

UNIVERSITY OF CALIFORNIA

Los Angeles

**Portfolio Selection and Risk Management:
An Introduction, Empirical Demonstration
and R-Application for Stock Portfolios**

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Science in Statistics

by

Angela Hei-Yan Leung

2009

© Copyright by
Angela Hei-Yan Leung
2009

The thesis of Angela Hei-Yan Leung is approved.

Nicolas Christou

Hongquan Xu

Jan de Leeuw, Committee Chair

University of California, Los Angeles

2009

*I dedicate this thesis to my
family, friends, and companions along
all stages in my life who have given me
unconditional support and faith in
each and every step I choose to pursue.*

TABLE OF CONTENTS

| | | |
|----------|---|----------|
| 1 | Introduction | 1 |
| 1.1 | Background on Statistical Finance | 1 |
| 1.1.1 | Investment Types | 2 |
| 1.1.2 | Choosing Investments | 3 |
| 1.1.3 | Managing Investments | 4 |
| 1.2 | A Preview of Discussion Topics | 5 |
| 1.2.1 | Portfolio Selection: Modern Portfolio Theory | 5 |
| 1.2.2 | Risk Management: Value-at-Risk | 5 |
| 2 | Portfolio Selection | 6 |
| 2.1 | Modern Portfolio Theory | 6 |
| 2.1.1 | Risk and Return | 6 |
| 2.1.2 | Diversification | 8 |
| 2.1.3 | Short Sales | 8 |
| 2.1.4 | Portfolio Possibilities | 10 |
| 2.1.5 | The Model and the Algorithm - An Overview | 10 |
| 2.2 | Portfolio Selection Methods - Simplified Alternatives | 13 |
| 2.2.1 | Equal Allocation | 13 |
| 2.2.2 | Single Index Model | 14 |
| 2.2.3 | Constant Correlation Model | 18 |
| 2.2.4 | Other Models | 20 |

| | | |
|----------|--|-----------|
| 2.3 | Critiques on Modern Portfolio Theory | 20 |
| 3 | Risk Management | 22 |
| 3.1 | Background and History | 22 |
| 3.1.1 | Types of Risks | 22 |
| 3.1.2 | Consequences of the Lack of Risk Management | 23 |
| 3.2 | Measuring Risk: Value-at-Risk (VaR) | 25 |
| 3.2.1 | Parametric - RiskMetrics | 29 |
| 3.2.2 | Historical Simulation | 30 |
| 3.2.3 | Monte Carlo Simulation | 32 |
| 3.3 | Model Validation | 34 |
| 3.3.1 | Back Testing | 34 |
| 3.3.2 | Stress Testing - Scenario Analysis | 35 |
| 3.4 | Critiques on VaR | 36 |
| 4 | Concept Demonstration - Portfolio Selection | 37 |
| 4.1 | The Data | 37 |
| 4.2 | Portfolio Selection Results | 40 |
| 4.3 | Portfolio Performance Review | 44 |
| 4.4 | Some Critiques | 47 |
| 5 | Concept Demonstration - VaR Management | 48 |
| 5.1 | The Data | 48 |
| 5.2 | Value-at-Risk Calculations | 48 |

| | | |
|----------|--|-----------|
| 5.3 | VaR Model Performance Review using Back Testing | 54 |
| 5.4 | Some Critiques | 59 |
| 6 | The R-Application | 60 |
| 6.1 | The Motivation | 60 |
| 6.2 | The Source Codes | 60 |
| 6.2.1 | Data Management | 61 |
| 6.2.2 | Portfolio: Equal Allocation | 62 |
| 6.2.3 | Portfolio: Single Index (With Short Sales) | 62 |
| 6.2.4 | Portfolio: Single Index (No Short Sales) | 64 |
| 6.2.5 | Portfolio: Constant Correlation (With Short Sales) | 65 |
| 6.2.6 | Portfolio: Constant Correlation (No Short Sales) | 67 |
| 6.2.7 | VaR: Parametric (Equally Weighted) | 68 |
| 6.2.8 | VaR: Parametric (Exponentially Weighted) | 69 |
| 6.2.9 | VaR: Historical (Equally Weighted) | 70 |
| 6.2.10 | VaR: Historical (Bootstrapping) | 70 |
| 6.2.11 | VaR: Monte Carlo Simulation | 71 |
| 7 | Concluding Remarks | 72 |
| | References | 73 |

LIST OF FIGURES

| | | |
|------|---|----|
| 2.1 | Portfolio Possibilities Curve - Three Assets | 11 |
| 3.1 | Value-at-Risk Visual Demonstration | 27 |
| 3.2 | Value-at-Risk Extreme Case | 27 |
| 4.1 | Daily Price Plot for DJIA Stocks June 01 - Dec 07 | 38 |
| 4.2 | Daily Return Plot for DJIA Stocks June 01 - Dec 07 | 38 |
| 4.3 | Monthly Price Plot for DJIA Stocks June 01 - Dec 07 | 39 |
| 4.4 | Monthly Return Plot for DJIA Stocks June 01 - Dec 07 | 39 |
| 4.5 | Portfolio Risk and Return using data from 2001-2007 | 43 |
| 4.6 | Portfolio Risk and Return using data from 2008-2009 | 46 |
| 4.7 | Portfolio Performance in 2008-2009 | 47 |
| 5.1 | Equal Allocation Portfolio Return Distribution | 49 |
| 5.2 | Single Index Model (Standard Short Sales) Distribution | 50 |
| 5.3 | Single Index Model (Lintner's Short Sales) Distribution | 50 |
| 5.4 | Single Index Model (No Short Sales) Distribution | 51 |
| 5.5 | Constant Correlation (Standard Short Sales) Distribution | 51 |
| 5.6 | Constant Correlation (Lintner's Short Sales) Distribution | 52 |
| 5.7 | Constant Correlation (No Short Sales) Distribution | 52 |
| 5.8 | Back Testing: Parametric VaR Equally Weighted | 56 |
| 5.9 | Back Testing: Parametric VaR Exponentially Weighted | 57 |
| 5.10 | Back Testing: Historical VaR Equally Weighted | 57 |

| | |
|---|----|
| 5.11 Back Testing: Historical VaR Bootstrapping | 58 |
| 5.12 Back Testing: Monte Carlo Simulated VaR | 58 |

LIST OF TABLES

| | | |
|-----|---|----|
| 4.1 | Summary Statistics of DJIA Stocks in Portfolio | 41 |
| 4.2 | Portfolio Allocation Summary | 42 |
| 4.3 | Portfolio Risk and Return Snapshot 2001-2007 | 43 |
| 4.4 | Portfolio Risk and Return Snapshot 2008-2009 | 46 |
| 5.1 | Summary Statistics of Portfolio Returns | 49 |
| 5.2 | Next-Day VaR: 95% C.I., 1-Year Window | 54 |
| 5.3 | Back Testing Results for VaR Models - Failure Rates | 56 |

ACKNOWLEDGMENTS

This thesis will not be complete without the following very important people's professional and generous support and advice. Firstly, I need to thank Professor Jan de Leeuw for being the chair of my thesis committee and making it such a smooth process. Next, I have to thank Professor Hongquan Xu for being a committee member. From the time I told him about my wildest interests in Six Sigma to Chinese Medicine and Horoscope, until now when I finally chose to do my thesis on the topic of statistical finance, he has shown great support. Furthermore, I cannot thank Professor Nicolas Christou enough for his 10+ years of mentorship and friendship. He has played the most important role in my pursuit of graduate studies and my appreciation for his advice and kindness is simply beyond words. Finally, I need to thank Szilard Pafka for his generous advice on the topic of mathematical finance. His tremendous help is truly unconditional and his enthusiasm in the subject has deeply moved me.

ABSTRACT OF THE THESIS

**Portfolio Selection and Risk Management:
An Introduction, Empirical Demonstration
and R-Application for Stock Portfolios**

by

Angela Hei-Yan Leung

Master of Science in Statistics

University of California, Los Angeles, 2009

Professor Jan de Leeuw, Chair

This paper serves as an introduction to Portfolio Selection and Risk Management theories founded upon Harry M. Markowitz's Nobel prize winning Modern Portfolio Theory. Simplified portfolio selection algorithms including the Single Index and Constant Correlation Models, along with widely-used risk management tool Value-at-Risk calculation methods are described in detail in the introduction sections of this paper. In the concept demonstration sections, the above theories are applied and evaluated using historical data of the Dow Jones Industrial Average stocks. A specially designed application in statistical software R is included in the application section of this paper to enable beginning investors and statistical finance students to put these ideas in action using their investment data of choice.

CHAPTER 1

Introduction

1.1 Background on Statistical Finance

Financial health plays a pivotal role in the overall well-being of an economy, organization, or individual. This can certainly be assessed qualitatively, but in order to make comparisons both vertically and horizontally, it makes sense to quantify this notion with the use of numbers and statistics. Therefore, it is vitally important to have standards and means to manage, monitor, maintain, and grow wealth. This leads to the emergence of statistical finance.

There are two streams in quantitative financial analysis or investment valuation: fundamental and technical, both of which make use of mathematical models for valuation. Fundamental analysis is an evaluation method that attempts to measure an asset's or investment's intrinsic value by examining related economic, financial and other qualitative and quantitative factors. Fundamental analysts attempt to study everything that can affect the security's value, including macroeconomic factors (like the overall economy and industry conditions) and organizational-specific factors (like financial condition and management of the entity of interest). The goal of performing fundamental analysis is to produce a value that an investor can compare with the security's current price, with the aim of figuring out what sort of position to take with that security.

On the other hand, technical analysis is a method of evaluating assets or

securities by analyzing statistics generated by market activity, such as past prices and volume. Technical analysts do not attempt to measure a security's intrinsic value, but instead use charts and other tools to identify patterns that can suggest future activity. Technical analysts believe that the historical performance of stocks and markets are good indications of future performance.

In this paper, we introduce and demonstrate some investment management methods in the technical analysis branch, which involve the use of well known statistical theories and methods.

1.1.1 Investment Types

Investing can be described as the redirection of resources from being consumed today to creating benefits in the future. [19] In other words, it is the use of assets to earn income or profit. Although we are no longer in a bartering society where goods were often more perishable, it is preferable, if not essential, to invest instead of keeping assets idle, so that investments can grow to fight against inflation and future uncertainties. This gives birth to the different types of financial instruments and investments we see today. A financial investment is the current commitment of dollars for a period of time in order to derive future payments that will compensate the investor for the time the funds are committed, the expected rate of inflation, and the uncertainty of the future payments. [15] The following are examples of some common types of investments:

Equities Title of ownership to a company which gives the right to receive a dividend, with the amount of the dividend being calculated according to the company's earnings.

Bonds Loans issued by a company or by a Government/State that give rise to

regular payment of coupons, which constitute the interest on the loan, and redemption of the security at maturity.

Money Market Instruments Short-term borrowing and lending for managing the cash in a portfolio.

Derivatives Investments whose values are “derived” from underlying assets or instruments. Examples are options, futures, and swaps. The use of derivatives allows portfolio managers to be more flexible and effective in developing and applying investment strategies due to the use of leverage. For example, derivative instruments allow the portfolio exposure to be modified in terms of assets and currencies, without modifying the real composition of the underlying portfolio. They also allow portfolio risk to be hedged and performance to be improved by using a leverage effect on the return. [1]

In this paper, we focus on equity investments. Specifically, we will be using stocks as the sole investment in all examples since such data are widely accessible. And since the idea behind this type of investment is intuitive and easy to understand, it is an ideal tool to be used as an introduction and demonstration of the statistical methods we will be discussing.

1.1.2 Choosing Investments

We focus our attention on stock investments in this paper. In the previous section we mentioned the two school of thoughts on investment analysis, fundamental and technical. In terms of choosing equity investments, meaning which companies' stocks to purchase or sell, it is intuitive for the company's past performance as well as future business plans to be considered. In order to do so, analyses on the company's financial statements and accounting or financial records need

to be conducted, and the intrinsic value of the company can be estimated and compared with its current stock price. A decision on whether to purchase or sell shares of this company can then be made. This is the fundamentalist approach of choosing investments. As we can see, most of the analyses focus on the business side.

The technical school of thought, however, focus on the behavior of the stock prices against factors such as market movements and other economic factors like interest rates and unemployment patterns. In this case, the internal matters of the companies is less of a concern. As you can see, in this paper, we lean towards the technical school of thought and arbitrarily determine our investment portfolio to be the 30 stocks currently listed under the Dow Jones Industrial Average list.

1.1.3 Managing Investments

Once securities are chosen and portfolio formed, it needs to be managed in order to maximize the benefits (minimize the losses). This gives rise to the idea of portfolio management. Portfolio management consists of constructing portfolios with proper allocation of assets and then making them evolve in order to reach the return objectives defined by the investor, while respecting the investor's constraints in term of risk and asset allocation. The investment methods used to reach the objectives range from quantitative investment, which originated from modern portfolio theory, to more traditional methods of financial analysis as we previously defined. [1] In this paper, we introduce and demonstrate methods of managing portfolios using modern portfolio theory.

1.2 A Preview of Discussion Topics

1.2.1 Portfolio Selection: Modern Portfolio Theory

Modern Portfolio Theory is Harry Markowitz's theory of portfolio choice in an uncertain future. In this theory, he quantified the difference between the risk of portfolio assets taken individually and the overall risk of the portfolio. The theory offers a solution to the problem of portfolio choice for a risk-averse investor: the optimal portfolios, from the rational investor's point of view, are defined as those that have the lowest risk for a given return. These portfolios are said to be mean-variance efficient. This theory is the foundation of our Portfolio Selection discussion.

1.2.2 Risk Management: Value-at-Risk

After a so-called optimal portfolio is selected, there comes the need to monitor and manage the return and risk components of this portfolio of choice. While it is desirable to maximize expected returns, it is equally important to keep the risk component under control, especially for investments that depend heavily on the volatility factor such as Options. For institutions that invest heavily in different types of financial instruments, such as banks or hedge funds, there is a strong need to integrate and measure the risks involved firm-wide. This gives rise to the development of Value-at-Risk, which summarizes the worst loss over a target horizon with a given level of confidence. [12] In this paper, although we only focus on stock portfolio which are less complex than investments held by financial institutions mentioned above, the demonstration of VaR calculations will give readers a flavor of how the methods work and what kind of statistical ideas are utilized.

CHAPTER 2

Portfolio Selection

2.1 Modern Portfolio Theory

Together with William Sharpe and Merton Miller, Harry Markowitz was awarded the Nobel Prize for their research in 1990. In the research, Markowitz demonstrated that the portfolio risk came from the covariances of the assets that made up the portfolio. The marginal contribution of a security to the portfolio return variance is therefore measured by the covariance between the security's return and the portfolio's return rather than by the variance of the security itself. Markowitz thus established that the risk of a portfolio is lower than the average of the risks of each asset taken individually and gave quantitative evidence of the contribution of diversification. [1]

In this paper, we apply Markowitz's idea and methodologies to real life data and assess the observed results.

2.1.1 Risk and Return

It is essential to discuss the basic characteristics of the return and its dispersion (going forward, we define risk as the standard deviation of the asset/portfolio of interest) of an asset as well as a portfolio before we proceed to the discussion of portfolio selection methodologies. For any asset, the expected return, variance,

and covariance are expressed as follows:

$$\bar{R}_i = \sum_{t=1}^N \frac{R_{it}}{N} \quad (2.1)$$

$$\sigma_i^2 = \sum_{t=1}^N \frac{(R_{it} - \bar{R}_i)^2}{(N-1)} \quad (2.2)$$

$$\text{cov}(R_i, R_j) = \sigma_{ij} = \sum_{t=1}^N \frac{(R_{it} - \bar{R}_i)(R_{jt} - \bar{R}_j)}{(N-1)} \quad (2.3)$$

where

N denotes the numbers of observed returns of asset i , and R_{it} denotes the t^{th} observed return of asset i .

Further, the correlation coefficient ρ between assets A and B is always between -1 and 1 and is expressed as follows:

$$\rho = \frac{\text{cov}(R_A, R_B)}{\sigma_A \sigma_B} \quad (2.4)$$

For a portfolio with assets that are linear combinations of one another, the expected return and variance are expressed as follows:

$$\bar{R}_p = E(R_p) = \sum_{i=1}^N x_i \bar{R}_i \quad (2.5)$$

$$\sigma_p^2 = \sum_{j=1}^N x_j^2 \sigma_j^2 + \sum_{j=1}^N \sum_{\substack{k=1 \\ k \neq j}}^N (x_j x_k \sigma_{jk}) \quad (2.6)$$

where

x_i denotes the proportion of asset i held in the portfolio.

2.1.2 Diversification

The variance expression in Equation 2.6 reveals the usefulness of diversification in reducing risk attributed to the correlation that exists between asset returns. Before going into the detail of choosing the optimal diversification weights, we present a simple example of why diversification works. [3]

Assume we have a portfolio of N assets with equal amounts of capital invested in each. With this assumption, our variance expression becomes:

$$\sigma_p^2 = \sum_{j=1}^N \left(\frac{1}{N}\right)^2 \sigma_j^2 + \sum_{j=1}^N \sum_{\substack{k=1 \\ k \neq j}}^N \left(\frac{1}{N} \frac{1}{N} \sigma_{jk}\right) \quad (2.7)$$

$$= \frac{1}{N} \sum_{j=1}^N \left[\frac{\sigma_j^2}{N}\right] + \frac{N-1}{N} \sum_{j=1}^N \sum_{\substack{k=1 \\ k \neq j}}^N \left[\frac{\sigma_{jk}}{N(N-1)}\right] \quad (2.8)$$

$$= \frac{1}{N} \bar{\sigma}_j^2 + \frac{N-1}{N} \bar{\sigma}_{jk} \quad (2.9)$$

From the above expression, we can easily see that as we increase the number of assets N , say to infinity, the first term approaches zero, while the second term approaches $\bar{\sigma}_{jk}$, which is the covariance term. This means that the individual risk of securities can be diversified away, and what is left to contribute to total portfolio risk is the covariance, which can be interpreted as the systemic market risk.

2.1.3 Short Sales

Now we are convinced that diversification is beneficial in reducing risks, we go into the discussion of short sales. As noted previously, x_i denotes the weights of each asset within the portfolio, and the sum of these weights must equal to the total available capital. In reality, capital can be obtained by a person's own savings, or by means of lending or borrowing. Therefore, these weights can be positive

(savings) as well as negative (loans). In the case when borrowing is not possible, we call it “Short Sales not allowed”, and this translate to the constraint where all the weights x_i must be positive. On the other hand, if borrowing is possible, we call this scenario “Short Sales Allowed”, and the x_i can be either positive or negative. Short Sales simply means to borrow a security (or commodity futures contract) from a broker and selling it, with the understanding that it must later be bought back (hopefully at a lower price) and returned to the broker. This is a technique used by investors who try to profit from the falling price of a stock.

The constraint of the standard definition of Short Sales is

$$\sum_{i=1}^N x_i = 1$$

which makes sense intuitive. Simply put, this constraint ensures that the total lending and borrowing of capital allocated to each asset in the portfolio sum to the total capital available. This is the constraint to be used when solving the optimization problem we will discuss in the next section.

Alternatively, Lintner has advocated a different definition of short sales, one that is more realistic. He assumes that when an investor sells stock short, cash is not received but rather is held as collateral. Furthermore, the investor must put up an additional amount of cash equal to the amount of stock he or she sells short. The investor generally does not receive any compensation on these funds. However, if the investor is a broker-dealer, interest can be earned on both the money put up and the money received from the short sales of securities [3]. This leads to the constraint:

$$\sum_{i=1}^N |x_i| = 1$$

It is beyond the scope of this paper to go into the technicalities of how this

constraint is derived. However, both definitions of short sales will be used in our concept demonstration and application section of this paper.

2.1.4 Portfolio Possibilities

If we plot the risks and returns of different portfolios consisting of different combination of weights (with the standard definition of short-sales as our constraint), we get the so-called portfolio possibilities curve. As an example, let's explore a portfolio with 3 assets. Generating different weight combinations, we get a picture resembling Figure 2.1. Here, the green portion represent possibilities where short-sales are not allowed (no negative weights), and the blue portion otherwise. As we can see, by having leverage (ability to short sell), investors are opened up to a lot more possibilities of higher returns (and also higher risks). In both cases, we realize that for the same amount of risk, there are always portfolios that yield higher returns than others. These portfolios are called "efficient portfolios" and they lie on the so-called "efficient frontier". Intuitively, these portfolios are more desirable. In this picture, the efficient frontier corresponds to the top half portion of the portfolio possibilities area.

2.1.5 The Model and the Algorithm - An Overview

To achieve optimal diversification, a mathematical portfolio selection model was developed by Markowitz. This model finds the composition of all the portfolios that correspond to the efficiency criterion defined for a given set of securities, and construct the corresponding efficient frontier. Simply put, the portfolio selection model involves minimizing the variance (risk) for a given return or maximizing the return for a given risk, which can be written as follows:

Portfolio Possibilities – Three Assets

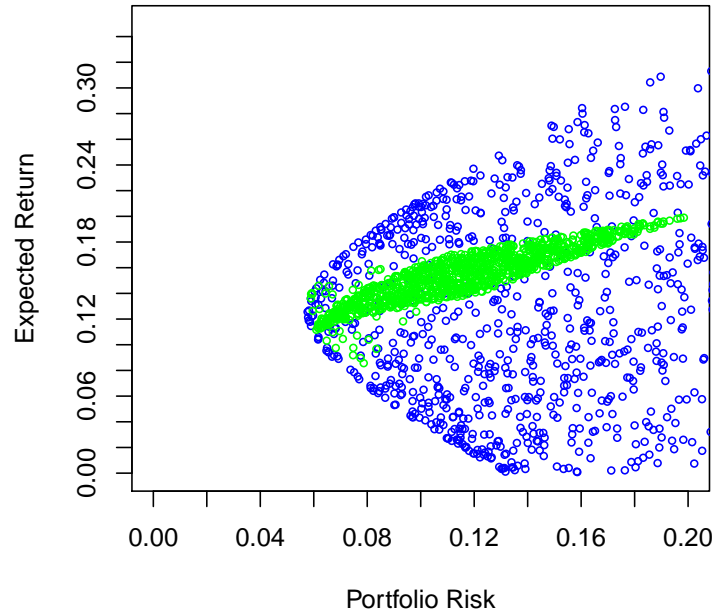


Figure 2.1: Portfolio Possibilities Curve - Three Assets

$$\text{Minimize } \sigma_p^2$$

under the conditions:

$$E(R_p) = E \tag{2.10}$$

$$\sum_{i=1}^N x_i = 1 \tag{2.11}$$

The efficient frontier calculation involves finding the weights of the assets that make up each portfolio. This is equivalent to the computation of the efficient frontier in Figure 2.1, with specified risk or return conditions to come up with an optimal set of weights. As we can imagine, this optimization problem becomes more computationally complicated to solve as we increase the number of assets and constraints. We will not go into the mathematical detail of how this algorithm

is derived in this paper, instead, the essence of the algorithm is described in the following paragraphs.

First, the investor's objective is defined as a function of the λ parameter, which measures the investor's aversion to risk. By varying λ , the characterization of the complete efficient frontier can be obtained. Without constraints, the resolution is achieved by defining the Lagrangian associated with the problem and equating all its partial derivatives to zero. We then have a linear system that gives us the proportions of the optimal portfolio (for each value of λ). If there are constraints on the weights, then for each value of λ , the solution will contain x_i values that will reflect the different x_i limits. These variables will have down, in and up status, respectively.

Since there are constraints in our optimization problem, we do not set all the partial derivatives of the Lagrangian to zero, but instead, we assume that the derivatives are strictly positive compared with the variables with down status, and strictly negative compared with variables with up status. This is a specific case of the Kuhn Tucker conditions. Assuming that the status of each variable is known, it is easy to write the system of linear equations that allows the solution to the problem to be calculated. The problem actually involves determining the status of each of the variables. The status depends on the value of λ . For each variable there is a critical value of λ for which the status of the variable changes. This value is calculated according to the parameters of the problem. We can then sequence the critical λ obtained for all the variables. The resolution algorithm involves giving an initial status to the variables in order to calculate a first solution that is valid up to the first critical λ . We then proceed sequentially by carrying out the necessary modifications each time a variable changes status. The portfolios that correspond to the critical λ are called "corner" portfolios.

They are sufficient for calculating the whole efficient frontier. [1]

In the book *Modern Portfolio Theory and Investment Analysis* by Elton et al., techniques to solve for the solution to the portfolio problem under different assumptions (short sales allowed and riskless lending/borrowing exists, short sales allowed no riskless lending/borrowing, short sales not allowed with riskless lending/borrowing, and short sales and riskless lending/borrowing not allowed) are listed in detail. The main issues with these techniques are that huge amounts of input data and computation time are needed. Furthermore, the input data are in a form to which the security analyst and portfolio manager cannot easily relate. For this reason, it is difficult to get estimates of the input data or to get practitioners to relate to the final output. [3] Keeping this thought, we now move onto the focus of this chapter where simplified alternative selection algorithms are presented. These methods are less computationally demanding with more manageable data collection requirements.

2.2 Portfolio Selection Methods - Simplified Alternatives

2.2.1 Equal Allocation

As we saw in Section 2.1.2, diversification is important in order to decrease risk for the same amount of return, or increase return for the same amount of risk. In our demonstration we used a simple equal allocation where

$$x_i = \frac{1}{N}, \text{ N = Number of Assets in Portfolio}$$

This is a simple way of diversifying a portfolio with no optimization needed. As mentioned in the previous sections, for any kinds of portfolio, the respective

expected return and variance are:

$$E(R_p) = \sum_1^N x_i R_i \quad (2.12)$$

$$\sigma_p = \sqrt{\sum_{i=1}^N x_i^2 \sigma_i^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N x_i x_j \sigma_{ij}} \quad (2.13)$$

Equal Allocation will be one of our demonstration portfolios in Chapter 4.

2.2.2 Single Index Model

William Sharpe (1963) studied Markowitz's research and worked on simplifying the calculations in order to develop a practical use of the model. This signifies the birth of the Single Index Model, which is written as follows:

$$R_{it} = \alpha_i + \beta_i R_{Mt} + \epsilon_{it} \quad (2.14)$$

where

R_{it} denotes the return on asset i ;

R_{Mt} denotes the return on market index;

ϵ_{it} denotes the specific return on asset i ; and

α_i and β_i are the coefficients to be determined.

The coefficients of the line, α_i and β_i , are obtained by linear regression of the market returns on the asset returns for the same period. The method used is the so-called ordinary least squares. The beta coefficient is given by:

$$\beta_i = \frac{cov(R_{it}, R_{Mt})}{var(R_{Mt})} \quad (2.15)$$

Following on from the model definition, the residual terms ϵ_{it} are non-correlated with the market return. The total risk of an asset is therefore broken down into a term for systematic risk (or market risk) and a term for non-systematic risk (or

diversifiable risk), or

$$\text{var}(R_{it}) = \beta_i^2 \text{var}(R_{Mt}) + \text{var}(\epsilon_{it}) \quad (2.16)$$

This relationship, when applied to a portfolio containing N securities in the same proportions, highlights the usefulness of diversification in reducing risk. The risk of this portfolio is written as follows:

$$\text{var}(R_{pt}) = \beta_p^2 \text{var}(R_{Mt}) + \frac{1}{N^2} \sum_{i=1}^N \text{var}(\epsilon_{it}) \quad (2.17)$$

The second term of this relationship tends towards 0 for a sufficiently large N . Therefore, the risk of a broadly diversified portfolio is only made up of the market risk. For all portfolios, the correlation coefficient between the index and the portfolio, raised to the square, gives the percentage of the total variance of the portfolio returns that is explained by the index movements.

The use of this model allows the Markowitz problem correlation matrix calculation to be simplified considerably. We have:

$$\text{cov}(R_{it}, R_{jt}) = \sigma_{ij} = \beta_i \beta_j \sigma_M^2, i \neq j \quad (2.18)$$

and since $\text{cov}(\epsilon_i, \epsilon_j) = 0$, the residual returns are non-correlated, and:

$$\text{var}(R_{it}) = \sigma_i^2 = \beta_i \sigma_M^2 + \text{var}(\epsilon_{it}) \quad (2.19)$$

Instead of having to calculate the $\frac{N(N+1)}{2}$ terms from the complete matrix, only $(2N+1)$ terms remain to be determined. For a broadly diversified portfolio, therefore, the value of its return and risk are simply:

$$\sigma_p = \beta_p \sigma_M = \sum_{i=1}^N x_i \beta_i \sigma_M \quad (2.20)$$

$$E(R_p) = \alpha_p + \beta_p E(R_M) = \sum_{i=1}^N x_i \alpha_i + \sum_{i=1}^N x_i \beta_i E(R_M) \quad (2.21)$$

since the residual returns have a null mean. We therefore come back to a completely linear problem.

Elton and Gruber [5] [6] [7] further proposed techniques for determining optimal portfolios that give a better understanding of the choice of securities to be included in a portfolio. The calculations are easy to carry out and lead to similar results to those given by the Markowitz model with the complete matrix. These techniques are founded upon the Single Index Model. The method is based on an optimal ranking of the assets, established with the help of the simplified correlation representation model. We determine a threshold in the list. All the securities located above the threshold will be part of the optimal portfolio. Those that are below the threshold will be excluded. The problem is to determine the value of the threshold. We subsequently determine the percentage to be assigned to each security, according to the characteristic parameters of each security and the value of the threshold. The ratio used to rank the assets is given by:

$$\frac{E(R_i) - R_F}{\beta_i} \tag{2.22}$$

where

$E(R_i)$ denotes the expected return of asset i ;

R_F denotes the risk-free rate; and

β_i denotes the β of asset i defined by the Single Index Model.

The ratio is calculated for all the assets from which the investor is liable to choose. The results are classified from the highest value to the lowest value. The higher the value of the ratio, the more desirable it is to hold the security in the portfolio. As a result, if a security is held in an optimal portfolio, then all the securities for which this ratio is higher are also held in the portfolio. In the same way, if a security is excluded from the portfolio, then all the securities with a lower ratio will also be excluded. There is therefore a threshold value above which we

select securities, and below which we exclude them. This value is denoted by C^* .

C^* is calculated by an iterative procedure, by successively introducing the securities from the list into the portfolio and by calculating the C_i value associated with a portfolio containing i securities. C_i is expressed as follows:

$$C_i = \frac{\sigma_M^2 \sum_{j=1}^i \frac{(E(R_i) - R_F)\beta_j}{\text{var}(\epsilon_j)}}{1 + \sigma_M^2 \sum_{j=1}^i \frac{\beta_j^2}{\text{var}(\epsilon_j)}} \quad (2.23)$$

with the notation being the same as that used to present the Single Index Model.

We stop adding securities when the ratio associated with the candidate security is lower than the value of C_i that was calculated for the portfolio. At the end of this procedure we know the list of assets that will figure in the optimal portfolio. What remains is to determine the proportion to attribute to each of them. Denoting by x_i the proportion of asset i to hold, we have:

$$x_i = \frac{Z_i}{\sum_{i=1}^N Z_i} \quad (2.24)$$

where N denotes the number of securities in the portfolio, and

$$Z_i = \frac{\beta_i}{\text{var}(\epsilon_i)} \left(\frac{E(R_i) - R_F}{\beta_i} - C^* \right) \quad (2.25)$$

In the case where short-sales are not allowed, C^* is simply $\text{Max}(C_i)$. When short sales are allowed, we include all securities in our portfolio and hold the previously excluded securities short. C^* then becomes the last C_i in our portfolio. Under the standard definition, the weights is calculated the same way in Equation 2.24. However, if we use the Linter's definition of short sales as discussed in Section

2.2.4, the Equation becomes:

$$x_i = \frac{Z_i}{\sum_{i=1}^N |Z_i|} \quad (2.26)$$

Our R-application algorithm include both definitions of short-sales based upon the above descriptions.

2.2.3 Constant Correlation Model

Similar to the Single Index Model algorithm shown above, another way of portfolio selection is the Constant Correlation Model method. Under this method, the pairwise correlation coefficients of securities are assumed to be constant. The estimate for each pair of stocks is expressed as follows:

$$\rho = \frac{\sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N \rho_{ij}}{N(N-1)} \quad (2.27)$$

where

ρ_{ij} denotes the covariance between security i and j ; and
 N denotes the number of securities in the portfolio.

The selection algorithm of the Constant Correlation Model is similar to that of the Single Index Model. The only difference is that instead of ranking by excess return to β , we now rank the securities by the excess return to σ , which is written as follows:

$$\frac{(R_i - R_F)}{\sigma_i} \quad (2.28)$$

where

R_i denotes the expected return on security i ;

R_F denotes the risk-free rate; and

σ_i denotes the standard deviation of the return of security i .

After securities are ranked using the above ratio, the securities with greater excess return to σ than the cut-off point will be included into the optimal portfolio. The cut-off point is computed as follows:

$$C^* = \frac{\rho}{1 - \rho + ip} \sum_{j=1}^i \frac{R_j - R_F}{\sigma_j} \quad (2.29)$$

where

R_j denotes the expected return on security j ;

R_F denotes the risk free rate;

σ_j denotes the standard deviation of the return of security j , and

ρ denotes the constant correlation coefficient.

Same as before, we include securities with C_i greater than C^* , and the Z_i is expressed as follows:

$$Z_i = \frac{1}{\sigma_i(1 - \rho)} \left[\frac{R_i - R_F}{\sigma_i} - C^* \right] \quad (2.30)$$

and the weights of each security is expressed as:

$$x_i = \frac{Z_i}{\sum_{i=1}^N Z_i} \quad (2.31)$$

where N denotes the number of securities in the portfolio. We can see that this algorithm is very similar to the Single Index Model shown in the previous section. In our R-Application, we program this as our calculation. The same regarding short-sales apply here: when short-sales are not allowed, C^* in this problem is simply the $\text{Max}(C_i)$. When short-sales are allowed, all securities are included and C^* is simply the last C_i in the iterative process. Optimal weights are

then calculated using Equation 2.24 and 2.26 respectively, depending on which definition of short-sales we decide to use.

Once again, our R-application algorithm include both definition of short-sales based upon the above descriptions.

Finally, the expected return and risk of this portfolio can be expressed as:

$$R_p = \sum_{i=1}^N x_i(R_i - R_F) + R_F \quad (2.32)$$

$$\sigma_p = \sqrt{\sum_{i=1}^N x_i^2 \sigma_i^2 + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N x_i x_j \sigma_{ij}} \quad (2.33)$$

2.2.4 Other Models

Other than the methods discussed above, there are also the Multi-Group Model as well as Multi-Index Model that can similarly be used to find optimal portfolios using iterative algorithms. However, since these models require information that are not as utterly available when we download data from open sources, we do not include them into our R-Application and discussions in this paper.

2.3 Critiques on Modern Portfolio Theory

Theoretically impressive and applicable once simplified, modern portfolio theory has nevertheless drawn severe criticisms. The principle objection is with the concept of β : while it is possible to measure the historical β for an investment, it is not possible to forecast it accurately. Without an accurate forecast, it is impossible to build a theoretically perfect portfolio. The theory also assumes it is possible to select investments whose performance is independent of other investments in the portfolio. However, in times of financial distress or market

turmoil, correlation of securities tend to increase and the undiversifiable risk of every portfolio increases, making the optimal portfolio not as valid as what theoretically holds. We will later see also that similar criticisms are made to our next discussion topic of Risk Management based on Value-at-Risk.

CHAPTER 3

Risk Management

3.1 Background and History

From our discussion of portfolio selection in the previous chapter, so far, we have been defining “risk” as a portfolio’s standard deviation or volatility. Recall that an investor’s goal is to achieve a certain predetermined range of return (normally high) with a predetermined level of risk (usually low) involved. Therefore, besides looking at a portfolio’s performance in generating returns, it is also important to monitor and manage its associated risk regularly in order to be in line with the aforementioned goal. Risk management gets even more crucial for entities whose investments are heavily dependent on the movements of financial markets, as well as other factors that potentially have impacts on such movements.

3.1.1 Types of Risks

Market Risk arises from movements in the level or volatility of market prices.

This is the risk we focus on in our application section.

Credit Risk originates from counterparties unwillingness or inability to fulfill their contractual obligations. This party can be an individual, corporation, or government. Credit risk has been historically more difficult to measure and manage due to the relatively lower frequency of default of corporations

or governments, which makes it harder to find sufficient data to model such behavior.

Liquidity Risk can be further broken down into market liquidity and funding liquidity risk. The former refers to the inability to conduct a transaction at prevailing market prices due to the size of the position relative to normal trading lots. The latter, also called cash flow liquidity risk, refers to the inability to meet payments obligations. These are risks that closely relate to the timing of transactions occurred.

Operational Risk arises from human and technological errors or accident. This kind of risk is extremely difficult to model and manage.

Other risks include legal risk, environmental risk, etc.

Progresses and advancements are made everyday in the management of each of the risks described above. Management and model of market risk is undoubtedly the most mature among all due to the availability of data that are involved in the modeling of such risk. Hence, in our application section we will be focusing on the measure of market risk, specifically for portfolios that consist of only stocks.

3.1.2 Consequences of the Lack of Risk Management

Although this paper serves as an introduction to the measurement and management market risk related to stock portfolios, it is worthwhile to briefly mention some historical (and current) events that caught the the world's attention to better manage financial risks. The following are some financial disasters that arose mainly due to the lack of risk management.

Portfolio Insurance [22] Around 1980, a dynamic hedging technology was de-

veloped by Leland, Obrien, and Rubinstein in which the principle came from option pricing theory. The principle of this product is to hedge a stock portfolio against market risk by selling stock index futures short or buying stock index put options. This product worked well until the market crash of October 1987, which was later blamed to have aggravated the disastrous outcome of the crash. “Portfolio Insurance” in the face of the 1987 market meltdown shows that even theoretically sophisticated models face the risk of market and need to be more closely monitored by users as well as governmental agencies to prevent catastrophic outcomes.

Long Term Capital Management [11] In 1994, hedge fund Long Term Capital Management (LTCM) which consisted of an all star team of traders and academics was formed in an attempt to create a fund that would profit from the combination of the academics’ quantitative models and the traders’ market judgment and execution capabilities. LTCM primarily uses trading strategy of convergence to profit from arbitrage opportunities. Not surprisingly, it generated stellar performance until swap spreads started widening unexpectedly after the Russia’s default on its government obligations along with other global events. The 1998 failure of LTCM created great turmoil in the world’s financial system. Many lessons were learned from the LTCM case at great cost. At the minimum, we learned that when evaluating risks of any investment portfolios, one should not rely solely on mathematical models, but instead should use them as aids along with judgment and common sense. Also, one needs to keep in mind of unexpected correlations or the breakdown of historical correlations, and perform stress testing and extreme scenario analysis to any investment profile.

Financial Crisis of 2007-2009 The “financial meltdown” we are currently ex-

periencing (at the time this paper is written) has a global effect very much like the previous two events but at an even greater magnitude due to increased globalization. The immediate cause of the crisis was the bursting of the United States housing bubble. Contributed to the growth of the housing bubble is primarily due to the lax of lending (credit risk), deregulation, and overleveraging. It is beyond the scope of this paper to decipher each of the components that make up this crisis. But a lesson learned, undoubtedly, is that proper risk management and education in financial products could have given more obvious alerts to regulators and investors alike, and the outcome of the disaster could have been mitigated.

These very few examples all involved the lack of risk management and the short of investors' and even governments' knowledge in very complicated financial instruments. It is therefore essential that not only financial institutions, but investors and governments, be well educated in the qualitative and quantitative aspects of investing in order to prevent more disasters from happening.

3.2 Measuring Risk: Value-at-Risk (VaR)

Due to the complexity of portfolios held by institutional investors these days, there is a strong need to have a standardized way to measure the risk that is exposed by the entity, especially those that play major roles in the world's financial market (e.g. financial institutions). This is especially necessary now that complex financial instruments such as options, swaps, and different types of structured loans like collateralized mortgage obligations (CMO) and collateralized debt obligations (CDO) abound. Different portfolio classes have different ways of risk assessments (e.g. duration for a fixed income portfolio, β for an

equity business, etc.) and these risks are difficult to be combined in a meaningful way. It is also difficult to aggregate results available at different times, such as positions marked in different time zones, or a high frequency trading desk with a business holding relatively illiquid positions. But since every business unit in a financial institution contributes to profit and loss in an additive fashion, and many of them mark-to-market daily, it is natural to define firm-wide risk using the distribution of possible losses at a fixed point in the future. It is based on the desire to have a meaningful measure of firm-wide risk that gave birth to Value-at-Risk (VaR).

In the simplest term, VaR attempts to make a statement regarding the risk of a firm or a particular combination of portfolios similar to the following:

*We are **95%** certain that we will not lose more than **\$1,300** in the next **1** day.*

Figure 3.1 shows pictorially the statement we just made. In this figure, we can see that the value-at-risk amount of \$1.3K is the cut-off point where there is 5% of worse possible losses lying below this point.

As we can see, three parameters come into play: a confidence level, a time horizon, and an amount of loss. It is important **NOT** to perceive VaR as the “worst case scenario”, since the amount is only the worst loss based on a certain confidence level. This means that there is still a 5% chance that the loss can get worse, or a lot worse, than the amount stated. An extreme case can be seen in Figure 3.2. Here, we can see that although the VaR is still \$1.3K, the potential loss is now greater. Therefore, VaR can be perceived as the “best of the worst cases” instead.

In practice, the time horizon N is almost always set to $N = 1$. This is because there is not enough data to estimate directly the behavior of market variables

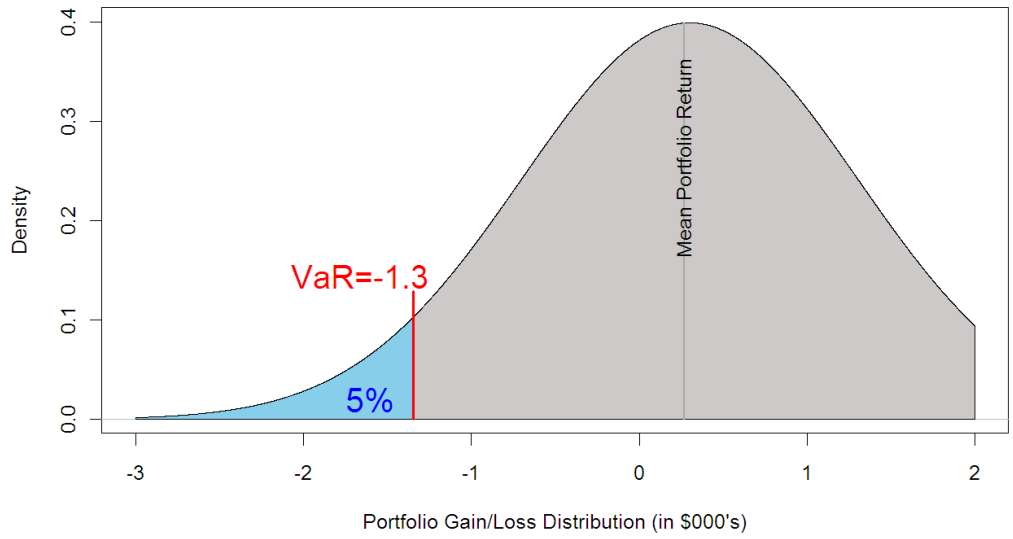


Figure 3.1: Value-at-Risk Visual Demonstration

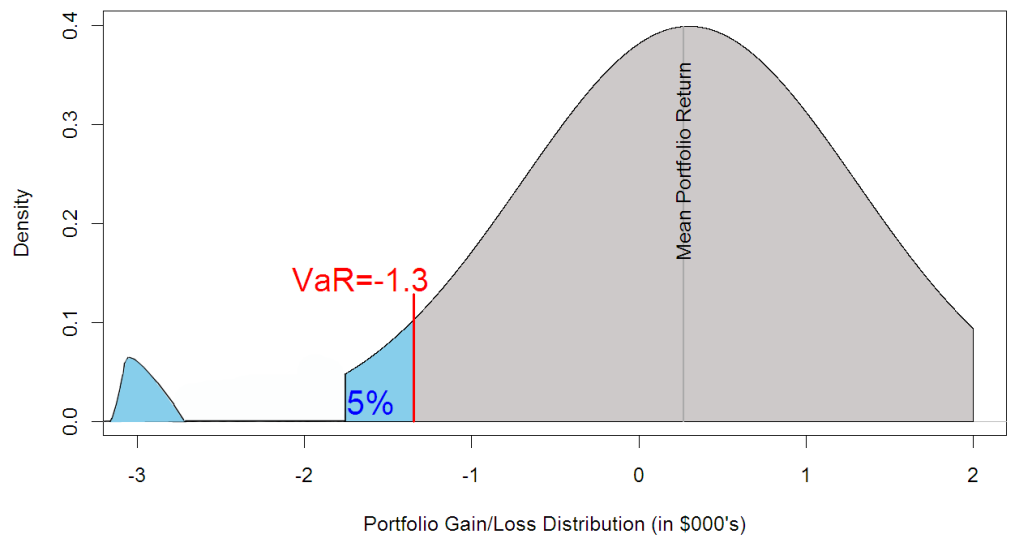


Figure 3.2: Value-at-Risk Extreme Case

over periods of time longer than 1 day. The usual assumption is:

$$\text{N-day VaR} = \text{1-day VaR} \times \sqrt{N}$$

This formula is exactly true when the changes in the value of the portfolio on successive days have independent identical normal distributions with mean zero. In other cases, it is an approximation.

In general, the algorithm for computing VaR is as follows:

1. Mark-to-Market of the current portfolio (i.e., calculate value of portfolio)
2. Measure the variability of the risk factors
3. Set the time horizon
4. Set the confidence level
5. Report the worst loss by processing all the preceding information

In simple terms, VaR can be derived from the probability distribution of the future portfolio value $f(w)$. At a given confidence level c , we want to find the worst possible realization W^* such that the probability of exceeding this value is c :

$$c = \int_{W^*}^{\infty} f(x)dw \quad (3.1)$$

or such that the probability of a value lower than W^* , $p = P(w \leq W^*)$, is $1 - c$:

$$1 - c = \int_{-\infty}^{W^*} f(x)dw = P(x \leq W^*) = p \quad (3.2)$$

In other words, the area from $-\infty$ to W^* must sum to $p = 1 - c$. The number W^* is called the quantile of the distribution, which is the cut-off value with a fixed probability of being exceeded. $(1 - c)$ corresponds to the area shaded in light blue in Figure 3.1. In the following sections we describe three specific methods of calculating VaR.

3.2.1 Parametric - RiskMetrics

VaR computation is simple if the distribution of returns can be assumed to belong to a parametric family such as the normal distribution. In this case, VaR can be derived directly from the portfolio's standard deviation using a multiplicative factor that depends on the confidence level. First, we translate the general distribution $f(w)$ into a standard normal distribution $\Phi(\epsilon)$, where ϵ has mean zero and standard deviation of unity. We associate W^* with the cut-off return R^* such that $W^* = W_0(1 + R^*)$, where W_0 is the value of the portfolio at the time of measurement. Further, we can associate R^* with a standard normal deviate $\alpha \geq 0$ by setting

$$-\alpha = \frac{-|R^*| - \mu}{\sigma} \quad (3.3)$$

This is equivalent to setting

$$1 - c = \int_{-\infty}^{W^*} f(x)dw = \int_{-\infty}^{-|R^*|} f(r)dr = \int_{-\infty}^{-\alpha} \Phi(\epsilon)d\epsilon \quad (3.4)$$

Now we can easily find the solution by using the cumulative standard normal distribution table, which gives us the area to the left of a standard normal variable with value equals to d :

$$N(d) = \int_{-\infty}^d \Phi(\epsilon)d\epsilon \quad (3.5)$$

With this, we can easily see that with parameters μ, σ expressed on daily basis (as what we will have in our example in Chapter 5), VaR can be expressed as follows:

$$VaR(mean) = W_0(R^* - \mu) = W_0\alpha\sigma\sqrt{\Delta t} \quad (3.6)$$

where Δt simply equals to 1 since we our parameters are expressed in terms of days. If they are expressed as annual terms, Δt will be 250 (estimated number of trading days per year).

Further, we can break down this parametric approach into two methods which involves different definitions of our σ term. In the ***Equally Weighted Moving Average*** approach, the calculation of portfolio standard deviation is:

$$\sigma_t = \sqrt{\frac{1}{k-1} \sum_{s=t-k}^{t-1} (x_s - \mu)^2} \quad (3.7)$$

where

- σ_t denotes the estimated standard deviation of the portfolio;
- k denotes the number of days included in “observation period”;
- x_s denotes the change in portfolio value on day s ; and
- μ denotes the mean change in portfolio value.

It is not difficult to see that under this method, each observed data point is given the same weight. If we want to give higher importance to more recent data, we should consider using the ***Exponentially Weighted Moving Average*** approach instead. Under this approach, different weights are attached to past observations, with the most recent observation receiving much higher weights than earlier observations. The expression for the standard deviation is as follows:

$$\sigma_t = \sqrt{(1-\lambda) \sum_{s=t-k}^{t-1} \lambda^{t-s-1} (x_s - \mu)^2} \quad (3.8)$$

where λ denotes the decay factor that determines the rate at which the weights on past observations decay as they become more distant. This approach aims to capture short-term movements in volatility. In our R-application, this can be easily calculated by the `cov.wt()` function.

3.2.2 Historical Simulation

While the parametric approach assumes a normal distribution for the observed data which is almost always not the case, historical simulation method does not

require such an assumption. Under this approach, there is no need to calculate the expected value and standard deviations. Instead, the actual percentiles of the observation period are used as the VaR measures. For example, for an observation period of 1000 days, the 95th percentile historical simulation VaR is the 51th largest loss observed in the sample of 1000 outcomes (because the 5% of the sample that should exceed the risk measure equals 50 losses). This can be very easily implemented in our R-application by using the `quantile()` function. Historical simulation approaches do not make the assumptions of normality or serial independence which can be viewed as a strength of this method. However, relaxing these assumptions also implies that this approach does not easily accommodate translations between multiple percentiles and holding periods. [8]

A variant of this approach is the **Historical Bootstrapping** method, which introduces elements of randomness into the historical method above. Under this approach, we sample from all observations available and create a window (e.g., 250 days) of portfolio values. From this point, VaR is calculated the same way as the historical method by simply obtaining the quantile of this new window of portfolio values. In other words, the bootstrap randomizes the selection of a historical period, rather than using a fix window of data as before. Again, under this method, no distributional assumptions (e.g., normality) are made which means it can include fat tails, jumps, or any departure from the normal distribution. Also, the cross-sectional correlation between different assets in the portfolio is preserved (as we are using the portfolio returns rather than returns of each asset of the portfolio). However, extensive amount of data is needed to get a good approximation. Finally, by resampling at random, any pattern of time variation is broken which means this method relies heavily on the assumption that returns are independent over time. We can clearly realize these shortcomings in this model in our concept demonstration section in Chapter 5.

3.2.3 Monte Carlo Simulation

When we introduce more complicated assets into our portfolio such as options, a more sophisticated method than the parametric approach is needed. Financial instruments such as options do not exhibit normal distribution in their returns, so when introduced into an equity-only portfolio, they change the distribution of the portfolio's return. As a result, tracking error and other mean/variance statistics no longer describe the portfolio's risk. To meet this need, the Monte Carlo Simulation approach is used.

Under this approach, different factor return outcomes are simulated. While the equity part of the portfolio stays the same as parametric, the option returns are computed through the appropriate options pricing model (e.g. Barone-Adesi, Whaley for American options and Black-Scholes for European options). The resulting security returns are aggregated into a single, simulated portfolio return for each outcome in the Monte Carlo Simulation. Such a distribution of portfolio returns more accurately describes a nonlinear portfolio than the normal distribution assumed by the parametric method.

The basic use of this approach in this paper can be summarized as follows:

1. From a data set in a fixed time period, calculate portfolio value $V = f(x_1, x_2, \dots)$, where x_1, x_2, \dots are risk factors that determine the value of the portfolio. For instance, for a two-stock portfolio, the expression will be $V = w_1x_1 + w_2x_2$ where w_i denotes the weights of asset i . If options are included in the portfolio, this function will be non-linear and we will need to use valuation models such as the Black-Scholes formula, where risk factors such as strike price, interest rate, volatility, time to maturity, etc. are used. In this paper, we consider only linear portfolios (stocks).

2. Measure the variance-covariance matrix of all risk factors (the x_i 's in step one) similar to what is done in the parametric approach.
3. Generate as many trials of artificial returns as needed assuming normal distribution with the variance-covariance matrix obtained above.
4. To obtain the above, the Cholesky matrix is used. This can be done conveniently in R using the `mvrnorm()` command in the `MASS` package, with specified number of iterations.
5. For each scenario (iteration), portfolio return/value is calculated using the $f(x_i)$ function specified in step one.
6. Now, an empirical distribution of the portfolio values is generated. Similar to the historical approach, we then measure VaR as the corresponding quantile.

It is not difficult to see that for linear portfolios (with no options) the VaR obtained from this method will be the same as that from the parametric approach. Therefore, this method is of best value when we have a portfolio with mixed classes (with options). This method is often seen as superior since it does not make distributional assumptions on the risk factors (valuation model is used to value the portfolio instead of approximation using normal assumptions). However, it depends on the assumptions made on the valuation model used and thus is prone to model risk. Moreover, it requires high computational power.

In this paper, since we use stock portfolio (which is linear) the superiority of the Monte Carlo simulation method is not obvious. However, we include the use of this method for demonstration purposes. The exact algorithm in our R-application can be found in Chapter 6.

3.3 Model Validation

3.3.1 Back Testing

Model validation is an integral component of a systematic risk-management process. Back testing VaR numbers provides valuable feedback to users about the accuracy of their models. The procedure also can be used to search for possible improvements. Literally, back testing involves testing how well the VaR estimates would have performed in the past.

Suppose that we are calculating a 1-day 95% VaR. Back testing would involve looking at how often the loss in a day exceeded the 1-day 95% VaR that would have been calculated for that day. If this happened on about 5% of the days, we can feel reasonably comfortable with the methodology for calculating VaR. If it happened on, say, 8% of days, the methodology is suspicious. If it happened on 2% of days, we also suspect the method to be overly conservative.

Currently, regulatory guidelines have required banks with substantial trading activity to set aside capital to insure against extreme portfolio losses. This is called the market risk capital requirement. [2] The requirement is expressed as follows:

$$MRC_t = \max \left(VaR_t(0.01), S_t \frac{1}{60} \sum_{i=0}^{59} VaR_{t-1}(0.01) \right) \quad (3.9)$$

where S_t denotes the multiplication factor that is applied to the average of previously reported VaR estimates which varies with back testing results. Specifically, S_t is determined by classifying the number of 1% VaR violations in the previous

250 trading days, N , into three distinct categories as follows:

$$S_t = \begin{cases} 3.0 & \text{if } N \leq 4, \text{ Green} \\ 3 + 0.2(N - 4) & \text{if } 5 \leq N \leq 9, \text{ Yellow} \\ 4.0 & \text{if } 10 \leq N, \text{ Red} \end{cases} \quad (3.10)$$

As we can see, the current regulatory framework for financial institution put VaR and back testing results on very high priority when it comes to capital requirements. It is thus worthwhile for us to demonstrate the ideas using a simple stock portfolio in Chapter 5.

3.3.2 Stress Testing - Scenario Analysis

Stress testing involves estimating how the portfolio would have performed under some of the most extreme market moves seen in the last 10 to 20 years. For example, to test the impact of an extreme movements in the US equity prices, a company might set the percentage changes in all market variables equal to those on October 19, 1987 (S&P moved by 22.3 σ 's). If this is considered to be too extreme, the company might choose January 8, 1988 (S&P moved by 6.8 σ 's). These scenarios can also be artificially generated.

Stress testing can be considered as a way of taking into account extreme events that do occur from time to time but that are virtually impossible according to the probability distributions assumed for market variables. For example, under the assumption of a normal distribution, a 5- σ daily move in a market variable happens about once every 7000 years, but, in practice, it is not uncommon to see such move once or twice every 10 years. [9]

3.4 Critiques on VaR

In his bestsellers *The Black Swan* and *Foiled by Randomness*, Nassim Taleb summarizes the shortcomings of value-at-risk (and portfolio theory) pretty convincingly. He stated that the field of statistics is based upon the law of large numbers, where “as you increase the sample size, no single observation is going to hurt you.” [21] [20] However, in risk management, it is the tail events that are of interest. While VaR gives us the “best of the worst” case scenario based upon some confidence level, there is no indication of how bad it can get below that cut-off point. Also, distributional assumptions are made in the calculation of VaR, and that does not always apply. But when we use historical simulation (thus no distributional assumptions), we are limited to assuming the future resembling the past. Value-at-risk methodology is therefore only part of the full story about the risk of a particular portfolio.

CHAPTER 4

Concept Demonstration - Portfolio Selection

4.1 The Data

Application of the portfolio selection algorithms discussed in Chapter 2 is made simple with the use of computer softwares such as Matlab, R, or Microsoft Excel. For this paper, we use statistical software R. Data needed to compute efficient portfolios are daily/monthly return information of the securities we are interested in, as well as the corresponding market data over the same period. In this paper, we use the 30 stocks currently listed under the Dow Jones Industrial Average as our portfolio securities, and the DJIA index as our market. Names of the DJIA companies are listed in Table 4.1. Such data is openly available at online sources such as www.finance.yahoo.com and www.google.com/finance. In this project, we download daily price information from Yahoo as *.csv files and convert them into monthly returns to proceed with our portfolio analysis.

The convention is to convert daily data into monthly for the purpose of portfolio selection analysis. We can see how the prices and returns of our data look before conversion in Figure 4.1 and Figure 4.2. We observe violent fluctuations on the daily graphs.

After we convert the daily data into monthly, we are ready to employ the portfolio selection algorithms discussed in Chapter 2. We can compare Figure 4.3 and Figure 4.4 to what we had before and we will notice the change in magnitude

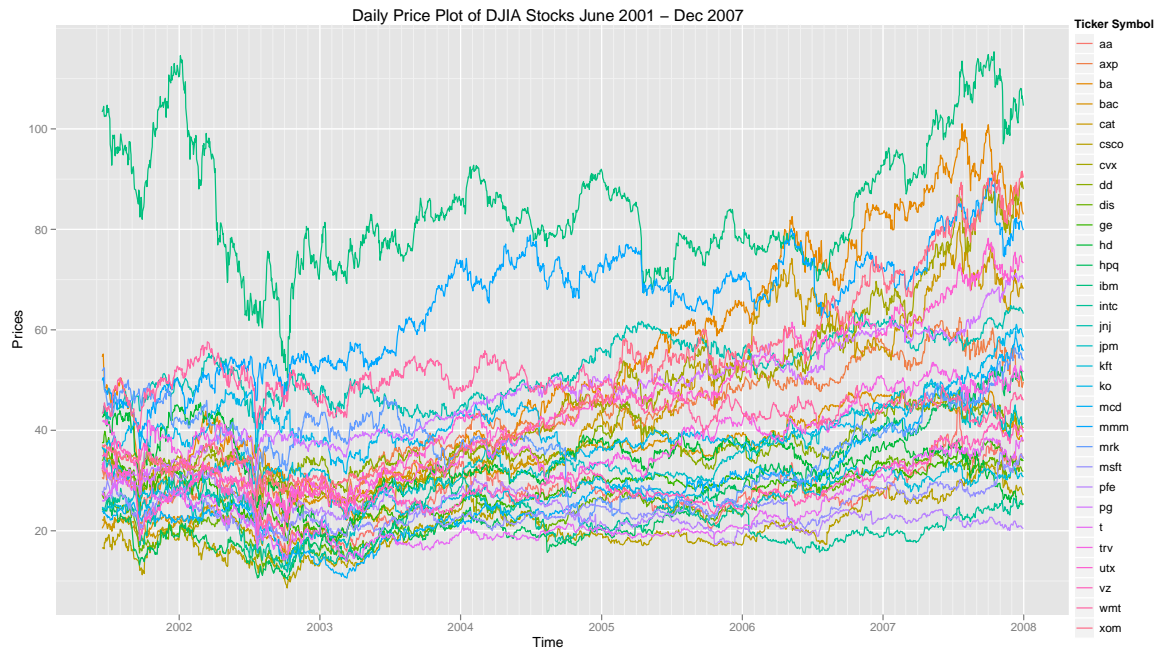


Figure 4.1: Daily Price Plot for DJIA Stocks June 01 - Dec 07

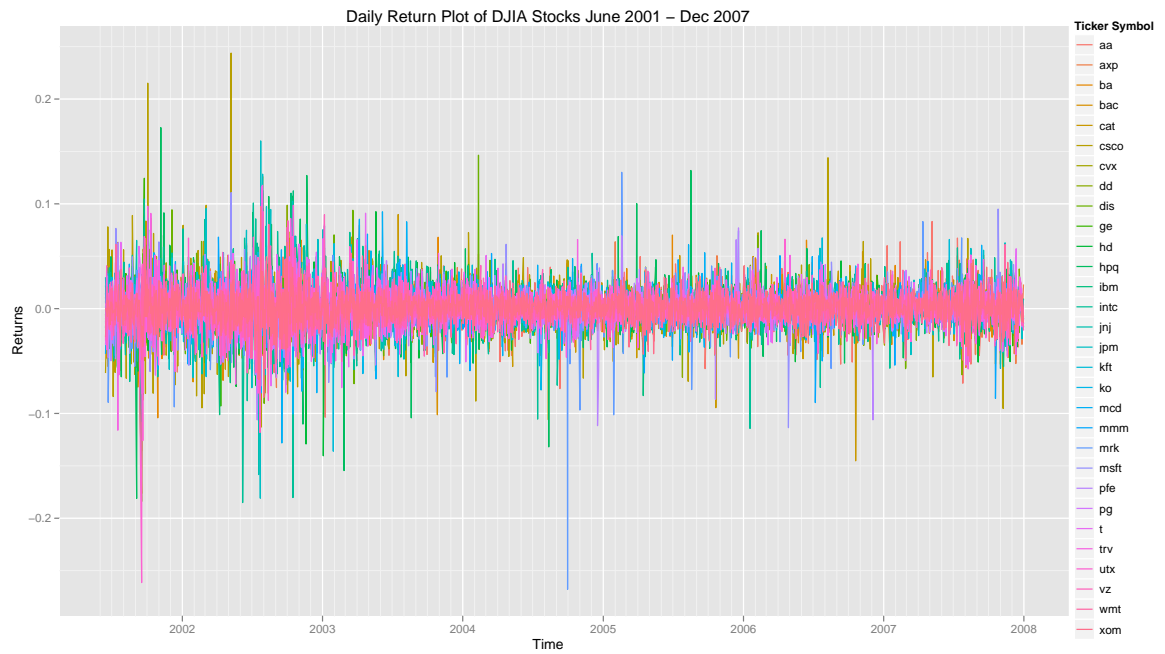


Figure 4.2: Daily Return Plot for DJIA Stocks June 01 - Dec 07

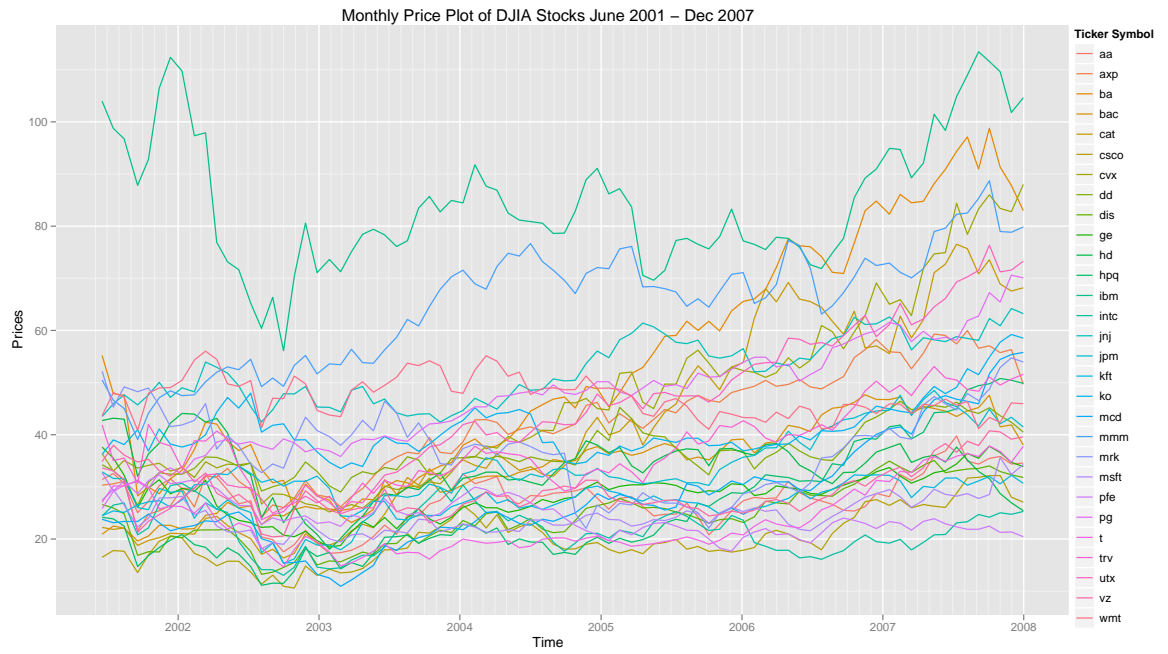


Figure 4.3: Monthly Price Plot for DJIA Stocks June 01 - Dec 07

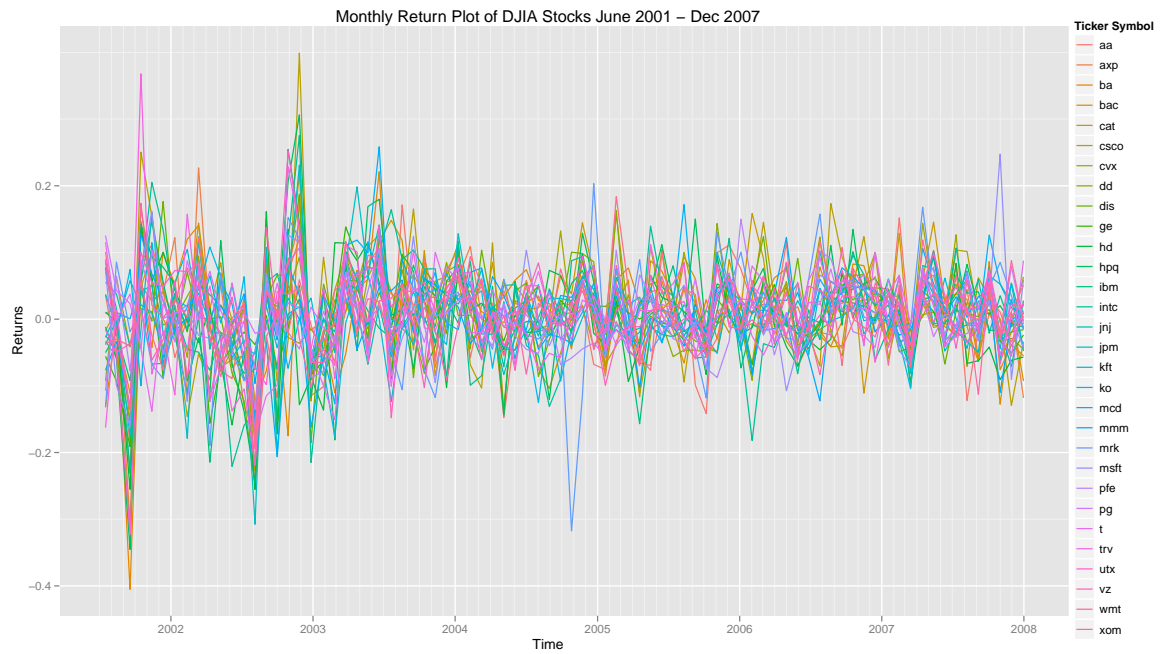


Figure 4.4: Monthly Return Plot for DJIA Stocks June 01 - Dec 07

of fluctuations from daily to monthly.

We hereby include the summary statistics of the monthly return information for each stock in our portfolio in Table 4.1. The original data downloaded from www.finance.yahoo.com was daily data, but we can easily convert them into monthly with the help of statistical software such as R. Please note that the conversion algorithm is not included in this paper.

4.2 Portfolio Selection Results

We include the 30 DJIA stocks in our portfolios and use the 5 methods (Equal Allocation, Single Index Models with and without short sales, and Constant Correlation Models with and without short sales) discussed in Chapter 2 to come up with the optimal weights for each stock. The reason we do not use other methods such as the Multi-Group and Multi-Index is because data required for those methods are not as readily available to download from open sources such as Yahoo and Google. Since we want our R-program to be very scalable, we therefore only choose to use the 5 methods we discussed. Once run through our data management and portfolio selection R-application, we get the portfolio allocation recommendations as shown in Table 4.2.

Using the allocation recommendations obtained, we calculate what would have been the risk and return for each of the portfolios and plot them on the same graph. This way, we can easily see the effect of diversification. See Table 4.3 and Figure 4.5 for the results.

In Figure 4.5, we easily notice the effect of diversification. Here, all the portfolios, regardless of method used, are in the range of relatively lower risk, whereas individual stocks (the black dots) all yield relatively low returns with

| Company | Stock | Min | 1st Quartile | Median | Mean | 3rd Quartile | Max |
|---------------------------|-------|------------|--------------|-----------|-----------|--------------|-----------|
| Alcoa | AA | -0.254299 | -0.042748 | 0.009531 | 0.004391 | 0.069700 | 0.171374 |
| American Express | AXP | -0.304894 | -0.019522 | 0.012208 | 0.008627 | 0.043840 | 0.226855 |
| Boeing | BA | -0.405133 | -0.030016 | 0.022791 | 0.008648 | 0.056331 | 0.221098 |
| Bank of America | BAC | -0.137381 | -0.013164 | 0.011398 | 0.008468 | 0.035512 | 0.116978 |
| Caterpillar | CAT | -0.15145 | -0.04116 | 0.01616 | 0.01646 | 0.06025 | 0.21784 |
| Cisco | CSCO | -0.23116 | -0.05032 | 0.01042 | 0.01065 | 0.05575 | 0.39906 |
| Chevron Corp. | CVX | -0.22721 | -0.02627 | 0.00848 | 0.01268 | 0.05435 | 0.15713 |
| DuPont | DD | -0.155382 | -0.032493 | -0.005052 | 0.003686 | 0.034071 | 0.166352 |
| The Walt Disney Co. | DIS | -0.320853 | -0.027942 | 0.004151 | 0.005385 | 0.046886 | 0.187500 |
| General Electric | GE | -0.1913772 | -0.0340702 | 0.0026978 | 0.0003953 | 0.0350410 | 0.1419256 |
| The Home Depot | HD | -0.255339 | -0.044527 | 0.002392 | -0.002989 | 0.055045 | 0.161264 |
| Hewlett-Packard | HPQ | -0.34535 | -0.02182 | 0.01051 | 0.01343 | 0.05544 | 0.30642 |
| IBM | IBM | -0.214505 | -0.030033 | -0.003263 | 0.002441 | 0.040637 | 0.254500 |
| Intel | INTC | -0.226682 | -0.048394 | 0.006228 | 0.004802 | 0.074917 | 0.275531 |
| Johnson & Johnson | JNJ | -0.081191 | -0.021614 | 0.002174 | 0.005403 | 0.036661 | 0.118623 |
| JPMorgan Chase | JPM | -0.307570 | -0.033201 | 0.003083 | 0.006124 | 0.041057 | 0.231007 |
| Kraft | KFT | -0.176827 | -0.027273 | 0.009926 | 0.003912 | 0.036952 | 0.104689 |
| Coca-Cola | KO | -0.155722 | -0.019319 | 0.007847 | 0.007091 | 0.044338 | 0.115929 |
| McDonald's | MCD | -0.20640 | -0.01947 | 0.01332 | 0.01276 | 0.05247 | 0.25854 |
| 3M | MMM | -0.132297 | -0.019864 | 0.010458 | 0.006893 | 0.039181 | 0.128166 |
| Merck | MRK | -0.317521 | -0.041799 | -0.003134 | 0.003823 | 0.057461 | 0.203267 |
| Microsoft | MSFT | -0.168508 | -0.033058 | 0.006211 | 0.005106 | 0.048024 | 0.247459 |
| Pfizer | PFE | -0.135432 | -0.042851 | -0.020110 | -0.004221 | 0.030991 | 0.149972 |
| Procter & Gamble | PG | -0.06054 | -0.01353 | 0.01100 | 0.01284 | 0.02921 | 0.12529 |
| AT&T | T | -0.174304 | -0.033179 | 0.012807 | 0.006367 | 0.056241 | 0.228513 |
| Travelers | TRV | -0.218853 | -0.039282 | -0.003141 | 0.005504 | 0.049754 | 0.367872 |
| United Technologies Corp. | UTX | -0.322611 | -0.009724 | 0.017743 | 0.011420 | 0.047628 | 0.100141 |
| Verizons Communications | VZ | -0.217910 | -0.033800 | 0.009384 | 0.003697 | 0.034799 | 0.252432 |
| Wal-mart | WMT | -0.179833 | -0.026173 | -0.002478 | 0.002357 | 0.031443 | 0.173722 |
| ExxonMobil | XOM | -0.19478 | -0.01997 | 0.01736 | 0.01249 | 0.04650 | 0.18374 |

Table 4.1: Summary Statistics of DJIA Stocks in Portfolio

| Stock | Eq Alloc. | SIM (SS Std.) | SIM (SS Lint.) | SIM (No SS) | CC (SS Std.) | CC (SS Lint.) | CC (No SS) |
|-------|-----------|---------------|----------------|-------------|--------------|---------------|------------|
| AA | 0.0333 | -0.0755 | -0.0228 | 0 | -0.0744 | -0.0155 | 0 |
| AXP | 0.0333 | 0.0209 | 0.0063 | 0 | 0.0494 | 0.0103 | 0 |
| BA | 0.0333 | 0.022 | 0.0066 | 0 | 0.0176 | 0.0037 | 0 |
| BAC | 0.0333 | 0.1668 | 0.0504 | 0.0047 | 0.214 | 0.0447 | 0 |
| CAT | 0.0333 | 0.1971 | 0.0596 | 0.0639 | 0.2393 | 0.0499 | 0.0885 |
| CSCO | 0.0333 | 0.0233 | 0.007 | 0 | 0.0214 | 0.0045 | 0 |
| CVX | 0.0333 | 0.1751 | 0.0529 | 0.1009 | 0.2538 | 0.053 | 0.0728 |
| DD | 0.0333 | -0.0967 | -0.0292 | 0 | -0.0929 | -0.0194 | 0 |
| DIS | 0.0333 | -0.041 | -0.0124 | 0 | -0.0506 | -0.0106 | 0 |
| GE | 0.0333 | -0.1839 | -0.0556 | 0 | -0.2357 | -0.0492 | 0 |
| HD | 0.0333 | -0.1839 | -0.0556 | 0 | -0.2406 | -0.0502 | 0 |
| HPQ | 0.0333 | 0.06 | 0.0181 | 0 | 0.0741 | 0.0155 | 0 |
| IBM | 0.0333 | -0.0873 | -0.0264 | 0 | -0.1311 | -0.0273 | 0 |
| INTC | 0.0333 | -0.0472 | -0.0143 | 0 | -0.0663 | -0.0138 | 0 |
| JNJ | 0.0333 | 0.0911 | 0.0276 | 0 | 0.0805 | 0.0168 | 0 |
| JPM | 0.0333 | -0.0533 | -0.0161 | 0 | -0.0353 | -0.0074 | 0 |
| KFT | 0.0333 | 0.0472 | 0.0143 | 0.012 | -0.0684 | -0.0143 | 0 |
| KO | 0.0333 | 0.1235 | 0.0373 | 0.0834 | 0.1142 | 0.0238 | 0 |
| MCD | 0.0333 | 0.1364 | 0.0412 | 0.0495 | 0.1858 | 0.0388 | 0.0272 |
| MMM | 0.0333 | 0.0789 | 0.0239 | 0 | 0.0921 | 0.0192 | 0 |
| MRK | 0.0333 | -0.008 | -0.0024 | 0 | -0.0867 | -0.0181 | 0 |
| MSFT | 0.0333 | -0.0223 | -0.0067 | 0 | -0.0514 | -0.0107 | 0 |
| PFE | 0.0333 | -0.2273 | -0.0687 | 0 | -0.463 | -0.0966 | 0 |
| PG | 0.0333 | 0.5668 | 0.1714 | 0.5507 | 1.0032 | 0.2093 | 0.6492 |
| T | 0.0333 | 0.0286 | 0.0087 | 0 | -0.01 | -0.0021 | 0 |
| TRV | 0.0333 | -0.0217 | -0.0066 | 0 | -0.0483 | -0.0101 | 0 |
| UTX | 0.0333 | 0.1936 | 0.0586 | 0.0156 | 0.2325 | 0.0485 | 0.046 |
| VZ | 0.0333 | -0.0429 | -0.013 | 0 | -0.0935 | -0.0195 | 0 |
| WMT | 0.0333 | -0.0625 | -0.0189 | 0 | -0.1479 | -0.0309 | 0 |
| XOM | 0.0333 | 0.2223 | 0.0672 | 0.1194 | 0.3182 | 0.0664 | 0.1163 |

Table 4.2: Portfolio Allocation Summary

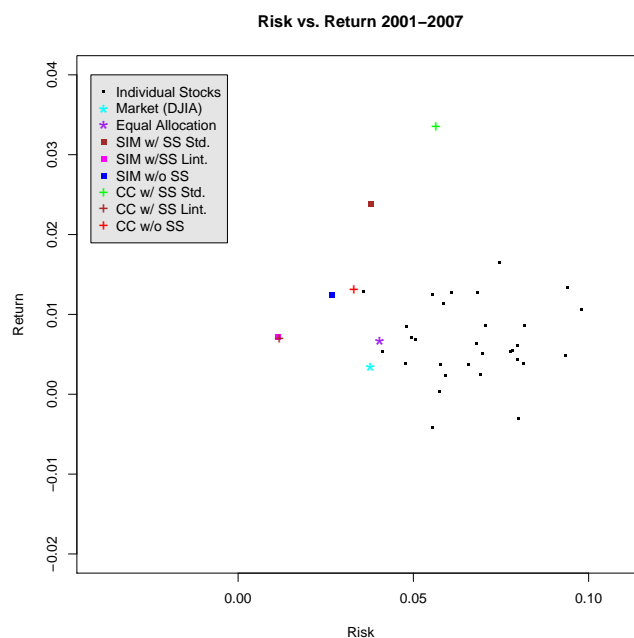


Figure 4.5: Portfolio Risk and Return using data from 2001-2007

| Portfolio | Return | Risk |
|---|--------|--------|
| Market (DJIA) | 0.0034 | 0.0377 |
| Equal Allocation | 0.0066 | 0.0404 |
| Single Index with Short Sales (Standard) | 0.0238 | 0.0379 |
| Single Index with Short Sales (Lintner's) | 0.0072 | 0.0115 |
| Single Index no Short Sales | 0.0124 | 0.0268 |
| Constant Correlation with Short Sales (Standard) | 0.0334 | 0.0564 |
| Constant Correlation with Short Sales (Lintner's) | 0.0070 | 0.0118 |
| Constant Correlation no Short Sales | 0.0130 | 0.0330 |

Table 4.3: Portfolio Risk and Return Snapshot 2001-2007

higher risks. Comparing all portfolios, we see that the Constant Correlation with standard definition of short sales yields the highest return in the period 2001-2007. Following closely is the Single Index Model with standard definition of short sales. This result makes sense since in the standard definition of short sales, as discussed in Chapter 2, we are assumed to have unlimited leverage. And with this unlimited access to capital, it is not a surprise that a higher level of return can be achieved. Back to Figure 4.5, the second set of “winners” are the SIM and CCM without short sales, followed by the models with Lintner’s definition of short sales. Finally, the Equal Allocation portfolio performed the worst with high risk and low returns. However, it is worthwhile to note that all of these portfolios performed better than the market (DJIA) during this period. This is a reassuring observation. In other words, these portfolios are able to “beat the market” in the 2001-2007 observation period. We now proceed with testing these portfolios in years 2008-2009.

4.3 Portfolio Performance Review

With the allocation and corresponding risk and returns, we now compare performance on returns of each portfolio using 2008-2009 data. See Table 4.4. Here we see that the overall market and most of all portfolios, as well as a lot of the stocks in the DJIA index, yielded negative returns (loss) during the period. This is a solid demonstration of undiversifiable risk. All of our portfolios generate losses with the exceptions of the 4 portfolios with short sales which give slightly positive returns. We can see visually how the portfolios performed based upon their relative risk and return on the same plot in Figure 4.6. Once again, the two winners in this picture are the portfolios with standard definition of short sales, meaning unlimited leverage. Following are the two portfolios with Lintner’s

definition of short sales which is slightly more restrictive than the previous two, but we can see clearly that they involve a much lower level of risk.

Finally, we plot a time series of return vs. time for our testing period and the result can be seen in Figure 4.7. In this picture, we can see that all portfolios move along the direction of the market, again another strong evidence of undiversifiable systemic risk. The reason we focus on returns in this exercise is because that is the tangible quantity investors often use to evaluate performance of an investment. On the other hand, if we were holding options which gets priced by σ , it is very important to look at also (or only) the change in dispersion. In this exercise, however, since we are demonstrating portfolio selection concept by a simple stock portfolio, we focus on returns.

It is important to take a step back and recall the purpose of this exercise. One of the main reasons for investors to employ statistical theories on financial data is to try to come up with tools that can help maintain and grow wealth. One way to assess whether the portfolio selection algorithm and allocations work to achieve our financial goal is to compare the portfolio returns against the market, or our initial investment goals. In Table 4.4, we notice that the market and all portfolio generate negative returns. Among all of the portfolios, the portfolios that involved short sales were able to “beat the market”. By looking at Figure 4.6, we realize the effect of diversification: most of all the individual stocks all generated losses on average during 2008-2009, with some that performed way below -2%. With diversification, we can see clearly that all of the portfolios perform in the higher range of the spectrum. 2008-2009 is a time of financial turmoil, and it is reasonable that losses are generated due to “undiversifiable risk”, or “systemic risk”. “None of the portfolios perform poorly in the year 2008-2009” - this is a subjective and qualitative statement. Therefore, whether

these tools are useful depends heavily on the investor's personal preference and tolerance in risk and return.

| Portfolio | Return | Risk |
|---|---------|--------|
| Market (DJIA) | -0.0106 | 0.0760 |
| Equal Allocation | -0.0016 | 0.0858 |
| Single Index with Short Sales (Standard) | 0.0024 | 0.0403 |
| Single Index with Short Sales (Lintner's) | 0.0007 | 0.0122 |
| Single Index no Short Sales | -0.0031 | 0.0379 |
| Constant Correlation with Short Sales (Standard) | 0.0033 | 0.0486 |
| Constant Correlation with Short Sales (Lintner's) | 0.0007 | 0.0101 |
| Constant Correlation no Short Sales | -0.0037 | 0.0375 |

Table 4.4: Portfolio Risk and Return Snapshot 2008-2009

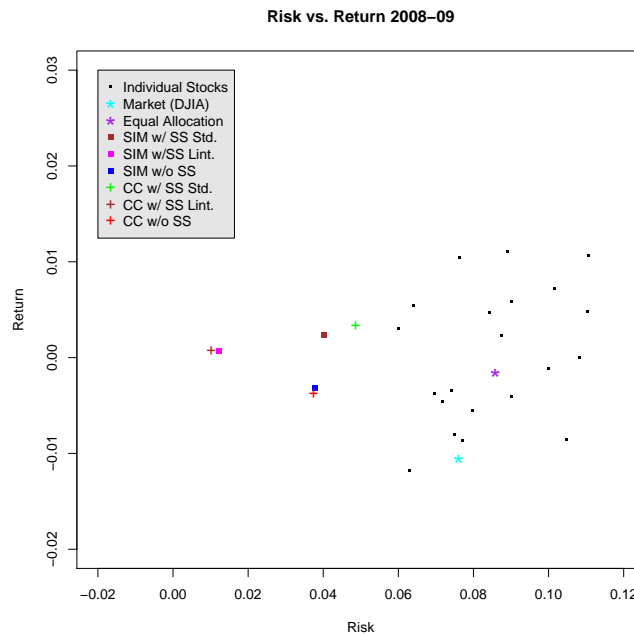


Figure 4.6: Portfolio Risk and Return using data from 2008-2009

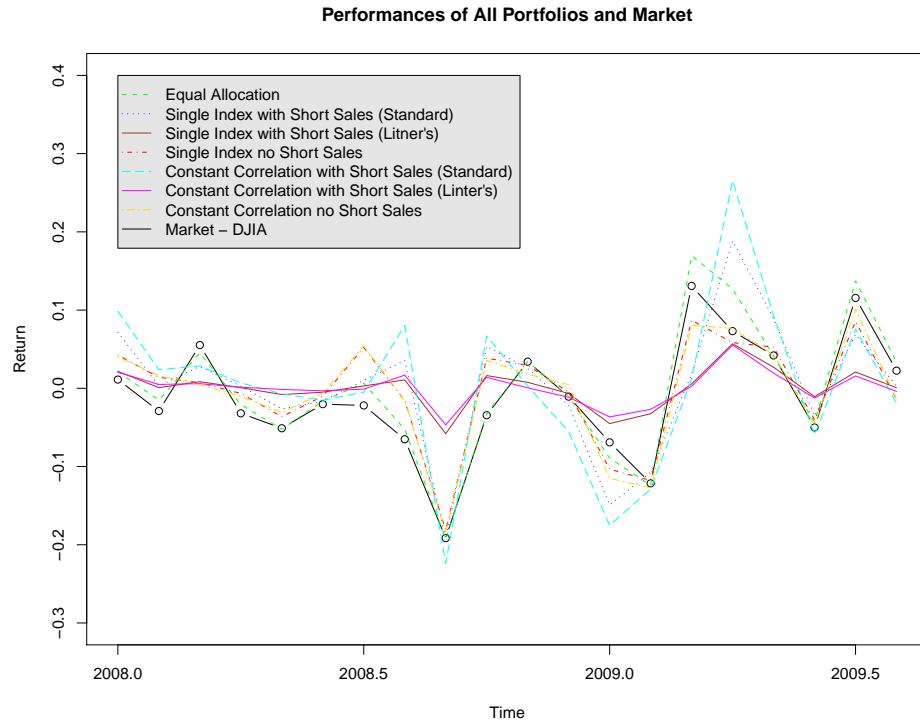


Figure 4.7: Portfolio Performance in 2008-2009

4.4 Some Critiques

The critiques on Modern Portfolio Theory mentioned in Section 2.3 also apply here. Most notably is the use of historical data in constructing future investment portfolios of being backward-looking. Indeed, there is no guarantee that the future is going to resemble the past. Therefore, the portfolio selection methods discussed in this chapter should be viewed as suggestions only and should be considered and used with due care. Investors should be equipped with also fundamental knowledge in finance in order to make sound investment decisions.

CHAPTER 5

Concept Demonstration - VaR Management

As discussed in Chapter 3, Value-at-Risk is a risk management measure used in banks for their complex portfolios that include diverse classes of financial instruments. In this chapter, we demonstrate how the different methods are used in managing risk of our linearly dependent stock portfolio.

5.1 The Data

We use the same data set (2001-2007) as our Chapter 4 demonstration so to be consistent. Here, we use daily returns instead of monthly since we are interested in 1-day value-at-risk (VaR). We already saw the daily price and return movements of the 30 DJIA stocks in Figure 4.1 and 4.2 in the previous chapter. Summary statistics of the daily returns of our demonstrated portfolios are listed in Table 5.1. We also show distribution patterns of the different portfolios in Figure 5.1 to 5.7.

5.2 Value-at-Risk Calculations

From the histograms, we feel confident making the claim that distributions of all our portfolio returns revolve around the mean, which for all cases are somewhere close to zero, as we can see from the “bell shape” of the histograms. However, the overlay of the normal curve (in blue) shows that there might be a problem of the

| Portfolio | Min | 1st Quartile | Median | Mean | 3rd Quartile | Max |
|------------------------------|------------|--------------|-----------|-----------|--------------|-----------|
| Equal Alloc. | -0.0651265 | -0.0050814 | 0.0005460 | 0.0003535 | 0.0056717 | 0.0597972 |
| Single Index (SS Standard) | -0.082545 | -0.005958 | 0.001486 | 0.001182 | 0.008131 | 0.039522 |
| Single Index (SS Lintner's) | -0.0249610 | -0.0018016 | 0.0004494 | 0.0003576 | 0.0024589 | 0.0119512 |
| Single Index (No SS) | -0.0636476 | -0.0043466 | 0.0005896 | 0.0006180 | 0.0058793 | 0.0491483 |
| Constant Corr.(SS Standard) | -0.109954 | -0.008050 | 0.002214 | 0.001650 | 0.011207 | 0.061177 |
| Constant Corr.(SS Lintner's) | -0.0229442 | -0.0016798 | 0.0004620 | 0.0003443 | 0.0023386 | 0.0127659 |
| Constant Corr. (No SS) | -0.0653077 | -0.0044116 | 0.0006845 | 0.0006533 | 0.0059953 | 0.0496990 |

Table 5.1: Summary Statistics of Portfolio Returns

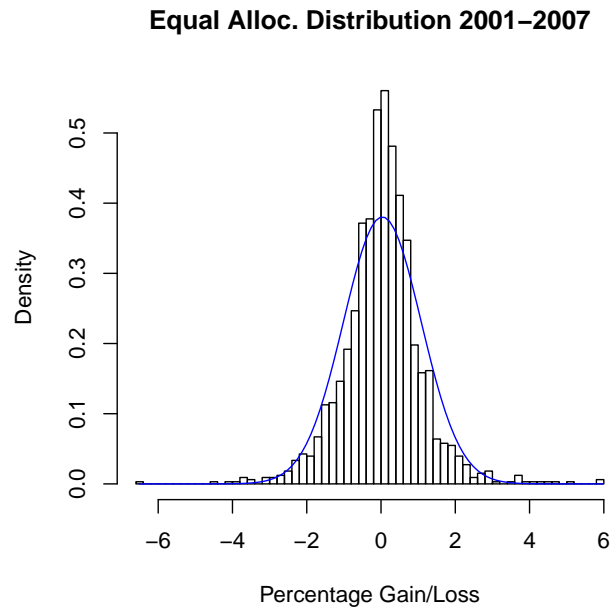


Figure 5.1: Equal Allocation Portfolio Return Distribution

Single Index (SS Std.) Distribution 2001–2007

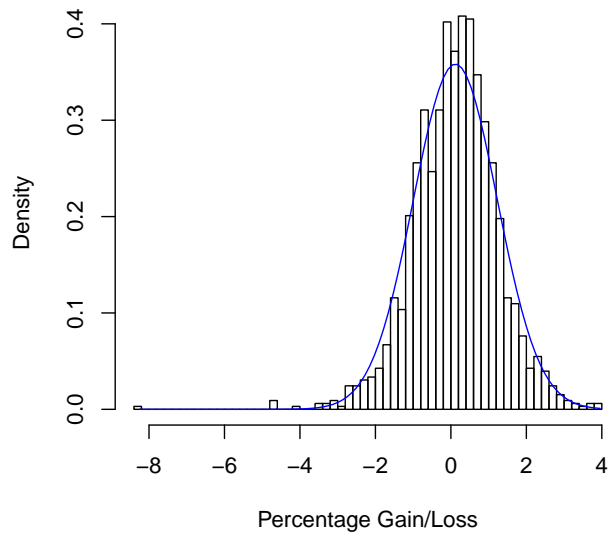


Figure 5.2: Single Index Model (Standard Short Sales) Distribution

Single Index (SS Lint.) Distribution 2001–2007

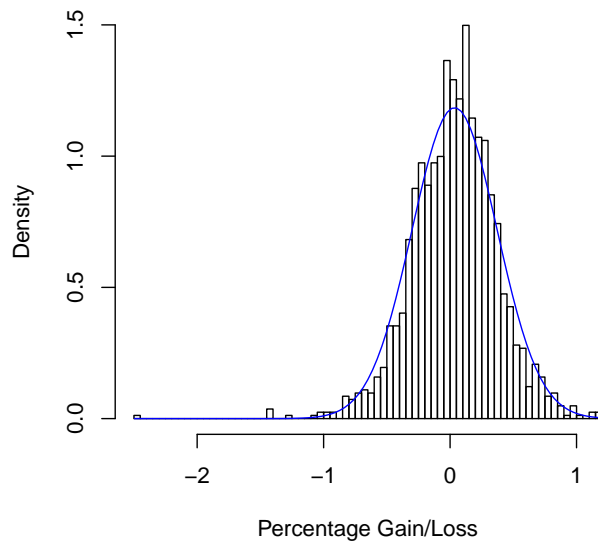


Figure 5.3: Single Index Model (Lintner's Short Sales) Distribution

Single Index (No SS) Distribution 2001–2007

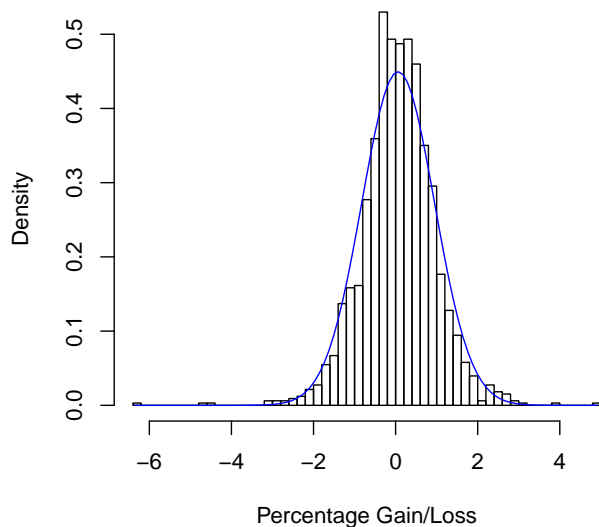


Figure 5.4: Single Index Model (No Short Sales) Distribution

Constant Corr. (SS Std.) Distribution 2001–2007

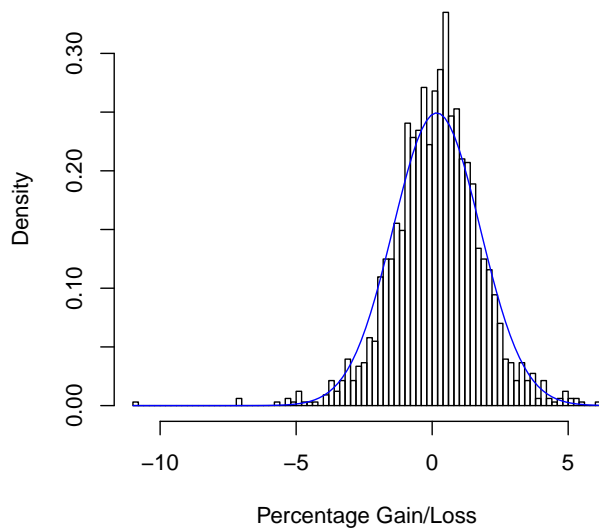


Figure 5.5: Constant Correlation (Standard Short Sales) Distribution

Constant Corr. (SS Lint.) Distribution 2001–2007

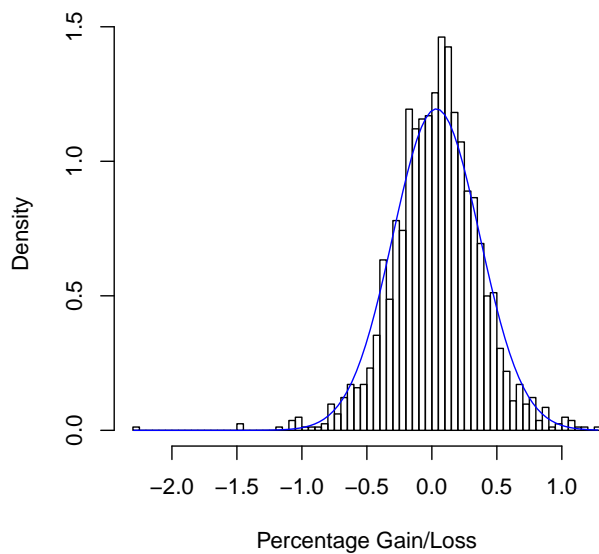


Figure 5.6: Constant Correlation (Lintner's Short Sales) Distribution

Constant Corr. (No SS) Distribution 2001–2007

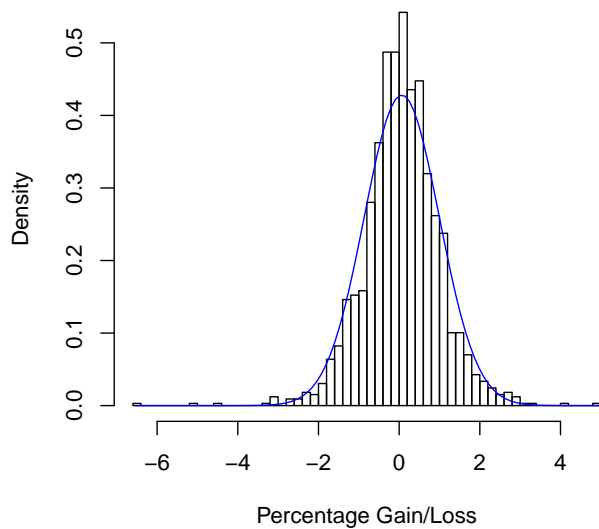


Figure 5.7: Constant Correlation (No Short Sales) Distribution

assumptions of normality (which is used in parametric method as well as Monte Carlo). This is one of the major critiques of VaR methods as a risk measure. Also, we can see the problem of “fat tail” in almost all of our portfolios. Since the method for calculating VaR using all approaches do not change regardless of which portfolio we use, for demonstration purposes, therefore, we will choose the Equal Allocation Portfolio as our portfolio.

We calculate next day VaR’s for all data points we have using the Parametric, Historical, and Monte Carlo methods discussed in Chapter 3. Hence we will have a long series of VaR’s values from the 7 years worth of daily returns data. The reason we want that many VaR’s is because we want to validate the coverage by each of the models and to find out which model yields the best results.

Notice that since our data starts from June 2001 to Dec 2007, the most recent “Next-Day VaR” means the one-day VaR for the trading day after our last data point, December 31, 2007. To get a series of VaR’s retrospectively, we simply calculate next-day VaR using previous data under different specified conditions. For example, to calculate the VaR of 12/30/2007, we use data from the previous 250 trading days prior to 12/30/2007 to come up with the VaR. And then we proceed to calculating the VaR for 12/29/2007 in a similar fashion. We keep doing this until the amount of data get exhausted. For starter, we can see a summary of next-day VaR (Jan 2, 2008) using the 3 methods (total of 5 approaches) discussed, all with a one-year data length window under 95% confidence. Refer to Table 5.2. For example, looking at the Parametric (Equally Weighted) value, we make the claim that we are 95% confident that the loss encountered by this portfolio tomorrow (January 2, 2008) will not exceed 1.54%.

| | Next-Day VaR (1/2/2008) |
|----------------------|-------------------------|
| Parametric Eq. | -0.0154 |
| Parametric Ex. | -0.0190 |
| Historical | -0.0175 |
| Historical Bootstrap | -0.0153 |
| Monte Carlo | -0.0151 |

Table 5.2: Next-Day VaR: 95% C.I., 1-Year Window

5.3 VaR Model Performance Review using Back Testing

Recall that by calculating next-day VaR we are interested in making the following statement:

We are $C\%$ certain that we will not lose more than $\$V$ in the next X day.

By now we have 3 methods (5 approaches total) of calculating this at-risk amount, but which method should be our ultimate choice? In the following sections we will review the results of these methods using our equal allocation portfolio using back testing. In simple terms, we calculate a series of next-day VaR and compare these calculated numbers with what the actual loss was in history. With this comparison, we are able to back test whether we indeed have only 5% of the times that the predicted VaR's (losses) are of greater magnitude than the actual loss. We use this failure rate to assess which model is the "best".

In Table 5.3 we have a summary of all of our models' back test results under a 95% confidence level over different data window lengths. We use our R-application (details in Chapter 6) to calculate these series of VaR's during year 2001-2007 using the above criteria, and compared them with what the actual

gains/losses were in history. With such comparisons, we come up with the back test summary table. The percentage amounts shown on Table 5.3 represent the percentage of calculated VaR that underestimated the actual loss (actual loss greater than what as “predicted”).

Skimming briefly, all the failure rates are within reasonable range which assures us that the models are reasonably constructed. Next, we proceed to the determination of which model yields the most ideal results. First of all, we need to take a step back and recall the purpose of back testing. By back testing, we are trying to assess which model we can trust when coming up with a next-day value-at-risk amount of our particular portfolio. We want something that will give us an amount that is 95% of the time correct. In other words, we want the model to give us a failure rate no more than 5%, but at the same time not extremely below the threshold either, since that will mean that the model is overly conservative and that resources might be wasted for unnecessarily high capital requirement (recall current regulatory framework discussed in Chapter 3). Having this in mind, we can reasonably say that the historical bootstrapping method is overly conservative. In Figure 5.11, we see an interesting pattern when compared to the graphs of all other methods. As discussed in Chapter 3, the random movements we observe here is related to the breakdown of time variation, which means total return independence are assumed. In investment language, we can see it as a breakdown of “time memory” of the trend. On the other hand, going back to Table 5.3, the historical simulation method yields failure rates that are all above the 5% threshold, so we can conclude that this method is too lax for our purpose. Similarly, the Monte Carlo method also gives us coverage rate that are unsatisfactory (all above 5%). We are then left with the parametric models. Recall that we want failure rates that are close to but not over 5%, with this criteria, the “winner” in this model competition is parametric with equally

| Window | Parametric (Eq) | Historical (Eq) | Historical (BS) | Monte Carlo | Lambda | Parametric (Exp.) |
|--------|-----------------|-----------------|-----------------|-------------|-------------|-------------------|
| 12-mo | 0.0482 | 0.0561 | 0.0331 | 0.0568 | 0.97 | 0.0510 |
| 6-mo | 0.0489 | 0.0568 | 0.0309 | 0.0554 | 0.94 | 0.0482 |
| 3-mo | 0.0510 | 0.0633 | 0.0331 | 0.0575 | 0.90 | 0.0510 |
| 1-mo | 0.0518 | 0.0891 | 0.0618 | 0.0640 | 0.80 | 0.0611 |

Table 5.3: Back Testing Results for VaR Models - Failure Rates

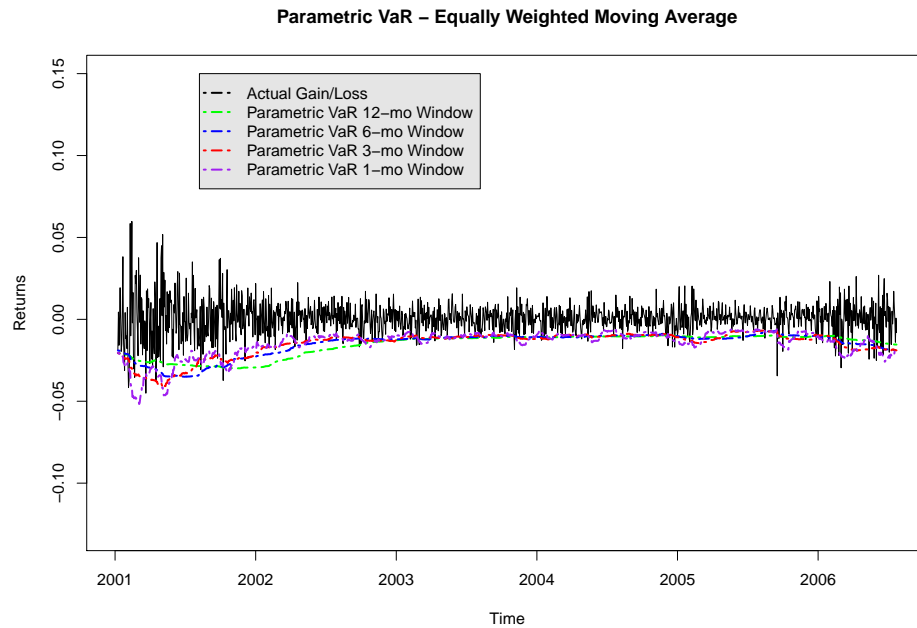


Figure 5.8: Back Testing: Parametric VaR Equally Weighted

weighted moving average using a 6-month data window.

Of course, this is a very rudimentary exercise of demonstrating how each method is used and validated. More data and portfolio varieties will be needed if we were doing this exercise for a real financial institution with complex investment portfolios. And with more complicated portfolios that consist of complex financial instruments, we would definitely be able to see the advantage of using the Monte Carlo method over the linear parametric method.

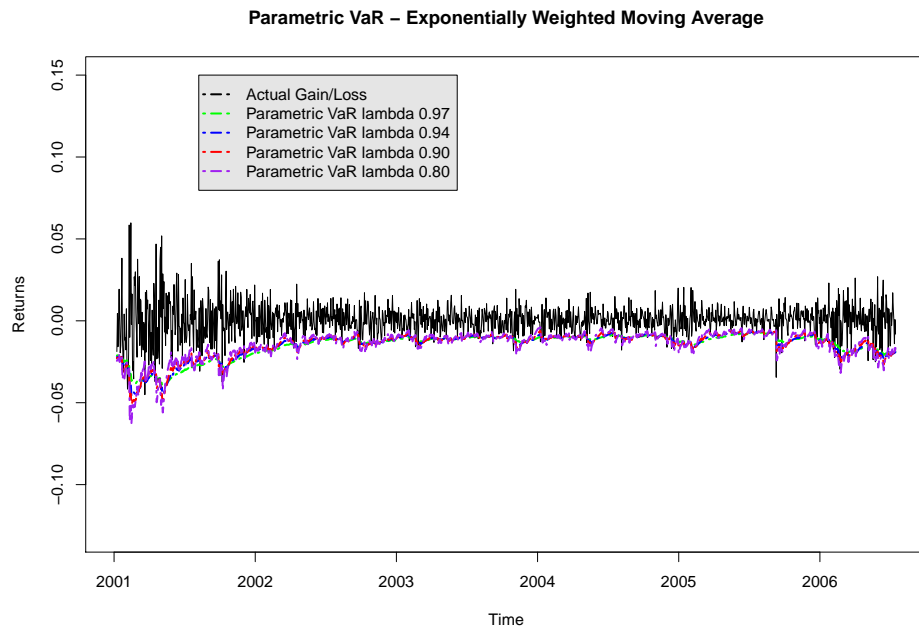


Figure 5.9: Back Testing: Parametric VaR Exponentially Weighted

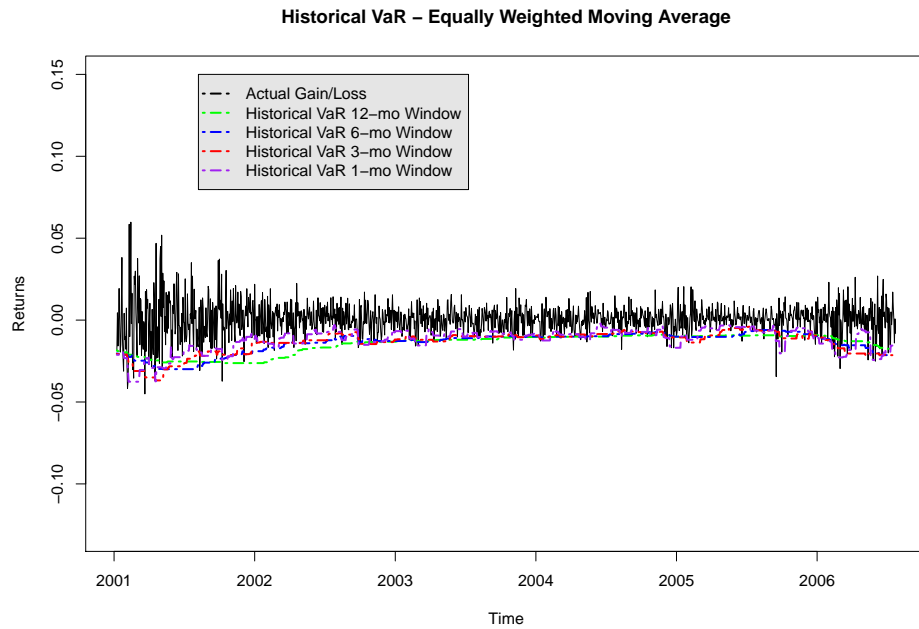


Figure 5.10: Back Testing: Historical VaR Equally Weighted

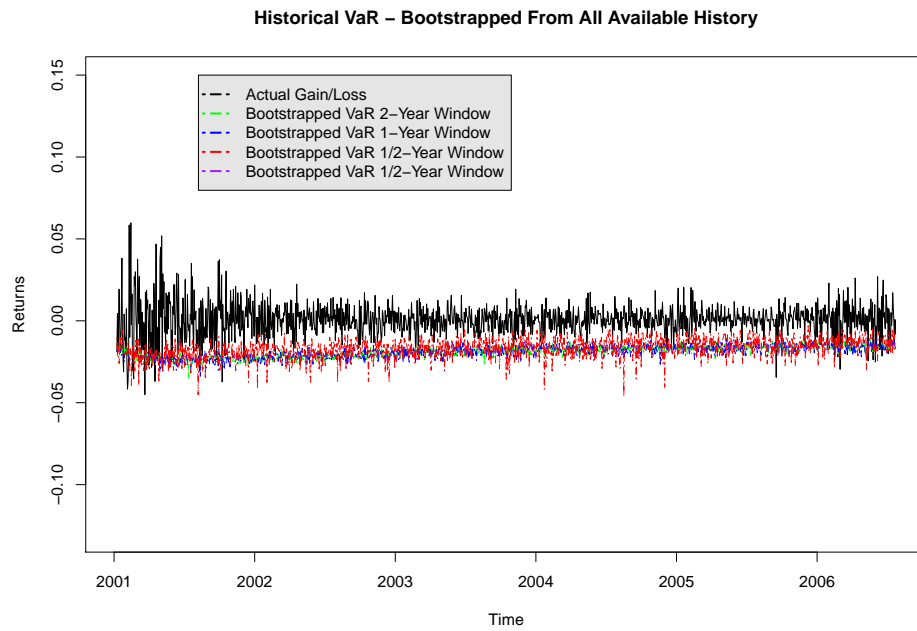


Figure 5.11: Back Testing: Historical VaR Bootstrapping

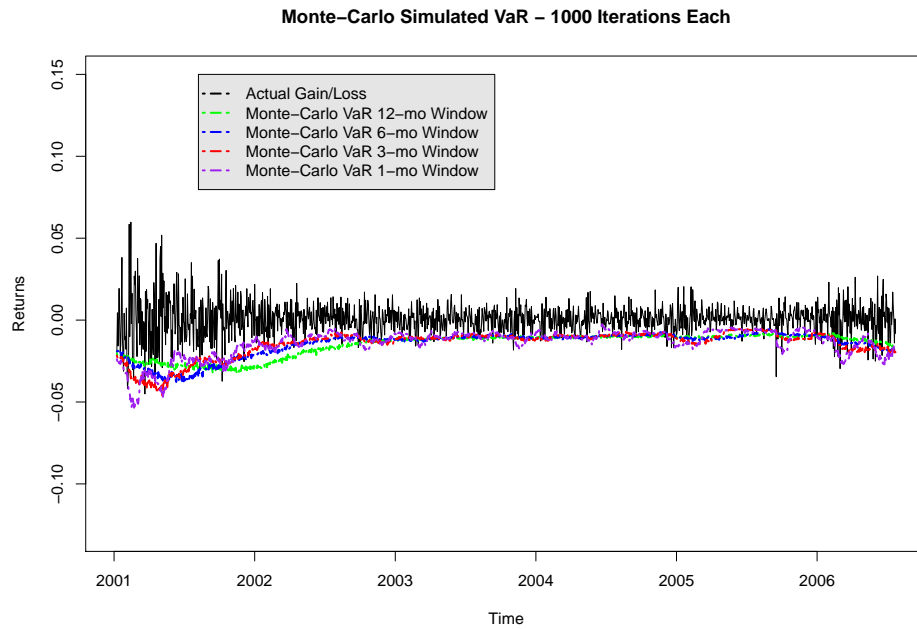


Figure 5.12: Back Testing: Monte Carlo Simulated VaR

5.4 Some Critiques

The criticisms made by Nassim Taleb in his international best sellers as mentioned in Section 3.4 also apply here. In our exercise, we found that using the parametric method with equally weighted moving average is the most reasonable way to calculate VaR's for our stock portfolio. With that, we conclude that we are 95% confident that the next-day loss (the trading day following 12/31/2007) will be no worse than -1.54%. However, what happens to the remaining 5%? For this remaining 5%, we can experience losses of -2%, or -20%, or even -200% and below! In risk management, we are interested in knowing and preparing for the worst possible scenario, and it is these tail events that we should pay attention to. In other words, value-at-risk method gives us only the beginning of the story. The detail and possible ending of this story will need to be further explored. Nevertheless, value-at-risk is a value-adding tool in terms of making investors or portfolio managers aware of possible danger, and it is certainly playing an important role in risk management for financial institutions today.

CHAPTER 6

The R-Application

6.1 The Motivation

One main goal of this paper is to develop a user-friendly tool for beginning investors or students in statistical finance to apply the ideas discussed in the previous sections onto real life data. This is made possible by statistical software R. We hereby include all the R-functions used to calculate the related results shown in Chapter 4 and 5. Soft copies of these codes are posted on <http://theses.stat.ucla.edu/>. They can be downloaded and run on your local machine so that anyone can start constructing and analyzing their stock portfolios in R once data are downloaded and working directory properly set.

6.2 The Source Codes

Here we include source codes of each of the functions used in this paper so that one can see clearly how each function is structured and notice their similarities and differences. Users can also extend and modify these codes to suit their specific needs.

6.2.1 Data Management

To use this function, download daily/monthly stock price data from Yahoo Finance and save them into your working directory as *.csv (where * is the name of the file). This function can easily merge data into usable data frames to be called out by R. The command for this function is `DataMerge()` and once run, the function will spit out a `total.Rdata` file and a `total.csv` file in the working directory you previously specified. Make sure to download both Stock data and Index data (if Single Index Model is to be used. See next section). For example, in our case, we type:

```
DataMerge(num=30,idx="dji",stk=c("aa","axp","ba","bac","cat","csc","cvx",
"dd","dis","ge","hd","hpg","ibm","intc","jnj","jpm","kft","ko","mcd","mmm",
"mrk","msft","pfe","pg","t","trv","utx","vz","wmt","xom"))
```

where `num` is the number of stocks in the portfolio, `idx` is the name of the index we are using, and `stk` is the character string of the names of the stock *.csv files downloaded from Yahoo. Source codes as follows:

```
DataMerge <- function(num=num1, idx=idx1, stk=stk1){
  if(class(idx)!="character") stop("Input names of stocks as characters")
  if(class(stk)!="character") stop("Input names of index as characters")
  a <- read.csv(paste(idx, ".csv", sep=""), header = TRUE)
  ddot <- as.Date(a$Date)
  pdot <- a$Adj.Close
  rdot <- (a$Adj.Close[-length(a$Adj.Close)]-a$Adj.Close[-1])/a$Adj.Close[-1]
  adot <- data.frame(ddot[-length(a$Date)], pdot[-length(a$Date)], rdot)
  names(adot) <- c("Date", paste("p", idx, sep=""), paste("r", idx, sep=""))
  total <- adot
  for(k in 1:num){
    a <- read.csv(paste(stk[k], ".csv", sep=""), header = TRUE)
    d <- as.Date(a$Date)
    r <- (a$Adj.Close[-length(a$Adj.Close)]-a$Adj.Close[-1])/a$Adj.Close[-1]
    p <- a$Adj.Close
    a <- data.frame(d[-length(a$Date)], p[-length(a$Date)], r)
    names(a) <- c("Date", paste("p", k, sep=""), paste("r", k, sep=""))
    total <- merge(total, a, all=FALSE, sort=FALSE)
  }
  write.csv(total, "new_total.csv")
  save(total, file="new_total.RData")
}
```

6.2.2 Portfolio: Equal Allocation

Suppose that we have merged our data using the `DataMerge()` discussed above where the downloaded data are daily prices. Then in order to use the portfolio selection functions, based upon convention, we will need to convert the data from daily into monthly returns (algorithm not included in this paper). Or, to avoid doing that, we can download monthly data at the beginning and run `DataMerge()` directly on monthly data. Once the data frame is merged and set up, we can run the portfolio functions which will give us with the optimal weights of each stock to be allocated, as well as the expected return and risk of this portfolio based upon historical data supplied into the function. For example, in our paper, we type the following:

```
port.eq(data=ret.mon)
```

where `ret.mon` is the name of my data frame or matrix with monthly returns of each stock in each column. Source codes as follows:

```
port.eq <- function(data=ret.mon){
  mean.ret <- apply(data, 2, mean)
  vcov.ret <- cov(data)
  num.stock <- ncol(data)
  e.weight <- (1/num.stock)*rep(1,num.stock)
  ret.equal <- t(mean.ret) %*% e.weight
  var.equal <- t(e.weight) %*% vcov.ret %*% e.weight
  return(list(stock=names(data),weights=e.weight,
             return=ret.equal,risk=var.equal^.5))
}
```

6.2.3 Portfolio: Single Index (With Short Sales)

Once we have a return data frame or matrix set up, similar to the Equal Allocation Portfolio case above, we can easily calculate our Single Index Model optimal portfolio weights, return, and risk using the `port.sim.ss()` function if short sales are allowed. For example, in our paper, we type the following:

```
port.sim.ss(data=ret.mon,idx=dji.mon,rf=0.001,lint=TRUE)
```

where `ret.mon` is the name of my data frame or matrix with monthly returns of each stock in each column, `dji.mon` is the market return data (in this case we use the DJIA index as our market), `rf` is the risk free rate (in this case we use the 3 year treasury bill rate published on www.finance.yahoo.com), and `lint` is the logical argument of whether we want to apply the Lintner's definition of short sales. Source codes as follows:

```
port.sim.ss <- function(data=ret.mon, idx=dji.mon, rf=0.001, lint=FALSE){
  if(length(data[,1])!=length(idx[,1]))
    stop("The number of Portfolio and Index returns must equal.")
  num.stock <- ncol(data)
  mean.ret <- apply(data, 2, mean)
  mean.idx <- apply(idx, 2, mean)
  var.idx <- var(idx)
  vcov.ret <- cov(data)
  sigma.ret <- rep(0,num.stock)
  for (i in 1:ncol(vcov.ret)){
    sigma.ret[i]<-sqrt(vcov.ret[i,i])
  }
  a <- rep(0,num.stock)
  b <- rep(0,num.stock)
  for(i in 1:num.stock){
    q <- lm(data[,i]~idx)
    a[i]<-q$coef[1]
    b[i]<-q$coef[2]
  }
  stock <- rep(0,num.stock); mse <- rep(0,num.stock); Rbar <- rep(0,num.stock)
  Ratio <- rep(0,num.stock); col1 <- rep(0,num.stock); col2 <- rep(0,num.stock)
  col3 <- rep(0,num.stock); col4 <- rep(0,num.stock); col5 <- rep(0,num.stock)
  for(i in 1:num.stock){
    Rbar[i] <- a[i]+b[i]*mean.idx
    mse[i] <- sum(lm(data[,i]~idx)$residuals^2)/(nrow(data)-2)
    Ratio[i] <- (Rbar[i]-rf)/b[i]
    stock[i] <- i
  }
  xx <- (cbind(stock, a, b, Rbar, mse, Ratio))
  aaa <- xx[order(-Ratio),]
  col1 <- (aaa[,4]-rf)*aaa[,3]/aaa[,5]
  col3 <- aaa[,3]^2/aaa[,5]
  for(i in(1:num.stock)) {
    col2[i] <- sum(col1[1:i])
    col4[i] <- sum(col3[1:i])
  }
}
```

```

    }
  for(i in (1:num.stock)) {
    col5[i] <- var.idx*col2[i]/(1+var.idx*col4[i])
    }
  z.short <- (aaa[,3]/aaa[,5])*(aaa[,6]-col5[length(col5)])
  lint <- lint
  ifelse(lint, x.short <- (z.short)/sum(abs(z.short)), x.short <- z.short/sum(z.short))
  a.single.short <- t(aaa[,2]) %*% x.short
  b.single.short <- t(aaa[,3]) %*% x.short
  ret.single.short <- a.single.short + b.single.short * mean.idx
  sd.single.short <- sqrt(b.single.short^2 * var.idx +(t(x.short^2)) %*% aaa[,5])
  table <- cbind(aaa,z.short,x.short)
  return(list(table=table,stock=names(data[(aaa[,1])]),
    weights=x.short,return=ret.single.short,risk=sd.single.short))
}

```

6.2.4 Portfolio: Single Index (No Short Sales)

The Single Index Model without short sales works the same way as the previous function where short sales are allowed. The only difference is that it spits out the weights assuming that there no borrowing/lending opportunities available, and so the Lintner's definition logical argument in no longer needed. In our case, we input our criteria as follows:

```
port.sim.ss(data=ret.mon,idx=dji.mon,rf=0.001)
```

Source codes as follows:

```

port.sim.nss <- function(data=ret.mon, idx=dji.mon, rf=0.001){
  if(length(data[,1])!=length(idx[,1]))
    stop("The number of Portfolio and Index returns must equal.")
  num.stock <- ncol(data)
  mean.ret <- apply(data, 2, mean)
  mean.idx <- apply(idx, 2, mean)
  var.idx <- var(idx)
  vcov.ret <- cov(data)
  sigma.ret <- rep(0,num.stock)
  for (i in 1:ncol(vcov.ret)){
    sigma.ret[i]<-sqrt(vcov.ret[i,i])
  }
  a <- rep(0,num.stock)
  b <- rep(0,num.stock)
}

```

```

for (i in 1:num.stock){
  q <- lm(data[,i]~idx)
  a[i]<-q$coef[1]
  b[i]<-q$coef[2]
}
stock <- rep(0,num.stock); mse <- rep(0,num.stock); Rbar <- rep(0,num.stock)
Ratio <- rep(0,num.stock); col1 <- rep(0,num.stock); col2 <- rep(0,num.stock)
col3 <- rep(0,num.stock); col4 <- rep(0,num.stock); col5 <- rep(0,num.stock)
for(i in 1:num.stock){
  Rbar[i] <- a[i]+b[i]*mean.idx
  mse[i] <- sum(lm(data[,i]~idx)$residuals^2)/(nrow(data)-2)
  Ratio[i] <- (Rbar[i]-rf)/b[i]
  stock[i] <- i
}
xx <- (cbind(stock, a, b, Rbar, mse, Ratio))
aaa <- xx[order(-Ratio),]
col1 <- (aaa[,4]-rf)*aaa[,3]/aaa[,5]
col3 <- aaa[,3]^2/aaa[,5]
for(i in(1:num.stock)) {
  col2[i] <- sum(col1[1:i])
  col4[i] <- sum(col3[1:i])
}
for(i in (1:num.stock)) {
  col5[i] <- var.idx*col2[i]/(1+var.idx*col4[i])
}
table1 <- cbind(aaa, col1, col2, col3, col4, col5)
table2 <- table1[1:which(col5==max(col5)), ]
z.no.short <- (table2[,3]/table2[,5])*(table2[,6]-max(col5))
x.no.short <- abs(z.no.short)/abs(sum(z.no.short))
a.single.no.short <- t(table2[,2]) %*% x.no.short
b.single.no.short <- t(table2[,3]) %*% x.no.short
ret.single.no.short <- a.single.no.short + b.single.no.short * mean.idx
sd.single.no.short <- sqrt(b.single.no.short^2 *
  var.idx + (t(x.no.short^2)) %*% table2[,5])
table <- cbind(table2[,1:6],z.no.short,x.no.short)
return(list(table=table,stock=names(data[(table2[,1])]),
  weights=x.no.short,return=ret.single.no.short,risk=sd.single.no.short))
}

```

6.2.5 Portfolio: Constant Correlation (With Short Sales)

Constant Correlation works the same way as the Single Index Model functions only that it is not required to input the market information since they are not part of the calculation. In our case where short sales are allowed, we type the

following:

```
port.cc.ss(data=ret.mon,rf=0.001,lint=TRUE)
```

Source codes as follows:

```
port.cc.ss <- function(data=ret.mon, rf=0.001, lint=FALSE){
  num.stock <- ncol(data)
  mean.ret <- apply(data, 2, mean)
  vcov.ret <- cov(data)
  sigma.ret <- rep(0,num.stock)
  for (i in 1:ncol(vcov.ret)){
    sigma.ret[i]<-sqrt(vcov.ret[i,i])
  }
  corr.ret <- cor(data)
  rho <- (sum(corr.ret)-num.stock)/(num.stock*(num.stock-1))
  x <- rep(0,5*num.stock)
  xx <- matrix(x,ncol=5,nrow=num.stock)
  stock <- seq(1,num.stock,1); Rbar <- rep(0,num.stock); Rbar_f <- rep(0,num.stock)
  sigma <- rep(0,num.stock); Ratio <- rep(0,num.stock); col1 <- rep(0,num.stock)
  col2 <- rep(0,num.stock); col3 <- rep(0,num.stock)
  Rbar <- mean.ret
  Rbar_f <- Rbar-rf
  for (i in 1:num.stock){
    sigma[i] <- sqrt(vcov.ret[i,i])
  }
  Ratio <- Rbar_f/sigma.ret
  xx <- (cbind(stock, Rbar, Rbar_f, sigma, Ratio))
  aaa <- xx[order(-Ratio),]
  for(i in (1:num.stock)){
    col1[i]<-rho/(1-rho+i*rho)
    col2[i]<-sum(aaa[,5][1:i])
  }
  for(i in (1:num.stock)){
    col3[i]<-col1[i]*col2[i]
  }
  table <- cbind(aaa, col1, col2, col3)
  z_cons <- (1/((1-rho)*table[,4]))*(table[,5]-table[,8][nrow(table)])
  lint <- lint
  ifelse(lint, x_cons <- (z_cons)/sum(abs(z_cons)), x_cons <- z_cons/sum(z_cons))
  R_cons_short <- t(aaa[,2]) %*% x_cons
  tempcorr <- matrix(rho, nc=num.stock, nr=num.stock)
  for(i in (1:num.stock)){
    tempcorr[i,i]<-1
  }
  sd_cons_short <- sqrt((t(x_cons*table[,4])) %*%
    tempcorr %*% (x_cons*table[,4]))
  table <- cbind(table[,1:5],z_cons,x_cons)
  return(list(table=table,stock=names(data[(table[,1])]),
```

```

    weights=x_cons,return=R_cons_short,risk=sd_cons_short))
}

```

6.2.6 Portfolio: Constant Correlation (No Short Sales)

Finally, the Constant Correlation short sales not allowed function works the same way as the one with short sales allowed. The only difference is that we get weights assuming there is no short sales opportunities available. In our case, we type the following:

```
port.cc.nss(data=ret.mon,rf=0.001)
```

Source codes as follows:

```

port.cc.nss <- function(data=ret.mon, rf=0.001){
  num.stock <- ncol(data)
  mean.ret <- apply(data, 2, mean)
  vcov.ret <- cov(data)
  sigma.ret <- rep(0,num.stock)
  for (i in 1:ncol(vcov.ret)){
    sigma.ret[i]<-sqrt(vcov.ret[i,i])
  }
  corr.ret <- cor(data)
  rho <- (sum(corr.ret)-num.stock)/(num.stock*(num.stock-1))
  x <- rep(0,5*num.stock)
  xx <- matrix(x,ncol=5,nrow=num.stock)
  stock <- seq(1,num.stock,1); Rbar <- rep(0,num.stock); Rbar_f <- rep(0,num.stock)
  sigma <- rep(0,num.stock); Ratio <- rep(0,num.stock); col1 <- rep(0,num.stock)
  col2 <- rep(0,num.stock); col3 <- rep(0,num.stock)
  Rbar <- mean.ret
  Rbar_f <- Rbar-rf
  for (i in 1:num.stock){
    sigma[i] <- sqrt(vcov.ret[i,i])
  }
  Ratio <- Rbar_f/sigma.ret
  xx <- (cbind(stock, Rbar, Rbar_f, sigma, Ratio))
  aaa <- xx[order(-Ratio),]
  for(i in (1:num.stock)){
    col1[i]<-rho/(1-rho+i*rho)
    col2[i]<-sum(aaa[,5][1:i])
  }
  for(i in (1:num.stock)){
    col3[i]<-col1[i]*col2[i]
  }
}

```

```

xxx <- cbind(aaa, col1, col2, col3)
table <- xxx[1:which(xxx[,8]==max(xxx[,8])),]
z_cons_no <- (1/((1-rho)*table[,4]))*(table[,5]-table[,8][nrow(table)])
x_cons_no <- abs(z_cons_no)/sum(abs(z_cons_no))
R_cons_no <- t(table[,2]) %*% x_cons_no
tempcorr <- matrix(rho,nc=nrow(table),nr=nrow(table))
for(i in (1:nrow(table))){
  tempcorr[i,i] <- 1
}
sd_cons_no <- sqrt((t(x_cons_no*table[,4]))
  %*% tempcorr %*% (x_cons_no*table[,4]))
table <- cbind(table[,1:5],z_cons_no,x_cons_no)
return(list(table=table,stock=names(data[(table[,1])]),
  weights=x_cons_no,return=R_cons_no,risk=sd_cons_no))
}

```

6.2.7 VaR: Parametric (Equally Weighted)

For our value-at-risk and the corresponding back testing functions, there are more input criteria. Similar as before, we need a merged data frame or matrix of returns (daily in the example of this paper) to carry out the computations. For example, we type the following to get 1-day parametric VaR's (equally weighted moving average):

```
MyVarParEq(port=port,wei=e.weight,comp.days=1,size.days=251,conf=0.95)
```

where `port` is the data frame or matrix of daily returns of all the stocks in our portfolio, `e.weight` is the vector of stock allocation of our portfolio, `comp.days` is the number of next day VaR we want to compute (if we put 1, it give us the VaR of the day after our last day in our data set), `size.days` is the window length, meaning the number of days we use to compute the variance-covariance information for each VaR, and `conf` is the confidence level. Source codes as follows:

```
MyVarParEq <- function(port=port,wei=e.weight,comp.days=1,size.days=251,
  conf=0.95){
  if(nrow(port)-size.days<comp.days)
```



```

    stop("Not enough data. Try lowering comp.days or size.days")
    val <- port %*% wei
    pvar <- c(rep(0,comp.days))
    for(i in 1:comp.days){
      a <- 1+i; b <- size.days+i
      data <- port[a:b,]
      var_cov <- cov(data)
      var <- t(wei) %*% var_cov %*% wei
      sd <- var^.5
      pvar[i] <- qnorm(conf,0,1) * sd * -1
    }
    pfail <- sum((val[0:comp.days]<0)&(val[0:comp.days]<pvar))/length(pvar)
    return(list(VaR=pvar,Fail=pfail))
}

```

6.2.8 VaR: Parametric (Exponentially Weighted)

Parametric VaR with exponentially weighted moving average works very similarly to its equally weighted counterpart. Only this time, we need to also specify the decay factor lambda, `lamb`.

```
MyVarParEx(port=port,wei=e.weight,comp.days=1,size.days=251,conf=0.95,lamb=0.97)
```

Source code as follows:

```

MyVarParEx <- function(port=port,wei=e.weight,comp.days=1,size.days=251,
  conf=0.95,lamb=0.97){
  if(nrow(port)-size.days<comp.days)
    stop("Not enough data. Try lowering comp.days or size.days")
  val <- port %*% wei
  pvar <- c(rep(0,comp.days))
  for(i in 1:comp.days){
    a <- 1+i; b <- size.days+i
    data <- port[a:b,]
    j <- 0:(nrow(data)-1)
    ewma.wt <- lamb^j; ewma.wt <- ewma.wt/sum(ewma.wt)
    cov.ewma <- cov.wt(data, wt=ewma.wt)
    var_cov <- cov.ewma$cov
    var <- t(wei) %*% var_cov %*% wei
    sd <- var^.5
    pvar[i] <- qnorm(conf,0,1) * sd * -1
  }
  pfail <- sum((val[0:comp.days]<0)&(val[0:comp.days]<pvar))/length(pvar)
  return(list(VaR=pvar,Fail=pfail))
}

```

6.2.9 VaR: Historical (Equally Weighted)

For historical VaR, we need to specify our portfolio return data frame as well as the number of days we want to calculate and use for the final percentile (which will become our VaR). For example, in our exercise we type the following:

```
MyVarHistEq(port=port,wei=e.weight,comp.days=1,size.days=251,conf=0.95)
```

Source codes as follows:

```
MyVarHistEq <- function(port=port, wei=e.weight, comp.days=1, size.days=251,
  conf=0.95){
  if(nrow(port)-size.days<comp.days)
    stop("Not enough data. Try lowering comp.days or size.days")
  val <- port %*% wei
  hvar <- c(rep(0,comp.days))
  for(i in 1:comp.days){
    a <- 1+i; b <- size.days+i
    data <- val[a:b,]
    hvar[i] <- quantile(data,1-conf)
  }
  hfail <- sum((val[0:comp.days]<0)&(val[0:comp.days]<hvar))/length(hvar)
  return(list(VaR=hvar, Fail=hfail))
}
```

6.2.10 VaR: Historical (Bootstrapping)

Historical Bootstrapping is similar to the previous method, but this time we bootstrap from the historical data instead of taking the actual historical sequence as is. Therefore, we are using all available history, from first day in data frame until the day before the calculated VaR to bootstrap a sample of data of length “size.days”. From this sample, the function draws the percentile based on the specified confidence level. For example, in our exercise we type the following:

```
MyVarHistBS(port=port,wei=e.weight,comp.days=1,size.days=251,conf=0.95)
```

Source codes as follows:

```

MyVarHistBS <- function(port=port, wei=e.weight, comp.days=1, size.days=251,
  conf=0.95){
  if(nrow(port)-size.days<comp.days)
    stop("Not enough data. Lower comp.days.")
  val <- port %*% wei
  hbvar <- c(rep(0,comp.days))
  for (i in 1:comp.days){
    a <- i+1
    data <- port[a:nrow(port),]
    portboot <- data[sample(nrow(data),size.days,replace=TRUE),]
    valboot <- portboot %*% wei
    hbvar[i] <- quantile(valboot, 1-conf)
  }
  hbfail <- sum((val[0:comp.days]<0)&(val[0:comp.days]<hbvar))/length(hbvar)
  return(list(VaR=hbvar, Fail=hbfail))
}

```

6.2.11 VaR: Monte Carlo Simulation

Finally, for Monte Carlo Simulated VaR we specify the portfolio, weight, compute days, the size of the window of the simulated sample, and the number of artificial returns to generate (iterations) when simulating based on the mean and variance of the window length of portfolio data. For example, we type the following:

```
MyVarMC(port=port,wei=e.weight,comp.days=1,size.days=251,conf=0.95,iter=1000)
```

Source code as follows:

```

MyVarMC <- function(port=port, wei=e.weight, comp.days=1, size.days=251,
  conf=0.95, iter=1000){
  val <- port %*% wei
  mcvar <- c(rep(0,comp.days))
  for(i in 1:comp.days){
    a <- 1+i; b <- size.days+i
    data <- port[a:b,]
    mu <- apply(data, 2, mean)
    var_cov <- cov(data)
    mc_ret <- mvrnorm(iter, mu, var_cov)
    mc_port_ret <- mc_ret %*% wei
    mcvar[i] <- quantile(mc_port_ret, 1-conf)
  }
  mcfail <- sum((val[0:comp.days]<0)&(val[0:comp.days]<mcvar))/length(mcvar)
  return(list(VaR=mcvar,Fail=mcfail))
}

```

CHAPTER 7

Concluding Remarks

This paper serves as an introduction and demonstration of statistical finance theories and their applications with a simple user-friendly stock investment portfolio. As emphasized previously, these tools are a good start to learn about investment management but in no way are they all-encompassing. Hence, the ideas and applications should be considered “work in progress” and it is of an investor’s and statistician’s best interest to keep updating oneself with knowledge and techniques in order to fine-tune these tools. The reason financial crises happen from time to time is a sound proof of the ever-changing financial environment that calls for some ever-improving theories and applications.

This paper opens doors to many future possibilities. The R-Application presented in Chapter 6 can be customized and modified into an R-Package for educational purposes in the statistical finance area, or as an introductory investment tool for beginning investors and students. Further, to answer to the criticisms encountered by Modern Portfolio Theory and Value-at-Risk methods discussed in Section 2.3, 3.4, 4.4 and 5.4, investment management topics such as Semi-Variance, Geometric Mean Returns, Safety First Portfolios, and Extreme Value Theory are a few among many to be explored, studied, and evaluated. Hence, the portfolio selection and risk management framework laid out in this paper is only the beginning to a very broad array of future explorations, and it is our job as scholars to continue to contribute and improve existing theories and methods.

REFERENCES

- [1] Amenc, N., & Le Sourd, V. (2003). *Portfolio theory and performance analysis*. Wiley finance series. Chichester, England: Wiley.
- [2] Campbell, S. D. (2006). *A review of backtesting and backtesting procedures*. *JOURNAL OF RISK*. 9 (2), 1-18.
- [3] Elton, E. J. (2007). *Modern Portfolio Theory and Investment Analysis*. Hoboken, NJ: J. Wiley & Sons.
- [4] Elton, Edwin J., Gruber, Martin M. and Padberg, Manfred W. (1977). *Simple Rules for Optimal Portfolio Selection: The Multi Group Case* *The Journal of Financial and Quantitative Analysis*. 12(3), 329-345.
- [5] Elton, Edwin J., Gruber, Martin M. and Padberg, Manfred W. (1976). *Simple Criteria for Optimal Portfolio Selection*. *The Journal of Finance*. 31(5), 1341-1357.
- [6] Elton, Edwin J., Gruber, Martin M. and Padberg, Manfred W. (1978). *Simple Criteria for Optimal Portfolio Selection: Tracing Out the Efficient Frontier*. *The Journal of Finance*. 33(1), 296-302.
- [7] Elton, Edwin J., Gruber, Martin M. and Padberg, Manfred W. (1978). *Optimal Portfolios from Simple Ranking Devices*. *Journal of Portfolio Management*. 4(3), 15-19.
- [8] Hendricks, D. (1996). *Evaluation of value-at-risk models using historical data*. *FRBNY Economic Policy Review*. Apr., 39-69.
- [9] Hull, J. (2000). *Options, futures & other derivatives*. Upper Saddle River, NJ: Prentice Hall.
- [10] Lopez, J. A. (1998). *Methods for evaluating value-at-risk estimates*. New York, N.Y.: Federal Reserve Bank of New York.
- [11] Lowenstein, R. (2000). *When genius failed: The rise and fall of Long-Term Capital Management*. New York: Random House.
- [12] Jorion, P. (2001). *Value at risk: The new benchmark for managing financial risk*. New York: McGraw-Hill.
- [13] Kwan, Clarence C.Y. (1984). *Portfolio Analysis Using Single Index, Multi-Index, and Constant Correlation Models: A Unified Treatment*. *The Journal of Finance*. 39(5), 1469-1483.

- [14] Pritsker, M. (2006). *The hidden dangers of historical simulation*. *Journal of Banking & Finance*. 30 (2), 561.
- [15] Reilly, F. K., & Brown, K. C. (2003). *Investment analysis and portfolio management*. Mason, Ohio [u.a.]: Thomson South-Western.
- [16] *RiskMetricsTM - Technical Document* (1996), J.P. Morgan, New York
- [17] Spector, P. (2008). *Data manipulation With R. Use R!*. New York: Springer Verlag. <http://dx.doi.org/10.1007/978-0-387-74731-6>.
- [18] Stulz, R. M. (2003). *Risk management & derivatives*. Mason, Ohio: Thomson/South-Western.
- [19] Sullivan, arthur; Steven M. Sheffrin (2003). *Economics: Principles in action*. Upper Saddle River, New Jersey 07458: Pearson Prentice Hall. pp. 271. ISBN 0-13-063085-3.
- [20] Taleb, N. (2005). *Foiled by randomness: The hidden role of chance in life and in the markets*. New York: Random House.
- [21] Taleb, N. (2007). *The black swan: The impact of the highly improbable*. New York: Random House.
- [22] Tufano, Peter and Kyrillos, Barbara B. (1994) *Leland O'Brien Rubinstein Associates, Inc.: Portfolio Insurance*. Harvard Business Publishing.
- [23] Vinod, H. D., & Reagle, D. P. (2005). *Preparing for the worst: Incorporating downside risk in stock market investments*. Wiley series in probability and statistics. [Hoboken, N.J.]: Wiley-Interscience.
- [24] Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis*. New York: Springer.