UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2012:7

An Empirical Comparison of Different Approaches in Portfolio Selection

Jiaqun Hu

Examensarbete i matematik, 30 hp
Handledare och examinator: Maciej Klimek
Maj 2012

Department of Mathematics
Uppsala University

Abstract

There are two parts in this paper, the first one is about using factor analysis method to evaluate the performance of individual asset. Afterwards, according to the composite score, ten stocks based on Shanghai Securities Composite Index 50 will be chosen for the portfolio optimization. The second part focuses on three methods for optimal portfolio selection, i.e. Mean-Variance method, MCD Robustified Mean-Variance Method and Mean-CVaR method. The purpose of this paper is to discuss and compare the portfolio compositions and performance of three different approaches in portfolio optimization.

Acknowledgements

Firstly, I would like to thank my project supervisor Professor Maciej Klimek for all his help and useful suggestions during my thesis period.

Secondly, I deeply appreciate all the help of staff, instruction of teachers and encouragement of classmates during my time of studying at Uppsala University.

Finally, many thanks for my family's supports.

Contents

1 Introduction

When an investor makes investment decisions, there are many problems he will face. How to choose the assets? How to allocate the assets? etc. Nonetheless, one of the most important things an investor may be concerned about is how to get an optimal portfolio.

Since 1952, Harry Markowitz [26] published his seminal work on portfolio selection, Modern Portfolio Theory (MPT) was unveiled. MPT is a theory which tries to maximize portfolio expected return for a given amount of portfolio risk, or minimize risk for a given level of expected return, by carefully selecting the proportions of various assets.

Doubtlessly, Markovitz Mean-Variance (MV) model has a significant effect on financial market. However, there exit some limitations using the variance as the risk measure. Such as due to the symmetrical measure of variance, it doesn't consider the direction of movement. Therefore, it maybe results that an asset which has better expected return is deemed to be as risky as an asset from lower expected return. Thus, to overcome the limitations, alternative risk measures such as Value at Risk (VaR) and Conditional Value at Value at Risk (CVaR) have been presented to replace the variance.

Moreover, MV model often performs poorly out-of-sample due to estimation errors in the mean vector and covariance matrix. As a consequence, minimum-variance portfolios may produce unstable weights that affect stability over time. It may also lead to extreme portfolio weights and dramatic fluctuation in weights with only unimportant changes in expected returns or the covariance matrix because of the loss of stability. Therefore, for the sake of achieving better stability properties, using robust methods to compute the mean and (or) covariance matrix has been introduced to reduce the estimation error. Two different methods are implemented: robust mean and covariance estimators, and the shrinkage estimator [25].

There are two parts in this paper, the first one is about using factor analysis method to evaluate the performance of single asset. Then, according to the composite score, ten stocks based on Shanghai Securities Composite Index 50 (SSE 50) will be chosen for the portfolio optimization. The second part focuses on three approaches for optimal portfolio selection, i.e. Mean-Variance method, MCD Robustified Mean-Variance Method and Mean-CVaR method, including the comparison of efficient frontier, performance and so on.

2 Background

In this chapter, some key concepts used throughout this thesis will be introduced, including financial statement, multivariate statistical analysis, portfolio optimization, etc.

2.1 Financial Statement

A financial statement is a written report which records the financial activities of a business, person, or other entities. Through investigating the financial statement, an investor can know the financial circumstance of a company. Consulting [31], some definitions about financial statement will be introduced shortly.

2.1.1 The Form of the Financial Statement

The form of the financial statement describes the way in which the statements and their components relate to each other. A form is obtained by a set of accounting relations which express the various components of financial statements in terms of other parts. Firms must publish three primary financial statements. They are the balance sheet, the income statement and the cash flow statement.

- **The Balance Sheet**
  The balance sheet lists assets, liabilities, and stockholders' (shareholders') equity.
- **The Income Statement**
  The income statement reports how shareholder's equity increased or decreased as a result of business activities.
- **The Cash Flow Statement**
  The cash flow statement describes how the firm generated and used cash during the period. Cash flows are divided into three types in the statement, that is cash flows from operating activities, cash flows from financing activities and cash flows from investing activities.

2.2 Multivariate Statistical Analysis

According to [17], multivariate statistical analysis refers to multiple advanced techniques for investigating relationships among multiple variables at the same time. Researchers use multivariate method to analyse more than one dependent (independent) variable.

There are various statistical approaches for implementing multivariate analysis. The common procedures are multiple regression analysis, factor analysis, path analysis, and multiple analysis of variance (MANOVA).

- **Multiple Regression Analysis**
  Regression analysis is the statistical method for predicting values of one or more dependent variables from a collection of independent variable values. It can also be used for evaluating the effects of the predictor variables on the responses.
- **Factor Analysis**
  The crucial purpose of factor analysis is to describe the covariance relationships among many variables in terms of a few underlying, but unobservable random quantities called factors.
- **Path Analysis**

Path analysis is an approach that employs simple bivariate correlations to estimate relationships in Structural Equation Modeling (SEM) model. Path analysis seeks to determine the strength of paths shown in path diagrams.

- **MANOVA**

MANOVA is an extension of analysis of variance (ANOVA) to accommodate more than one dependent variable. It is a dependence technique measuring the differences for two or more metric dependent variables based on a set of categorical (non-metric) variables acting as independent variables.

In this thesis, factor analysis will be implemented to evaluate the financial statement. The details will be showed in section 3.1.

2.3 Portfolio Optimization

In finance, the portfolio optimization is the process of finding an optimal asset allocation which can satisfy the given risk preferences of an investor.

2.3.1 Terminology

- **Expected Return**

The most basic concept concerning any investment is that of a return (or a rate of return) corresponding to a given investment period. Referring to Wikipedia, the rate of return can be calculated by two approaches over a single period, that is, arithmetic method and logarithmic method.

Therefore, in a single period, the arithmetic return can be obtained by the following formula:

$$r_{\text{arith}} = \frac{p_t - p_0}{p_0}$$

where p_0 is the amount invested at the beginning of the period and p_t is the amount received by the investor at the end of the period.

Correspondingly, the logarithmic return is given by

$$r_{\log} = \frac{\ln\left(\frac{p_t}{p_0}\right)}{t}$$

where t is the time period.

In effect, the rate of return is always related to a particular investment period which can be regarded as a random variable R. To know its expected value $E[R]$ (or the expected return) is vital in any kind of portfolio optimization. In the discrete case, the expected return is the weighted-average outcome in different areas, such as probability theory, finance and so on. Let R_{ij} denotes the jth possible outcome for the return on asset i and $\bar{R}_i$ indicates the expected return. Then, the expected value of the M equally likely returns for asset i can be calculated by the following formula:

$$E[R_i] = \bar{R}_i = \sum_{j=1}^{M} \frac{R_{ij}}{M}$$

Furthermore, if the outcomes are not equally likely and using the symbol P_{ij} to denote the probability of the jth return on the ith asset, then expected return becomes

$$E[R_i] = \bar{R}_i = \sum_{j=1}^{M} P_{ij} R_{ij}$$

Therefore, the expected return on a portfolio of assets is simply a weighted average of the expected return on the individual assets. The weight employed to each return is the fraction of the portfolio invested in that asset. Assuming that $\bar{R}_P$ describes the expected return on the portfolio and X_i indicates the fraction of the investor's funds invested in the ith asset, and N is the number of assets, then the expected value of a constant times is given by

$$E[R_P] = \bar{R}_P = \sum_{i=1}^{N} X_i E[R_i]$$

- **Risk**

According to Wikipedia, the definition of risk in finance is the probability that an investment's real return will be unlike as expected. It indicates the possibility of losing some or all of the original investment. Damodaran [9] claimed that risk included not only "downside risk" but also "upside risk" which meant that returns were exceeding expectations.

In the classical mean-variance model, Markowitz used variance as the measure of risk. The formula for the variance of the return on ith asset when each of M possible outcome is

$${\sigma_i}^2 = \sum_{j=1}^{M} \frac{(R_{ij} - \bar{R}_i)^2}{M}$$

If the outcomes are not equally likely, then, the formula for the variance of the return on the ith asset is obtained through multiplying by the probability with which they occur is obtained by

$${\sigma_i}^2 = \sum_{j=1}^{M} [P_{ij}(R_{ij} - \bar{R}_i)^2]$$

Moreover, the definition of variance of a portfolio P, designated by σ_P^2 is

$$\sigma_P^2 = E[(R_P - \bar{R}_P)^2] = \sum_{i=1}^{n} \sum_{j=1}^{n} X_i X_j \sigma_{ij}$$

where σ_{ij} is the covariance between assets i and j, i.e. how much these two assets change together, X_i is the amount invested in asset i.

- **Efficient Frontier**

In terms of Wikipedia, Harry Markowitz, Elton [10] and other researchers presented the idea of efficient frontier. They advocate that a combination of assets, i.e. a portfolio, is introduced as "efficient" if it had the best possible expected level of return for its risk level. The "efficient frontier" is the upward-sloped part of the left boundary in a region that collects all possible portfolios which combine every feasible risky assets, without including any holdings of the risk-free asset and can be plotted in a risk-expected return space. Figure 2.1 shows an example of an efficient frontier.

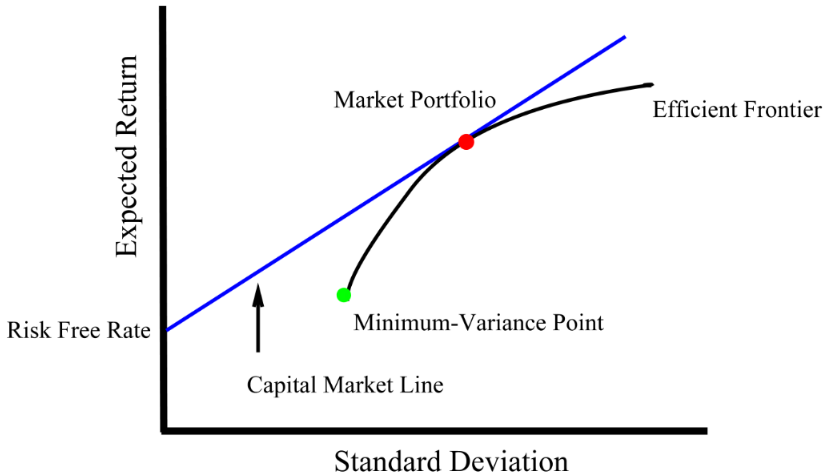

Figure 2.1: Efficient Frontier

Remarkably, referring to the descriptions of Luenberger [24], the market portfolio is the summation of all assets if everyone invests the same amount of risky asset and assigns to the same means, variances as well as covariances for the returns, namely, it must incorporate every asset in proportion to that asset's representation in the total market. Meanwhile, the single straight line drawn from the risk-free point passing through the market portfolio is named capital market line (CML), which can be viewed in figure 2.1. Apparently, this line displays the relationship between the expected return and the risk measured by standard deviation for efficient assets or portfolios. Moreover, in this diagram, the upper curve of the minimum-variance point represents the efficient frontier.

Furthermore, under certain assumptions, Fama [13] demonstrates that the tangency portfolio which is the feasible point that maximizes the angle between a line drawn from the risk-free asset to a point in the feasible region and the horizontal axis must contain all assets available to investors, and each asset must be included in proportion to its market value associating with the entire market. Therefore, the tangency portfolio is usually considered as the market portfolio.

Meantime, if the angle is denoted by θ, then, according to [24] the $\tan\theta$ of tangency portfolio can be obtained by

$$\tan\theta = \frac{E(R_{\mathrm{TP}})-R_f}{\sigma_{\mathrm{TP}}}$$

where $E(R_{\mathrm{TP}})$ means the expected return of tangency portfolio, R_f is the risk-free rate, σ_{TP} is the standard deviation of tangency portfolio. In fact, this is also equal to the Sharpe ratio (which will be introduced later) of the tangency portfolio. In particular, connecting with the efficient market hypothesis which describes that no one can beat the market, all portfolios should have a Sharpe ratio no more than the market portfolio's. Therefore, the tangency portfolio is often treated as the portfolio of risky assets with the highest Sharpe ratio.

2.3.2 Classical Framework for Mean-Variance Optimization

- **Minimum Variance Portfolio**

According to Markowitz theory, the investor's problem is a constrained minimization problem in the sense that the investor must seek

$$\min \quad \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}$$

subject to

$$\sum_{i=1}^{n} w_i = 1$$

$$\sum_{i=1}^{n} r_i w_i = R$$

$$0 \leq w_i \leq 1 \ for \ i = 1, 2, ..., n$$

where σ_{ij} is the covariance between assets i and j, w_i is the weight of asset i, r_i is the expected return on asset i and R is the expected return required by an investor.

- **Maximum Expected Returns Portfolio**

For an investor that wishes to attain the maximum expected return, the following problem maximizes the expected return of the portfolio.

$$\max \quad \sum_{i=1}^{n} w_i E[R_i]$$

subject to

$$\sum_{i=1}^{n} w_i = 1$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} = \sigma_0$$

$$X_i \geq 0 \ for \ i = 1, 2, ..., n$$

where R_i is the expected return on asset i, w_i is the weight on asset i, σ_{ij} is the covariance between assets i and j and σ_0 is the target risk given by an investor.

2.3.3 Robust Portfolio

Robust portfolio is one of the major selections in the investment management process, and the decision making is regularly based on the recommendations of risk-return optimization procedures. Several major issues are worthy of attention. The first one is to carefully consider how portfolio risk and return are defined, checking whether these definitions are suitable for given investigated or predicted asset return distributions and underlying investor preferences. It is extremely critical since this concerns giving rise to alternative theories of risk measures and asset allocation frameworks greater than the classical mean-variance optimization. The next one is how the optimization problem is produced and resolved in practice. It is very pivotal, especially for larger portfolios because working knowledge of the state-of-the-art abilities of quantitative software for managing portfolio is crucial. The last but not the least thing is to evaluate the sensitivity of portfolio optimization models to inaccuracies in input estimates.

- **Robust Mean and Covariance Estimators**

In traditional mean-variance model, it uses the sample mean and covariance estimators for the purpose of estimating the expected returns and covariances among assets. However, one problem with the standard estimators is that they are sensitive to outliers. In some cases, even a single outlier with an extreme return can significantly influence the resulting means and covariances. This is an undesirable property of many classical estimators. So, robust estimation has been introduced. There are many different robust approaches and related estimators for the mean and covariance, such as the Minimum Volume Ellipsoid (MVE) estimator, the minimum covariance determinant (MCD) estimator. In this study, the MCD method will be used. The details see section 4.2.

2.3.4 Portfolio Risk Measures

Since the introduction of MV model, variance becomes the most common risk measures in portfolio optimization. However, this model has some limitations, such as it relies on the assumption that the returns of assets are multivariate normally distributed. Nevertheless, many researches show that asset returns are not normal, such as Brooks and Kat [6]. Therefore, the mean and the variance alone do not fully describe the characteristics of the joint asset return distribution. In other words, a portfolio manager will face many risks and undesirable scenarios cannot be captured merely by the variance of the portfolio. As a consequence, especially in cases of significant non-normality, the classical mean-variance approach will not be a satisfactory portfolio allocation model.

Beginning with about the middle 1990s, substantial thought and innovation in the financial market have been directed towards generating a better comprehension of risk and its measurement, and contributed to improve the management of risk in financial portfolios. From the statistical area of view, a major innovation is to concentrate on the ratio between the bulk of the risk and the risk of the tails. Actually, the latter has become a key statistical determinant of risk management policies. Changing conditions and various portfolios may require different and new risk measures. The following presents some popular different portfolio risk measures used in practice for asset allocation:

- **Deviation Mean-Absolute Deviation**

In 1988, Konno [22] reported using mean-absolute deviation (MAD) instead of squared deviations in the mean-variance approach, here the scatter measure is based on the absolute deviations from the mean. Therefore, MAD can be defined by:

$$\mathrm{MAD}(R_P) = E(\left|\sum_{i=1}^{N} w_i R_i - \sum_{i=1}^{N} w_i \mu_i\right|) = E(|R_P - E(R_P)|)$$

where

$$R_P = \sum_{i=1}^{N} w_i R_i$$

$$E(R_P) = \sum_{i=1}^{N} w_i \mu_i$$

R_i and μ_i are the portfolio return, the return on asset i, and the expected return on asset i respectively. Through using the case of the mean-absolute deviation method , the computation of optimal portfolios is very simplified since the optimization problem becomes linear and can be resolve by common linear programming techniques.

- **Semi-variance**

  Markowitz [27] suggested using semi-variance to rectify the fact that variance penalizes both overperformence and underperformence in his original book. Afterwards, he [28] claimed that "...Semi-variance seems more plausible than variance as a measure of risk, since it is concerned only with adverse deviation.".

  The portfolio semi-variance is defined as

$$\sigma^2_{P,\min} = E\left[\left(\min\left(\sum_{i=1}^{N} w_i R_i - \sum_{i=1}^{N} w_i \mu_i, 0\right)\right)^2\right] = E[(\min(R_P - E(R_P), 0))^2]$$

where

$$R_P = \sum_{i=1}^{N} w_i R_i$$

$$E(R_P) = \sum_{i=1}^{N} w_i \mu_i$$

R_i and μ_i are the portfolio return, the return on asset i, and the expected return on asset i, respectively.

- **Value-at-Risk (VaR)**

  Value-at-Risk (VaR) maybe is one of most popular downside risk measures which is first developed by JP Morgan. Later, it was applied feasible through the RiskMetrics^{TM} software in October 1994 [18]. VaR is connected with the percentiles of loss distributions, and estimates the possible maximum loss at a particular probability level (for example, 95%) over a certain time horizon (for example, 10 days). The Basel Committee presented several revisions to the original so-called 1988 Basel Accord that regulated the minimal capital requirements for banks in April 1993. Comparing with the previously Basel Accord which had solely covered credit risk (deposits and lending), the new proposal that took effect in 1998 includes not only the market risk but also the organization-wide commodities exposures (measured by 10 day 95% VaR) [5]. Now, most financial institutions use VaR to both trace and report the market risk exposure of their trading portfolios.

  Formally, VaR is defined as

$$\text{VaR}_\alpha(R_P) = -\inf\{R : 1-\alpha \leq P(R_P) \leq R)\}$$

where P denotes the probability function. Typical values for α are 90 %, 95 %, 99 % and the minus in front makes the value positive. Some of the practical and computational issues associate with using VaR can be found in Gordon [14], Alexei [2], and Mittnik, Rachev, and Schwartz [29].

- **Condition Value-at-Risk (CVaR)**

  Artzner et al. [4] proposed a set of valuable properties for a risk measure due to the deficiencies of VaR. They claimed that the risk measures satisfying these properties were coherent risk measures [1]. Conditional Value-at-Risk (CVaR) is a coherent risk measure which is defined by the following formula

$$\text{CVaR}_\alpha(R_P) = E(-R_P | -R_P \geq \text{VaR}_\alpha(R_P))$$

where the vertical stroke | denotes the conditional expectation. Therefore, this expression means that CVaR estimates the expected amount of losses if it exceeds portfolio VaR. Some alternative names of this risk measure are referred to as expected shortfall [4], expected tail loss (ETL), and tail VaR. Similar to VaR, the most typically considered values for α are 90 %, 95%, and 99%.

  In this project, the CVaR will be used as one of the measures of risk. The particulars see section 4.3.

2.3.5 Portfolio Performance Measures

- **Sharpe Ratio**

Consulting Wikipedia, the Sharpe ratio was developed by William Forsyth Sharpe [35] in 1996 which can be also called some alternative names, such as Sharpe index, reward-to-variability ratio and so on. This ratio focus on measuring the additional return (or risk premium) per unit of dispersion in an investment asset or a trading strategy which is considered as risk, that is, a variance risk measure. The definition of Sharpe ratio in a portfolio is:

$$S = \frac{E(R_P) - R_f}{\sigma_P}$$

where R_P is the portfolio return, R_f is the risk free return. σ_P is the standard deviation of the excess of the portfolio return.

- **Sortino ratio**

Referring to Wikipedia, the Sortino ratio measures the acutal return beyond the investor's target return per unit of downside risk. Actually, this ratio is a revision of the Sharpe ratio which evaluates the risk-adjusted return of an investment asset, portfolio or trading strategy. Nonetheless, dissimilar to Sharpe ratio, Sortino ratio penalizes solely those returns below a particular target given by an investor whereas Sharpe ratio penalizes both overperformance and underperformance equally. The mathematical definition of Sortino ratio is:

$$S = \frac{E(R_P) - T}{\sqrt{E[(\min(R_p - T, 0))^2]}}$$

where R is the return of an investment asset or portfolio, T is typically known as the minimum acceptable return, or MAR, which is absolutely higher than the risk free rate, and $f(\cdot)$ is the probability density function of the returns. The denominator can be treated as the root mean squared amount by which a return is less than the target, i.e. the underperformance. Comparatively, returns beyond target are considered as under-performance of 0.

Remarkably, although both these two ratios measure the risk-adjusted returns, various conclusions will be given as the real characteristic of the investment's return-generating competence since they are implemented in considerably different approaches.

3 Stock Selection

Broadly speaking, an optimal portfolio can reduce the risk. Therefore, the first problem an investor will face is how to choose the specific assets for a portfolio when he wants to make an investment decision. In this part, an useful way to select stocks for asset mix will be presented. At the beginning, some details about factor analysis will be described. Subsequently, some definitions of financial indicators will be introduced as well. Eventually, an empirical result will show the comparison of using three methods to implement the factor analysis. Meanwhile, according the composite score, ten stocks will be picked out for the portfolio optimization.

3.1 Factor Analysis

Factor analysis is an statistical technique whose main purpose is to detect and describe the underlying structure among the variables in the analysis. Generally, factor analysis supplies the tools for analysing the structure of the interrelationships among a large number of variables by defining a set of variables that can be highly interconnected which are called factors. These groups of variables (factors), which can be highly correlated, are considered to describe dimensionality of the data.

3.1.1 The Factor Analysis Model and The Orthogonal Factor Model

According to [17], commonly, the factor analysis model is

$$\begin{aligned} X_1 - \mu_1 &= l_{11}F_1 + l_{12}F_2 + \cdots + l_{1m}F_m + \epsilon_1 \\ X_2 - \mu_1 &= l_{21}F_1 + l_{22}F_2 + \cdots + l_{2m}F_m + \epsilon_2 \\ &\vdots \\ X_p - \mu_p &= l_{p1}F_1 + l_{p2}F_2 + \cdots + l_{pm}F_m + \epsilon_p \end{aligned} \tag{3.1}$$

or, in matrix notation,

$$\underset{(p\times 1)}{X - \mu} = \underset{(p\times m)}{L}\,\underset{(m\times 1)}{F} + \underset{(p\times 1)}{\epsilon} \tag{3.2}$$

where X is an observable random vector, with p components, has mean μ and covariance matrix $\Sigma = (\sigma_{ij})_{p\times p}$. The factor model supposes that X is linearly dependent upon a few unobservable random variables F_1, F_2, $\cdots$, F_m which are called common factors, and p additional sources of variation ϵ_1, ϵ_2, $\cdots$, ϵ_p are designated errors or, sometimes, specific factors. The coefficient l_{ij} are named the loadings of the ith variable on the jth specific factor, so the matrix L indicates the matrix of factor loadings. The ith specific factor ϵ_i is related only to the ith response X_i. The p deviations $X_1 - \mu_1$, $X_2 - \mu_2$,$\cdots$, $X_p - \mu_p$ are described in terms of $p+m$ random variables $F_1, F_2, \cdots, F_m, \epsilon_1, \epsilon_2, \cdots, \epsilon_p$ which are unobservable.

Usually, assume that

$$E(F) = \underset{(m\times 1)}{0}, \quad \mathrm{Cov}(F) = E[FF'] = \underset{(m\times m)}{I}$$

$$E(\epsilon) = \underset{(p\times 1)}{0}, \quad \mathrm{Cov}(\epsilon) = E[\epsilon\epsilon'] = \underset{(p\times p)}{\Psi} = \begin{bmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_p \end{bmatrix} \tag{3.3}$$

Since F and ϵ are independent, so

$$\text{Cov}(\epsilon, F) = E(\epsilon F') = \underset{(p\times m)}{0} \tag{3.4}$$

These suppositions and the connection in (3.2) are regarded as the orthogonal factor model.

Note that, the orthogonal factor model intimates a covariance structure for X. From the model in (3.2), it is simple to get

$$\begin{aligned}(X-\mu)(X-\mu)' &= (LF+\epsilon)(LF+\epsilon)' \\ &= (LF+\epsilon)((LF)'+\epsilon') \\ &= LF(LF)' + \epsilon(LF)' + LF\epsilon' + \epsilon\epsilon'\end{aligned}$$

associating with (3.3) and (3.4), then the covariance matrix becomes

$$\begin{aligned}\Sigma = \text{Cov}(X) &= E(X-\mu)(X-\mu)' \\ &= LE(FF')L' + E(\epsilon F')L' + LE(F\epsilon') + E(\epsilon\epsilon') \\ &= LL' + \Psi\end{aligned}$$

Moreover, in terms of the model in (3.2), it is easy to obtain that $(X-\mu)F' = (LF+\epsilon)F' = LFF' + \epsilon F'$, thus, $\text{Cov}(X, F) = E(X-\mu)F' = LE(FF') + E(\epsilon F') = L$.

Therefore, the covariance structure for the orthogonal factor model is

1. $\text{Cov}(X) = LL' + \Psi$

or

$$\begin{aligned}\text{Var}(X_i) &= l_{i1}^2 + \cdots + l_{im}^2 + \psi_i \\ \text{Cov}(X_i, X_k) &= l_{i1}l_{k1} + \cdots + l_{im}l_{km}\end{aligned} \tag{3.5}$$

2. $\text{Cov}(X, F) = L$

or

$$\text{Cov}(X_i, F_j) = l_{ij} \tag{3.6}$$

Furthermore, the fraction of the variance of the ith variable contributed by the m common factors is designated the ith communality. That division of $\text{Var}(X_i) = \sigma_{ii}$ due to the specific factor is often named the uniqueness, or specific variance. Let the ith communality is denoted by h_i^2, referring to (3.5), it is no difficult to attain the following expression:

$$\underbrace{\sigma_{ii}}_{\text{Var}(X_i)} = \underbrace{h_i^2}_{\text{communality}} + \underbrace{\psi_i}_{\text{specific variance}}$$

where

$$h_i^2 = l_{i1}^2 + l_{i2}^2 + \cdots + l_{im}^2 \tag{3.7}$$

Namely, the ith communality is given by totalling the squares of ith variable's loadings on the m common factors.

3.1.2 Methods of Estimation

For the sake of estimating the parameters including the loadings, specific errors and so on, three different approaches of estimation will be presented soon. They are the principal component method, the principal factor method and the maximum likelihood method as well.

- **The Principal Component Method**

Let Σ have eigenvalue-eigenvector pairs (λ_i, e_i) with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$ in terms of [17]. Then, the covariance matrix Σ can be expressed by:

$$\Sigma = \sum_{i=1}^{p} \lambda_i e_i e_i' = \left[\sqrt{\lambda_1} e_1 \,\vdots\, \sqrt{\lambda_2} e_2 \,\vdots\, \cdots \,\vdots\, \sqrt{\lambda_p} e_p \right] \begin{bmatrix} \sqrt{\lambda_1} e_1' \\ \hdashline \sqrt{\lambda_2} e_2' \\ \hdashline \vdots \\ \hdashline \sqrt{\lambda_p} e_p' \end{bmatrix} = LL' \tag{3.8}$$

Namely, the covariance matrix can be obtained by the sum of p eigenvalues λ_i multiplying their eigenvectors e_i and their transposes. Therefore, the basic idea of the principal component approach is to approximate (3.8) through totalling from 1 to m instead of the original sum without considering the last $p-m$ fractions in the sum. In particular, the m means the number of common factors, usually given by the experience or the number of eigenvalues greater than 1.

Then, connecting with (3.5), the approximation of Σ is

$$\Sigma = LL' + \Psi = \left[\sqrt{\lambda_1} e_1 \,\vdots\, \sqrt{\lambda_2} e_2 \,\vdots\, \cdots \,\vdots\, \sqrt{\lambda_m} e_m \right] \begin{bmatrix} \sqrt{\lambda_1} e_1' \\ \hdashline \sqrt{\lambda_2} e_2' \\ \hdashline \vdots \\ \hdashline \sqrt{\lambda_m} e_m' \end{bmatrix} + \begin{bmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_p \end{bmatrix} \tag{3.9}$$

where $\psi_i = \sigma_{ii} - \sum_{j=1}^{m} l_{ij}^2$ for $i = 1, 2, \ldots, p$.

So, suppose that there exist n observable random vector X_i with p components which are denoted by:

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{np} \end{bmatrix} \tag{3.10}$$

Meanwhile, their covariance matrix is expressed by

$$S = \begin{bmatrix} 1 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & 1 & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{pp} & \cdots & 1 \end{bmatrix} \tag{3.11}$$

Then, for the the sample covariance matrix S is specified in terms of its eigenvalue-eigenvector pairs $(\hat{\lambda}_1, \hat{e}_1), (\hat{\lambda}_2, \hat{e}_2), \ldots, (\hat{\lambda}_p, \hat{e}_p)$, where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \ldots \geq \hat{\lambda}_p$. Let $m < p$ be the number of common factors. Then the matrix of estimated factor loadings $\tilde{l}_{ij}^2$ is described by

$$\tilde{L} = \left[\sqrt{\hat{\lambda}_1} \hat{e}_1 \,\vdots\, \sqrt{\hat{\lambda}_2} \hat{e}_2 \,\vdots\, \cdots \,\vdots\, \sqrt{\hat{\lambda}_m} \hat{e_m} \right] \tag{3.12}$$

Connecting to (3.9), the estimated specific variances can be obtained by the diagonal elements of the matrix $S - LL'$, so

$$\tilde{\Psi} = \begin{bmatrix} \tilde{\psi}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\psi}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\psi}_p \end{bmatrix} \text{ with } \tilde{\psi}_i = s_{ii} - \sum_{j=1}^{n} \tilde{l}_{ij}^2 \tag{3.13}$$

Meanwhile, in terms of (3.7), the communalities are estimated as

$$\tilde{h}_i^2 = \tilde{l}_{i1}^2 + \tilde{l}_{i2}^2 + \cdots + \tilde{l}_{im}^2 \tag{3.14}$$

Particularly, in cases in which the units of the variables are not commensurate, it is usually recommendable to work with the standardized variables since standardization prevents the problems that there exists one variable with large variance inappropriately affects the determination of factor loadings. The formula of standardization is:

$$Z_j = \begin{bmatrix} \frac{(x_{j1} - \bar{x}_1)}{\sqrt{s_{11}}} \\ \frac{(x_{j2} - \bar{x}_2)}{\sqrt{s_{22}}} \\ \vdots \\ \frac{(x_{jp} - \bar{x}_p)}{\sqrt{s_{pp}}} \end{bmatrix} \qquad j = 1, \cdots, n \tag{3.15}$$

whose sample covariance matrix is the same as its sample correlation matrix R of the observations $X_1, X_2, \cdots, X_n$.

- **The Principal Factor Method**

Sometimes, the principal factor approach which is a modification of the principal component method will be considered. The distinctness between this two approaches is for the principal factor, $h_i^{*2} = 1 - \psi_i^*$ is used instead of the diagonal elements of sample covariance matrix S which are equal to 1 (referring to (3.11)), and the ψ_i^* is the initial approximation ψ_i. Therefore, a "reduced" sample covariance matrix will be given by:

$$S^* = \begin{bmatrix} h_1^{*2} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & h_2^{*2} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & h_p^{*2} \end{bmatrix} \tag{3.16}$$

In other words, the principal factor method of factor analysis uses the estimates

$$\tilde{L_r^*} = \left[\sqrt{\hat{\lambda}_1^*}\hat{e}_1^* \;\vdots\; \sqrt{\hat{\lambda}_2^*}\hat{e}_2^* \;\vdots\; \cdots \;\vdots\; \sqrt{\hat{\lambda}_m^*}\hat{e}_m^* \right] \tag{3.17}$$

$$\psi_i^* = 1 - \sum_{j=1}^{m} l_{ij}^{*2}$$

where $(\hat{\lambda}_i^*, \hat{e}_i^*), i = 1, 2, \ldots, m$ are the (biggest) eigenvalue-eigenvector pairs determined form S^*.

Correspondingly, the communalities would then be (re)estimated by

$$\tilde{h}_i^{*2} = \sum_{j=1}^{m} l_{ij}^{*2} \tag{3.18}$$

In particular, the principal factor solution can be obtained iteratively, with communality estimates of (3.18) becoming the initial estimates for the next step.

- **The Maximum Likelihood Method**

Before introducing the maximum likelihood method, some concepts of likelihood should first be explained. Referring to [17], if there exist $p \times 1$ vectors $X_1, X_2, \cdots, X_n$ describe a random sample from a multivariate normal population with mean vector μ and covariance matrix Σ. Since $X_1, X_2, \cdots, X_n$ are interactively independent and each has distribution $N_p(\mu, \Sigma)$, thus, the joint density function of all the observations is the product of the marginal normal densities which is given by:

$$\left\{ \begin{array}{c} \text{Joint density of} \\ X_1, X_2, \cdots, X_n \end{array} \right\} = \prod_{j=1}^{n} \left\{ \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-(x_j-\mu)'\Sigma^{-1}(x_j-\mu)/2} \right\} \tag{3.19}$$

When the numerical values of the observations become feasible, they may replace the X_j in Equation (3.19). The resulting expression, which is treated as function of μ and Σ for fixed sets of observations $X_1, X_2, \cdots, X_n$, is named the likelihood.

Many great statistical approaches apply values to the population parameters that "best" explain the observed data. Indeed, one of the best methods is to choose the parameter values that maximize the joint density evaluated at the observations, which is called maximum likelihood estimation, and the maximizing parameter values are named maximum likelihood estimates.

Therefore,for the factor analysis, if the common factors F and the specific factors ϵ can be supposed to be normally distributed, the maximum likelihood estimates of the factor loadings and specific variances may be attained. When F_j and ϵ_j are jointly normal, the observations $X_j - \mu = LF_j + \epsilon_j$ are also normal, the likelihood is

$$\begin{aligned} L(\mu, \Sigma) &= (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} e^{-(\frac{1}{2})\mathrm{tr}\left[\Sigma^{-1}\left(\sum_{j=1}^{n}(X_j-\bar{X})(X_j-\bar{X})'+n(\bar{X}-\mu)(\bar{X}-\mu)'\right)\right]} \\ &= (2\pi)^{-\frac{(n-1)p}{2}} |\Sigma|^{-\frac{(n-1)}{2}} e^{-(\frac{1}{2})\mathrm{tr}\left[\Sigma^{-1}\left(\sum_{j=1}^{n}(X_j-\bar{X})(X_j-\bar{X})'\right)\right]} \\ &\qquad \times (2\pi)^{-\frac{p}{2}} |\Sigma|^{-(\frac{1}{2})} e^{-\frac{n}{2}(\bar{X}-\mu)'\Sigma^{-1}(\bar{X}-\mu)} \end{aligned} \tag{3.20}$$

which depends on L and Ψ through $\Sigma = LL' + \Psi$. Meantime, it is worthy to make L well defined through imposing the computationally convenient uniqueness condition

$$L'\Psi^{-1}L = \triangle \quad \text{a diagonal matrix} \tag{3.21}$$

Then,the maximum likelihood estimates $\hat{L}$ and $\hat{\Psi}$ must be given by numerical maximization of (3.20).

Although a simple analytical expression cannot be obtained for the maximum likelihood estimators $\hat{L}$ and $\hat{\Psi}$, they can be shown to satisfy certain equations. Lawley and Maxwell [23] and Jöreskog [19], use the unbiased estimate $S = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})'$ of the covariance matrix replaced the maximum likelihood estimate $S_n = \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X})(X_j - \bar{X})'$. They ignored the contribution to the likelihood in (3.20) from the second term involving $(\mu - \bar{X})$, then maximizing the reduced likelihood over L and Ψ is equivalent to maximizing the Wishart likelihood and the log-likelihood function is given by

$$\log L = -\frac{1}{2}\ln|\Sigma| - \frac{1}{2}\mathrm{tr}(\Sigma^{-1}S) \tag{3.22}$$

Meantime, let

$$F(\hat{\mu}, \Psi, L) = \ln|\Sigma| + \mathrm{tr}(\Sigma^{-1}S) - \ln|S| - p \tag{3.23}$$

Therefore, the maximization problem of log L is equivalent to minimizing F.

Johnson, R.A. and Wichern, D.W. [17] introduced one computational scheme which is listed in the following:

1. Compute initial estimates of the specific variances $\psi_1, \cdots, \psi_p$. In terms of the report of Jöreskog [20], the specific variances are suggested setting

$$\hat{\psi}_i = \left(1 - \frac{1}{2} \cdot \frac{m}{p}\right)\left(\frac{1}{S_{ii}}\right) \tag{3.24}$$

where S_{ii} is the ith diagonal element of S^{-1}.
2. Given $\hat{\Psi}$, compute the first m distinct eigenvalues, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_m \geq 1$ as well as corresponding eigenvectors, $\hat{e}_1, \hat{e}_2, \cdots, \hat{e}_m$, of the covariance matrix

$$S^* = \hat{\Psi}^{-1/2} S \hat{\Psi}^{-1/2} \tag{3.25}$$

Let $\hat{E} = \left[\ \hat{e}_1 \ \vdots \ \hat{e}_2 \ \vdots \ \cdots \ \vdots \ \hat{e}_m \ \right]$ be the $p \times m$ matrix of normalized eigenvectors and $\hat{\Lambda} = diag[\hat{\lambda}_1, \cdots, \hat{\lambda}_m]$ be the $m \times m$ diagonal matrix of eigenvalues. Because $\hat{\Lambda} = I + \hat{\Delta}$ and $\hat{E} = \hat{\Psi}^{-1/2} \hat{L} \hat{\Delta}^{-1/2}$. Thus, the estimates can be attained by

$$\hat{L} = \hat{\Psi}^{1/2} \hat{E} \hat{\Delta}^{1/2} = \hat{\Psi}^{1/2} \hat{E} (\hat{\Lambda} - I)^{1/2} \tag{3.26}$$

3. Substitute $\hat{L}$ obtained in (3.26), and minimize the result with respect to $\hat{\psi}_1, \cdots, \hat{\psi}_p$. The values $\hat{\psi}_1, \cdots, \hat{\psi}_p$ obtained from this minimization are employed at Step 2 to create a new $\hat{L}$ and repeat step2, step3 until convergence.

For more details,the original work of Lawley and the more recent work of Jöreskog are recommended.

3.1.3 Factor Rotation

Since the original loadings may not easy to be understood, it is usual to rotate them until a simple structure is achieved. Most of the rationale for rotating factors was presented by Thurstone [36] and Cattell [7] in 1947 and 1978. They advocated that using rotation could simplify the procedure of the factor structure and hence made it easier and more reliable to be interpreted. Meanwhile, in terms of the reports by Thurstone, he suggested five criteria to identify a simple structure which is list below:
A loadings' matrix is straightforward if
1. Each row includes no less than one zero.
2. There exist many zeros for each column.
3. There exist some variables with zero loadings on one factor whereas larger loadings on other factor for any pair of factors.
4. There exists a fairly large proportion of zero loadings for any pair of factors.
5. There exist merely a few large loadings for any pair of factors.

- **Orthogonal Rotation**

Consulting [17], from matrix algebra, an orthogonal transformation correlate with rigid rotation (or reflection) of the coordinate axes. To this end, an orthogonal transformation of the factor loadings, implied orthogonal transformation of the factors as well, is named factor rotation.

If $\hat{L}$ is the $p \times m$ matrix of estimated factor loadings given by any method then

$$\hat{L}^* = \hat{L}T, \qquad where \qquad TT' = T'T = I \tag{3.27}$$

is a $p \times m$ matrix of "rotated" loadings. In addition, the estimated covariance matrix keeps unchanged, because

$$\hat{L}\hat{L}' + \hat{\Psi} = \hat{L}TT'\hat{L} + \hat{\Psi} = \hat{L}^*\hat{L}^{*'} + \hat{\Psi} \tag{3.28}$$

There are two methods for determining an orthogonal rotation to simple structure, graphical and analytical method. In this study, analytical method will be used.

- **Varimax**

In 1958, Kaiser [15] presented an analytical measure of simple structure known as the varimax (or normal varimax) criterion. Generally, through using varimax, a simple solution with a great amount of zeros (or small) loadings and a few large loadings for each factor can be obtained. The main idea of varimax is defining $\tilde{l}^*_{ij} = \hat{l}^*_{ij}/\hat{h}_i$ to be the rotated coefficients scaled by the square root of the communalities. Then the (normal) varimax procedure chooses the orthogonal transformation T that makes

$$V = \frac{1}{p}\sum_{j=1}^{m}\left[\sum_{i=1}^{p}\tilde{l}^{*4}_{ij} - \left(\sum_{i=1}^{p}\tilde{l}^{*2}_{ij}\right)^2 /p\right] \tag{3.29}$$

as big as possible.

Although scaling the rotated coefficients $\hat{l}^*_{ij}$ has the influence of giving variables with small communalities comparatively more weight in the determination of simple structure, nevertheless, after fixing the transformation T, the loadings $\tilde{l}^*_{ij}$ are multiplied by $\hat{h}_i$ so that the original communalities are remained.

In spite of the fact that (3.29) seems very hard to understand, it has a simple explanation referring to the interpretation of Johnson, R.A. and Wichern, D.W. [17]:
"...looks rather...In words,

$$V \propto \sum_{j=1}^{m}(\text{variance of squares of (scaled) loadings for jth factor})$$

..., maximizing V corresponds to 'spreading out' the squares of the loadings on each factor as much as possible. Therefore, we hope to find groups of large and negligible coefficients in any column of the rotated loadings matrix $\hat{L}^*$."

3.1.4 Factor Scores

According to [17], in factor analysis, analysts generally concentrate on the parameters in the factor model. Nonetheless, the estimated values of the common factors which are named factor scores, may also be needed. There two approaches to obtain the factor scores $\hat{f}_j$, the weighted least squares method and the regression method. For this project, factor scores will be obtained by regression method.

Beginning again with the original factor model $X - \mu = LF + \epsilon$, firstly, considering the loadings matrix L and specific variance matrix Ψ as known when the common factors F and specific factors (or errors) ϵ are jointly normally distributed with means and covariances given by (3.3), then, the linear combination $X - \mu = LF + \epsilon$ has an $N_p(0, LL' + \Psi)$ distribution. In addition, the joint distribution of (X - μ) and F in $N_{m+p}(0, \Sigma^*)$, where

$$\underset{(m+p)\times(m+p)}{\Sigma^*} = \left[\begin{array}{c:c} \Sigma = LL'_{(p\times p)} + \Psi & L_{(p\times m)} \\ \hdashline L'_{(m\times p)} & I_{(m\times m)} \end{array}\right] \tag{3.30}$$

and 0 is an $(m + p) \times 1$ vector of zeros. Meantime, the conditional distribution of $F|X$ is multivariate normal with

$$\text{mean} = E(F|X) = L'\Sigma(X - \mu) = L'(LL' + \Psi)^{-1}(X - \mu) \tag{3.31}$$

and

$$\text{covariance} = \text{Cov}(F|X) = I - L'\Sigma^{-1}L = I - L'(LL' + \Psi)^{-1}L \tag{3.32}$$

The amount $L'(LL' + \Psi)^{-1}$ in (3.31) are the coefficients in a (multivariate) regression of the factors on the variables. Therefore, factor scores can be obtained by estimating these coefficients.

Then, the jth factor score vector is given by

$$\hat{f}_j = \underset{(m\times p)}{\hat{L}'} \underset{(p\times p)}{(\hat{L}\hat{L}' + \hat{\Psi})^{-1}} \underset{(p\times 1)}{(x_j - \bar{X})} = \underset{(m\times m)}{(I + \hat{L}'\hat{\Psi}^{-1}\hat{L})^{-1}} \underset{(m\times p)}{L'} \underset{(p\times p)}{\hat{\Psi}^{-1}} \underset{(p\times 1)}{(x_j - \bar{X})} \qquad j = 1, 2, \cdots, n \tag{3.33}$$

In order to decrease the effects of a (possibly) incorrect determination of the number of factors, researchers tend to calculate the factor scores in (3.31) by using S (the original sample covariance matrix) instead of $\Sigma = \hat{L}\hat{L}' + \hat{\Psi}$. Then, the factor score vector is given by

$$\hat{f}_j = \hat{L}'S^{-1}Z_j \qquad j = 1, 2, \cdots, n \tag{3.34}$$

where, Z_j are standardized variables. (See (3.15).)

3.2 Financial Indicators

Many indicators are used in financial statements. In this project, ten indicators depicting distinct capabilities of an stock will be applied for the purpose of actor analysis, that is, the value of these indicators for each stock will be considered as original data set to be investigated. Afterwards, the relationships between the factors and the indicators can be displayed by using various approaches. Subsequently, through analysing these connections, the hidden meanings of factors can be revealed. Eventually, the performance of individual stock will be evaluated by the factor scores instead of these indicators. The following itemize the definitions of these indicators in terms of [31]:

- **Assets**
  Assets are investments that are expected to generate payoff.
- **Shareholder's Equity**
  Stockholder's Equity is the claim by the owners. Liabilities are claims to the payoffs by claimants other than owners. The three parts of the balance sheet are tied together in the following accounting relation

$$\text{Shareholder's Equity} = \text{Assets} - \text{Liabilities}$$

- **Net Income**
  Net income measures the value added to shareholder's equity. Expenses are the value go out in earning revenue. The accounting relation which determines net income is

$$\text{Net Income} = \text{Revenues} - \text{Expenses}$$

where revenues is the value coming in from selling products.

- **Return on Equity (ROE)**
  Return on Equity (ROE) measures the rate of return on the ownership interest (shareholders' equity) of the common stock owners. The formula of ROE is

$$\text{ROE} = \frac{\text{Net Income (after tax)}}{\text{Shareholder's Equity}}$$

- **Operating Margin**

  Operating Margin is a measurement of what proportion of a company's revenue is left over, before taxes and other indirect costs (such as rent, bonus, interest, etc.), after paying for variable costs of production as wages, raw materials, etc. The formula of Operating Margin is

$$\text{Operating Marigin} = \frac{\text{Operating Income}}{\text{Revenues}}$$

- **Earnings Per Share (EPS)**

  EPS is the amount of earnings per each outstanding share of a company's stock. The basic formula of EPS is

$$\text{EPS} = \frac{\text{Profit}}{\text{Weighted Average Common Shares}}$$

where

$$\text{Profit} = \text{Earnings (after tax)} - \text{Payable Dividends}$$

- **Net Asset Value Per Share (NAVPS)**

  NAVPS is an expression for net asset value that represents a fund's (mutual, exchange-traded, and closed-end) or a company's value per share. The formula of NAVPS is

$$\text{NAVPS} = \frac{\text{Net Asset Value}}{\text{The Number of Shares Outstanding}}$$

  Usually, net asset value is equal to shareholders' equity. Furthermore, shares outstanding are shares that have been authorized, issued, purchased and are held by investors.

- **The Growth Rate of Net Income (GRONI)**

  Net income is the profit of a firm for a period. The growth rate of net income describe how much the net income changes from one period to another. This period can be any period of time, such as months, seasons, years. Managers should inspect the growth rate of net income to determine whether their firm is growing at a sustainable rate.

- **The Growth Rate of Main Business Revenue (GROMBR)**

  Taking the derivative of the main business revenue growth tells you how much the main business revenue is changing. This indictor can indicate the growth potentiality of a company.

- **The Growth Rate of Shareholder's Equity (GROSE)**

  The growth rate of Shareholder's Equity shows how much the Shareholder's equity increase from one period to another. This indictor can reflect the enterprises' development capability.

3.3 Empirical Results

3.3.1 Data Set

The data consists of the annual report (2011) based on Shanghai Securities Composite Index 50 (SSE 50) which can be found online, such as the website of Shanghai Stock Exchange or Yahoo Finance. More information is listed in Appendix A.

3.3.2 Factor Analysis Test

Usually, the data set should be tested whether it is suitable for factor analysis before using. There are mainly two ways to carry out the factor analysis test, namely, Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity.

- **Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy**

The KMO measure of sampling adequacy [21] is an index applied to test the appropriateness of factor analysis. High values (between 0.5 and 1) shows factor analysis is appropriate. Relatively, value below 0.5 hints that factor analysis may not be appropriate.

- **Bartlett's Test of Sphericity**

Bartlett's test of sphericity [3] is a statistical method used to examine the hypothesis that the variables are uncorrelated in the population. Namely, the population correlation matrix is an identity matrix, each variable correlates perfectly with itself, i.e. the diagonal elements are equal to 1, while it has no correlation with other variable, that is, other elements are 0. Usually, if the p-value of Bartlett's Test of Sphericity Approx. Chi-Square is less than 0.05, the factor analysis is appropriate and vice versa.

3.3.3 Results

Before we implement the factor analysis, the data should be standardized according to equation (3.15) since the variables are not commensurate. Meantime, six stocks will be excluded because their annual reports haven't published or there exists some extreme numeric in their financial report. Therefore, actually, there are merely 44 stocks which will be investigated with the help of factor analysis. The details of the data set (after standardization) can be found in Appendix B.

Table 2.1 shows the factor test for the data set. It is clearly to see the value of KMO measure is 0.6607339(>0.5) and the p-value of Bartlett's test is equal to zero. Consequently, this data set is suitable for factor analysis.

Table 3.1: The KMO Measure and Bartlett's Test

KMO measure of sampling adequacy	0.66073
Bartlett's Test of Sphericity Approx. Chi-Square	405.371
Df	45
P-value	0

The comparison of three methods are presented in Table 3.2. All approaches produced identical loadings, as expected. Since the original loadings are not readily interpretable, therefore, varmix method is used to rotate them until a simple structure is attained. After rotation, we can obtain the following information.

1. Factor 1 reflects the enterprise's development capability because it is obviously to see the values of GRONI,GROMBR,GROSE are larger and positive. As we mentioned in 3.2, these indicators can indicate the growth potentiality of a enterprise. Stocks with higher growth ability are typically desirable to invest than others.

2. Factor 2 shows the scale of firms since this factor has a higher positive relationship with net income, asset and shareholder's equity. Usually, a firm with bigger scale, the stronger power to resist risk it will have.

3. Factor 3 displays the ability of per share due to large numeric in EPS and NAVPS. Both this two indicators can represent the efficiency and operating performance of a company. Basically, a stock with higher EPS and NAVPS is worthy to be considered as investment.

4. Factor 4 relates to the profitability based on larger value of Operating Margin. This indicator is one of the major indexes to evaluate the profitability of a corporation. Definitely, investors prefer to put funds on a business which has great earning capacity.

Otherwise, through table 3.2, we can also view there is no apparent difference among the three methods, in other words, it doesn't matter which approaches selected to implement the factor analysis. Nevertheless, principal factor procedure is recommended since it has the smallest value of summing the error square. (See table 3.3.) Actually, in practice, we prefer using loadings to "explain" the covariance (referring to (3.5)) as much as possible accompanying with small errors, that is, higher "Cumulative Var", lower "The Sum of Error Square"[1]. Therefore, associating with table 3.2 as well as 3.3, it is easy to see the performance of principle factor method is relatively better.

[1]Suppose $\Sigma - (LL' + \Psi) = (\sigma_{ij})_{p\times p}$, then the sum of error square is $\sum_{i=1}^{P}\sum_{j=1}^{P}\sigma_{ij}^2$

Table 3.2: The Comparison of Three Methods

Before Rotation	Principal Component				Principal Factor				Maximum Likelihood			
	Fact 1	Factor2	Factor3	Factor4	Factor1	Factor2	Factor3	Factor4	Factor1	Factor2	Factor3	Factor4
Net Income	0.181	0.920	-0.190	-0.234	0.192	0.944	-0.200	-0.178	0.582	0.796	0.018	0.151
GRONI	0.817	-0.254	-0.312	0.006	0.780	-0.247	-0.297	-0.012	0.379	-0.298	0.692	-0.147
Return on Equity (ROE)	0.888	-0.160	-0.168	-0.088	0.865	-0.163	-0.180	-0.090	0.533	-0.298	0.703	0.014
Asset	0.176	0.807	-0.217	0.330	0.168	0.705	-0.165	0.242	0.511	0.531	-0.103	-0.213
Shareholder's Equity	-0.004	0.912	-0.191	-0.310	0.004	0.919	-0.188	-0.234	0.420	0.851	-0.068	0.196
Operating Margin	0.666	0.350	0.092	0.592	0.666	0.323	0.078	0.580	0.854	-0.135	-0.027	-0.496
Earnings Per Share (EPS)	0.718	0.116	0.641	-0.088	0.734	0.105	0.633	-0.109	0.771	-0.471	-0.017	0.422
NAVPS	0.437	0.240	0.826	-0.118	0.438	0.222	0.761	-0.108	0.622	-0.339	-0.299	0.454
GROMBR	0.819	-0.182	-0.195	-0.253	0.770	-0.174	-0.185	-0.190	0.384	-0.256	0.677	0.079
GROSE	0.849	-0.267	-0.292	-0.018	0.833	-0.271	-0.304	-0.038	0.386	-0.318	0.761	-0.147
SS loadings	4.060	2.717	1.470	0.704	3.882	2.590	1.336	0.550	3.213	2.339	2.117	0.786
Proportion Var	0.406	0.272	0.147	0.070	0.388	0.259	0.134	0.055	0.321	0.234	0.212	0.079
Cumulative Var	0.406	0.678	0.825	0.895	0.388	0.647	0.781	0.836	0.321	0.555	0.767	0.845
After Rotation	Fact 1	Factor2	Factor3	Factor4	Factor1	Factor2	Factor3	Factor4	Factor1	Factor2	Factor3	Factor4
Net Income	-	0.972	-	0.124	-	0.992	-	-	-	0.99	-	-
GRONI	0.892	-	-	0.165	0.855	-	-	0.15	0.837	-	-	0.163
Return on Equity (ROE)	0.89	-	0.207	0.128	0.871	-	0.2	0.122	0.897	-	0.224	0.112
Asset	-	0.689	-	0.6	-	0.661	-	0.416		0.669	-	0.389
Shareholder's Equity	-0.111	0.975	-	-	-0.116	0.959	-	-	-0.118	0.963	-	-
Operating Margin	0.334	0.146	0.307	0.835	0.344	0.214	0.29	0.801	0.337	0.221	0.287	0.866
Earnings Per Share (EPS)	0.354	-	0.893	0.155	0.36	-	0.9	0.148	0.361	-	0.919	0.141
NAVPS	-	-	0.966	-	-	-	0.904	-	-	-	0.885	-
GROMBR	0.877	-	0.18	-	0.814	-	0.174	-	0.806	-	0.163	0.149
GROSE	0.92	-	-	0.148	0.914	-	-	0.135	0.906	-	-	-
SS loadings	3.453	2.411	1.915	1.172	3.251	2.403	1.796	0.909	3.236	2.421	1.8	0.998
Proportion Var	0.345	0.241	0.191	0.117	0.325	0.24	0.18	0.091	0.324	0.242	0.18	0.1
Cumulative Var	0.345	0.586	0.778	0.895	0.325	0.565	0.745	0.836	0.324	0.566	0.746	0.845

Table 3.3: **The Sum of Error Square**

Principal Component	Principal Factor	Maximum Likelihood
0.1168274	0.02444969	0.03648019

Afterwards, for the sake of evaluating the compound performance of each stock, the composite score is given by

$$\text{The Composite Score} = \sum_{i=1}^{4} p_i \hat{f}_i$$

where p_i is the community rate of each factor, i.e. the proportion of each factor's community to total variance which can be found in table 3.2, called "Proportion Var ". $\hat{f}_i$ is the factor score on factor i in term of section 3.1.4.

Table 3.4: The Composite Score of Principal Factor Method

Stock	Score	Stock	Score	Stock	Score	Stock	Score	Stock	Score
600519	1.604685	600015	0.34421	601699	-0.09481	601111	-0.36914	601958	-0.50622
600111	0.807335	601169	0.304241	601628	-0.12221	600058	-0.38135	600050	-0.52686
601166	0.761795	600028	0.293051	600383	-0.13513	600019	-0.38515	601600	-0.61605
600036	0.653073	600585	0.223475	600837	-0.18575	601601	-0.4266	600010	-0.62808
601328	0.638723	600030	0.211008	601898	-0.20559	601766	-0.46239		
600048	0.597853	601818	0.158095	600348	-0.20635	601299	-0.46496		
600000	0.597497	600104	0.067217	600547	-0.2143	601989	-0.4739		
601318	0.553414	600188	0.051229	600489	-0.24137	600089	-0.47647		
601088	0.51066	600362	0.019318	601668	-0.31396	600068	-0.48032		
600016	0.398447	600031	-0.04622	601899	-0.33587	601168	-0.49627		

As we mentioned before, the principal factor method performs comparatively better than the others. Consequently, in this case, we solely calculate the composite score of this approach. Eventually, according to the order of table 3.4, the top 10 stocks are employed to the portfolio selection which will be presented shortly.

To sum up, factor analysis is an utilitarian approach in portfolio selection since it can help an investor to know the situations of an individual stock, including the stock's growth potentiality, profitability and so on. Meanwhile, according to the factor scores, an investor can predict the future behaviour of an asset and whether it is worthy to be selected for a portfolio. In a word, through using factor analysis, it can supply a set of assets which perform better than others to an investor for considering as a portfolio to be optimised.

4 Comparison of Three Approaches in Portfolio Selection

After choosing the assets for portfolio, an investor may meet another puzzle, how to allocate these assets. For the sake of solving this problem. In this section, the details of three different portfolio selections will be introduced. Meanwhile, some comparison of the three portfolio optimization will be presented, including the efficient frontier, the portfolio compositions and performance such as the Sharpe ratio and Sortino ratio.

4.1 The Mean-Variance (MV) Portfolio

The MV method is the classical approach to use variance as the measure of risk while mean return as expected return proposed by Markovitz in 1952. The definition of MV model is :

$$\min \quad \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}$$

subject to

$$\sum_{i=1}^{n} r_i w_i = R$$

$$\sum_{i=1}^{n} w_i = 1$$

$$0 \leq w_i \leq 1 \; for \; i = 1, 2, ..., n$$

where σ_{ij} is the covariance between assets i and j, w_i is the weight of asset i, r_i is the expected return on asset i and R is the expected return given by an investor.

The merit of Markowitz model is simplicity. This model is very easy to be understood and used since it is just composed of two summary statistics which are variance and mean. Nonetheless, this model has some demerits. Such as it is extremely monotonous to calculate the covariance of assets for constructing this model and hard to resolve this quadratic programming model if the problem is in a large scale. Furthermore, MV model often performs poorly out-of-sample due to estimation errors in the mean vector and covariance matrix. Consequently, it may affect the stability of the model. Therefore, in order to improve this downside, robust method will be introduced shortly.

4.2 The MCD Robustified Mean-Variance Portfolio

In order to achieve better stability properties compared to traditional minimum variance portfolios, the minimum covariance determinant (MCD) robustified method was used to compute the mean and (or) covariance matrix for reducing the estimation error.

4.2.1 The MCD estimator

The MCD method is a highly robust estimator of location and scatter [33]. Given an $n \times p$ data matrix $X = (x_1, \cdots, x_n)'$ with $x_i = (x_{i1}, \cdots, x_ip)'$, it is focus on to find the h (with $\lfloor (n+p+1)/2 \rfloor \leq h \leq n$) observations whose classical covariance matrix has the lowest possible determinant. Then, the MCD estimate of location is the average of these h points, whereas the MCD estimate of scatter is their covariance matrix.

Although the MCD was already introduced in 1984, its application became feasible until the efficient algorithm was presented by Rousseeuw and Van Driessen [34] since there exited an obstruction, that is, the previous algorithms were limited to a few hundred objects in a few dimensions. Therefore, in order to deal with such problems, Rousseeuw and Van Driessen developed a new algorithm, called fast MCD.

4.2.2 The Fast MCD Algorithm

- **The Concentration Step (C-Step)**

A key component of fast MCD algorithm is the concentration step (C-step). The C-step can be described as follows [34]:

Consider a data set $X = \{x_1, x_2, \cdots, x_n\}$ of p-variant observations. Given the initial approximation $(\hat{\mu}_{\text{old}}, \hat{\Sigma}_{\text{old}})$ for the center and scatter matrix.

1. Computer the distance

$$d_{\text{old}}(i) = D(x_i, \hat{\mu}_{old}, \hat{\Sigma}_{\text{old}}) = \sqrt{(x_i - \hat{\mu}_{\text{old}})' \hat{\Sigma}^{-1} (x_i - \hat{\mu}_{\text{old}})}$$

2. Sort these distances, yielding a permutation π for which

$$d_{\text{old}}(\pi(1)) \leq d_{\text{old}}(\pi(2)) \leq \cdots \leq d_{\text{old}}(\pi(n))$$

3. Set $H = \{\pi(1), \pi(2), \cdots, \pi(h)\}$.
4. Compute $\hat{\mu}_{\text{new}} = \frac{1}{h} \sum_{i \in H} x_i$ and $\hat{\Sigma}_{\text{new}} = \frac{1}{h} \sum_{i \in H} (x_i - \hat{\mu}_{\text{new}})(x_i - \hat{\mu}_{\text{new}})'$

- **The Process of Fast MCD Algorithm**

If C-steps were applied iteratively, the sequence of determinants obtained in this way must converge in a finite number of steps since there are only finitely a few h-subsets. Afterwards, the determinant is no longer reduced by running the C-step. Therefore, there is no guarantee that the value of the iteration procedure satisfies the global minimum of the MCD objective function. Nevertheless, it is an essential condition. Fortunately, Rousseeuw and Van Driessen [34] have proved that $\det(\hat{\Sigma}_{\text{new}}) \leq \det(\hat{\Sigma}_{\text{old}})$, with equality only if $\hat{\Sigma}_{\text{new}} = \hat{\Sigma}_{\text{old}}$. It hence provides a partial idea of an algorithm, i.e. fast MCD algorithm. The concept of this algorithm is taking many initial h-subsets $H_1 \subset \{1, 2, \cdots, n\}$, implementing C-steps to each until convergence, and retaining the solution with the overall lowest determinant.

Consequently, when $n < 600$, the process of the fast MCD algorithm is:
1. Set an initial h-subset H_1, that is, beginning with a a random $(p+1)$-subset J.
2. Compute the $\hat{\mu}_0 = \frac{1}{p+1} \sum_{i \in J} x_i$ and $\hat{\Sigma}_0 = \frac{1}{p+1} \sum_{i \in J} (x_i - \hat{\mu}_0)(x_i - \hat{\mu}_0)'$. If $\det(\Sigma_0) = 0$, random observations are added to J until $\det(\Sigma_0) > 0$.
3. Apply the C-step to the initial h-subset H_1, and obtain the $(\hat{\mu}_1, \hat{\Sigma}_1)$. If $\det(\hat{\Sigma}_1) = 0$ or $\det(\hat{\Sigma}_2) = \det(\hat{\Sigma}_1)$, stop; otherwise, running another C-step produces $\det(\hat{\Sigma}_3)$, and so on, until convergence is reached.

Particularly, since reducing the number of C-steps would improve the speed, the fast MCD algorithm only employs two C-steps to each initial subset, and just select ten subsets with lowest determinant whose further C-steps are taken until convergence. Then, the solution $(\hat{\mu}, \hat{\Sigma})$ with lowest $\det(\hat{\Sigma})$ is selected as the raw fast MCD estimates, i.e. $\hat{\mu}_{\text{rawMCD}}$ and $\hat{\Sigma}_{\text{rawMCD}}$ which relate to the empirical mean and covariance matrix of the h-subset.

4.2.3 The MCD Robustified Mean-Variance Model

Using the $\hat{\Sigma}_{\text{rawMCD}} = (\tilde{\sigma}_{ij})_{(n\times n)}$ to replace the sample covariance matrix into MV model, then, the MCD Robustified Mean-Variance Model is given by :

$$\min \quad \sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{w}_i\tilde{w}_j\tilde{\sigma}_{ij}$$

subject to

$$\sum_{i=1}^{n}\tilde{r}_i\tilde{w}_i = R$$

$$\sum_{i=1}^{n}\tilde{w}_i = 1$$

$$0 \leq \tilde{w}_i \leq 1 \ for \ i = 1, 2, ..., n$$

where $\tilde{\sigma}_{ij}$ is the raw fast MCD covariance between assets i and j, $\tilde{w}_i$ is the weight on asset i, $\tilde{r}_i$ is the expected return of asset i per period and R is the expected return required by an investor.

4.3 The Mean-CVaR Portfolio

Unlike the MV model, the mean-CVaR uses the Conditional Value at Risk as a risk measure. Since it is similar to VaR but it estimates “how bad things can get” if the VaR loss is exceeded, CVaR has become a popular risk measure. Moreover, CVaR has an additional advantage is that it is easy to optimise the portfolio weights through linear programming to minimise CVaR.

4.3.1 The Mean-CVaR Model

Before formulating the mean-CVaR optimization problem, some mathematical properties of CVaR measure will be described. In terms of [30], let w denote the portfolio vector of weights such that each component w_i equals the weight of asset i. At the same time, a random vector is marked to describe the uncertain outcomes by y. Therefore, the loss associated with the portfolio vector w is represented by the function $f(w, y)$. Meanwhile, let $p(y)$ be the probability associated with scenario y.

Now, assuming the probability that the loss function does not exceed a certain value γ is given by the cumulative probability

$$\Psi(x, y) = \int_{f(w,y)\leq\gamma} p(y)dy$$

Using this cumulative probability, the VaR can be describe as follow

$$\text{VaR}_\alpha(w) = \min\{\gamma|\Psi(x, \gamma) \geq \alpha\}$$

Since CVaR of the losses of portfolio w is the expected value of losses conditioned on the losses being in excess of VaR, the CVaR is give by

$$\begin{aligned}\text{CVaR}_\alpha(w) &= E(f(w, y)|f(w, y) \geq \text{VaR}_\alpha(w)) \\ &= \frac{1}{1-\alpha}\int_{f(w,y)\geq\text{VaR}_\alpha(w)} f(w, y)p(y)dy\end{aligned}$$

Moreover, it is easy to see

$$\begin{aligned}\text{CVaR}_\alpha(w) &= \frac{1}{1-\alpha}\int\limits_{f(w,y)\geq \text{VaR}_\alpha(w)} f(w,y)p(y)dy \\ &\geq \frac{1}{1-\alpha}\int\limits_{f(w,y)\geq \text{VaR}_\alpha(w)} \text{VaR}_\alpha(w)p(y)dy \\ &= \text{VaR}_\alpha(w)\end{aligned}$$

because

$$\frac{1}{1-\alpha}\int\limits_{f(w,y)\geq \text{VaR}_\alpha(w)} p(y)dy = 1$$

Namely, CVaR is always at least as large as VaR. Nonetheless, CVaR is a coherent risk measure while VaR is not. It can also be shown that CVaR is a concave function, so that it has a unique minimum. Nonetheless, working directly with the above formulas turns out to be rather tricky in practice as they involve the VaR function. Fortunately, in 2000, Rockefellar and Uryaserv [32][37] presented a new simple approach to optimize the CVaR.

Their idea is to denote the α-VaR and α-CVaR values for the loss random variable connected with w and particular probability level α in $(0,1)$ by $\zeta_\alpha(w)$ and $\phi_\alpha(w)$. Meanwhile, they are given by

$$\zeta_\alpha(w) = \min\{\zeta \in \mathbb{R} : \Psi(w,\zeta) \geq \alpha\}$$

and

$$\phi_\alpha(w) = \frac{1}{1-\alpha}\int\limits_{f(w,y)\geq \zeta_\alpha(w)} f(w,y)p(y)dy$$

The key to their approach is use the function F_α which defined by

$$F_\alpha(w,\zeta) = \zeta + \frac{1}{1-\alpha}\int_{y\in R^m} [f(w,y)-\zeta]^+ p(y)dy$$

to replace the characterization of $\zeta_\alpha(w)$ and $\phi_\alpha(w)$. Specifically, they proved the following two important theorem.

Theorem 1. $F_\alpha(w,\zeta)$ is a convex and continuously differentiable function of ζ.

Theorem 2. Minimizing the α-CVaR of the loss associated with w over all $w \in W$ is equivalent to minimizing $F_\alpha(w,\zeta)$ over all $(x,\zeta) \in X \times \mathbb{R}$ in the sense that

$$\min_{w\in W} \text{CVaR}_\alpha(w) = \min_{(w,\zeta)\in W\times\mathbb{R}} F_\alpha(w,\zeta).$$

Furthermore, the integral in $F_\alpha(w,\zeta)$ can be estimated in different ways. One of them is sampling the probability distribution of y in term of its density p(y) to do the integration. If the sampling generate a set of vectors $y_1, y_2, \cdots, y_k$, then the corresponding approximation is

$$\tilde{F}_\alpha(w,\zeta) = \zeta + \frac{1}{1-\alpha}\sum_{j=1}^{k} \pi_j [f(w,y)-\zeta]^+$$

where π_j are probabilities of scenarios y_j.

Then, by using auxiliary real variables $z_j,\ j = 1,\cdots,k$, the function $\tilde{F}_\alpha(w,\zeta)$ can be described by the linear function $\zeta + \frac{1}{1-\alpha}\sum_{j=1}^{k}\pi_j z_j$ and the collocation of linear constraints is

$$z_j \geq f(w,y_j) - \zeta, \qquad z_j \geq 0, j = 1,\cdots,k, \qquad \zeta \in \mathbb{R}$$

Now, consider the case where the decision vector w represent the portfolio vector of weights in the sense that $W = (w_1, \cdots, w_n)$ with w_i being the weight of asset i. Meantime, using y_i to denote the return on asset i, then, $Y = (y_1, \cdots, y_n)$. It is easy to get the loss function is given by

$$f(w, y) = -[w_1 y_1 + \cdots + w_n y_n] = -WY^T$$

In practice, the probability density function $p(y)$ is often not available, or is very difficult to estimate. Instead, K different scenarios $Y_1, \cdots, Y_K$ are sampled from the probability distribution or that have been obtained from computer simulations. In this paper, π_j is defined by K^{-1} for $j = 1, \cdots, K$.

Therefore the mean-CVaR model is obtained according to Theorem 2, that is

$$\min_{(w,\zeta)\in W\times\mathbb{R}} \quad \zeta + \frac{1}{1-\alpha} K^{-1} \sum_{j=1}^{K} z_j$$

subject to

$$z_j \geq -WY_j^T - \zeta$$

$$\sum_{i=1}^{n} r_i w_i = R$$

$$\sum_{i=1}^{n} w_i = 1$$

$$0 \leq w_i \leq 1 \ for \ i = 1, 2, ..., n$$

$$z_j \geq 0, \ for \ j = 1, \cdots, K$$

where w_i is the weight of asset i, ζ is the value of $\alpha - \text{VaR}$. z_j is the auxiliary variables in scenario j, r_i is the expected return on asset i and R is the expected return required by an investor.

As a consequence, CVaR_α can be optimized via optimization of the function $F_\alpha(w, \zeta)$ with respect to the weights w and ζ. In terms of theorem 1, $F_\alpha(w, \zeta)$ is a convex function of ζ. Meanwhile, when $f(x, y)$ is convex, the $F_\alpha(w, \zeta)$ is also a convex function of w. Therefore, the optimization problems become smooth convex optimization problems that can be resolved using common optimization approaches for such problems.

4.4 Empirical Results

4.4.1 Scenario Generation

Since the CVaR function is estimated by the weighed total over all scenarios, so first of all we should to generate the scenarios. There are many approaches to create the scenarios, such as Monte Carlo simulation and historical scenario generation. In our project, we will use the latter one which is introduced by Uryaserv [30] to simulate the scenarios. Meanwhile, in this case, we want to optimize the portfolio merely for one day period. Therefore, the set of K historical Logarithmic returns are given by

$$r_{ij} = \ln \left(\frac{p_i^{t_j+1}}{p^{t_j}} \right) \qquad j = 1, \cdots, K$$

where p_i^t is the historical price of stock i in time t. Particularly, in this case, we use the closing price to compute the Logarithmic returns.

Moreover, assuming that all scenarios r_{ij} are equally probable, i.e. $\pi_j = 1/K$, then, the expected end-of-period return of stock i is

$$E(r_i) = \sum_{j=1}^{K} \pi_j r_{ij} = K^{-1} \sum_{j=1}^{K} r_{ij}$$

4.4.2 Results

The data consisting of 120 historical daily Logarithmic returns (November 14,2011 - May 03, 2012) of ten stocks chosen by section 3 is used as scenarios to compare the performance of three portfolios for one day period without considering any supposition about the distribution. Meanwhile, according to the Chinese benchmark interest rate, the risk free rate denoted by R_f is 0.01% , the minimum acceptable rate of return (MAR) is 0.1% , $\alpha = 95\%$ and $h = 0.75n = 90$ for the fast MCD algorithm since this yields a higher finite-sample efficiency in terms of [8]. Furthermore, for our case, "Short Selling" is not allowed. Appendix C lists the stocks' daily closing prices and their Logarithmic returns can be found in Appendix D as well.

Especially, in this project, Sharpe ratio and Sortino Ratio are employed to evaluate the performance of the three portfolios. In practice, the higher Sharpe (Sortino) ratio it has, the better performance a portfolio will have.

Table 4.1 presents the contrast of three portfolios for different given expected returns, including the standard deviation, CVaR, Sharpe ratio and Sortino ratio respectively, while table 4.2 shows their compositions. Meanwhile, their performance can be directly viewed through the figure 4.1, 4.2, 4.3 and 4.4 separately.

As expected, associating with figure 4.1, 4.2, it is easy to find that the MCD robustified mean-variance portfolio has the lowest standard deviation coinciding with the smallest CVaR of mean-CVaR portfolio when the target return is given. Apparently, when the return is given, the MCD portfolio will have the largest Sharpe ratio due to the minimal divergence which also can be seen in terms of figure 4.3. In effect, through using robust method, it usually can reduce the variance. However, it maybe leads to more bias, especially when the sample size is relatively small since fast MCD algorithm just chooses the "best" case to compute the covariance. For the sake of dealing with this problem, commonly, the raw MCD covariance will be multiplied by a consistency factor given through computer simulation. Meanwhile, in order to get a fixed subset with the lowest determinant, we implement the fast MCD algorithm 10000 times through programming. In addition, according to figure 4.2, we can find although this method can decrease the variance, it will result the larger value of CVaR than other approaches.

Moreover, no surprise that mean-CVaR portfolio has the higher standard deviation because the aim of CVaR optimization is to reshape one tail of loss distribution. It doesn't consider the opposite tail corresponding to high profit. One the contrary, both MV and MCD-MV approaches treat the variance as risk measure which incorporate information from both tails. However, the difference between MV and CVaR portfolio is not significant, especially when the return increases, it indicates these two portfolios become similar linking up all figures.

Otherwise, connecting with the portfolio compositions (table 4.2), we can find that the front stocks have more weights than the tail parts. Consequently, it maybe implies that using the factor analysis method to evaluate single stock performance is feasible.

Table 4.1: Portfolio Performances for Given Returns

Expected Return	0.0012			0.0014			0.0016		
Portfolio	MV	MCD	CVaR	MV	MCD	CVaR	MV	MCD	CVaR
Standard Deviation	0.0101	0.0091	0.0107	0.0110	0.0097	0.0116	0.0120	0.0104	0.0125
CVaR	0.0179	0.0195	0.0164	0.0195	0.0211	0.0177	0.0210	0.0228	0.0191
Sharpe Ratio	0.1089	0.1209	0.1028	0.1182	0.1340	0.1121	0.1250	0.1442	0.1200
Sortino Ratio	0.0210	0.0194	0.0207	0.0390	0.0371	0.0397	0.0546	0.0520	0.0537
Expected Return	0.0018			0.002			0.0022		
Portfolio	MV	MCD	CVaR	MV	MCD	CVaR	MV	MCD	CVaR
Standard Deviation	0.0131	0.0112	0.0135	0.0144	0.0121	0.0148	0.0159	0.0132	0.0160
CVaR	0.0224	0.0249	0.0211	0.0243	0.0275	0.0233	0.0264	0.0301	0.0256
Sharpe Ratio	0.1298	0.1518	0.1259	0.1319	0.1570	0.1284	0.1321	0.1591	0.1313
Sortino Ratio	0.0675	0.0641	0.0664	0.0785	0.0739	0.0762	0.0861	0.0832	0.0876
Expected Return	0.0024			0.0026			0.0028		
Portfolio	MV	MCD	CVaR	MV	MCD	CVaR	MV	MCD	CVaR
Standard Deviation	0.0174	0.0144	0.0174	0.0190	0.0156	0.0191	0.0208	0.0169	0.0208
CVaR	0.0287	0.0329	0.0284	0.0321	0.0362	0.0319	0.0358	0.0395	0.0357
Sharpe Ratio	0.1322	0.1597	0.1322	0.1316	0.1603	0.1309	0.1298	0.1598	0.1298
Sortino Ratio	0.0949	0.0904	0.0940	0.1000	0.0963	0.0990	0.1038	0.0997	0.1038

Table 4.2: Portfolio Compositions

Expected Return	0.0012			0.0014			0.0016		
Portfolio	MV	MCD	CVaR	MV	MCD	CVaR	MV	MCD	CVaR
600519	0.1765	0.3083	0.1614	0.1664	0.3119	0.1591	0.1531	0.3156	0.1094
600111	0.1192	0.1385	0.0502	0.1538	0.1784	0.0835	0.1964	0.2184	0.1287
601166	0.399	0.2764	0.4182	0.4217	0.2794	0.4584	0.435	0.2825	0.4516
600036	-	-	-	-	-	-	-	-	-
601328	0.1926	0.2297	0.1406	0.0652	0.1243	0.0315	-	0.0189	-
600048	0.0491	0.0131	0.2296	0.0736	0.0335	0.2676	0.1067	0.054	0.3103
600000	-	-	-	-	-	-	-	-	-
601318	-	-	-	-	-	-	-	-	-
601088	-	-	-	-	-	-	-	-	-
600016	0.0636	0.034	-	0.1194	0.0724	-	0.1088	0.1107	-
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Expected Return	0.0018			0.002			0.0022		
Portfolio	MV	MCD	CVaR	MV	MCD	CVaR	MV	MCD	CVaR
600519	0.1365	0.309	0.0691	0.1025	0.3003	0.017	0.0588	0.279	-
600111	0.2472	0.2712	0.1922	0.3029	0.3269	0.2541	0.3614	0.3848	0.3335
601166	0.4387	0.2644	0.4238	0.418	0.2418	0.39	0.3836	0.206	0.3447
600036	-	-	-	-	-	-	-	-	-
601328	-	-	-	-	-	-	-	-	-
600048	0.1489	0.0806	0.3148	0.1766	0.1087	0.3241	0.1962	0.1302	0.2809
600000	-	-	-	-	-	-	-	-	-
601318	-	-	-	-	-	-	-	-	-
601088	-	-	-	-	-	-	-	-	-
600016	0.0288	0.0747	0.0002	-	0.0223	0.0148	-	-	0.0409
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Expected Return	0.0024			0.0026			0.0028		
Portfolio	MV	MCD	CVaR	MV	MCD	CVaR	MV	MCD	CVaR
600519	0.0150	0.2484	-	-	0.2178	-	-	0.1872	-
600111	0.4199	0.4444	0.4019	0.4814	0.504	0.4667	0.5444	0.5636	0.5376
601166	0.3492	0.1604	0.3276	0.2915	0.1148	0.2614	0.2216	0.0692	0.2076
600036	-	-	-	-	-	-	-	-	-
601328	-	-	-	-	-	-	-	-	-
600048	0.2158	0.1467	0.2705	0.2271	0.1633	0.2719	0.234	0.1799	0.2548
600000	-	-	-	-	-	-	-	-	-
601318	-	-	-	-	-	-	-	-	-
601088	-	-	-	-	-	-	-	-	-
600016	-	-	-	-	-	-	-	-	-
Total	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

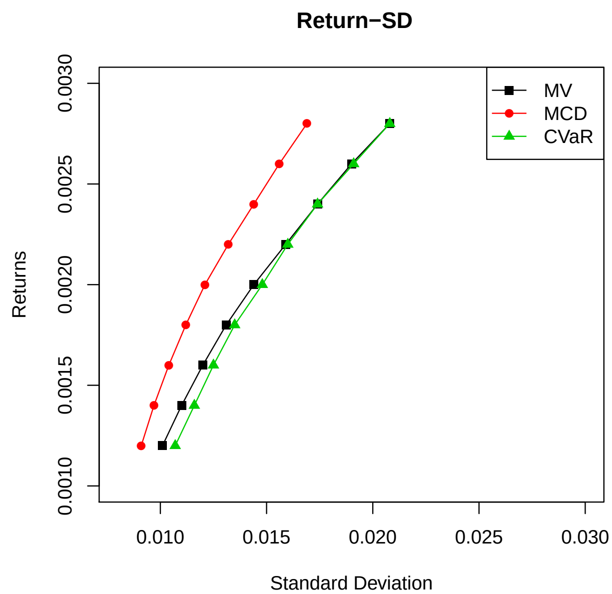

Figure 4.1: Return-Standard Deviation

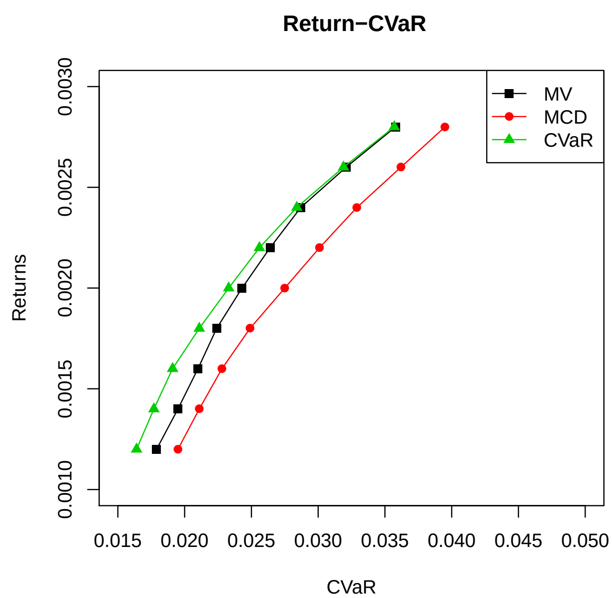

Figure 4.2: Return-CVaR

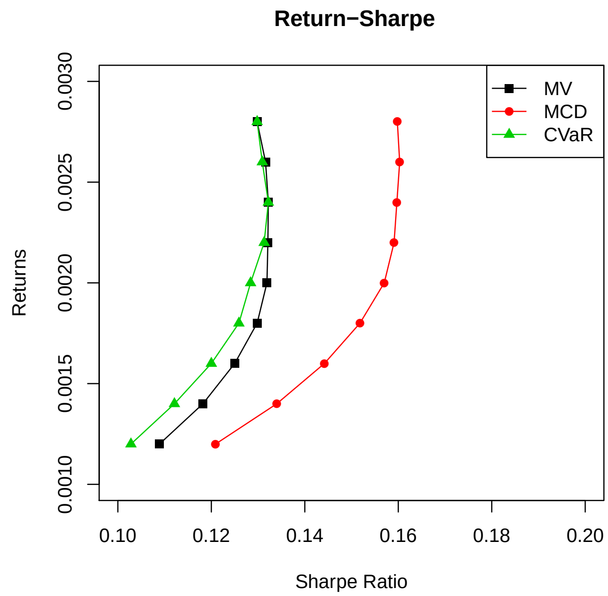

Figure 4.3: Return-Sharpe Ratio

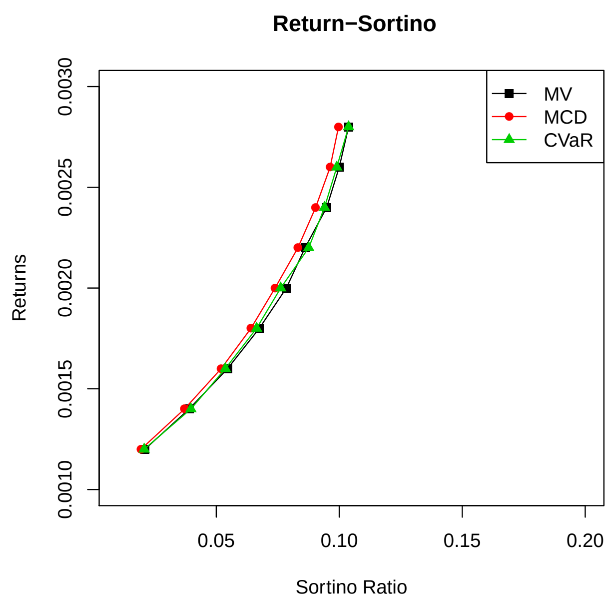

Figure 4.4: Return-Sortino Ratio

Table 4.3: Portfolio Performance in Different Conditions

Condition	Global Minimum Risk			Tangency Portfolio		
Portfolio	MV	MCD	CVaR	MV	MCD	CVaR
Expect Return	0.0006	0.0007	0.0009	0.0022	0.0025	0.0022
Standard Deviation	0.0087	0.0085	0.0096	0.0160	0.0151	0.0158
CVaR	0.0157	0.0167	0.0152	0.0266	0.0347	0.0251
Sharpe Ratio	0.0575	0.0706	0.0833	0.1313	0.1589	0.1329
Sortino Ratio (MAR=0.05%)	0.0130	0.0241	0.0418	0.1261	0.1262	0.1239

Table 4.4: Portfolio Compositions

Condition	Global Minimum Risk			Tangency Portfolio		
Portfolio	MV	MCD	CVaR	MV	MCD	CVaR
600519	0.1836	0.2811	0.2182	0.0553	0.2315	-
600111	-	0.023	0.0043	0.3661	0.4774	0.3237
601166	0.3046	0.2615	0.2852	0.3809	0.1352	0.3413
600036	-	-	-	-	-	-
601328	0.5118	0.4344	0.1804	-	-	-
600048	-	-	0.0713	0.1978	0.1559	0.2764
600000	-	-	0.0027	-	-	-
601318	-	-	-	-	-	-
601088	-	-	-	-	-	-
600016	-	-	0.2379	-	-	0.058
Total	1.00	1.00	1.00	1.00	1.00	1.00

Now, let us turn to table 4.3 and 4.4. These two tables list the performance and portfolio compositions of the global minimum risk portfolio which means this portfolio with the lowest possible risk and the tangency portfolio which is introduced before. (See the efficient frontier in section 2.) Remarkably, for CVaR portfolio, the risk intimates CVaR. Through the comparison of the Sharpe ratio and Sortnio ratio, it seems that MCD-MV portfolio and mean-CVaR portfolio perform comparatively better than mean-variance portfolio due to the higher expect return. In particular, relating to figure 4.3, it also hints that the tangency portfolio has the highest Sharpe ratio mentioned in section 2.

Figure 4.5 shows the efficient frontiers of three portfolios, from this graph, it is clearly to see the MCD portfolio has the global minimum standard deviation as well as the largest tangency portfolio's expect return.

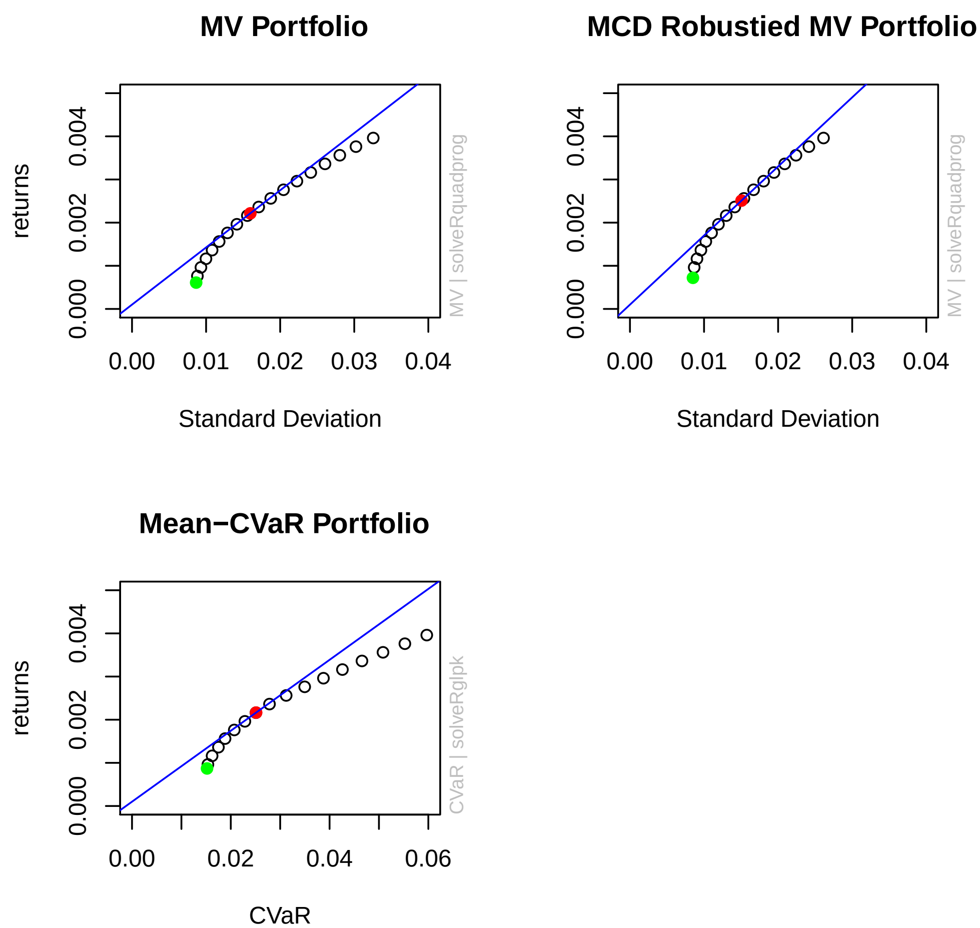

Figure 4.5: Efficient Frontier

5 Conclusions and Future Work

There are two parts in this project. The first one is using factor analysis to evaluate the performance of individual stock. Empirical results shows that through factor analysis, it can effectively reduce the dimension of variables. In this case, 4 factors are used instead of 10 indicators to estimate the individual stock's capacity. Meantime, consequences also indicate that there is no apparent difference among the principal component method, the principal factor method and maximum likelihood method. Nevertheless, due to having relatively smallest errors, the principal factor procedure is recommended. Eventually, according to the composite score, 10 stocks were chosen for the portfolio selection.

In the second parts, we mainly compare the performance of three different portfolios, i.e. Mean-Variance portfolio, MCD Robustified Mean-Variance portfolio and Mean-CVaR portfolio. Practical outcomes present for a set of given returns, the MCD method has the minimum dispersion while CVaR approaches can minimise the CVaR. Meanwhile, it is no noticeable distinctness between MV portfolio and mean-CVaR portfolio when the return grows. However, comparing with the MV and MCD-MV methods, the CVaR approach has greater practicability since this method can be used in different situations no matter whether the returns follow the normal distribution. Contrarily, in effect, both MV and MCD-MV portfolios rely on the assumption that the returns of assets are multivariate normally distributed.

Furthermore, it seems that using MCD and CVaR approaches maybe can get a comparatively optimal portfolio in global minimum risk portfolio and tangency portfolio conditions because their performance is relatively better than MV method. Nonetheless, there also exit some drawbacks of these two portfolios, such as although the MCD procedure can decrease the variance, it always will increase the bias and value of CVaR. In addition, CVaR method just focus on the loss situation without considering the profit state. Actually, recent years, many researches suggested using the semi-variance method to investigate the Chinese financial market. Therefore, future work maybe should include more alternative risk measures, for instance, combining with the robust method and semi-variance measure or CVaR approach and so on.

References

[1] Acerbi, C. and Tasch, D. *On the Coherence of Expected Shortfall*. Journal of Banking and Finance, 26(7)(2002), 1487-1503.

[2] Alexei, A.G. and Pflug, G. *Value-at-Risk in Portfolio Optimization: Properties and Computational Approach*. Journal of Risk , 7(2)(2005), 1-31.

[3] Armstrong, J. S. and Soelberg, P. *On the Interpretation of Factor Analysis*. Psychological Bulletin, 70(1968), 361-364.

[4] Artzner, P., Delbaen, F., Eber, J-M. and Heath, D. *Coherent Measures of Risk*. Mathematical Finance, 9(3)(1999), 203-228.

[5] Basel Committee on Banking Supervision. *Amendment to the Capital Accord to Incorporate Market Risks*, 1996.

[6] Brooks, C. and Kat, H. *The statistical properties of hedge fund index returns and their implications for investors*. Journal of Alternative Investments, 5(2)(2002), 26-44.

[7] Carroll, J. B. *The Scientific Use of Factor Analysis in the Behavioral and Life Sciences*. New York, Plenum Press, 1978.

[8] Croux, C. and Haesbroeck, G. *Influence Function and Efficiency of The Minimum Covariance Determinant Scatter Matrix Estimator*. Journal of Multivariate Analysis, 71(2)(1999), 161-190.

[9] Damodaran, A. *Investment Philosophies: Successful Investment Philosophies and the Greatest Investors Who Made Them Work*. Wiley, 2003.

[10] Elton, E.J. and Gruber, M.J. *Investments and Portfolio Performance*. World Scientific, 2011, 382-383.

[11] Elton, E. J., Gruber, M.J., Brown, S. J. and Goetzmann, W.N. *Modern portfolio theory and investment analysis(8 ed.)*. New York, J. Wiley & Sons, cop, 2011.

[12] Fabozzi, F.J. *Robust portfolio optimization and management*. Hoboken, N.J., Wiley, cop, 2009.

[13] Fama, E.F. *Efficient Capital Markets: A Review of Theory and Empirical Work*. Journal of Finance, 25(2)(1970), 383-417.

[14] Gordon, J.A. and Alexandre M.B. *Economic Implications of Using a Mean-VaR Model for Portfolio Selection: A Comparison with Mean-Variance Analysis*. Journal of Economic Dynamics & Control , 26(7)(2002), 1159-1193.

[15] Henry, F. K. *The Varimax Criterion for Analytic Rotation in Factor Analysis*. Psychometrika, 23(3)(1958), 187-200.

[16] Hubert, M. and Van Driessen, K. *Fast and Robust Discriminant Analysis*. Computational Statistics & Data Analysis, 45(2)(2004), 301-320.

[17] Johnson, R.A. and Wichern, D.W. *Applied multivariate statistical analysis*. Upper Saddle River, N.J., Pearson Prentice Hall, cop, 2007.

[18] JP Morgan/Reuters. *RiskMetrics ^{TM} -Technical Document (4 ed)*. New York, Morgan Guaranty Trust Company of New York, 1996.

[19] Jöreskog, K. *A general approach to confirmatory maximum likelihood factor analysis.* Psychometrika, 34(2)(1969), 183-202.

[20] Jöreskog, K. *Factor Analysis by Least Squares and Maximum Likelihood.* In *Statistical Methods for Digital Computers*, edited by Enslein, K., Ralston, A. and Wilf, H.S., New York, John Wiley , 1975.

[21] Kaiser, H.F. *An index of factorial simplicity.* Psychometrika, 39(1974), 31-36.

[22] Konno, H. *Portfolio Optimization Using L1 Risk Function.* IHSS Report 889, Institute of Human and Social Sciences, Tokyo Institute of Technology, 1988.

[23] Lawley, D. N. and Maxwell, A. E. *Factor analysis as a statistical method.* London, Butterworth, 1963.

[24] Luenberger, D.G. *Investment Science.* New York, Oxford University Press, 1998.

[25] Marazzi, A. *Algorithms, Routines and S Functions for Robust Statistics(2 ed.).* Chapman & Hall, 1993.

[26] Markowitz, H.M. *Portfolio Section.* The Journal of Finance, 7(1)(1952), 77-91.

[27] Markowitz, H.M. *Portfolio Selection-Efficient Diversification of Investment.* New York, Wiley, 1959.

[28] Markowitz, H.M. *Foundations of Portfolio Theory.* Journal of Finance, 46(2)(1991), 469-477.

[29] Mittnik, S., Rachev, S. and Schwartz, E. *Value At-Risk and Asset Allocation with Stable Return Distributions.* Allgemeines Statistisches Archiv, 86(1)(2003), 53-68.

[30] Palmquist, J. and Uryasev, S. *Portfolio optimization with conditional Value-At-Risk objective constraints.* The Journal of Risk, 4(2)(2002).

[31] Penman, S. H. *Financial statement analysis and security valuation(4 ed.).* Boston, Mass, McGraw-Hill, 2009, 33-40.

[32] Rockafellar, T.R. and Uryasev, S. *Optimization of Conditional Value-at-Risk.* Journal of Risk, 2(2)(2000), 21-41.

[33] Rousseeuw, P.J. *Least Median of Squares Regression.* Journal of the American Statistical Association, 79(388)(1984), 871-880.

[34] Rousseeuw, P.J. and Van Driessen, K. *A Fast Algorithm for the Minimum Covariance Determinant Estimator.* Technometrics , 41(3)(1999), 212-223.

[35] Sharpe, W. F. *Mutual Fund Performance.* Journal of Business, 39(S1)(1966), 119-138.

[36] Thurstone, L. L. *Multiple Factor Analysis: A Development and Expansion of Vectors of The Mind.* Chicago, University of Chicago, 1947.

[37] Uryasev, S. *Conditional Value-at-Risk: Optimization Algorithms and Applications.* Computational Intelligence for Financial Engineering, 2000, 49-57.

[38] Würtz, D., Chalabi, Y., William, C. and Ellis, A. *Portfolio Optimization with R/Rmetrics.* Zurich, Rmetrics Association and Finance Online Publishing, 2009.

Appredix A Shanghai Securities Composite Index 50

Stock Code	Company Name
600000	Shanghai Pudong Development Bank Co., Ltd
600010	Inner Mongolia BaoTou Steel Union Co.,Ltd.
600015	Hua Xia Bank Co.,Limited
600016	China Merchants Bank Co.,Limited
600019	Baoshan Iron & Steel Co., Ltd.
600028	China Petroleum & Chemical Corporation
600030	CITIC Securities Company Limited
600031	Sany Heavy Industry Co.,Ltd
600036	China Merchants Bank Co.,Limited
600048	Poly Real Estate Group Co.,Ltd
600050	China United Network Communications Limited
600058	Minmetals Development Co., Ltd.
600068	China Gezhouba Group Company Limited
600089	Tbea Co.,Ltd
600104	SAIC Motor Corporation Limited
600111	Inner Mongolia Baotou Steel Rare-Earth(Group) Hi-Tech Co.,Ltd.
600188	Yanzhou Coal Mining Company Limited
600348	Yang Quan Coal Industry (Group) Co.,Ltd.
600362	Jiangxi Copper Co., Ltd
600383	Gemdale Corporation
600489	Zhongjin Gold Corporation,Limited
600519	Kweichow Moutai Co.,Ltd.
600547	Shandong Gold Mining Co.,Ltd.
600585	Anhui Conch Cement Co., Ltd.
600837	HAITONG Securities Company Limited
600900	China Yangtze Power Co.,ltd.
601006	Daqin Railway Co.,Ltd
601088	China Shenhua Energy Company Limited
601111	Air China Limited
601118	China Hainan Rubber Lndustry Group Co.,LTD
601166	Industrial Bank Co., Ltd
601168	Western Mining Co., Ltd.
601169	Bank of Beijing Co.,Ltd
601288	Agricultural Bank of China Limited
601299	China CNR Corporation Limited
601318	Ping An Insurance (Group) Company of China, Ltd.
601328	Bank of Communication Co.,Ltd.
601398	Industrial and Commercial Bank of China Limited
601600	Aluminum Corporation of China Limited
601601	China Pacific Insurance (Group) Co.,Ltd.
601628	China Life Insurance Company Limited
601668	China State Construction Engineering Corporation Limited
601699	Shanxi Lu'an Environmental Energy Development Co.,Ltd.
601766	CSR Corporation Limited
601818	China Everbright Bank Company Limited
601857	PetroChina Company Limited
601898	China Coal Energy Company Limited
601899	ZiJin Mining Group Co.,Ltd
601958	Jinduicheng Molybdenum Co.,Ltd
601989	China Shipbuilding Industry Company Limited

Appredix B The Data Set of Financial Indicators

Stock Code	NI	GRONI	ROE	ASSET	SE	OM	EPS	NAVPS	GRMBR	GRSE
600000	0.976	0.211	0.076	1.869	0.817	1.370	0.199	0.434	0.283	0.010
600016	1.018	0.480	0.369	1.440	0.598	1.025	-0.109	-0.300	0.895	0.147
600030	0.001	-0.294	-0.271	-0.518	0.108	1.713	0.026	0.406	-1.712	0.085
600048	0.976	0.211	0.076	1.869	0.817	1.370	0.201	0.435	0.283	0.010
600362	-0.399	0.068	-0.076	-0.593	-0.430	-0.785	0.521	1.225	1.046	-0.226
600519	-0.252	0.718	1.601	-0.625	-0.593	2.043	5.431	4.207	1.232	0.602
600837	-0.628	-0.732	-0.965	-0.564	-0.365	1.006	-0.611	-0.154	-1.493	-0.783
601088	2.139	-0.176	0.237	-0.283	1.663	0.386	0.793	1.194	0.102	-0.419
601166	0.858	0.137	0.423	1.609	0.434	1.527	0.873	1.068	0.350	0.177
601328	2.531	0.011	0.107	3.681	2.217	1.299	-0.281	-1.247	-0.341	0.048
601601	-0.282	-0.522	-0.609	-0.121	-0.004	-0.806	-0.169	-1.237	-0.787	-1.009
601699	-0.579	-0.287	0.754	-0.625	-0.708	-0.167	0.356	0.080	-1.082	0.173
601958	-0.784	-0.669	-1.095	-0.643	-0.725	-0.539	-0.723	-0.459	-1.118	-0.822
600010	-0.800	0.600	-1.242	-0.611	-0.731	-1.018	-0.836	-0.965	-0.990	-0.797
600019	-0.345	-1.170	-0.963	-0.440	0.335	-0.906	-0.581	-0.011	-0.848	-0.767
600050	-0.740	-0.242	-1.411	-0.226	-0.069	-0.991	-0.845	-0.651	-0.320	-0.827
600089	-0.752	-0.860	-0.721	-0.626	-0.732	-0.801	-0.546	-0.292	-1.188	-0.477
600188	-0.261	-0.545	0.268	-0.566	-0.397	0.075	0.419	0.575	0.444	-0.237
600383	-0.633	-0.280	-0.262	-0.572	-0.580	0.747	-0.393	-0.353	-0.334	0.146
600547	-0.707	0.427	1.462	-0.646	-0.813	-0.769	0.109	-0.504	-0.200	1.004
601111	-0.337	-1.105	-0.136	-0.494	-0.346	-0.651	-0.438	-0.536	-0.419	-0.345
601168	-0.776	-0.690	-0.913	-0.632	-0.747	-0.848	-0.626	-0.302	-0.639	-0.697
601628	0.382	-1.214	-0.721	0.833	1.303	-0.842	-0.408	0.153	-1.321	-1.164
601766	-0.577	0.385	-0.033	-0.570	-0.621	-0.804	-0.648	-0.989	-0.250	-0.149
601898	-0.203	0.135	-0.526	-0.508	0.046	-0.382	-0.356	-0.002	-0.215	-0.412
600015	-0.222	0.401	-0.279	0.513	-0.150	0.654	0.214	0.751	1.138	2.375
600028	3.921	-0.453	-0.217	0.406	4.523	-0.903	-0.276	-0.154	0.055	-0.328
600036	1.563	0.178	0.401	1.972	1.001	1.175	0.356	0.357	0.217	0.093
600104	0.508	-0.096	0.207	-0.358	0.287	-0.642	0.479	0.739	-0.477	-0.122
600489	-0.713	0.353	0.336	-0.640	-0.780	-0.649	-0.176	-0.412	0.941	1.262
601169	-0.240	0.036	0.046	0.243	-0.304	1.469	0.184	0.460	0.125	-0.096
601318	0.458	-0.272	-0.238	1.493	0.612	-0.530	0.978	2.439	0.074	-0.160
601600	-0.818	-1.603	-1.551	-0.510	-0.288	-1.072	-0.881	-0.539	-0.394	-0.816
601668	0.064	0.280	-0.210	-0.181	0.137	-0.848	-0.558	-0.740	0.025	-0.218
601899	-0.454	-0.178	0.487	-0.608	-0.593	0.038	-0.701	-1.167	0.417	-0.251
600068	-0.730	-0.271	-0.324	-0.595	-0.751	-0.888	-0.562	-0.689	-0.107	-0.479
600111	-0.603	5.426	4.030	-0.644	-0.814	1.856	1.257	-0.346	3.872	4.505
600031	-0.260	0.402	2.413	-0.609	-0.654	-0.185	-0.042	-0.830	0.858	2.102
601989	-0.521	-0.331	-0.503	-0.507	-0.430	-0.698	-0.656	-0.808	-0.969	-0.552
601299	-0.635	0.421	-0.504	-0.566	-0.594	-0.925	-0.626	-0.731	0.399	-0.591
600058	-0.798	0.170	-1.053	-0.611	-0.775	-1.068	-0.520	0.548	-0.278	-0.704
600348	-0.647	-0.207	0.647	-0.631	-0.748	-0.745	-0.019	-0.323	2.241	0.526
600585	-0.065	0.951	0.761	-0.578	-0.367	0.346	0.746	0.547	0.489	0.292
601818	0.365	0.195	0.120	0.974	0.215	1.360	-0.558	-0.879	-0.001	-0.112

Appredix C The Stock Price

Date	600519	600111	601166	600036	601328	600048	600000	601318	601088	600016
2011-11-11	204.17	47.65	12.93	12.11	4.71	10.31	9.21	38.63	26.99	6.12
2011-11-14	207.18	46.9	13.12	12.15	4.69	9.95	9.17	39.22	26.79	6.11
2011-11-15	209.59	46.4	13.15	12.37	4.79	10.16	9.36	40.09	27.22	6.27
2011-11-16	207.17	47.93	13.28	12.2	4.78	9.98	9.24	40.34	27.21	6.18
2011-11-17	207.83	49.1	13.51	12.26	4.79	9.97	9.31	40.4	27.28	6.22
2011-11-18	206.81	48.9	13.3	12.08	4.8	10.03	9.18	39.78	27.16	6.1
2011-11-21	204.6	48.09	13.29	12.11	4.85	9.87	9.22	39.92	27.44	6.11
2011-11-22	202.5	49.17	13.56	12.2	4.87	9.82	9.27	40.56	27.59	6.15
2011-11-23	205	47.71	13.36	11.96	4.78	9.58	9.05	38.86	26.96	6.05
2011-11-24	204	48.35	13.4	11.88	4.79	9.65	9.08	38.8	27.05	6.06
2011-11-25	207.27	49.15	13.14	12.07	4.85	9.93	9.21	39.92	28.15	6.1
2011-11-28	203.66	50.98	13.15	11.95	4.82	9.85	9.17	39.5	28.47	6.03
2011-11-29	200.7	49.2	13.26	11.64	4.73	9.4	8.91	37.75	27.42	5.9
2011-11-30	201.81	48.42	12.97	11.56	4.7	9.35	8.87	37.24	27.38	5.9
2011-12-01	200.53	46.2	12.96	11.42	4.68	9.06	8.8	36.39	26.62	5.87
2011-12-02	200.7	46.44	13.18	11.42	4.67	9.14	8.73	36.64	26.2	5.85
2011-12-05	206.05	47.07	13.08	11.55	4.65	9.08	8.76	36.33	26.07	5.91
2011-12-06	209.02	46.18	12.65	11.49	4.59	8.9	8.7	35.6	25.59	5.82
2011-12-07	207.67	46.3	12.63	11.59	4.6	9.17	8.71	36.15	25.76	5.86
2011-12-08	209.96	46.13	12.52	11.49	4.54	9.08	8.63	35.74	25.55	5.82
2011-12-09	213.75	46.05	12.55	11.45	4.54	9.09	8.61	36.07	25.84	5.85
2011-12-12	212.88	46.9	12.58	11.57	4.59	9.32	8.73	36.68	26.15	5.94
2011-12-13	210	44.5	12.46	11.22	4.5	9.28	8.49	35.3	25.4	5.8
2011-12-14	212.95	45.59	12.51	11.65	4.55	9.75	8.81	37.18	26.06	5.99
2011-12-15	214.68	44.62	12.38	11.75	4.56	9.7	8.8	37.05	25.6	6.03
2011-12-16	206.72	42.92	12.39	11.88	4.59	9.76	8.81	37.1	25.59	6.05
2011-12-19	203.57	43.06	12.53	11.78	4.58	9.88	8.73	36.83	25.59	6.04
2011-12-20	205.88	42.95	12.2	11.9	4.57	9.92	8.79	37.93	25.78	6.12
2011-12-21	202.99	43.28	12.57	11.89	4.59	9.94	8.76	37.97	25.77	6.12
2011-12-22	199.81	42.56	12.59	11.9	4.6	10.12	8.7	37.98	25.51	6.05
2011-12-23	197.59	42.68	12.63	11.98	4.64	9.76	8.67	37.5	25.25	6.03
2011-12-26	196.22	40.96	12.58	11.66	4.6	9.58	8.59	36.02	25.04	5.94
2011-12-27	197.88	40.47	12.65	11.69	4.6	9.38	8.6	36	24.91	5.92
2011-12-28	196.95	37.95	12.65	11.42	4.5	9.48	8.5	35.22	24.04	5.82
2011-12-29	201.56	38.3	12.53	11.63	4.53	10	8.62	35.93	24.45	5.94
2011-12-30	200.98	37.78	12.51	11.63	4.5	10.06	8.61	36.16	24.35	6.02
2012-01-03	200.22	37.94	12.22	11.63	4.46	10.09	8.6	36.36	24.27	6.01
2012-01-04	197.25	36.29	12.26	11.6	4.44	9.91	8.53	34.43	24.25	5.98
2012-01-05	200	35.91	12.07	11.71	4.5	10.08	8.53	33.99	24.59	6.02

Date	600519	600111	601166	600036	601328	600048	600000	601318	601088	600016
2012-01-06	200.8	35.63	12.3	11.68	4.52	10.2	8.54	33.99	24.89	6.01
2012-01-09	193	36.32	12.33	11.67	4.5	10.08	8.46	33.58	24.56	5.9
2012-01-10	189	36.19	12.44	11.7	4.47	9.77	8.44	33.78	24.48	5.86
2012-01-11	190.11	36.79	12.39	11.55	4.5	9.67	8.39	34.12	24.69	5.83
2012-01-12	192	36.21	12.55	11.67	4.46	9.74	8.35	33.96	24.9	5.83
2012-01-13	192	37.62	12.53	11.88	4.48	10.06	8.49	34.45	25.4	5.89
2012-01-17	185.22	37.58	12.39	11.68	4.45	10	8.41	33.9	24.6	5.86
2012-01-18	183.14	35.64	12.34	11.9	4.56	9.8	8.66	33.92	24.3	5.97
2012-01-19	186.6	36.33	12.3	12	4.6	9.7	8.71	33.85	24.24	6.02
2012-01-20	188.07	39.05	12.36	12.38	4.7	10.18	8.96	34.73	26.05	6.18
2012-01-23	194.47	40.38	12.52	12.57	4.74	10.35	9.07	36.3	26.77	6.23
2012-01-24	189.62	42.62	12.5	12.49	4.72	10.31	9	35.83	26.5	6.17
2012-01-25	190.5	43.4	12.71	12.6	4.73	10.39	9.09	36.37	26.48	6.23
2012-01-26	188.75	44.96	12.8	12.49	4.75	10.27	9.04	36.02	26.32	6.23
2012-01-27	177.48	41.03	13.05	12.39	4.74	9.93	8.97	35.59	26	6.18
2012-01-30	180.3	45.19	13.15	12.79	4.84	10.35	9.2	37.81	27.39	6.33
2012-01-31	177.3	45.45	13.03	12.52	4.75	10.12	9.04	37.4	26.85	6.26
2012-02-01	180.79	46.58	13.3	12.7	4.84	10.67	9.2	38.52	27.1	6.44
2012-02-02	186	45.6	13.27	12.99	4.92	10.94	9.41	39.1	27.48	6.54
2012-02-03	183.72	46.05	13.32	12.7	4.84	10.44	9.22	38.6	26.62	6.38
2012-02-06	186.42	46.58	13.61	12.65	4.84	10.5	9.23	38.38	26.84	6.41
2012-02-07	186.2	44.71	13.44	12.47	4.79	10.31	9.09	37.42	26.56	6.3
2012-02-08	186.43	45.57	13.77	12.85	4.92	10.56	9.4	39.69	27.27	6.48
2012-02-09	186.49	45.62	14.07	12.98	5.02	10.68	9.44	40.14	27.36	6.55
2012-02-10	188.53	45.61	13.85	12.87	5.08	10.45	9.38	39.61	27.34	6.54
2012-02-13	185.9	44.56	13.86	12.73	4.95	10.13	9.23	38.94	26.89	6.43
2012-02-14	188.3	48.5	13.57	12.98	5.07	10.42	9.45	38.91	27.65	6.6
2012-02-15	190.66	47.58	14.14	12.99	5.03	10.56	9.41	39.88	27.54	6.55
2012-02-16	190.51	47.51	14.18	12.9	4.95	10.91	9.38	40.08	27.44	6.51
2012-02-17	193.35	47.66	14.17	12.69	4.89	10.58	9.27	40.28	27.17	6.43
2012-02-21	193.8	47.49	14.03	12.67	4.89	10.63	9.19	39.58	27.15	6.41
2012-02-22	195	48.11	14.45	12.68	4.89	10.63	9.23	40.22	27.32	6.42
2012-02-23	191.72	47.5	14.42	12.62	4.87	10.6	9.19	39.68	27.14	6.41
2012-02-24	191.44	47.31	14.24	12.72	4.93	10.72	9.25	39.77	27.08	6.44
2012-02-27	190.9	47.19	14.12	12.7	4.94	10.77	9.26	40.1	27.21	6.47
2012-02-28	191.98	47.4	14.08	12.77	4.96	10.98	9.29	40.72	27.35	6.47
2012-02-29	192.8	48.18	14.1	12.79	4.97	11.29	9.33	41.08	27.56	6.48
2012-03-01	194.36	49.77	14.08	12.71	4.96	11.29	9.33	41.03	27.76	6.46
2012-03-02	198.06	50.2	14.12	12.88	4.99	11.68	9.52	41.62	28.14	6.52
2012-03-05	198.5	53.01	14.06	12.82	5	11.55	9.48	41.5	28.01	6.51

Date	600519	600111	601166	600036	601328	600048	600000	601318	601088	600016
2012-03-06	200.55	54.65	14.12	12.95	5.07	11.5	9.66	41.68	27.95	6.59
2012-03-07	204.08	55.26	14.12	12.87	5.02	11.1	9.55	40.82	27.59	6.57
2012-03-08	202.8	55.06	14.12	12.9	5.02	11.14	9.55	40.55	27.5	6.58
2012-03-09	200.85	55.9	14.36	13.03	5.05	11.63	9.69	41.25	27.82	6.69
2012-03-12	199.85	56.98	14.31	12.88	5	11.52	9.59	40.16	27.65	6.6
2012-03-13	196.01	55.12	13.88	12.58	4.93	11.49	9.44	39.45	26.88	6.49
2012-03-14	193.12	54.3	13.75	12.41	4.9	11.29	9.34	39.15	26.55	6.42
2012-03-15	194.3	58.7	13.88	12.46	4.9	11.7	9.44	39.95	26.75	6.45
2012-03-16	199.8	62.86	13.82	12.51	4.91	11.68	9.47	40.07	27.08	6.46
2012-03-19	200.6	63.94	13.63	12.35	4.87	11.25	9.39	39.74	26.77	6.38
2012-03-20	201.36	63.96	13.84	12.48	4.91	11.33	9.48	40.63	27.15	6.45
2012-03-21	198.8	64.2	13.62	12.31	4.87	11	9.36	39.69	26.33	6.37
2012-03-22	198.8	63.85	13.6	12.23	4.87	10.72	9.18	39.57	26.17	6.36
2012-03-23	207.59	68.94	13.7	12.23	4.85	10.81	9.24	39.55	26.3	6.37
2012-03-26	207.52	71.93	13.65	12.16	4.79	10.72	9.29	39.03	26.3	6.32
2012-03-27	207.53	71.36	13.45	12.03	4.74	10.73	9.13	38.77	26.02	6.21
2012-03-28	205.14	70.3	13.46	12	4.74	10.64	9.08	38.95	26.3	6.2
2012-03-29	212.08	66.88	13.6	12.05	4.72	10.59	9.07	38.4	25.92	6.29
2012-03-30	215.1	65.52	13.73	12.1	4.74	10.53	9.13	38.25	26.18	6.3
2012-04-02	201.49	70.96	13.72	12.04	4.74	10.56	9.11	38.33	26.26	6.3
2012-04-03	201.4	68.28	13.38	11.81	4.69	10.54	8.96	37.43	25.5	6.2
2012-04-04	201.15	67.6	13.13	11.82	4.63	10.8	8.86	36.26	25.24	6.12
2012-04-05	196.98	66.79	13.34	11.89	4.71	11.29	8.94	36.58	25.6	6.27
2012-04-09	206.1	73.1	13.29	11.84	4.67	11.4	8.93	37.77	25.92	6.27
2012-04-10	207.88	71.67	13.24	11.81	4.67	11.36	8.94	38.11	26.01	6.29
2012-04-11	209.25	71.59	13.24	11.68	4.63	11.3	8.85	37.65	25.53	6.18
2012-04-12	208.13	70.01	13.2	11.83	4.67	11.66	8.93	38.09	25.69	6.18
2012-04-13	209	68.05	13.17	11.77	4.65	11.79	8.89	37.9	25.65	6.24
2012-04-16	212.4	68.71	13.47	12.01	4.73	11.97	9.13	39	26.25	6.39
2012-04-17	213	70.85	13.42	11.99	4.74	11.87	9.13	39.75	26.44	6.39
2012-04-18	212.84	69.84	13.29	11.92	4.72	11.99	9.05	39.5	26.37	6.38
2012-04-19	210.2	69.53	13.18	11.84	4.71	11.65	8.94	38.86	25.9	6.36
2012-04-20	213.86	69.53	13.44	12.01	4.79	12.18	9.12	39.99	26.52	6.44
2012-04-23	209.42	70.03	13.57	11.96	4.81	12.22	9.12	41.03	26.7	6.45
2012-04-24	212.5	69.7	13.93	12.2	4.84	12.19	9.31	41.37	26.8	6.53
2012-04-25	211.8	69	13.91	12.1	4.83	12	9.24	40.65	26.41	6.48
2012-04-26	211.54	67.71	14.12	12.29	4.86	12.02	9.39	40.8	26.42	6.58
2012-04-27	213	70.36	14.26	12.27	4.87	12.32	9.35	40.8	26.6	6.59
2012-04-30	216.83	69.55	14.24	12.21	4.86	12.31	9.38	41.17	26.77	6.6
2012-05-01	224.5	70.98	14.31	12.22	4.9	12.51	9.42	40.73	26.89	6.65
2012-05-02	226	77.1	14.53	12.41	4.93	13.01	9.55	42.17	27.63	6.8
2012-05-03	226.7	76.65	14.4	12.35	4.91	12.95	9.48	43.08	27.54	6.8

Appredix D The Logarithmic Return

Date	600519	600111	601166	600036	601328	600048	600000	601318	601088	600016
2011-11-14	0.0146	-0.0159	0.0146	0.0033	-0.0043	-0.0355	-0.0044	0.0152	-0.0074	-0.0016
2011-11-15	0.0116	-0.0107	0.0023	0.0179	0.0211	0.0209	0.0205	0.0219	0.0159	0.0258
2011-11-16	-0.0116	0.0324	0.0098	-0.0138	-0.0021	-0.0179	-0.0129	0.0062	-0.0004	-0.0145
2011-11-17	0.0032	0.0241	0.0172	0.0049	0.0021	-0.0010	0.0075	0.0015	0.0026	0.0065
2011-11-18	-0.0049	-0.0041	-0.0157	-0.0148	0.0021	0.0060	-0.0141	-0.0155	-0.0044	-0.0195
2011-11-21	-0.0107	-0.0167	-0.0008	0.0025	0.0104	-0.0161	0.0043	0.0035	0.0103	0.0016
2011-11-22	-0.0103	0.0222	0.0201	0.0074	0.0041	-0.0051	0.0054	0.0159	0.0055	0.0065
2011-11-23	0.0123	-0.0301	-0.0149	-0.0199	-0.0187	-0.0247	-0.0240	-0.0428	-0.0231	-0.0164
2011-11-24	-0.0049	0.0133	0.0030	-0.0067	0.0021	0.0073	0.0033	-0.0015	0.0033	0.0017
2011-11-25	0.0159	0.0164	-0.0196	0.0159	0.0124	0.0286	0.0142	0.0285	0.0399	0.0066
2011-11-28	-0.0176	0.0366	0.0008	-0.0100	-0.0062	-0.0081	-0.0044	-0.0106	0.0113	-0.0115
2011-11-29	-0.0146	-0.0355	0.0083	-0.0263	-0.0188	-0.0468	-0.0288	-0.0453	-0.0376	-0.0218
2011-11-30	0.0055	-0.0160	-0.0221	-0.0069	-0.0064	-0.0053	-0.0045	-0.0136	-0.0015	0.0000
2011-12-01	-0.0064	-0.0469	-0.0008	-0.0122	-0.0043	-0.0315	-0.0079	-0.0231	-0.0282	-0.0051
2011-12-02	0.0008	0.0052	0.0168	0.0000	-0.0021	0.0088	-0.0080	0.0068	-0.0159	-0.0034
2011-12-05	0.0263	0.0135	-0.0076	0.0113	-0.0043	-0.0066	0.0034	-0.0085	-0.0050	0.0102
2011-12-06	0.0143	-0.0191	-0.0334	-0.0052	-0.0130	-0.0200	-0.0069	-0.0203	-0.0186	-0.0153
2011-12-07	-0.0065	0.0026	-0.0016	0.0087	0.0022	0.0299	0.0011	0.0153	0.0066	0.0068
2011-12-08	0.0110	-0.0037	-0.0087	-0.0087	-0.0131	-0.0099	-0.0092	-0.0114	-0.0082	-0.0068
2011-12-09	0.0179	-0.0017	0.0024	-0.0035	0.0000	0.0011	-0.0023	0.0092	0.0113	0.0051
2011-12-12	-0.0041	0.0183	0.0024	0.0104	0.0110	0.0250	0.0138	0.0168	0.0119	0.0153
2011-12-13	-0.0136	-0.0525	-0.0096	-0.0307	-0.0198	-0.0043	-0.0279	-0.0383	-0.0291	-0.0239
2011-12-14	0.0139	0.0242	0.0040	0.0376	0.0110	0.0494	0.0370	0.0519	0.0257	0.0322
2011-12-15	0.0081	-0.0215	-0.0104	0.0085	0.0022	-0.0051	-0.0011	-0.0035	-0.0178	0.0067
2011-12-16	-0.0378	-0.0388	0.0008	0.0110	0.0066	0.0062	0.0011	0.0013	-0.0004	0.0033
2011-12-19	-0.0154	0.0033	0.0112	-0.0085	-0.0022	0.0122	-0.0091	-0.0073	0.0000	-0.0017
2011-12-20	0.0113	-0.0026	-0.0267	0.0101	-0.0022	0.0040	0.0068	0.0294	0.0074	0.0132
2011-12-21	-0.0141	0.0077	0.0299	-0.0008	0.0044	0.0020	-0.0034	0.0011	-0.0004	0.0000
2011-12-22	-0.0158	-0.0168	0.0016	0.0008	0.0022	0.0179	-0.0069	0.0003	-0.0101	-0.0115
2011-12-23	-0.0112	0.0028	0.0032	0.0067	0.0087	-0.0362	-0.0035	-0.0127	-0.0102	-0.0033
2011-12-26	-0.0070	-0.0411	-0.0040	-0.0271	-0.0087	-0.0186	-0.0093	-0.0403	-0.0084	-0.0150
2011-12-27	0.0084	-0.0120	0.0055	0.0026	0.0000	-0.0211	0.0012	-0.0006	-0.0052	-0.0034
2011-12-28	-0.0047	-0.0643	0.0000	-0.0234	-0.0220	0.0106	-0.0117	-0.0219	-0.0356	-0.0170
2011-12-29	0.0231	0.0092	-0.0095	0.0182	0.0066	0.0534	0.0140	0.0200	0.0169	0.0204
2011-12-30	-0.0029	-0.0137	-0.0016	0.0000	-0.0066	0.0060	-0.0012	0.0064	-0.0041	0.0134
2012-01-03	-0.0038	0.0042	-0.0235	0.0000	-0.0089	0.0030	-0.0012	0.0055	-0.0033	-0.0017
2012-01-04	-0.0149	-0.0445	0.0033	-0.0026	-0.0045	-0.0180	-0.0082	-0.0545	-0.0008	-0.0050
2012-01-05	0.0138	-0.0105	-0.0156	0.0094	0.0134	0.0170	0.0000	-0.0129	0.0139	0.0067
2012-01-06	0.0040	-0.0078	0.0189	-0.0026	0.0044	0.0118	0.0012	0.0000	0.0121	-0.0017

Date	600519	600111	601166	600036	601328	600048	600000	601318	601088	600016
2012-01-09	-0.0396	0.0192	0.0024	-0.0009	-0.0044	-0.0118	-0.0094	-0.0121	-0.0133	-0.0185
2012-01-10	-0.0209	-0.0036	0.0089	0.0026	-0.0067	-0.0312	-0.0024	0.0059	-0.0033	-0.0068
2012-01-11	0.0059	0.0164	-0.0040	-0.0129	0.0067	-0.0103	-0.0059	0.0100	0.0085	-0.0051
2012-01-12	0.0099	-0.0159	0.0128	0.0103	-0.0089	0.0072	-0.0048	-0.0047	0.0085	0.0000
2012-01-13	0.0000	0.0382	-0.0016	0.0178	0.0045	0.0323	0.0166	0.0143	0.0199	0.0102
2012-01-17	-0.0360	-0.0011	-0.0112	-0.0170	-0.0067	-0.0060	-0.0095	-0.0161	-0.0320	-0.0051
2012-01-18	-0.0113	-0.0530	-0.0040	0.0187	0.0244	-0.0202	0.0293	0.0006	-0.0123	0.0186
2012-01-19	0.0187	0.0192	-0.0032	0.0084	0.0087	-0.0103	0.0058	-0.0021	-0.0025	0.0083
2012-01-20	0.0078	0.0722	0.0049	0.0312	0.0215	0.0483	0.0283	0.0257	0.0720	0.0262
2012-01-23	0.0335	0.0335	0.0129	0.0152	0.0085	0.0166	0.0122	0.0442	0.0273	0.0081
2012-01-24	-0.0253	0.0540	-0.0016	-0.0064	-0.0042	-0.0039	-0.0077	-0.0130	-0.0101	-0.0097
2012-01-25	0.0046	0.0181	0.0167	0.0088	0.0021	0.0077	0.0100	0.0150	-0.0008	0.0097
2012-01-26	-0.0092	0.0353	0.0071	-0.0088	0.0042	-0.0116	-0.0055	-0.0097	-0.0061	0.0000
2012-01-27	-0.0616	-0.0915	0.0193	-0.0080	-0.0021	-0.0337	-0.0078	-0.0120	-0.0122	-0.0081
2012-01-30	0.0158	0.0966	0.0076	0.0318	0.0209	0.0414	0.0253	0.0605	0.0521	0.0240
2012-01-31	-0.0168	0.0057	-0.0092	-0.0213	-0.0188	-0.0225	-0.0175	-0.0109	-0.0199	-0.0111
2012-02-01	0.0195	0.0246	0.0205	0.0143	0.0188	0.0529	0.0175	0.0295	0.0093	0.0283
2012-02-02	0.0284	-0.0213	-0.0023	0.0226	0.0164	0.0250	0.0226	0.0149	0.0139	0.0154
2012-02-03	-0.0123	0.0098	0.0038	-0.0226	-0.0164	-0.0468	-0.0204	-0.0129	-0.0318	-0.0248
2012-02-06	0.0146	0.0114	0.0215	-0.0039	0.0000	0.0057	0.0011	-0.0057	0.0082	0.0047
2012-02-07	-0.0012	-0.0410	-0.0126	-0.0143	-0.0104	-0.0183	-0.0153	-0.0253	-0.0105	-0.0173
2012-02-08	0.0012	0.0191	0.0243	0.0300	0.0268	0.0240	0.0335	0.0589	0.0264	0.0282
2012-02-09	0.0003	0.0011	0.0216	0.0101	0.0201	0.0113	0.0042	0.0113	0.0033	0.0107
2012-02-10	0.0109	-0.0002	-0.0158	-0.0085	0.0119	-0.0218	-0.0064	-0.0133	-0.0007	-0.0015
2012-02-13	-0.0140	-0.0233	0.0007	-0.0109	-0.0259	-0.0311	-0.0161	-0.0171	-0.0166	-0.0170
2012-02-14	0.0128	0.0847	-0.0211	0.0194	0.0240	0.0282	0.0236	-0.0008	0.0279	0.0261
2012-02-15	0.0125	-0.0192	0.0411	0.0008	-0.0079	0.0133	-0.0042	0.0246	-0.0040	-0.0076
2012-02-16	-0.0008	-0.0015	0.0028	-0.0070	-0.0160	0.0326	-0.0032	0.0050	-0.0036	-0.0061
2012-02-17	0.0148	0.0032	-0.0007	-0.0164	-0.0122	-0.0307	-0.0118	0.0050	-0.0099	-0.0124
2012-02-21	0.0023	-0.0036	-0.0099	-0.0016	0.0000	0.0047	-0.0087	-0.0175	-0.0007	-0.0031
2012-02-22	0.0062	0.0130	0.0295	0.0008	0.0000	0.0000	0.0043	0.0160	0.0062	0.0016
2012-02-23	-0.0170	-0.0128	-0.0021	-0.0047	-0.0041	-0.0028	-0.0043	-0.0135	-0.0066	-0.0016
2012-02-24	-0.0015	-0.0040	-0.0126	0.0079	0.0122	0.0113	0.0065	0.0023	-0.0022	0.0047
2012-02-27	-0.0028	-0.0025	-0.0085	-0.0016	0.0020	0.0047	0.0011	0.0083	0.0048	0.0046
2012-02-28	0.0056	0.0044	-0.0028	0.0055	0.0040	0.0193	0.0032	0.0153	0.0051	0.0000
2012-02-29	0.0043	0.0163	0.0014	0.0016	0.0020	0.0278	0.0043	0.0088	0.0076	0.0015
2012-03-01	0.0081	0.0325	-0.0014	-0.0063	-0.0020	0.0000	0.0000	-0.0012	0.0072	-0.0031
2012-03-02	0.0189	0.0086	0.0028	0.0133	0.0060	0.0340	0.0202	0.0143	0.0136	0.0092
2012-03-05	0.0022	0.0545	-0.0043	-0.0047	0.0020	-0.0112	-0.0042	-0.0029	-0.0046	-0.0015
2012-03-06	0.0103	0.0305	0.0043	0.0101	0.0139	-0.0043	0.0188	0.0043	-0.0021	0.0122

Date	600519	600111	601166	600036	601328	600048	600000	601318	601088	600016
2012-03-07	0.0174	0.0111	0.0000	-0.0062	-0.0099	-0.0354	-0.0115	-0.0208	-0.0130	-0.0030
2012-03-08	-0.0063	-0.0036	0.0000	0.0023	0.0000	0.0036	0.0000	-0.0066	-0.0033	0.0015
2012-03-09	-0.0097	0.0151	0.0169	0.0100	0.0060	0.0430	0.0146	0.0171	0.0116	0.0166
2012-03-12	-0.0050	0.0191	-0.0035	-0.0116	-0.0100	-0.0095	-0.0104	-0.0268	-0.0061	-0.0135
2012-03-13	-0.0194	-0.0332	-0.0305	-0.0236	-0.0141	-0.0026	-0.0158	-0.0178	-0.0282	-0.0168
2012-03-14	-0.0149	-0.0150	-0.0094	-0.0136	-0.0061	-0.0176	-0.0106	-0.0076	-0.0124	-0.0108
2012-03-15	0.0061	0.0779	0.0094	0.0040	0.0000	0.0357	0.0106	0.0202	0.0075	0.0047
2012-03-16	0.0279	0.0685	-0.0043	0.0040	0.0020	-0.0017	0.0032	0.0030	0.0123	0.0015
2012-03-19	0.0040	0.0170	-0.0138	-0.0129	-0.0082	-0.0375	-0.0085	-0.0083	-0.0115	-0.0125
2012-03-20	0.0038	0.0003	0.0153	0.0105	0.0082	0.0071	0.0095	0.0221	0.0141	0.0109
2012-03-21	-0.0128	0.0037	-0.0160	-0.0137	-0.0082	-0.0296	-0.0127	-0.0234	-0.0307	-0.0125
2012-03-22	0.0000	-0.0055	-0.0015	-0.0065	0.0000	-0.0258	-0.0194	-0.0030	-0.0061	-0.0016
2012-03-23	0.0433	0.0767	0.0073	0.0000	-0.0041	0.0084	0.0065	-0.0005	0.0050	0.0016
2012-03-26	-0.0003	0.0425	-0.0037	-0.0057	-0.0124	-0.0084	0.0054	-0.0132	0.0000	-0.0079
2012-03-27	0.0000	-0.0080	-0.0148	-0.0107	-0.0105	0.0009	-0.0174	-0.0067	-0.0107	-0.0176
2012-03-28	-0.0116	-0.0150	0.0007	-0.0025	0.0000	-0.0084	-0.0055	0.0046	0.0107	-0.0016
2012-03-29	0.0333	-0.0499	0.0103	0.0042	-0.0042	-0.0047	-0.0011	-0.0142	-0.0146	0.0144
2012-03-30	0.0141	-0.0205	0.0095	0.0041	0.0042	-0.0057	0.0066	-0.0039	0.0100	0.0016
2012-04-02	-0.0654	0.0798	-0.0007	-0.0050	0.0000	0.0028	-0.0022	0.0021	0.0031	0.0000
2012-04-03	-0.0004	-0.0385	-0.0251	-0.0193	-0.0106	-0.0019	-0.0166	-0.0238	-0.0294	-0.0160
2012-04-04	-0.0012	-0.0100	-0.0189	0.0008	-0.0129	0.0244	-0.0112	-0.0318	-0.0102	-0.0130
2012-04-05	-0.0209	-0.0121	0.0159	0.0059	0.0171	0.0444	0.0090	0.0088	0.0142	0.0242
2012-04-09	0.0453	0.0903	-0.0038	-0.0042	-0.0085	0.0097	-0.0011	0.0320	0.0124	0.0000
2012-04-10	0.0086	-0.0198	-0.0038	-0.0025	0.0000	-0.0035	0.0011	0.0090	0.0035	0.0032
2012-04-11	0.0066	-0.0011	0.0000	-0.0111	-0.0086	-0.0053	-0.0101	-0.0121	-0.0186	-0.0176
2012-04-12	-0.0054	-0.0223	-0.0030	0.0128	0.0086	0.0314	0.0090	0.0116	0.0062	0.0000
2012-04-13	0.0042	-0.0284	-0.0023	-0.0051	-0.0043	0.0111	-0.0045	-0.0050	-0.0016	0.0097
2012-04-16	0.0161	0.0097	0.0225	0.0202	0.0171	0.0152	0.0266	0.0286	0.0231	0.0238
2012-04-17	0.0028	0.0307	-0.0037	-0.0017	0.0021	-0.0084	0.0000	0.0190	0.0072	0.0000
2012-04-18	-0.0008	-0.0144	-0.0097	-0.0059	-0.0042	0.0101	-0.0088	-0.0063	-0.0027	-0.0016
2012-04-19	-0.0125	-0.0044	-0.0083	-0.0067	-0.0021	-0.0288	-0.0122	-0.0163	-0.0180	-0.0031
2012-04-20	0.0173	0.0000	0.0195	0.0143	0.0168	0.0445	0.0199	0.0287	0.0237	0.0125
2012-04-23	-0.0210	0.0072	0.0096	-0.0042	0.0042	0.0033	0.0000	0.0257	0.0068	0.0016
2012-04-24	0.0146	-0.0047	0.0262	0.0199	0.0062	-0.0025	0.0206	0.0083	0.0037	0.0123
2012-04-25	-0.0033	-0.0101	-0.0014	-0.0082	-0.0021	-0.0157	-0.0075	-0.0176	-0.0147	-0.0077
2012-04-26	-0.0012	-0.0189	0.0150	0.0156	0.0062	0.0017	0.0161	0.0037	0.0004	0.0153
2012-04-27	0.0069	0.0384	0.0099	-0.0016	0.0021	0.0247	-0.0043	0.0000	0.0068	0.0015
2012-04-30	0.0178	-0.0116	-0.0014	-0.0049	-0.0021	-0.0008	0.0032	0.0090	0.0064	0.0015
2012-05-01	0.0348	0.0204	0.0049	0.0008	0.0082	0.0161	0.0043	-0.0107	0.0045	0.0075
2012-05-02	0.0067	0.0827	0.0153	0.0154	0.0061	0.0392	0.0137	0.0347	0.0271	0.0223
2012-05-03	0.0031	-0.0059	-0.0090	-0.0048	-0.0041	-0.0046	-0.0074	0.0213	-0.0033	0.0000