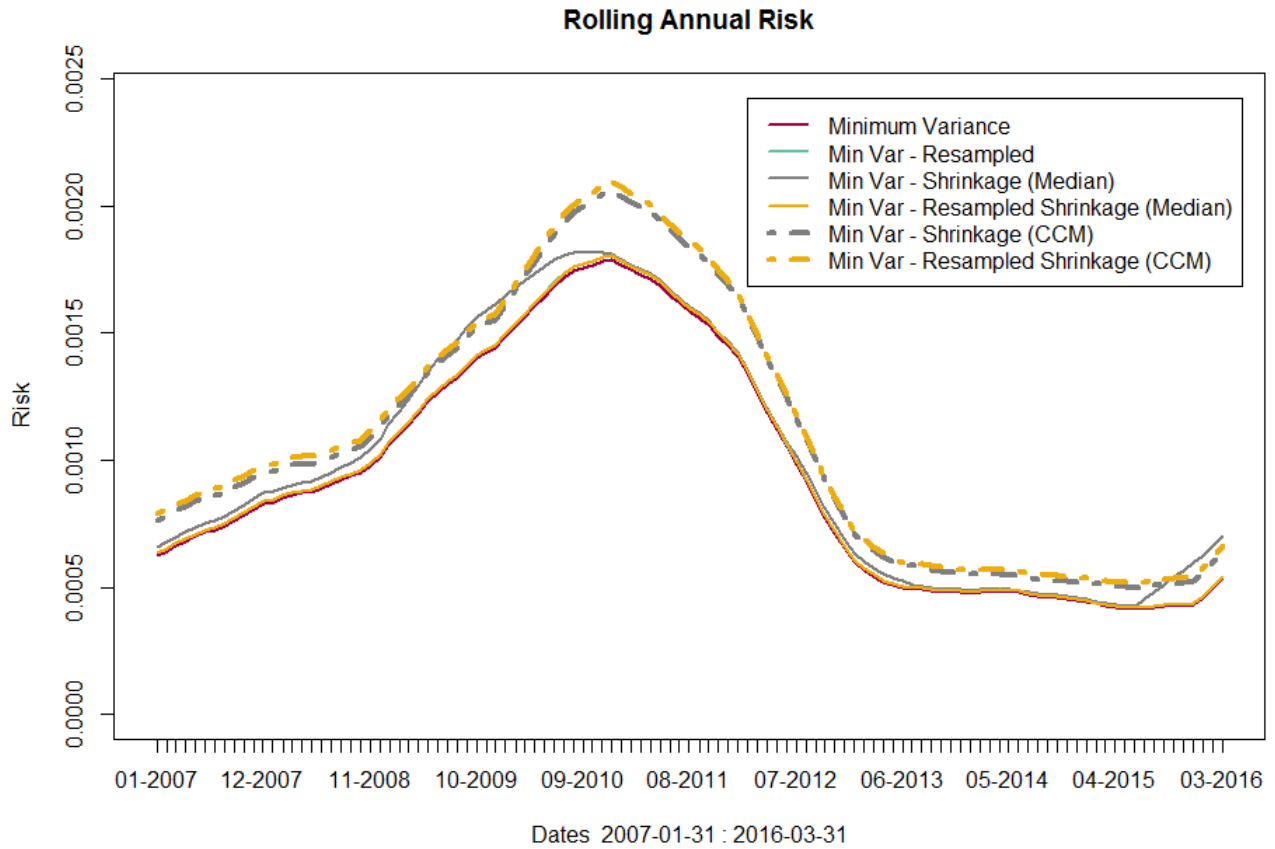


FIGURE 5.22: Annualised Risk under CCM

5.22 contrasts the risks of CCM shrinkage. We observe a similar set of behaviours from the CCM portfolio set as we did from the ACM. Both sufficiently outweigh median shrinkage in risk, with the CCM shrinkage with resampling empirically being 14.66%⁴ riskier than the minimum; at 0.001085 compared to the minimum at 0.000946. A similar, but slightly larger risk increase is seen from resampling, allowing a risk increase to 0.001106.

TABLE 5.22: CCM vs Median Shrinkage Risk

	Resampling	Average Risk	Risk at 3/2016
Minimum Variance		0.000946228	0.000536703
Minimum Variance	Yes	0.000955895	0.000543652
Min Var - Shrinkage Median		0.000998637	0.000702242
Min Var - Shrinkage Median	Yes	0.000955912	0.000543829
Min Var - Shrinkage CCM		0.0010849	0.000641795
Min Var - Shrinkage CCM	Yes	0.00110589	0.000662873

$$^4 \frac{ave_risk_portfolio_of_interest}{ave_risk_min_var} = \frac{0.0010849}{0.000946228} = 1.14655$$

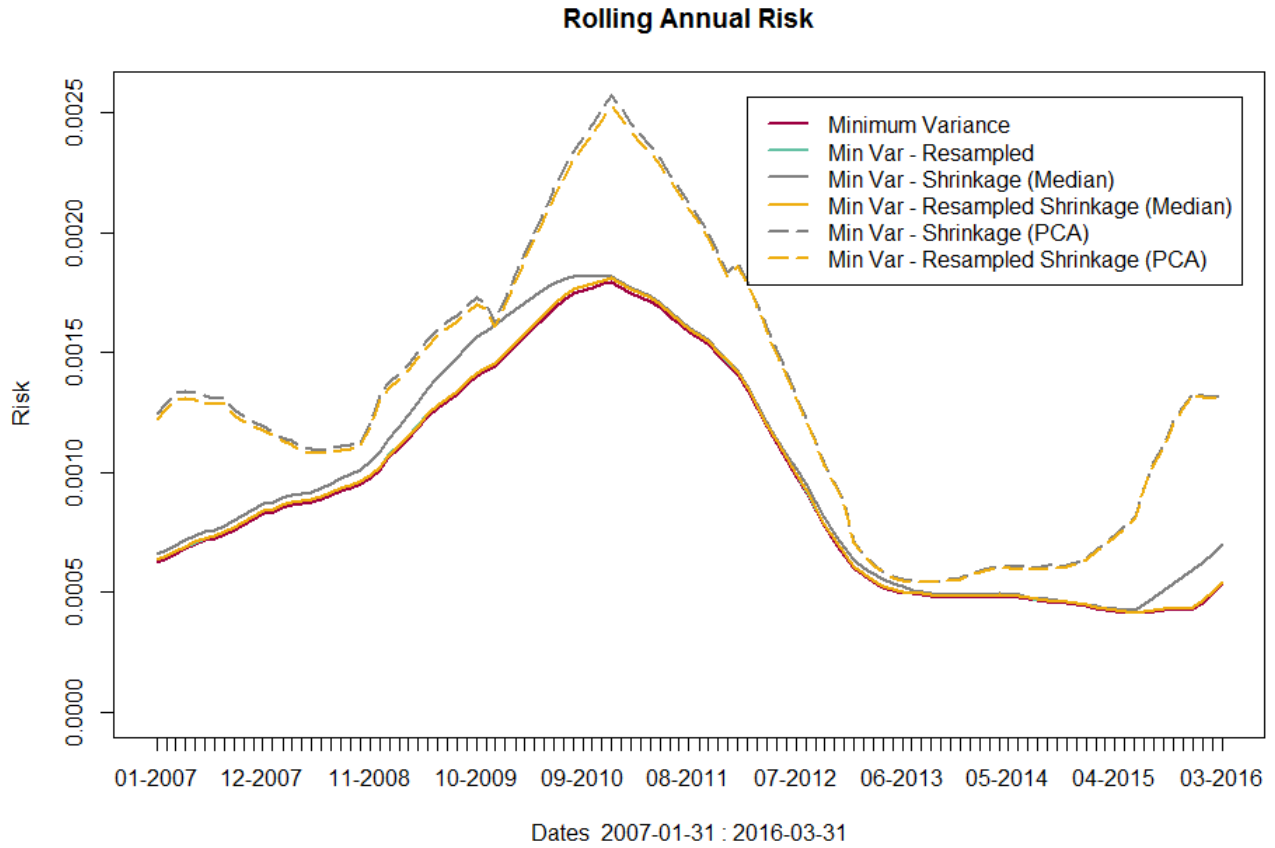


FIGURE 5.23: Annualised Risk under PCA

Lastly we view PCA risk for the minimum variance portfolios. 5.23 displays the effects: and we again observe a clear higher risk under the non-median shrinkage. PCA shrinkage with resampling stands empirically 39.54%⁵ riskier than the minimum; at 0.001320 compared to the minimum at 0.000946. A similar, but slightly lower risk is seen from resampling, allowing a risk decrease to 0.001303.

TABLE 5.23: CCM vs Median Shrinkage Risk

	Resampling	Average Risk	Risk at 3/2016
Minimum Variance		0.000946228	0.000536703
Minimum Variance	Yes	0.000955895	0.000543652
Min Var - Shrinkage Median		0.000998637	0.000702242
Min Var - Shrinkage Median	Yes	0.000955912	0.000543829
Min Var - Shrinkage PCA		0.00132041	0.00131412
Min Var - Shrinkage PCA	Yes	0.00130291	0.00130706

We view the table `tab:tableanalytics3x14a3a` comparing alternative risk metrics as we did for the MV portfolios. The empirical s.d. of the return vector is seen to be largest for the ALSI at 0.044, but of the shrinkage and resampling options, we see PCA obtaining the most volatility. PCA without resampling holds returns with s.d. 0.0328 and with resampling at 0.0326; with the only other 0.03's being ACM resampled shrinkage and median

⁵ $\frac{ave_risk_portfolio_f_interest}{ave_risk_min_var} = \frac{0.00132041}{0.000946228} = 1.39544$

TABLE 5.24: Performance Analytics 4

Portfolio	s.d. (σ)	Treynor	Return/Risk Ratio
Min Var	0.02952	0.12027	0.00626
Min Var - R	0.02944	0.12485	0.00648
Min Var - S (Median)	0.03018	0.12086	0.00652
Min Var - R/S (Median)	0.02947	0.12226	0.00633
Min Var - S (ACM)	0.03019	0.13588	0.00722
Min Var - R/S (ACM)	0.03102	0.13110	0.00724
Min Var - S (CCM)	0.02841	0.15183	0.00710
Min Var - R/S (CCM)	0.02847	0.15336	0.00729
Min Var - S (PCA)	0.03275	0.13414	0.00726
Min Var - R/S (PCA)	0.03264	0.13613	0.00736
Equally Weighted	0.03770	0.08845	0.00603
ALSI	0.04402	0.08399	0.00523

S Shrinkage
R Resampled
S/R Both applied

shrinkage. We see the lowest volatility of returns in the CCM shrinkage, at 0.0284; and 0.0285 without shrinkage; and with shrinkage respectively.

We look to the Treynor metric as an alternative measure of riskiness as before. We look to the resampled and shrinkage portfolios in a direct comparison to one another, observing a highest measure of 0.152 under the CCM shrinkage (and 0.153 under resampled CCM). The lowest score is seen between both the minimum variance (no shrinkage) and the median shrinkage, as with and without resampling, all four portfolios lie between 0.120 and 0.125. Similarly to the previous observations, we see a strong similarity between median shrinkage and no shrinkage. Recall the prior comparison of median shrinkage riskiness with a mere 1.02% increase in riskiness above minimum variance (seen in 5.21).

Lastly in the alternative risk metrics we assess return per unit of risk. We see similar results amongst the minimum variance portfolios, with and without techniques applied to them: all outperforming the index at a mere 0.0052. Of the portfolios the minimum variance does however get beaten by all portfolios with any kind of shrinkage and/ or resampling presence. The resampled min var is also beaten by any shrinkage-resampled portfolio, aside from median shrinkage. This leaves us to look at the top performers: ACM, CCM and PCA. The results are close between the three however the overall best performer is resampled PCA onto the minimum variance; allowing a retrun/risk ratio of 0.0074; seconded at 0.0073 by both the resampled CCM and from the non-resampled PCA.

5.2.8 Number of Constituents under Shrinkage

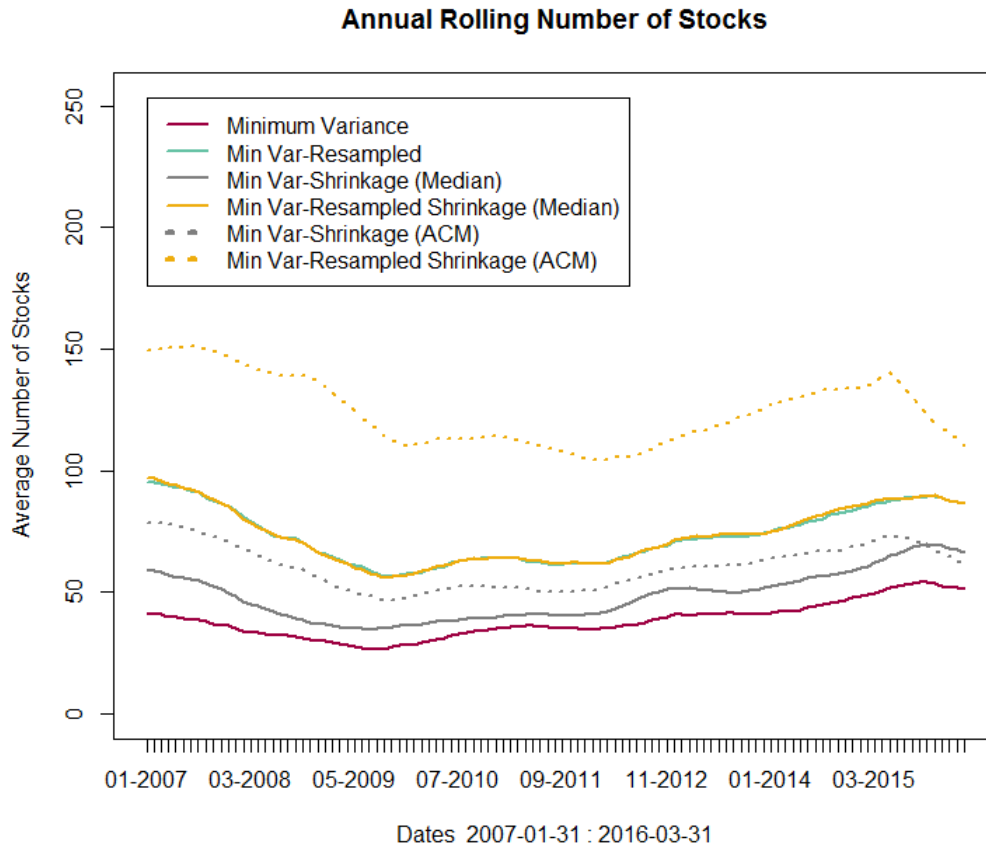


FIGURE 5.24: Number of Stocks under ACM

The final comparison looks to the number of constituents in the minimum variance portfolios under various adjustments. 5.24 outlines ACM vs. median shrinkage- and we observe a clear largest in ACM-resampled shrinkage at an average of 125.6, out-diversifying by far (second largest at 73.5). The lowest diversification comes at the minimum variance itself. A paradoxical concept is seen, in that the lowest variance comes from the least diversified of the minimum variance set.

TABLE 5.25: ACM vs Median Shrinkage Constituents

	Resampling	Mean	Constituents 3/15-3/16
Minimum Variance		38.3258	51.3333
Minimum Variance	Yes	73.268	86.3333
Min Var - Shrinkage Median		48.4174	66.5
Min Var - Shrinkage Median	Yes	73.539	86.5
Min Var - Shrinkage ACM		60.476	61.8333
Min Var - Shrinkage ACM	Yes	124.588	110.333

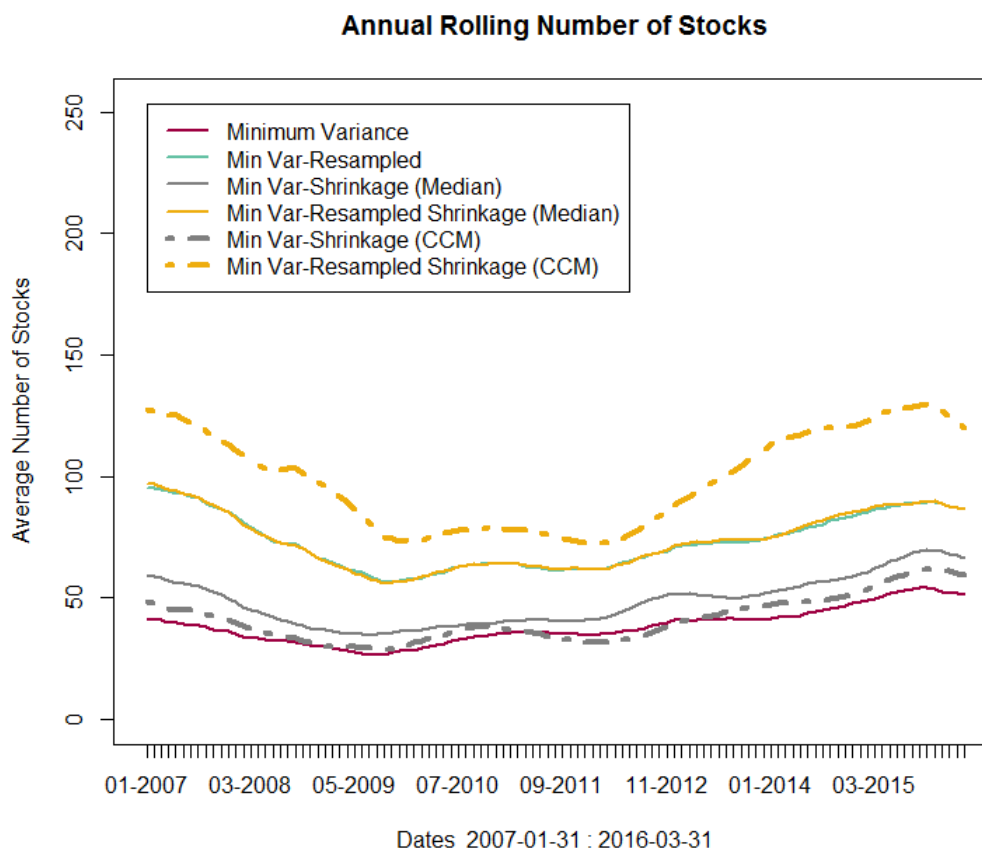


FIGURE 5.25: Number of Stocks under CCM

In the same vein we compare CCM to median shrinkage portfolios in number of stocks at 5.25. We observe a similar behaviour to the above ACM, in which there is a consistent out-stocking of all others. We see 99.00 stocks held on average in the resampled CCM portfolio.

TABLE 5.26: CCM vs Median Shrinkage Constituents

	Resampling	Mean	Constituents 3/15-3/16
Minimum Variance		38.3258	51.3333
Minimum Variance	Yes	73.268	86.3333
Min Var - Shrinkage Median		48.4174	66.5
Min Var - Shrinkage Median	Yes	73.539	86.5
Min Var - Shrinkage CCM		41.0998	59.4167
Min Var - Shrinkage CCM	Yes	98.9985	119.833

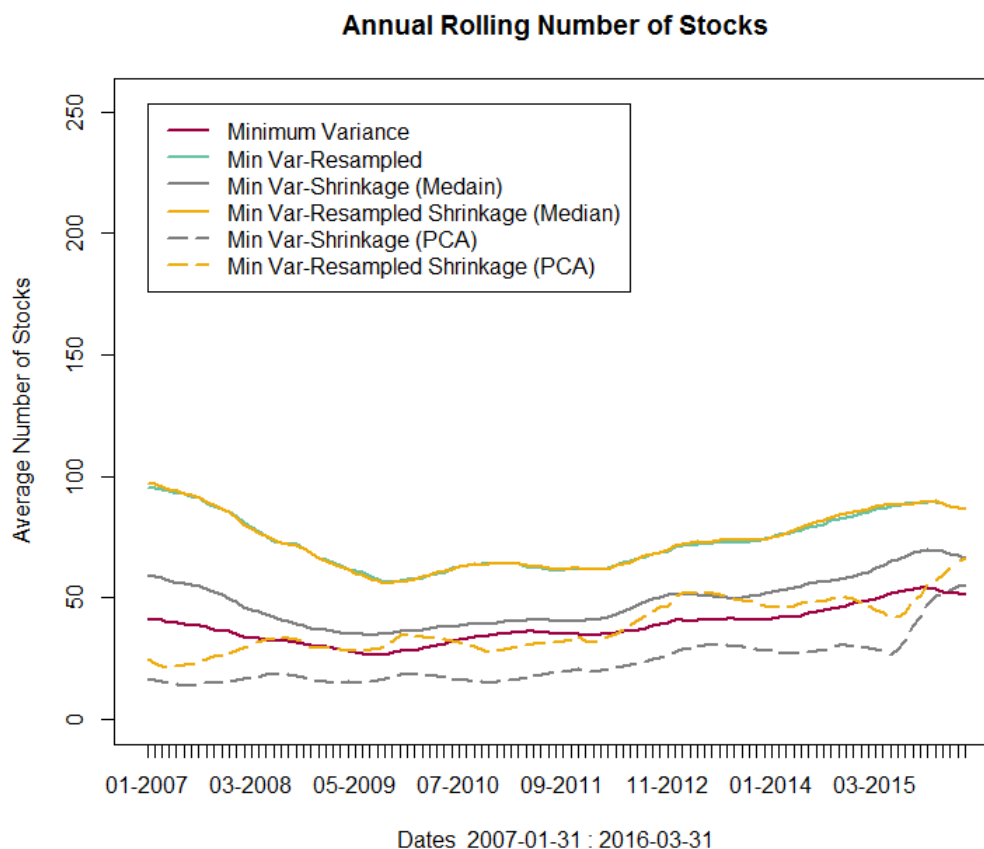


FIGURE 5.26: Number of Stocks under PCA

Finally we compare the PCA effect on the number of constituents in 5.26. Of interest this behaviour takes the opposite direction as the other shrinkage techniques did. Note again however that resampling, as always, increases the number- even in the same is not true of shrinkage. Recall 5.15 the same set of comparisons for the PCA shrinkage applied to the MV portfolio where this phenomenon is repeated from.

TABLE 5.27: PCA vs Median Shrinkage Constituents

	Resampling	Mean	Constituents 3/15-3/16
Minimum Variance		38.3258	51.3333
Minimum Variance	Yes	73.268	86.3333
Min Var - Shrinkage Median		48.4174	66.5
Min Var - Shrinkage Median	Yes	73.539	86.5
Min Var - Shrinkage PCA		23.2688	55.5833
Min Var - Shrinkage PCA	Yes	38.0143	66.25

Chapter 6

Sensistivity Analysis: Assessment of IC Level

This chapter serves to assess the independence of the IC effect and the shrinkage effect. Due to the choice of PCA as the preferred method in South Africa as an emerging market, and having obtained arguably the best performance thus far, the three interpretable levels of IC are used in conjunction with PCA-shrinkage. Recall the distinct IC levels: 0.1 (an excellent predictive ability); 0.05 (a good predictive ability) and 0.01 (a reasonable predictive ability); we wish to assess whether the admirable performance of PCA-shrinkage is sensitive to the IC level.

6.1 Return Comparative

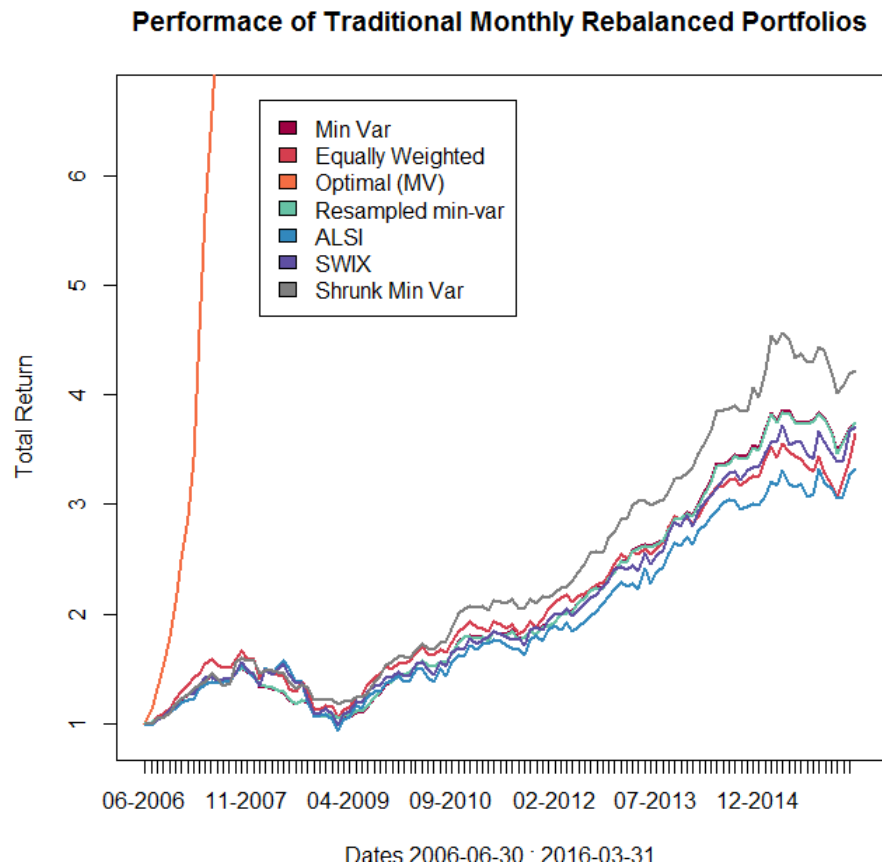


FIGURE 6.1: Resampling Risk under Reduced IC

6.1 demonstrates the performance of the traditional portfolios as a reference point. The author has coded multiple scenarios, under varying IC, and published the results which follow for this thesis. The IC reduction allows the MV portfolio (see the exponential orange line in the graph) to be brought into a comparative and more realistic setting alongside the other portfolios, and it is alongside the various levels of IC reduction that PCA shrinkage and resampling will be performed.

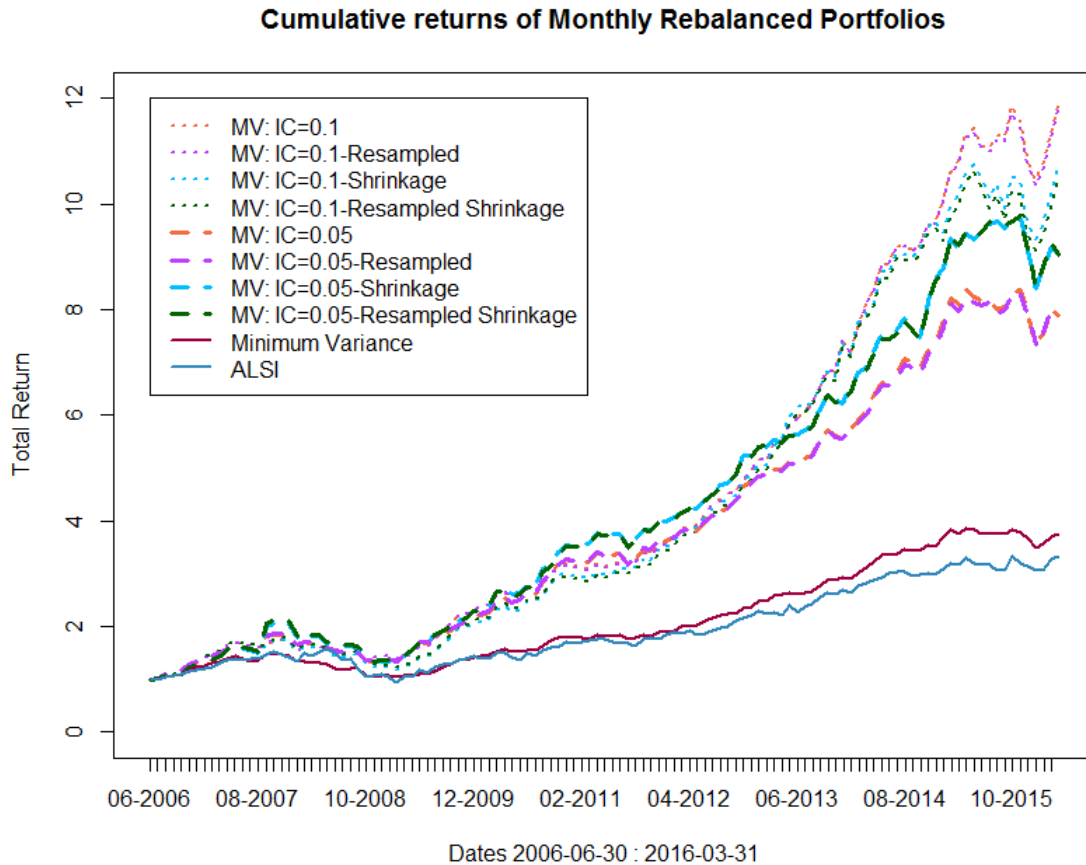


FIGURE 6.2: Resampling Risk under Reduced IC

6.2 provides a graphical comparative of the cumulative return obtained; by the IC=0.1-set of mean-variance portfolios and the IC=0.05-set of mean-variance portfolios. In both cases we obtain a relationship where the presence and absence of resampling provide a similar result. Shrinkage application resulted in a large increase in the cumulative return by March 2016 under an IC of 0.05, however this effect is not consistent across an IC of 0.1.

The table on the following page (6.1) details the difference between the portfolios with and without shrinkage. From the table we find that $MV_{0.1}$ holds a cumulative return of 11.87 and the shrunken $MV_{0.1}$ holds a cumulative return of 10.73, a decrease of 1.14. Compare this to $MV_{0.05}$ at 7.91 cumulative, and the shrunken $MV_{0.05}$ at 9.00 cumulative, an increase of 1.09, thus we observe an effect in the opposite direction through imposing shrinkage. Further, the resampling impact allows for even more of a discrepancy: in the comparison between $MV_{0.01}$ (11.87) and the resampled shrunken $MV_{0.01}$ (10.49), the decrease is 1.38. Compare this to the difference of $MV_{0.05}$ (7.91) to the resampled shrunken $MV_{0.05}$ (9.03), the difference between the two being 1.12 in the opposite direction.

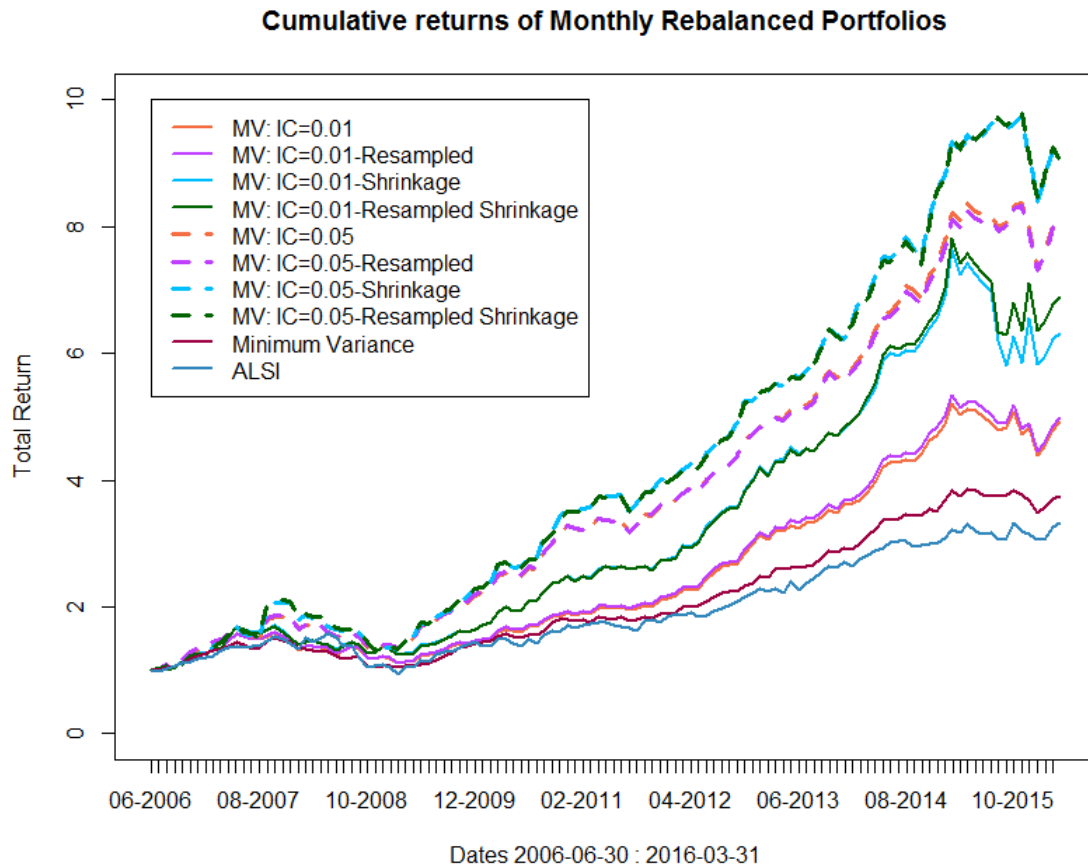


FIGURE 6.3: Resampling Risk under Reduced IC

We observe from 6.3 that the prior patterns continue, with shrinkage allowing an improvement in performance, as IC is lowered. In a comparison from 6.1 the difference between $MV_{0.01}$ at 4.93 and the shrunken $MV_{0.01}$ at 6.23, a difference of 1.29 in favour of the shrinkage. We similarly compare $MV_{0.01}$ at 4.93 to the shrunken-resampled MV at 6.90; a difference of 1.97 in favour of the resampled-shrinkage. This performance in fact, is a mere 2% per annum under the performance on an investor with an IC of 0.05. This imagery indicates that the independence assumption between the IC level and use of shrinkage is false, and results vary in accordance with both variables. The pattern naively suggests that the larger the IC level (provided it is non-negligible), the lower the relative benefit of allowing PCA shrinkage.

TABLE 6.1: Portfolio Comparative of IC by Return

	IC	50 th percentile	Cumulative	Annual Average
MV	0.1	3.16544	11.8684	0.27549
MV - R	0.1	3.14214	11.7873	0.27463
MV - S	0.1	3.01091	10.7331	0.26294
MV - R/S	0.1	2.95171	10.4914	0.26011
MV	0.05	3.28038	7.90723	0.22555
MV - R	0.05	3.28939	7.91974	0.22574
MV - S	0.05	3.63019	9.00193	0.24128
MV - R/S	0.05	3.61012	9.03557	0.24173
MV	0.01	1.97463	4.92804	0.16985
MV - R	0.01	2.00592	4.99554	0.17142
MV - S	0.01	2.59696	6.32014	0.19884
MV - R/S	0.01	2.58527	6.8988	0.20921
EQ		1.88423	3.64883	0.13578
ALSI		1.71924	3.32348	0.12539

S Shrinkage
R Resampled
S/R Both applied

In 6.1 we display the table from which the above quantitative values have been drawn. Comparison to the equally-weighted portfolio as well as the ALSI have been provided for ease of reference.

TABLE 6.2: Portfolio Analytics 5

	IC	Realized Beta	R^2	Correlation	Tracking Error
MV	0.1	0.54724	0.39764	0.63466	0.03706
MV - R	0.1	0.54753	0.39799	0.63493	0.03703
MV - S	0.1	0.52999	0.30216	0.55509	0.04184
MV - R/S	0.1	0.53647	0.30809	0.56036	0.04161
MV	0.05	0.58562	0.3894	0.62819	0.0375
MV - R	0.05	0.58745	0.39496	0.63256	0.03723
MV - S	0.05	0.61685	0.26437	0.52025	0.04848
MV - R/S	0.05	0.6205	0.26048	0.51653	0.04909
MV	0.01	0.49402	0.36619	0.6096	0.03613
MV - R	0.01	0.4965	0.36471	0.60839	0.03625
MV - S	0.01	0.52729	0.26229	0.51826	0.04394
MV - R/S	0.01	0.52876	0.27087	0.5264	0.0434
EQ		0.6477	0.56834	0.75633	0.02914
ALSI & Min Var		1	1	1	0
		0.39657	0.34414	0.59139	0.03567

S Shrinkage
R Resampled
S/R Both applied

6.2 displays the alternative performance metrics. We observe some interesting phenomena here: the IC=0.05 portfolio set under shrinkage and resampled shrinkage offer the higher realized beta; but the lowest R^2 and the lowest correlation to the index. The tracking error is similarly seen largest in the IC=0.05 portfolio under both shrinkage and resampled shrinkage. Note that these performances may differ under an alternative shrinkage metric aside from the median shrinkage. The peculiarity is that there appears to be a peak around IC=0.05; at least in the local maxima/minima sense. Further investigation into IC factor influence on performance is recommended to conclude whether this influence is consistent and, if so, what the optimal shrinkage type and optimal shrinkage factor are across varying IC.

6.2 Risk Comparative

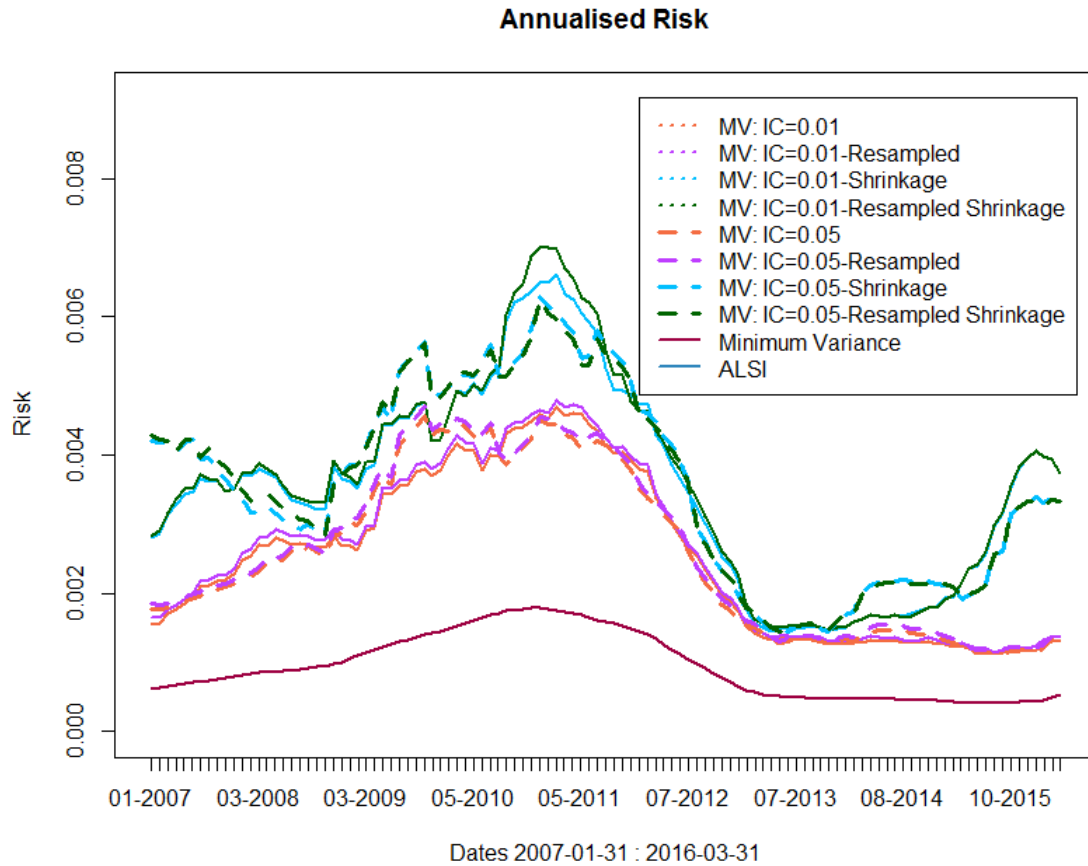


FIGURE 6.4: Resampling Risk under Reduced IC

In 6.4 we observe the risk comparison between an IC level of 0.1 and 0.05. We cannot discern an effect from resampling from this image alone, however we note that at an IC of 0.05 the resampling PCA risk is marginally higher at 0.0026, in comparison to the MV without resampling at 0.0025. A similar marginal difference is seen from the IC at 0.1. Through shrinkage however we observe a large increase in risk. At an IC level of 0.1; we have an increase in risk by 40% up to 0.0035. With an IC of 0.05, a very similar increase is seen at 41.1% up to 0.0036. In both case, a small additional increase is seen when resampling is used in addition to shrinkage.

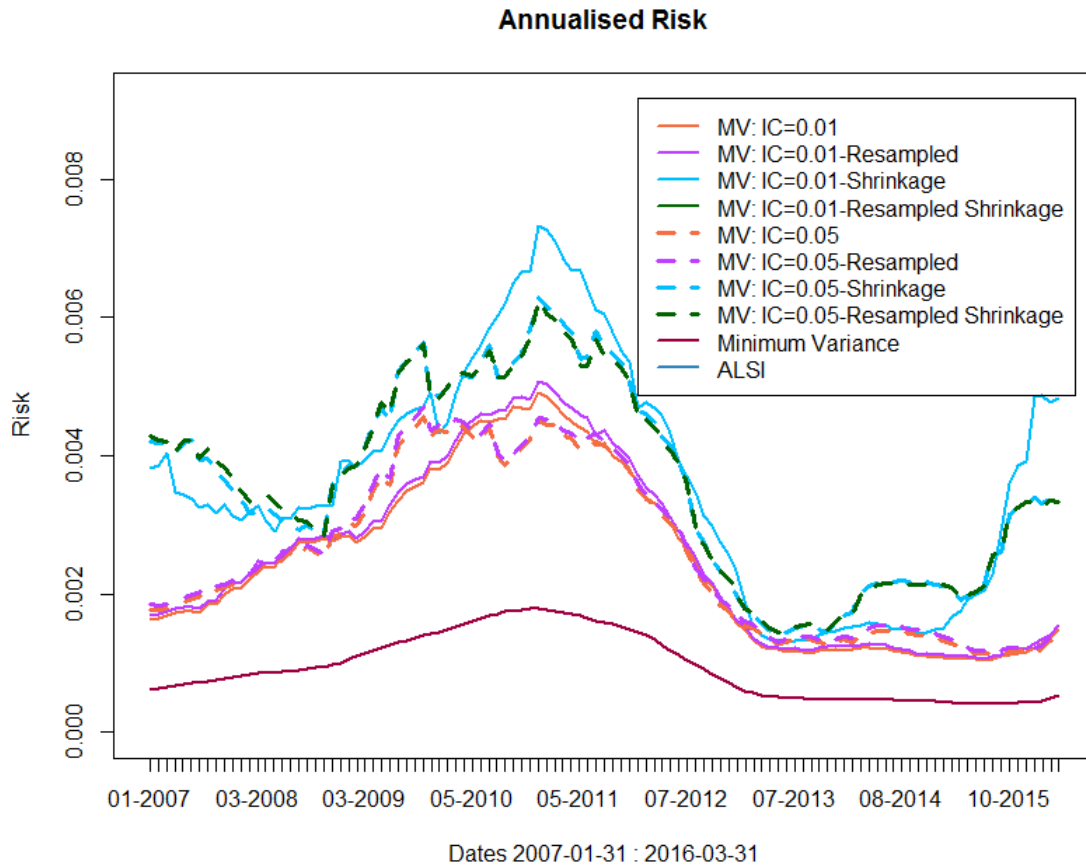


FIGURE 6.5: Resampling Risk under Reduced IC

In 6.5 we observe the risk comparison between an IC level of 0.01 and 0.05. The more erratic behaviour appears here, as opposed to with a higher IC level- note the risk level at 0.0048 during March 2016; over 45% as high as the PCA portfolio under IC 0.05. In general however, we are seeing little difference between the overall risk patterns under an IC of 0.01 as opposed to 0.05. Under IC=0.01 the PCA portfolio has a risk level of 47.1% (at 0.0036) above the non-shrunken portfolio (at 0.0025), as opposed to the aforementioned 41.1% increase when IC=0.05. Thus we observe no phenomena which are worthwhile to point the reader's attention to.

TABLE 6.3: Portfolio Comparative of IC by Risk

	IC	Average Risk	Risk at 3/16
MV	0.1	0.00251952	0.00132168
MV - R	0.1	0.00259153	0.00137144
MV - S	0.1	0.00354132	0.00374733
MV - R/S	0.1	0.00360666	0.00372852
MV	0.05	0.00253926	0.00139385
MV - R	0.05	0.00260423	0.0014406
MV - S	0.05	0.00357669	0.00333663
MV - R/S	0.05	0.00358092	0.00334132
MV	0.01	0.00246537	0.00149125
MV - R	0.01	0.00254346	0.00154059
MV - S	0.01	0.00362602	0.0048298
MV - R/S	0.01	0.0035946	0.00425028
EQ		0.00379678	0.00184975
ALSI			

S Shrinkage
R Resampled
S/R Both applied

We display the table 6.3 from which the above quantitative values have been drawn. We will now contrast the alternative metrics which indicate riskiness and return per unit of risk to determine if the consistency across IC levels holds in these metrics as it did in portfolio risk.

TABLE 6.4: Portfolio Analytics 6

	IC	s.d. (σ)	Treynor	Return/Risk Ratio
MV	0.1	0.03796	0.39024	0.01619
MV - R	0.1	0.03796	0.3884	0.01613
MV - S	0.1	0.04203	0.37826	0.01532
MV - R/S	0.1	0.04214	0.3682	0.01512
MV	0.05	0.04104	0.27582	0.01268
MV - R	0.05	0.04088	0.2753	0.01269
MV - S	0.05	0.05219	0.28841	0.0138
MV - R/S	0.05	0.05288	0.28748	0.01383
MV	0.01	0.03567	0.20969	0.00861
MV - R	0.01	0.03592	0.21193	0.00873
MV - S	0.01	0.04479	0.25361	0.01075
MV - R/S	0.01	0.04421	0.27333	0.0115
EQ		0.0377	0.10528	0.00603
ALSI &		0.04402	0.05741	0.00523
Min Var		0.02952	0.17964	0.00625

S Shrinkage
R Resampled
S/R Both applied

In 6.4 we observe the comparative of risks, for various IC levels. Note a trending relationship of local minima/ maxima from the return metrics coming into play here as well in some cases but not in others. In a like-for-like portfolio comparative (i.e. resampling vs resampling and shrinkage vs. shrinkage and resampled shrinkage vs. resampled shrinkage) we look to the differences in performance and attempt to look for consistent dynamics. We find the higher volatility of returns (s.d.) form an IC level of 0.05. Otherwise though, we see consistency in the declines of riskiness (Treynor measure) as the IC level is decreased; and we also see a lower return per unit of risk as IC level is reduced.

6.3 Overall Assessment of Possible IC Confounding

This dissertation has made use of a cut-off decision of the IC-value to 5% for the majority of the findings, which has been an arbitrarily selected figure. Although it has not been problematic in the research sphere, and is recommended in international literature (Munro & Bradfield 2016; Grinold & Khan, 1999); the assessment carried out within this thesis has indicated the presence of an interaction effect between the IC level and the effects of shrinkage. The need for further establishment of the credibility of the creation of a correlated set of returns should be considered strongly. The inconsistencies arising between the various IC levels are not discussed in depth within Grinold and Khan's paper, however this is an aspect which should requires further work for the purpose of robustness. The understanding of

the applicability to the South African market of these theoretical concepts may need further investigation, as this may be behaviour unique to South Africa or it may be a phenomenon which is common amongst IC reduction at various levels.

Chapter 7

Conclusion

7.1 Summary of Shrinkage Performances

This is the summary chapter, in which crucial findings are relayed. In this chapter we look to the effect of shrinkage, ergo we summarise the pertinent contents of the prior tables below. Each row in the table below corresponds to a shrinkage method applied either with or without resampling (indicated in column 2). This chapter displays a summarised set of results which are discussed within the concluding remarks.

TABLE 7.1: MV Shrinkage Returns Narrative

MV Portfolios	Resampling	50 th percentile	Cumulative	Annual Average
None		3.42207	8.35124	0.232151
None	Yes	3.23689	7.88684	0.225236
Median		3.12813	7.84105	0.224535
Median	Yes	3.14799	7.54975	0.219983
ACM		3.06319	7.43824	0.218199
ACM	Yes	3.12594	7.59807	0.220749
CCM		3.32796	8.32054	0.231705
CCM	Yes	3.41884	8.44076	0.233444
PCA		3.54494	9.00193	0.241278
PCA	Yes	3.51519	9.03557	0.241733

We found a strong argument for the superiority of PCA-shrinkage, assuming an IC of 0.05. The observation, as seen in 7.1 is a clear out-performance of the other shrinkage-types applied onto the portfolios. The effect of resampling is similarly seen to allow a consistent improvement for the MV portfolios, for all types of shrinkage aside from the median shrinkage.

For the subsequent analysis across the minimum variance portfolio set we observed (summarised in 7.2) a overall best performer of the PCA-shrinkage. This consistently allowed a strong performance, higher than any other shrinkage. One truly fantastic observation is that on average there has been an outperformance of the minimum variance portfolio by over 1% per annum. This is a return achievable, under the assumption that the PCA-shrinkage is able to capture the relationships in the shares to the same extent that happened over this period. In fact, if we compare the resampled

TABLE 7.2: Min Var Shrinkage Returns Narrative

Minimum Variance	Resampling	50 th percentile	Cumulative	Annual Average
None		1.79288	3.74607	0.138720
None	Yes	1.80125	3.84405	0.141616
Median		1.83385	3.86391	0.142195
Median	Yes	1.80045	3.77927	0.139709
ACM		1.92122	4.19365	0.151432
ACM	Yes	1.93863	4.20055	0.151618
CCM		1.98734	4.13359	0.149800
CCM	Yes	2.00143	4.22691	0.152327
PCA		2.06365	4.21375	0.151974
PCA	Yes	2.06788	4.2613	0.153246

PCA-shrinkage (15.32% p.a.) to the minimum variance as is, at 13.87%, we observe a full 1.35%. This is an achievement made under no assumption of forecasting ability, but merely applications onto historical covariance patterns.

We similarly observe the ACM and CCM portfolio to be performing phenomenally well relative to the minimum variance, and the median portfolio performed on-par. There is much scope for new discovery and further robustification into this work. We should note currently that under the selection of the best model *ex – post facto* may be causing us to fall into the trap of a type I error. The results appear significant, however this may merely be outlying observations. The law of large numbers suggests that our observations should be observed individually, and then aggregated—this is a possibly extension of the work. To either look at aggregations of the shrinkage methods, applied out of sample. And alternatively to assess the individual shares out-performance of the index and to analyse using non-parametric statistics whether the results are indeed a significant out-performance or whether statistical rigour labels this as the aforementioned fluke (type I error). These concerns are discussed in depth by Bailey *et al.* (2014) who warns that spurious relationships are common occurrences.

TABLE 7.3: MV Shrinkage Risks Narrative

MV Portfolios	Resampling	Average Risk	Risk at 3/2016
None		0.00267856	0.00156463
None	Yes	0.00261238	0.00145287
Median		0.00266116	0.00173307
Median	Yes	0.00262945	0.00149138
ACM		0.0028297	0.00142125
ACM	Yes	0.00307599	0.00158292
CCM		0.00276957	0.00153515
CCM	Yes	0.00282171	0.00162843
PCA		0.00357669	0.00333663
PCA	Yes	0.00358092	0.00334132

We observed a general increase in risk imposed through the use of resampling. Shrinkage as well is seen to increase risk, especially PCA shrinkage. Bearing in mind the increase in return described above, this trade off is perhaps one to be expected. 7.3 Indicates also that the median-shrinkage marginally decreases in risk through the use of resampling procedures; and is also seen in general as the least risky shrinkage procedure.

TABLE 7.4: Min Var Shrinkage Risks Narrative

Minimum Variance	Resampling	Average Risk	Risk at 3/2016
None		0.000946228	0.000536703
None	Yes	0.000955895	0.000543652
Median		0.000998637	0.000702242
Median	Yes	0.000955912	0.000543829
ACM		0.00107419	0.000550768
ACM	Yes	0.00117515	0.000578637
CCM		0.0010849	0.000641795
CCM	Yes	0.00110589	0.000662873
PCA		0.00132041	0.00131412
PCA	Yes	0.00130291	0.00130706

We observed a consistently higher risk imposed through the use of resampling on top of the shrinkage, as seen in 7.4. This effect of an increased risk is consistent with both minimum variance and mean variance portfolios.

TABLE 7.5: MV Shrinkage Constituents Narrative

MV Portfolios	Resampling	Average	Constituents 3/15-3/16
None		15.9737	19.5833
None	Yes	30.0848	37.6667
Median		19.6884	23.6667
Median	Yes	29.7508	37.3333
ACM		21.1817	25.6667
ACM	Yes	36.5916	40.5833
CCM		16.8844	23.25
CCM	Yes	32.8003	42.8333
PCA		9.2042	11.4167
PCA	Yes	14.0435	13.6667

In reference to both the MV stock levels in 7.5 we see lower average numbers than in the minimum variance portfolio set in 7.6. This is intuitive due to the assumptions about performances involved in the construction of the MV portfolio weightings. We further observed a large increase in the number of constituents within the minimum-variance portfolios through resampling (by approximately double) and a relatively small increase in number of constituents through allowing for shrinkage. The shrinkage in general is thus seen to target shares currently in the portfolios more so than bring new ones in. In contrast, resampling is seen to operate through diversification, as opposed to targeting.

TABLE 7.6: Min Var Shrinkage Constituents Narrative

Minimum Variance	Resampling	Average	Constituents 3/15-3/16
None		38.3258	51.3333
None	Yes	73.268	86.3333
Median		48.4174	66.5
Median	Yes	73.539	86.5
ACM		60.476	61.8333
ACM	Yes	124.588	110.333
CCM		41.0998	59.4167
CCM	Yes	98.9985	119.833
PCA		23.2688	55.5833
PCA	Yes	38.0143	66.25

If we look to the minimum-variance portfolios: 10 additional stocks are introduced in the median portfolio; and 20 new stocks in the ACM. This effect is dwarfed by the number introduced through resampling: In the case of ACM without resampling at 60.48 but with resampling at 124.59.

This effect suggests that the minimum-variance constituents are more substitutable for one another than those which constitute the MV portfolio.

7.2 Findings

This dissertation has looked to the intersection and interactions between shrinkage; resampling and IC factors. We have used shrinkage and resampling as techniques through which we intended to improve portfolio performance, in both risk and return. We based the performance measures on the out-of-period results, and we enhanced the robustness of the research by assessing the performance of the techniques under various investor skill profiles. This was done through an adjustment of the information coefficient. The results with an IC level of 0.05 have shown a significant improvement in performance through using PCA shrinkage as well as resampling. Although we were faced with an increase in riskiness of the MV portfolio through shrinkage, the out-of-sample measures of return; as well as return per unit of risk justify the use of resampling. The performance-enhancing techniques have been applied over both the minimum-variance portfolio as well as the mean-variance portfolio. Whilst the results have been seen to vary in accordance with the uses of the various practises, we note that in the South African market as an emerging market adheres to non-traditional patterns. However, we find that there are reliable methods none-the-less which allow for improvements of the returns provided by traditional portfolios through the implementation of the simple procedures of resampling and shrinkage.

The performance-enhancing techniques have been applied over both the minimum variance portfolio as well as the mean variance portfolio. Whilst the results have been seen to vary in accordance with the uses of the various practises, we note that in the South African market as an emerging market adheres to non-traditional patterns. However, we find that there are reliable methods none-the-less which allow for improvements of the returns provided by traditional portfolios through the implementations of the simple procedures of resampling and shrinkage.

Although there is clearly dependence between IC-levels and shrinkage-effect; there is power in shrinkage which allows for an increase in performance. The IC-level effect (beyond a certain threshold) and the presence of shrinkage have a negative interaction effect on the returns; i.e. the lower IC plus shrinkage has a higher return than changing the IC to be higher or removing the shrinkage effect. Resampling has also been shown to be able to increase the shrinkage effect, thus to increase the return, although not large an increase as the shrinkage itself has. Note, however, that an observed increase in return by 0.01% as seen in 6.1; provides a significant difference in monthly compounding. This can be seen by the difference between the $MV_{0.01}$ -shrunken portfolio (cumulative return 6.32) vs. the same portfolio with resampling implemented (cumulative return 6.90)- resulting in a difference of over 50% of the original investment amount (from 9.75 years prior).

In closing notes, we have observed the ability of resampling and shrinkage to improve both risk management as well as return acquisition. In particular, they appear powerful in conjunction on the minimum variance portfolio, and although this work is not conclusive of the general case, these results have been successful in achieving the dissertation's goals. The application of resampling on top of shrinkage managed to improve performance. Short of the (unrealistic) best manner to acquire high return (perfectly accurate forecasting), we have explained and tested techniques which have demonstrated being able to provide an empirically better-performing portfolio. These techniques can also be described as able to help to ensure that accurate return forecasts are not squandered in application. The resampled-shrinkage procedure can be understood through this dissertation and the procedure has been developed such that it can be trivially applied to any data set with minimal industry-expertise required.

7.3 Shortcomings and Further Research

7.3.1 IC level

The selection of $IC=0.05$ was arbitrary, however, it proved useful. It stands necessary that there is further research into the applicability into the South African market. The market is known to be more volatile (as an emerging economy market); and a higher variability may cause a lower ability for the same investor to reliably leverage a reasonable forecasting ability (i.e. 5% correlation to the market). To reliably compare the results between each technique, the same set of correlated returns was used in each strategy of shrinkage; resampling and resampled shrinkage. The intention of this was to allow the same market behaviour to occur under each strategy. This should be rerun using various correlated return sets (each correlated at 0.05) in case one of the strategies was unrealistically benefited or disbenefited by virtue of the specific market conditions. The process of repetition and aggregation is proposed for further research.

7.3.2 IC Dependence Modelling

We have seen that there exists a lack of consistency of resampling and especially of shrinkage, based on the IC level. Investigation into the IC factor as influential on the performance of shrinkage is a possible extension which may provide interesting results. IC influence on performance is recommended in order to conclude what the optimal shrinkage type and optimal shrinkage factor are across varying IC.

7.3.3 Transaction Costs

The immediate shortfall of this research in practical implementation is the assumption of zero-transaction costs. To allow this research to be applicable in the market, one needs to account for the following aspects:

1. Direct rebalancing costs: With every rebalancing of the portfolio, there is a cost to buy and sell shares which has been ignored
2. Willing buyer, willing seller: There has been an assumption of both,

availability of shares limitlessly at the current price; and also availability of buyers such that shares can be sold limitlessly at their current price

7.3.4 Shrinkage Techniques: SIM and Blends

We have assessed four separate shrinkage techniques (as well as none at all). One suggestion which was not reached in this paper is to increase the scope of the techniques through two options. Primarily, to look into SIM shrinkage and whether a form of this is applicable, or whether it can be developed to be able to deal with such a matrix of such large dimensions. Secondly: we can combine effects and performance through simple aggregations. The blended models have an ability to combine and outperform the originals frequently in variance reduction and sometimes even in return (Ledoit & Wolf, 2003). This is a highly recommended extension of this paper, and in the author's opinion, one such blended model should comprise of the following combination: PCA; CCM and the unadjusted covariance SC.

7.3.5 Shrinkage Techniques: DC

The use of blended models has been seen by Munro & Bradfield (2016) to outperform the other methods and, although it is beyond the scope of this thesis, this is an area for further research which has potential for further performance improvement. An additional point not covered in the paper is the applicability of the DC versus the SC matrix. In a large dimensional space, the DC matrix has the potential to allow for behaviour which differs from the SC matrix due to the much smaller number of parameters to be estimated. When in aggregation with other methods which have non-zero off-diagonal entries there is no dimension reduction, but there is a natural shrinkage applied. It is further worth investigation the difference between the SC and the DM matrices when using high frequency data from many variables.

7.3.6 Pseudo Relationships

Bailey warns against spurious high backtested performance (2014). Further research is required, as these results are not a robust proof of the methods being successful in general, but only successful in the environment of the JSE over the past decade.

Bibliography

Bailey, D., Borwein, J., López de Prado, M. and Zhu, Q. (2014). Pseudo-Mathematics and Financial Charlatanism: The Effects of Backtest Overfitting on Out-of-Sample Performance. *Notices of the American Mathematical Society*, 61(5), p.458.

Beaven, A., Gebbie, T. (2001) Active portfolio optimisation by re-sampled expected utility function maximising sets. https://www.academia.edu/2543-349/Active_Portfolio_Optimization_by_resampled_expected_utility_function_maximising_sets_2001_expired

Bengtsson, C. and Holst, J. (2002). On portfolio selection: Improved covariance matrix estimation for Swedish asset returns. *In 31st Meeting, Euro Working Group on Financial Modeling*. Hermes Center, Nov 2002. available as http://www.hermes.ucy.ac.cy/conferences/ewgfm/papers/bengtsson_holst_20021004.pdf.

Best, M. and Grauer, R. (1991). On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results. *Review of Financial Studies*, 4(2), pp.315-342.

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business*, 45(3), p.444.

Chan, L., Karceski, J. and Lakonishok, J. (1999). On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *Review of Financial Studies*, 12(5), pp.937-974.

Chellathurai, T. and Draviam, T. (2008). Markowitz principles for multi-period portfolio selection problems with moments of any order. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 464(2092), pp.827-854.

Clarke, R., de Silva, H. and Thorley, S. (2010). Minimum Variance Portfolio Composition. *SSRN Electronic Journal*.

DeMiguel, V., Garlappi, L., Nogales, F. and Uppal, R. (2009). A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms. *Management Science*, 55(5), pp.798-812.

Dempster, M. and Thompson, G. (2002) Dynamic portfolio replication using stochastic programming. *Op Cit*, 100-128

Disatnik, D. and Benninga, S. (2007). Shrinking the Covariance Matrix. *The Journal of Portfolio Management*, 33(4), pp.55-63.

Fama, E. and French, K. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), p.427.

Fernholz, R. and Shay, B. (1982). Stochastic Portfolio Theory and Stock Market Equilibrium. *The Journal of Finance*, 37(2), p.615.

Frazzini, A. and Pedersen, L. (2013). Betting Against Beta. *SSRN Electronic Journal*.

Frost, P. and Savarino, J. (1986). An Empirical Bayes Approach to Efficient Portfolio Selection. *The Journal of Financial and Quantitative Analysis*, 21(3), p.293.

Goetzmann, W., Grinold, R. and Kahn, R. (1996). Active Portfolio Management. *The Journal of Finance*, 51(4), p.1559.

Green, R. and Hollifield, B. (1992). When Will Mean-Variance Efficient Portfolios be Well Diversified?. *The Journal of Finance*, 47(5), p.1785.

Grinold, R. and Kahn, R. (1996). Active Portfolio Management. *The Journal of Finance*, 51(4), p.1559.

Hastie, T., Tibshirani, R. and Friedman, J. (2016). *The elements of statistical learning*. 1st ed. New York: Springer.

Jagannathan, R. and Ma, T. (2003). Risk Reduction in Large Portfolios: A Role for Portfolio Weight Constraints. *SSRN Electronic Journal*.

Jagannathan, R. and Ma, T. (2003). Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *The Journal of Finance*, 58(4), pp.1651-1683.

Jobson, J. and Korkie, B. (1980). Estimation for Markowitz Efficient Portfolios. *Journal of the American Statistical Association*, 75(371), p.544.

Jorion, P. (1986). Bayes-Stein Estimation for Portfolio Analysis. *The Journal of Financial and Quantitative Analysis*, 21(3), p.279.

Knill, O. (2011). Linear Algebra with Probability. *Lecture notes distributed in Math-19B*, Harvard

Kritzman, M. (2006). Are Optimizers Error Maximizers?. *The Journal of Portfolio Management*, 32(4), pp.66-69.

Lakhoo, L. (2017). Enhanced minimum variance optimisation: a pragmatic approach. 1st ed. *University of Cape Town*.

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(5), pp.603-621.

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

Ledoit, O. and Wolf, M. (2003*). Honey, I Shrunk the Sample Covariance Matrix. *SSRN Electronic Journal*.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), p.13.

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

Matoti, L. (2009). Building a statistical linear factor model and a global minimum variance portfolio using estimated covariance matrices. 1st ed.

Merton, R. (1969). Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case. *The Review of Economics and Statistics*, 51(3), p.247.

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

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), p.768.

Munro, B. and Bradfield, D. (2016). Putting the squeeze on the sample covariance matrix for portfolio construction. *Investment Analysts Journal*, 45(1), pp.47-62.

Munro, B. and Bradfield, D. (2016*). The number of stocks required for effective portfolio diversification: the South African case. *South African Journal of Accounting Research*, 31(1), pp.44-59.

Schäfer, J. and Strimmer, K. (2005). A Shrinkage Approach to Large-Scale Covariance Matrix Estimation and Implications for Functional Genomics. *Statistical Applications in Genetics and Molecular Biology*, 4(1).

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), p.425.

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(3), pp.273-282.

Treynor, J. and Black, F. (1973). How to Use Security Analysis to Improve Portfolio Selection. *The Journal of Business*, 46(1), p.66.

Ye, J. (2008). How Variation in Signal Quality Affects Performance. *Financial Analysts Journal*, 64(4), pp.48-61.

Appendix A

Data and External Inputs

A.1 Software

- The statistical language R version 3.1.1 has been used for all coding
- Microsoft Excel and R images have been used for graphics: charts, figures and images
- The data fed into the R has been extracted by the author from a *DataStream* terminal; courtesy of UCT Library

A.2 Data

The data used within this dissertation is weekly returns (in percentages), where at each period τ the return is calculated according to:

$$R_\tau = \frac{X_{\tau+1} + D_\tau}{X_\tau} - X_\tau$$

where : $X_t =$ share price at time t ; $D_t =$ dividend payment at time t

The series provided by *DataStream* named *RI* (Total Return) which, when differenced, provides a single index which corresponds exactly to the return outlined above. I.e. The resultant series *RI* under an ARIMA(0,1,0) is exactly the series R_t ; $t \in (1, \dots, T)$ where T is the length of the time-series.

Appendix B

Resampling Paper Replication in SA context

B.1 Identifying and dealing with survivor bias from *Datastream*

An investigation was carried out, to replicate the methodology of Jobson & Korkie (1981), the work off of which Michaud (1989) based his conclusions. Jobson & Korkie specify the use of the population 'mean' return vector and covariance matrix, as opposed to the 100% accuracy forecasted returns.

We see patterns reminiscent of those seen in the historical returns, within the literature. Bearing in mind, this the methodology applied onto historical data is not replicated within this paper. This is used as a confirmation that the methodology, when applied to historical returns, provides comparable results to the literature..

From original data- *Datastream*:

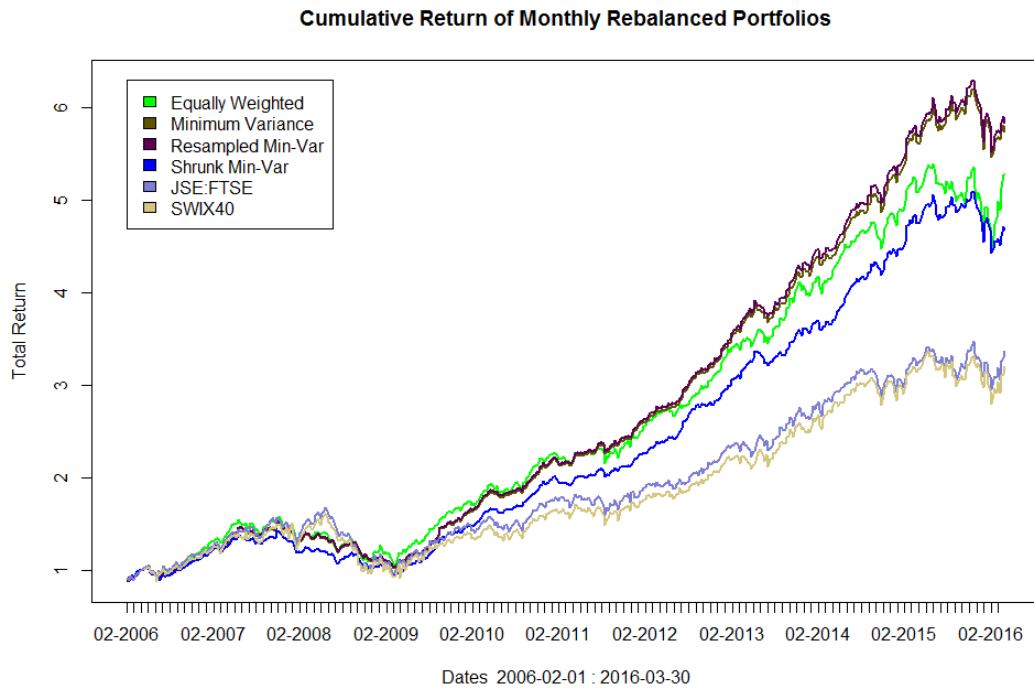


FIGURE B.1: Cumulative returns, optimised on Historical Aggregate. Problematic due to inconsistent measurements defining "total returns" provided by *Datastream* for the index compared to individual shares

Although no graphical display was provided in either of their papers, it was immediately evident that this data did not follow traditional patterns. South African markets seem to reward risky behaviour pre-2010, and reward conservative behaviour post-2011. It is evident that both the "minimum-variance" and the "MV" portfolios consistently outperform the index, and additionally that shrinkage provided no benefit what-so-ever. The question arose as to the integrity of the index data, as such behaviour leads to arbitrage opportunities.

The reality, is that the entirety of *datastream* economic data for individual constituents of any index, is **survival biased**. In relevance to this dissertation- this was not acceptable. The shrinkage algorithms are designed to reduce volatility, however under a volatility which is strongly correlated to positive performance, one will (as we did) continuously shrink the shares which are performing **best** and infact cause a worse overall performance (from the returns perspective) through the use of shrinkage, a globally recognised method to improve performance.

Thus the data was thrown out, and alternative (non-biased) financial data was sourced.