MISPECIFICATION BOOTSTRAP TESTS OF THE CAPITAL ASSET PRICING MODEL

A Thesis

by

NHIEU BO

BS, Texas A&M University-Corpus Christi, 2014

Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in

MATHEMATICS

Texas A&M University-Corpus Christi
Corpus Christi, Texas

December 2017

MISPECIFICATION BOOTSTRAP TESTS OF THE CAPITAL ASSET PRICING MODEL

A Thesis

by

NHIEU BO

This thesis meets the standards for scope and quality of
Texas A&M University-Corpus Christi and is hereby approved.

LEI JIN, PhD
Chair

H. SWINT FRIDAY, PhD, CFP
Committee Member

BLAIR STERBA-BOATWRIGHT, PhD
Committee Member

December 2017

ABSTRACT

The development of the Capital Asset Pricing Model (CAPM) marks the birth of asset pricing framework in finance. The CAPM is a simple and powerful tool to describe the linear relationship between risk and expected return. According to the CAPM, all pricing errors should be jointly equal to zero. Many empirical studies were conducted to test the validity of the model in various stock markets. Traditional methods such as Black, Jensen, and Scholes (1972), Fama-MacBeth (1973) and cross-sectional regression have some limitations and encounter difficulties because they often involve estimation of the covariance matrix between all estimated price errors. It becomes even more difficult when the number of assets becomes larger. Our research is motivated by the objective to overcome the limitations of the traditional methods. In this study, we propose to use bootstrap methods which can capture the characteristics of the original data without any covariance estimation.

The principle philosophy of bootstrap procedures is to treat the data sample as the population to draw bootstrap re-samples. The bootstrap methods comprise two general steps. First, we use historical monthly returns to estimate the parameters using both ordinary least square and the Cochrane-Orcutt method. Next, we implement model-based procedures to generate bootstrap samples. Following the idea of the block bootstrap, we consider all assets at a point in time as one block under different bootstrap schemes to capture the dependence structure between different assets. With the assumption of no serial correlation in the CAPM, we conduct the independent bootstrap over time scale. Furthermore, we introduce the block bootstrap with blocks over time to capture the temporal dependence. The bootstrap tests were applied to the CAPM in the US and Vietnam (VN) stock markets, providing some interesting results.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I: INTRODUCTION

Risky asset valuation is one of the significant quantitative problems in financial economics. The concept of investment returns measures the performance and profits of an investment. The question is how risky assets are priced to measure investment returns in financial markets. The development of many asset pricing models helps investors and portfolio managers with asset valuation. Theories on risk and return and modern portfolio theory have contributed to the development of many asset valuation models. This chapter provides a literature review in financial economics and how asset pricing models were developed. The most popular model is the Capital Asset Pricing Model (CAPM) introduced by William Sharpe (1964), John Litner (1965) and Jan Mossin (1966). Sharpe's contribution to the price information for financial assets won the 1990 Nobel Prize in Economics [26].

1.1 Theoretical Background

The risk-return trade-off is a well-known fundamental principle in finance. Rational investors expect to get higher returns when risks associated with the investment are higher to compensate the increased uncertainty. In competitive financial markets, this concept holds true universally. Markowitz (1959) pioneered a mean-variance theory in selecting investment portfolios to maximize the expected return for a given level of risk. Markowitz's model framework assumes investors are efficient and risk-averse and hence, the portfolio section depends on investor's risk-return utility function [24]. His modern portfolio theory mean-variance frontiers led to the 1990 Nobel Memorial Prize in Economic Sciences. Most rational investors will choose the less risky alternative. However, higher risk does not always equal higher realized returns since there are no guarantees. Bearing additional risk gives investors the possibility of higher expected returns. Sharpe (1964) and Lintner (1965) suggested a positive relationship between the market risk premium and the expected return of an asset or portfolio [32] [23]. Sharpe (1964) introduced two market prices: the price of time (pure interest rate over investment horizon) and the price of risk (additional expected return for bearing additional risk) [32].

Derived from Markowitz's modern portfolio theory (1959) and Tobin (1958), the CAPM of Sharpe (1964), Lintner (1965), and Mossin (1966) marks the birth of asset pricing and valuation.The CAPM describes the linear relationship between the systematic risks and expected returns as a function of the risk-free rate, the asset's beta, and the expected risk premium. Beta represents the slope of the regression line and is typically estimated using the linear regression analysis of investment returns against the market returns. Theoretically, the market portfolio has a beta of one. Stocks with betas greater than one indicate a higher level of risk relative to the market's movements. In contrast, stocks with betas less than one tend to be less volatile than the market. Perold (2004) discusses the four assumptions associated with the CAPM. Firstly, the key assumption of the CAPM is that the return on asset is positive. Secondly, investors are risk-adverse meaning they prefer a lower risk for a given level of return. They evaluate their investment portfolios and make decisions solely regarding expected return and risk measured by the variance over the same single holding period. Lastly, capital markets are perfect in several senses: no transaction costs, no short selling restrictions or taxes; an ability to diversify all assets; capability to lend and borrow at the risk-free rate; and availability of all information to investors [25].

The CAPM has many implications in financial practice. At equilibrium, the CAPM provides the basic estimate of the relationship between risks and returns known as the Security Market Line (SML), which helps investors evaluate and possibly identify mispricings of an asset. In corporate finance, the CAPM is used to determine the cost of equity as part of the Weighted Average Cost of Capital (WACC).Three components can fully determine the expected return on an asset: (1) risk-free rate, (2) asset's beta to measure the asset's price movements relative to the market itself and (3) the market risk premium (Perold,2004) [25]. The simple equation of the SML is given below:

$$E[R_i] = R_f + \beta_i(E[R_m] - R_f) \tag{1.1}$$

where $E[R_i]$ is the expected return on asset i, asset beta β_i, the risk-free rate R_f, and the market risk premium is $E([R_m] - R_f)$.

There are many factor pricing models used in financial economics. Linear factor pricing models are simple but most widely used in finance. The Arbitrage Pricing Theory (APT) is another linear factor asset pricing model used in financial economics. The APT was first introduced by Ross (1976)[29]. Ross's theory predicts a linear relationship between the returns on portfolios or individual securities by incorporating macroeconomic variables as independent explanatory factors such as unemployment rate, consumer price index (CPI), crude oil price, inflation, etc. One of the advantages of the APT over the CAPM is that the APT has more flexible assumptions while retaining the higher degree of generality [22]. An empirical study done by Cagnetti (2017) shows that the APT performs better than the CAPM as it allows multiple sources of systematic risks, in all tests considered for the Italian stock market (ISM) [3]. However, accordingly, to Cagnetti (2017), the CAPM theory is intuitively simple and easy to apply in practice. The APT can be tested by checking the pricing errors whiich are represented by the intercept α [3]. The APT k-factor model holds when α=0 for the exact k-factors. Roll (1977) argues that the CAPM is untestable as the true market portfolio is unobservable. Roll's critique states that the market portfolio should include all assets in the market [28]. In practice, the market proxies are used as the market portfolio in the CAPM. The S&P 500 is often used as the market portfolio for the US stock market.

The CAPM calculates required returns based on a proxy measurement of the market risk. According to Markowitz's theories, the central idea of mean-variance theory assumes that a risk-averse investor will choose a portfolio which maximizes his/her expected return. That is, investors will choose portfolios that: (1) maximize the expected return, for a given level of risk which is measured by the portfolio variance and (2) minimize the level of risk, for a given expected return (Markowitz 1959) [24]. The CAPM states that the expected return on an asset is the risk-free rate plus beta times the expected excess rate of return. The risk-free rate of return is the rate expected on an investments assumed to have no risks involved, typically the T-Bill or 10-year government bond yield are used as the risk-free rate. Beta is equal to the covariance of the rate of return on the asset and the market return divided by the variance of the market portfolio return [32]. The beta is a measure of non-diversifiable risk (systematic risk) which is correlated with the market portfolio's variance. In general, the market is efficient enough to incorporate and reflect all relevant information.

Many mathematical and statistical methods are applied to financial economics to test a model's specifications and its validity in the empirical world. The fact that no such a model that fits all; hence, mispricings are possible. As researchers and investors are always seeking an improved model to make sound investment decisions, this study reviews previous method in testing asset pricing models and proposes a modern method in testing validity of the CAPM using bootstrap methods (Efron, 1979) in the US and Vietnam stock market. Due to some statistical model assumptions, the bootstrap methods are preferred to the previous classical approaches.

1.2 Bootstrap Methods

Most classical approaches in statistics rely heavily on restrictive assumptions such as distributional assumptions and stationarity. Bootstrapping is a non-parametric statistical procedure with many practical applications to data analysis and statistical modeling. The bootstrap is a computer-based method used to carry out statistical inferences and relies on the empirical distribution function. Fox and Monette (2002) describe the general idea for the bootstrap method by the following analogy *"The population is to the sample as the sample is to the bootstrap samples"* [12]. The bootstrap method was first introduced by Efron (1979) inspired by earlier work on the *jackknife* as a modern alternative to the classical approach. Efron's bootstrap is basically a re-sampling with replacement from the original data method and is a more computationally intensive method. Bootstrapping is used to evaluate some properties of statistical parameters without making assumptions about the distribution of the data [8]. By using re-sampling and simulation method, the bootstrap can give an estimate of the distribution of test statistics based on the bootstrapped data with a high level of accuracy and it also fits nicely into the data mining paradigm (Guszcza, 2005) [14].

The general idea for the bootstrap is quite simple. Let $E_1, E_2 ...., E_N$ be an independent and identically distributed sample of size N from distribution Ψ and let $E_1^*, E_2^*, \cdots, E_N^*$ be a bootstrap sample of size N from the original data $(E_1, E_2...., E_N)$. To get an approximation of the estimator distribution, create many alternative versions of the data by repeating the data-generating procedure for B times. Then the estimator $\hat{\theta}$ is calculated for each of the B re-sampled data sets, where $\hat{\alpha} = s(E_1, E_2, \cdots, E_N)$, with s denoting some function associated with some parameter α. The bootstrap method allows us to approximate the empirical sampling distribution and use this approximation to make statistical inferences such as confidence intervals, hypothesis tests, and so on [12].

There are two general types of bootstrapping: parametric and non-parametric bootstrap. The parametric bootstrapping estimate assumes that the data comes from a distribution family with a known probability density function (pdf). The parameters can be estimated from the sample data then used to get the approximated empirical pdf. The parametric method is useful when some knowledge about of the underlying distribution of the population is available [12]. On the other hand, in the non-parametric bootstrapping approach, the empirical function is estimated for the data by obtaining a corresponding probability density function from the histogram of the estimator $\hat{\alpha}$. In fact, if the approximation is close to Ψ, the statistical inferences based on this estimation method become statistically reliable and accurate.

The general procedure for bootstrapping consists of three steps. Firstly, re-sampling the data for a given number of times (500 times or more). Secondly, from each sample, statistics estimating the unknown parameters of interest are calculated and stacked. Finally, the sample bootstrap statistics are used to estimate the sampling distribution (bootstrap distribution) and based on the sampling distribution, we can obtain the confidence intervals for the unknown parameters.

Bootstrap is also useful in regression analysis. Given an independent variable X_i, bootstrap response values, Y_i^b is $\hat{Y}_i$ plus a randomly re-sampled residuals E_i^b, where $\hat{Y}_i$ are the fitted values of the regression model, for $i = 1, 2, \cdots, N$ is the number of assets. Then a regression model is fit to the newly generated data to get the new estimators. We adopted the model-based bootstrapping schematic illustrated by Shalizi (2017) [31] in Figure (1.1). It shows the schematic of the model-based bootstrapping method used in this study. The bootstrap offers many advantages in modeling and hypothesis testing. A great advantage of the bootstrapping method is its simplicity. It may

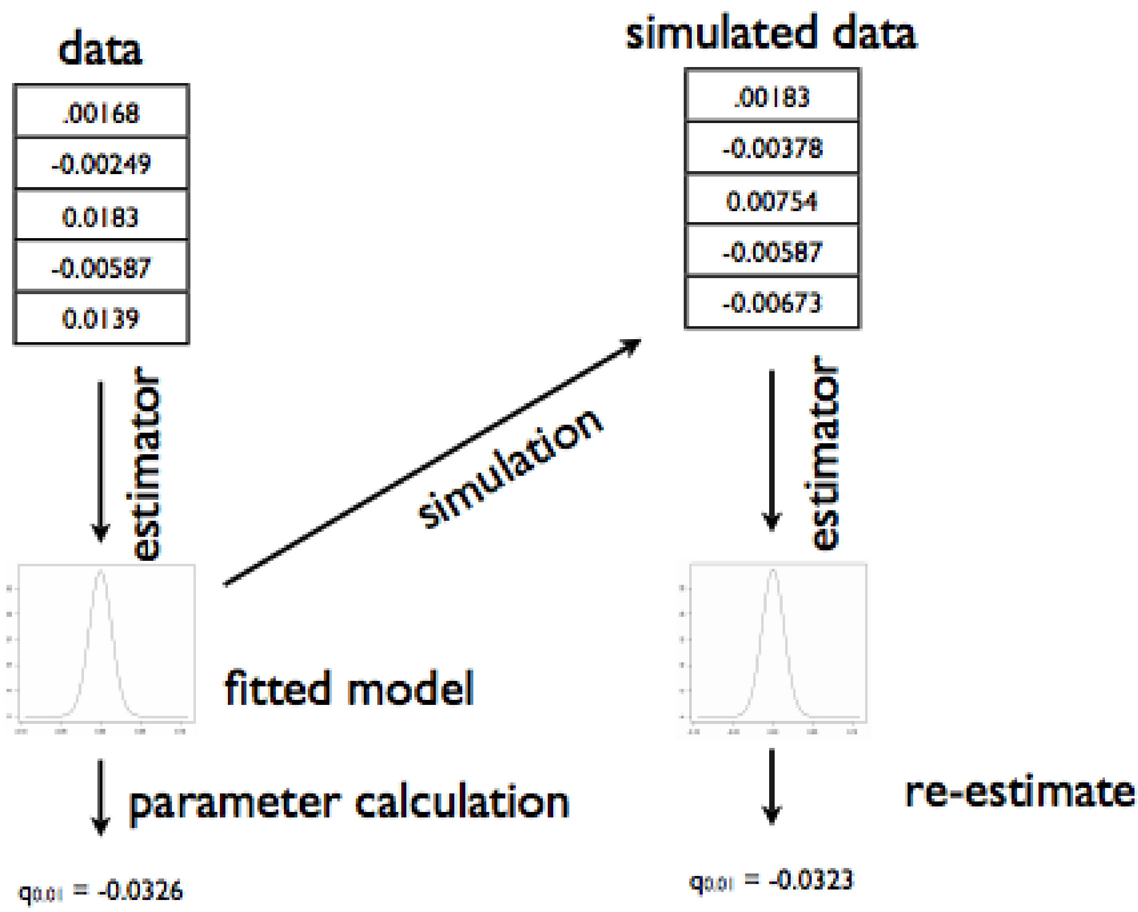

Figure 1.1
Schematic for the model-based bootstrapping: new bootstrap generated data used to re-estimate new estimators and compute test statistics. The diagram was adopted from *Advanced Data Analysis from an Elementary Point of View* (Shalizi, 2017) [31]

practically provide more accurate inferences including valid standard errors, confidence intervals, and hypothesis tests even when the sample size is small since it does not require distributional assumptions. Bootstrapping requires can be implemented in an automatic way and for complex data in which classical approaches may be difficult to apply. The estimators become more accurate for a larger sample size given that the sample is more likely to be a representative sample from the population.

Bootstrapping, on the other hand, has some disadvantages that the users need to be cautious when using. The bootstrap method relies on the quality of the original sample. As a result, if the original sample is not representative of the population, the simulated distribution of any statistics will not reflect the population accurately. It is possible that bootstrap may give results that are entirely wrong.Many previous studies have reviewed the applications of bootstrap methods in financial economics. Ruiz and Pascual (2002) point out two primary objectives for using bootstrap tests of time series data. Firstly, a bootstrap method can be used to estimate the distribution of an estimator or test statistic by re-sampling. Secondly, the probability distribution of returns can be estimated directly by applying the bootstrap technique [30].

1.3 Research Problem

Traditional methods such as Black, Scholes, and Jensen (1972), time series estimation, cross-sectional regression, and Fama and MacBeth (1973) were introduced and applied to test the pricing errors of the CAPM. However, they encountered difficulties as the number of assets grows because the number of covariances to be estimated grows quadratically with the number of assets. This task becomes more difficult when the number of assets becomes larger. Our study portfolio contains up to thirty individual assets; hence, testing all estimated pricing errors by calculating covariance matrices may not be reliable.

According to the CAPM, all pricing errors α_i should be indistinguishable from from zero. We propose to use bootstrap methods to estimate the sampling distributions of the pricing errors α. That is, under the null hypothesis H_0, all α_i should jointly be equal to zero. The bootstrap method provides some advantages over the traditional methods. By re-sampling from the original data and re-estimating estimators from bootstrap samples, we can preserve some characteristics of the data in the presence dependence between the assets and temporal dependence. Finally, the p-value under the null hypothesis H_0 can be calculated from the bootstrap test statistics. Following the idea of the block bootstrap, to capture the dependence structure between different assets, we consider all assets at one point of time as one block under different bootstrap schemes. With the assumption of no serial correlation in the CAPM, we conduct independent bootstrap over the time series. We also introduce the block bootstrap with blocks over time to capture the temporal dependence.

CHAPTER II: NOTATION AND METHODS

This chapter describes the CAPM's notations and different empirical methods used in testing the CAPM. Monthly prices of stocks cannot be used directly for analysis due to non-stationary process of stock prices. Thus, we are interested in the stock returns rather than the stock prices themselves. The first order logarithmic difference of the stock prices is applied to convert the non-stationary process in stock prices to a stationary process. Throughout this paper, R_i used in the asset pricing models denotes the monthly log return on asset i^{th}. The continuously compounded returns or log returns of the stocks are calculated using the equation below:

$$R_t = ln(\frac{P_t}{P_{t-1}}) \tag{2.2}$$

where, P_t is the adjusted closing price of stock at time t and P_{t-1} is the adjusted closing price of stock at time $(t-1)$.

One of the advantages of the continuously compounded multi-period return is the simplification of the additive process instead of multiplicative process. In modeling of financial time series returns, it is far easier to derive the properties of additive processes than multiplicative processes (Campbell et al., 1996) [4]. However, the continuously compounded return does have a disadvantage. For portfolio with w_i the simple return $R_s = (P_t/P_{t-1}) - 1$, is the weighted average of the simple returns for individual assets themselves. Unfortunately, this property is not applicable to the continuously compounded returns since the logarithm of a sum is not same as the sum of the logs. In empirical applications, this problem is minor since the continuously compounded return on a portfolio over short intervals of time is approximately close to the weighted average of the continuously compounded returns on the individual assets (Campbell et al., 1996) [4].

2.1 CAPM as One-factor model

Recall that the Capital Assets Pricing Model (CAPM) representing an asset's return is represented by three components:

$$R_i - R_f = \alpha_i + \beta_i(R_m - R_f) + \varepsilon_i \tag{2.3}$$

The three important parameters of the model are alpha (pricing error), beta (slope), and standard deviation for the ε_i terms (σ_i). The difference $(R_m - R_f)$ is called the market risk premium. Beta measures the asset's *systematic risks* associated with the market-risk part. It follows from the CAPM formula (2.3) that the error terms ε_i has mean *zero*, $E(\varepsilon_i)$=0, and are uncorrelated to with market portfolio, $cov(\varepsilon_i, R_m)$ is close to zero. The beta value can be computed as follows:

$$\beta_i = \frac{\sigma_{i,m}}{\sigma_m^2} = \frac{Cov(R_i, R_m)}{\sigma_m^2} \tag{2.4}$$

The beta value of an individual asset depends upon the choice of the *proxy* for the market portfolio. For the US based investors, one of the common proxies for market portfolio is the return on S&P 500 index could be used as the market portfolio. For Vietnam based investors, the return on Vietnam Index (VNI) is used as the proxy in the CAPM. The other source of risk comes from σ_i^2, which is known as *specific risk* associated with the asset's own fluctuations. Assets with higher betas indicate more sensitivity to the market., while a negative beta indicates the asset is negatively correlated with the market meaning that $cov(\varepsilon_i, R_m) < 0$. According to CAPM, alpha for all assets should equal zero. In practice, alpha is also referred as the pricing error or Jensen's alpha. Both beta and alpha are estimated based on historical performance and data. Sigma measures the *non-systematic* risk has little correlation to the market risk. This type of risk is also known as *idiosyncratic risk* and is *diversifiable* by diversifying assets in the portfolio.

Using the CAPM one-factor model (2.3), the total risk for an asset is decomposed into two primary components: systematic and unsystematic risk. The unsystematic risk component is related to the error term for ε_i of the model (2.3).

$$Var(R_i) = \beta_i^2 Var(R_m) + Var(\varepsilon_i) \tag{2.5}$$

Suppose a portfolio has N assets equally weighted $w = 1/N$. The return of the would be:

$$R_p = wR_1 + wR_2 + wR_3 + ... + wR_N = (\frac{1}{N})\sum_{i=1}^{N} R_i = w\sum_{i=1}^{N} R_i \tag{2.6}$$

The portfolio beta is the sum of N equally weighted of the individually betas β_i:

$$\beta_p = \frac{\sigma_{m,p}}{\sigma_m^2} = \sum_{i=1}^{n} w\beta_i \tag{2.7}$$

Similarly, the variance of the portfolio is:

$$Var(R_p) = \beta^2 Var(R_m) + \frac{1}{N} Var(\varepsilon_p) \tag{2.8}$$

As N approaches infinity, the unsystematic risk portion of (2.8) of the portfolio will disappear due to diversification and all that remains is the market risk of an asset. Thus, diversification help investors reduce overall investment risk for the same level of expected return.

2.2 Ordinary Least Squares (OLS) Regression

Let r_{it} be the excess return of the stock i^{th} and the time t and r_{mt} be the market risk premium at the time t ($r_{mt} = R_{mt} - R_{ft}$). Ordinary least squares (OLS) regression can be applied to evaluate the factor model $E(r_{it}) = \beta_i E(f)$, where f is the pricing factor for CAPM:

$$r_{it} = \alpha_i + \beta_i f_t + \varepsilon_{it} \tag{2.9}$$

where $t = 1, 2, \cdots, T$

Comparing the expected factor model and the time-series regression model (2.9), all regression intercepts α_i should be zero for the CAPM to hold. In the other words, the regression intercepts are equal to the pricing errors of the model. The standard OLS formula provides an estimate of α and β. Statistical t-tests are applied to check whether the individual pricing errors (α_i) are zero assuming that residuals of the model (2.9) are uncorrelated and have constant variance. We are interested in checking whether all pricing errors are jointly equal to zero. One of the classical approaches is to perform the Wald test to test the joint significance of the coefficients.

2.3 Time Series Estimation and Evaluation

The model (2.9) can be evaluated by following the time series approach proposed by Black, Jensen, and Schole (1972) [19]. The time series regression method can also be applied to test asset pricing models. The CAPM states that the differences in average returns for a cross-section of stocks depends linearly on the betas. Cochrane (2000) describes the time series method in testing CAPM below. With this approach, time series regression is run for each asset or portfolio over the period of time $(t = 1, 2, 3, \cdots, T)$ to estimate coefficients α_i and β_i. Then, the regression could be described as the cross section:

$$E([R_{it}] - R_{ft}) = \alpha_i + \beta_i(E[R_{mt}] - R_{ft}) \tag{2.10}$$

Let λ be the slope of the cross-sectional relationship between expected stock returns on assets $(E[R_i] - R_f)$ and β_i. The estimate of the slope λ is the expected market risk premium $\lambda = (E[R_m] - R_f)$. The time series method gives the estimates of pricing error $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\lambda}$. The CAPM can be tested under the null hypothesis that all α are jointly zero. Let Σ be the covariance matrix of regression residuals. The test can be done by looking at $\hat{\alpha}'\Sigma\hat{\alpha}$ [7].

2.4 Cross-sectional Estimation Method

Another empirical method for estimating and evaluating the CAPM is the cross-sectional regression approach. The general idea of this method is *across assets*. The cross-sectional regression is a two-step procedure. First, we run time series regressions on the model (2.3) to get estimates of β_i. We denote the intercept in this step as a to differentiate the interest estimate of pricing error α_i for the analysis throughout the paper. The equation (2.3) can be rewritten as:

$$R_{it} - R_{ft} = a_i + \beta_i(R_{mt} - R_{ft}) + \varepsilon_{it} \tag{2.11}$$

The second step is to run the cross-sectional regressions across assets on βi to get an estimate of λ:

$$(E[R_i] - R_f) = \beta_i\lambda + \alpha_i \qquad i = 1, 2, \cdots, N \tag{2.12}$$

Note that the equation (2.12) has no intercept to capture the pricing errors α as the intercept a were already included in the time series regressions. The cross-section regressions are run on the β_i as the dependent variable. The slope of the regression line is λ in this case. Overall, this method gives estimates of $\hat{\beta}$ from the time series regression (first step), the slope coefficient $\hat{\lambda}$ from the cross-section regression (step 2), and the pricing error $\hat{\alpha} = \frac{1}{T}\sum_{t=1}^{T}\alpha_i$. The model can be tested by checking by $\hat{\alpha}'Cov[\hat{\alpha},\hat{\alpha}']^{-1}\hat{\alpha} \sim \chi^2_{N-1}$, where $Cov[\hat{\alpha},\hat{\alpha}']$ is the covariance matrix of the sample pricing errors from the cross-sectional regression. The cross-sectional R^2 is used to evaluate whether a beta pricing model explains the cross-section of expected returns across assets well enough.The next stage is to test the linear relationship between betas and returns by adding a second power of the estimated beta, β_i^2. If the coefficient of the second power of the estimated beta is statistically different from zero, then the null hypothesis is rejected.

2.5 Fama and MacBeth Estimation Method

Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973) introduce a two-pass cross-sectional methodology to evaluate the asset pricing models including the CAPM [11]. The Fama and MacBeth method is a two-step procedure that regresses excess returns at each point in time on betas (from the first pass) to estimate parameters. This method is preferred to time series and cross-sectional regressions of the models in which the betas change over time. In the first pass, the betas are estimated from the OLS regressions of monthly returns on the market premium factor $(R_m - R_f)$ then estimate the risk factor $\hat{\lambda}$ from the cross-sectional regression of average returns on the β_i. In the second pass, the General Least Squares method (GLS) is applied to the second stage of cross-sectional regressions. Since returns on time series are cross-sectionally correlated, the GLS method is preferred to the OLS estimation method. This allows the error terms to be correlated or have unequal variances. The GLS estimator $\hat{\lambda}$ is given by:

$$\hat{\lambda} = (\beta'\Sigma^{-1}\beta)^{-1}\beta'\Sigma^{-1}\bar{r_p} \tag{2.13}$$

where, $\Sigma = Cov(\hat{\beta})$ and $\bar{r_p}$ is the vector of average excess monthly returns during the observed periods for the US and VN stock market.

Unlike time series regressions, standard errors for the Fama-MacBeth procedure are corrected for cross-sectional correlation between assets. The Fama-MacBeth process applied to the CAPM is given by:

- Step 1: Run regression on each asset's excess returns against market risk premium factor $(R_m - R_f)$ to determine how exposed it is to the risk factor, which is measured in the betas for N assets.

$$r_t^i = \alpha_i + \beta_i(R_{mt} - R_{ft}) + \varepsilon_t^i \qquad t = 1, 2, ..., T \tag{2.14}$$

- Step 2: Let r_i= $R_i - R_f$ be the excess return on i asset. Perform T cross-sectional regressions of the returns to obtain the estimates of betas $(\hat{\beta}s)$ calculated from Step 1. Each regression uses the same $\hat{\beta}$s found in Step 1. The goal is to determine how N expected returns are correlated to the market risk factor loading over the time T.

$$\begin{aligned} r_{i,1} &= \lambda_{1,0} + \lambda_{1,1}\hat{\beta}_i + \varepsilon_{i,1} \\ r_{i,2} &= \lambda_{2,0} + \lambda_{1,1}\hat{\beta}_i + \varepsilon_{i,2} \\ &\vdots \\ r_{i,T} &= \lambda_{T,0} + \lambda_{1,1}\hat{\beta}_i + \varepsilon_{i,T} \end{aligned} \tag{2.15}$$

where $i = 1, 2, ..., N$ for each t. The estimates of $\hat{\lambda}$ and $\hat{\alpha}$ are estimated as follows:

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_t \qquad \text{and} \qquad \hat{\alpha}_i = \frac{1}{T}\sum_{t=1}^{T}\hat{\alpha}_{it} \tag{2.16}$$

In fact, the ε_{it} in the Fama-MacBeth procedure above can be interpreted as the pricing errors, α_i. The model can be tested by checking the statistical significance of the intercepts. We want to know if all α_i are *jointly* zero for (2.16). This can be done by $\hat{\alpha}' Cov[\hat{\alpha}, \hat{\alpha}']^{-1}\hat{\alpha} \sim \chi^2_{N-1}$, where $Cov[\hat{\alpha}, \hat{\alpha}']$ is the covariance matrix of the sample pricing errors given by:

$$Cov[\hat{\alpha}, \hat{\alpha}'] = \frac{1}{T^2}\sum_{t=1}^{T}(\hat{\alpha}_t - \hat{\alpha})(\hat{\alpha}_t - \hat{\alpha})' \tag{2.17}$$

The equation (2.5) is used to calculate for the covariance matrix of the sample errors, which can be troublesome if the sample size is huge. Therefore, calculating the covariance matrix for a sample size of 30 assets over the 10-year period requires substantial computational effort and is not practical when the number of assets becomes very large. In addition, the β_i used in the first pass are unknown so they need to be estimated, which could lead to bias.

2.6 Three-factor Model: Fama-French (1992)

In an attempt to test the CAPM, the Fama-French (1992) three-factor model describes the relationship between risk and expected returns by adding two additional factors to the traditional CAPM. Following by Fama and Macbeth (1972) approach, Fama and French (1992) find that the relationship between the market betas (referred as β in the CAPM) and the average returns on NYSE, AMEX, and NASDAQ stocks are not significant during the test period of 1963-1990. However, when the variation of β that is related to size and book-to-market equity is allowed in the model, the average return is positively related to β [9]. In other words, their tests did not support the world of CAPM in which the market β is the only explanatory variable and positively related to the average stock's return. The Fama-French three-factor model argues that stock returns can be explained by the three factors: market risk premium, the out-performance of size (SMB), and the performance of Book-To-Market (BTM) ratio.

As many empirical studies have shifted to the anomalies in the CAPM framework, the Fama-French (1992) three-factor model is used as an alternative to test the CAPM in different stock markets. Their model has become one of the popular statistical models for asset pricing and valuation in finance. Amanda and Husodo (2015) find that stocks with small capitalization outperformed stocks with large capitalization for the Indonesian stock market from 2003 to 2013. Fama and French (1996) also show that firms with high Book-To-Market (BTM) are sensitive to small changes in business conditions[10].

The most widely used Fama-French three-factor model included two additional factors of SMB (small minus big) and HML (high minus low) is given by

$$R_i - R_f = \alpha_i + \beta_1 (R_m - R_f) + \beta_2 SMB + \beta_3 HML + \varepsilon_i \tag{2.18}$$

where R_i is the average stock return on the stock i, R_m is the return on a stock market index (S&P 500), SMB is defined as the difference in returns on well-diversified portfolios of small and big stocks, and HML is defined as the difference in returns on well-diversified portfolios of high and low book value-to market (B/M) stock (Fama and French, 1992). The development of the three-factor model by Fama and French (1992) is to better explain the systematic variation in average return of an asset by multiple factors rather than using the market risk alone. These factors are used in classifying stocks in terms of growth and value. Fama-French (1992) three-factor model also has some practical implications for portfolio evaluation. Portfolio managers of pension funds or mutual funds can compare their average returns of their portfolios with the average returns of the benchmark portfolios with similar size and BTM and decide whether or not making changes to achieve investment goals. Fama and French (1993, 2015) later extend the Fama-French (1992) three-factor model by adding profitability (RMW-robust minus weak) and investment (CMA-conservative minus aggressive) factors, which significantly raises the possibility for new anomalies in returns for the CAPM whose explanatory variable is the only market β.

2.7 Previous Empirical Results

The validity of CAPM is debated in the finance literature. Many empirical studies were conducted to test the model across the global financial markets. The empirical results are mixed and do not fully support CAPM. Banz (1981) finds the relationship between return and market capitalization (market size) of NYSE stocks, which suggests that the CAPM is misspecified. On average, small NYSE firms had higher risk-adjusted returns than larger companies for the 1926-1975 period [2]. A study conducted by Bajpai and Sharma (2015) show that CAPM is still significant in the Indian equity market during the testing period. Another study done by Abusharbeh and Sous (2016) for the Palestinian Exchange Market reveals that the intercept of the CAPM is insignificantly different from zero and the beta (regression slope) is equal to the excess return of market return [1].

Hasan et al (2011) investigate the risk-return relationship for the CAPM in the Dhaka Stock Exchange using monthly stock returns from 2005-2009. The results refute the CAPM hypothesis that the intercept term should be equal to zero, which provides empirical evidence against the CAPM in DSE market during the observed period [16].This result is consistent with Bilgin and Basti (2011) that there is no significant relationship between betas and risk premium in Istanbul Stock Exchange (ISE); hence, the CAPM is not valid for ISE over the sample period [1]. Koseoglu and Mercangoz (2013) also performed a study to test the validity of the standard and zero beta CAPM in ISE. They find that both models are still proper in ISE; however, the *zero beta* form of CAPM is more valid [20], which contradicts Bilgin and Bastin (2011). An analysis of the cross section of returns for Equity Real Estate Investment Trust (EREIT) done by Friday, Howton, and Conover (2000) found no significant relationship between EREIT returns and a constant beta. However, when betas are allowed to vary across bull markets, there exists a positive relationship between the betas and cross-sectional returns [13].

Model analysis and statistics testing provide many implications in financial economics. The pricing errors in the CAPM are referred to as Jensen's alpha (1967, 1968) [18], which is the abnormal rate of return (if any) of an asset or portfolio. Jensen's alpha measures the risk-adjusted performance of a portfolio or asset derived from the CAPM. In other words, Jensen's alpha is the excess return above the expected required return based on the CAPM. Investors seek to buy underpriced funds (positive alpha) to get abnormal returns on investment, which are above the efficient frontier). In contrast, negative alpha indicates that funds are overpriced.

CHAPTER III: DATA, METHODOLOGY AND ANALYSIS

3.1 Data

This section documents the data used in this study for the US and Vietnam financial market. Monthly closing prices of 20 selected non-financial stocks traded on the Ho Chi Minh City Stock Exchange (HOSE) are collected during the period from November 2007 through October 2017. We exclude financial stocks because they usually have higher leverage ratios. Economic distress in financial firms results in a higher leverage ratio, which does not have the same meaning as the leverage ratios for non-financial firms. The monthly returns on the Vietnam Index (VNI) are used as the market portfolio return R_m and the monthly 10-year government bond yields are used as the risk-free rate R_f of the CAPM. The data of monthly prices for the 20 selected VN stocks and the 10-year government bond were collected from Quandl Core Financial Data.

For the US stock market, the monthly closing price for the fifteen largest non-financial firms listed on S&P 500 (a value-weighted portfolio of S&P 500 stocks) by market capitalization.The top holdings for S&P 500 non-financial companies ranked by market capitalization were obtained from the Morningstar website. Similarly, the fifteen top holdings were selected from the Russell 2000 small market capitalization index. The monthly return for S&P 500 index was used as the market proxy for the market return in the regression model (2.3). The data of the monthly stock prices for the US stock market used were obtained from the Center for Research in Security Prices (CRSP) database. The data covers the period from January 2007 through December 2016. The interest rate on the three-month US T-Bill is used as the risk-free rate R_f used in the CAPM (2.3). The monthly yield data on the US T-Bill is obtained from the Federal Reserve Economic Data (FRED) during the period of January 2007 through December 2016.

Table 3.2 provides a summary of descriptive statistics of the 30 selected US stocks during the period 2007-2016. The second column summarizes the mean of monthly log-returns for each stock. The estimates of intercepts ($\hat{\alpha}_i$) and $\hat{\beta}_i$ were obtained from the OLS regressions for (2.3)

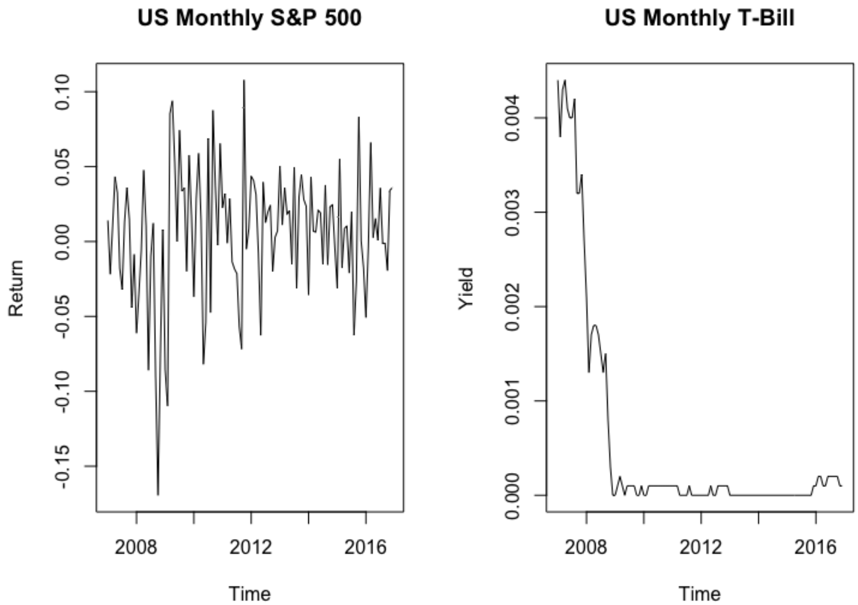

Figure 3.2
Monthly S&P 500 Market Return and 3-Month T-Bill Yield from January 2007 through December 2016. (Data: CRSP)

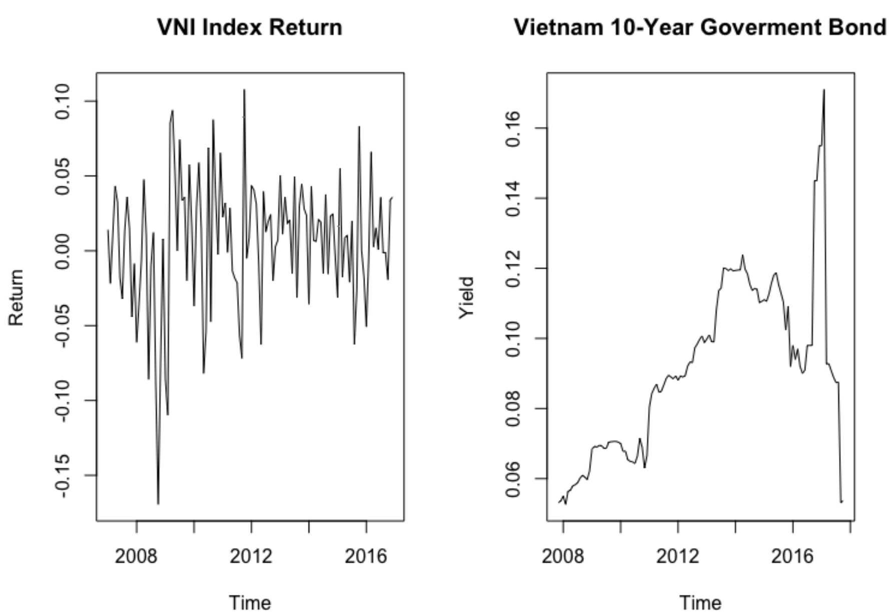

Figure 3.3
Monthly VNI Market Return and 10-Year Government Bond Yield November 2007 through December 2017 (Data: Quandl)

Table 3.1
Descriptive Summary Statistics of 20 Selected VN Stocks January 2008 through October 2017

Sticker	Mean	Std. Dev	Min	Max	Sknewness	Kurtosis	Intercept	t-value	Beta
GIL	-0.0009	0.1104	-0.3119	0.3109	-0.1478	1.0608	(0.0286)	(2.2192)	0.6914
ACL	-0.0144	0.1216	-0.5661	0.3611	-0.5221	5.2430	(0.0219)	(1.7342)	0.9091
ANV	-0.0126	0.1756	-0.7328	0.5348	-0.0733	2.9427	0.0089	0.4495	1.2219
BMC	-0.0184	0.1757	-0.5341	0.6812	0.7596	3.0295	0.0111	0.6242	1.3070
CDC	-0.0087	0.1884	-0.5987	0.7053	0.2183	2.1045	0.0401	2.1823	1.5156
DPM	-0.0103	0.0940	-0.2661	0.2952	0.0012	1.1645	(0.0239)	(2.9040)	0.8425
DPR	-0.0061	0.1092	-0.6096	0.3360	-1.0844	8.3614	(0.0135)	(1.2454)	0.9101
DXV	-0.0082	0.1313	-0.4234	0.3499	-0.0787	0.9862	(0.0278)	(1.7669)	0.7785
GTA	-0.0078	0.1033	-0.3216	0.3028	-0.1408	1.0212	(0.0266)	(2.3598)	0.7872
HAX	0.0059	0.1628	-0.3185	0.7115	0.9809	3.0374	0.0246	1.3631	1.1914
BMP	0.0084	0.1245	-0.3365	0.3462	-0.1769	0.8904	0.0175	1.5194	1.0878
KHP	-0.0087	0.0939	-0.3244	0.3762	0.1902	3.1599	(0.0277)	(2.9424)	0.7858
GMC	0.0028	0.1346	-0.3598	0.6187	1.1634	5.9075	0.0086	0.6268	1.0517
ICF	-0.0234	0.1332	-0.3836	0.5671	0.4682	3.0438	(0.0277)	(1.9000)	0.9443
KDC	0.0004	0.1307	-0.4533	0.3984	0.0099	1.8438	0.0202	1.7768	1.2038
PET	-0.0127	0.1360	-0.4768	0.4414	0.1547	2.0032	0.0098	0.8684	1.2328
PVD	-0.0120	0.1136	-0.2877	0.2744	-0.0092	-0.0113	(0.0237)	(2.0761)	0.8634
VSC	0.0092	0.1079	-0.3495	0.3259	0.2035	1.5935	(0.0086)	(0.7516)	0.7973
VNM	0.0165	0.0815	-0.2157	0.3145	0.1937	1.6637	(0.0172)	(1.9254)	0.6277
SSC	0.0034	0.1214	-0.3976	0.3971	0.3095	1.9398	(0.0120)	(0.8701)	0.8239

on the original historical data over the observed period. The residuals then are re-sampled in the bootstrap procedures to generate bootstrap samples.

Figure 3.2 provides the plots of historical monthly S&P 500 returns and 3-Month T-Bill yields over the period 2007-2016. The S&P 500 comprises of 500 large-cap US stocks. The S&P 500 the is value-weighted index as the performance indicator for the US stock market. We chose the S&P 500 index as the benchmark for the US stock market. The monthly 3-month T-Bill yield in Figure 3.2 is used as the risk-free rate R_f for the equation (2.3).

Table 3.2
Descriptive Summary Statistics of 30 Selected U.S. Stocks 2007-2016

Stock	**Mean**	**Std. Dev**	**Min**	**Max**	**Skewness**	**Kurtosis**	**Intercept**	**t-vale**	**Beta**
AAPL	0.00253	0.20127	-1.9186	0.2133	-7.57329	71.26956	(0.00394)	(0.21281)	0.07766
MSFT	0.00588	0.07210	-0.1811	0.2227	-0.02803	0.57864	(0.00004)	(0.00618)	0.35883
PM	0.00588	0.07210	-0.1580	0.2482	-0.02803	0.57864	0.00020	0.03621	0.11330
AMZN	0.02513	0.10048	-0.2931	0.4327	0.16478	2.18579	0.01892	2.05738	0.20778
JNJ	0.00458	0.04185	-0.1431	0.1036	-0.42257	1.02485	(0.00173)	(0.42366)	0.15963
XOM	0.00458	0.04185	-0.1431	0.1036	-0.42257	1.02485	(0.00173)	(0.42366)	0.15963
PG	0.00403	0.05974	-0.1590	0.1416	-0.18764	0.01536	(0.00239)	(0.43011)	0.10152
CVX	0.00218	0.04342	-0.1260	0.1017	-0.37023	0.43923	(0.00415)	(1.00795)	0.15101
T	-0.00111	0.08672	-0.3544	0.2241	-0.84195	2.86670	(0.00715)	(0.88963)	0.29906
PFE	0.00179	0.05511	-0.1945	0.1284	-0.53012	0.83451	(0.00433)	(0.82445)	0.25596
GE	0.00103	0.05201	-0.1692	0.0987	-0.65801	0.59966	(0.00547)	(1.09309)	0.06319
HD	0.00461	0.07206	-0.2339	0.1653	-0.58339	0.68405	(0.00167)	(0.24711)	0.17635
UNH	0.01001	0.06443	-0.1875	0.1489	-0.36904	0.50435	0.00385	0.61379	0.23442
INTC	0.00941	0.08120	-0.3658	0.2359	-1.59649	5.58302	0.00351	0.46094	0.37181
KO	-0.00121	0.08400	-0.7703	0.1327	-6.61913	59.98881	(0.00771)	(0.99350)	0.06020
EXAS	0.01310	0.22164	-0.7700	0.9609	0.16330	3.89403	0.00861	0.43037	1.09788
KNX	0.00474	0.07200	-0.1893	0.1795	0.08308	0.06165	(0.00169)	(0.24739)	0.09522
MKSI	0.00840	0.09198	-0.2595	0.3881	0.19384	2.21742	0.00212	0.24600	0.17636
CW	0.00795	0.07760	-0.2084	0.1698	-0.39885	0.23502	0.00172	0.23953	0.20104
LFUS	0.01324	0.10402	-0.4658	0.3997	-0.70643	4.59490	0.00795	0.85541	0.68325
IDA	0.00655	0.05112	-0.1790	0.1650	-0.08180	1.32117	0.00012	0.02376	0.09689
HLS	0.00477	0.10135	-0.3851	0.2340	-0.88196	2.36358	(0.00063)	(0.06863)	0.63226
MPWR	0.01593	0.11868	-0.5740	0.2759	-1.28710	5.13177	0.00997	0.91636	0.33951
FICO	0.00922	0.10130	-0.3910	0.2850	-0.51026	2.19668	0.00364	0.38979	0.53868
UMPQ	-0.00349	0.09568	-0.3897	0.2644	-0.51407	1.86583	(0.00979)	(1.08075)	0.16469
WGL	0.00740	0.05246	-0.1218	0.1402	-0.10159	-0.14987	0.00058	0.11640	(0.10220)
BLKB	0.00825	0.08542	-0.1956	0.2707	0.07073	0.59175	0.00243	0.30920	0.40958
VWD	0.00421	0.14000	-0.7873	0.5796	-1.34046	10.81195	(0.00053)	(0.04289)	0.97250
BRS	-0.00505	0.14011	-0.5360	0.4443	-0.40869	2.46010	(0.01082)	(0.84683)	0.43889
ROG	0.00333	0.12056	-0.4000	0.3115	-0.39453	0.79118	(0.00276)	(0.24599)	0.27044

3.2 Methodology

Literature in applied statistics reviews different empirical methods used to estimate an assets pricing model and its specification such as time series regressions, cross-sectional regressions, and the Fama-Macbeth two-pass method. However, when the sizes of portfolios increase, test statistics may be difficult to obtain. Since our data contains monthly returns of portfolios with large numbers of assets, bootstrap methods work nicely in testing the pricing errors of the CAPM. Scott and Westfall (2004), in testing of abnormal returns of assets, bootstrapping the multivariate regression model allows the model adjust the errors due to the presence of cross-sectional correlation and time series dependence in data structure. In addition, even when the residual distribution is heavy-tailed, bootstrapping in regression provides a high level of accuracy in making statistical inferences and test statistics become less biased for larger sample sizes [17]. To capture the cross-sectional correlation across assets in the portfolio and serial dependence over time of model residuals, we implement several of bootstrapping methods in the attempt to re-sample the original data to test the true alpha of the CAPM. The following sections describe different bootstrap methods used in this study: (1) independent bootstrap with OLS estimation, (2) independent bootstrap with Cochrane-Orcutt procedure, (3) circular block bootstrap with OLS, and (4) circular block bootstrap with Cochrane-Orcutt procedure.

3.3 Independent Bootstrap

Financial literature shows the Fama Macbeth (1973) method survived most of the empirical tests of asset pricing models including CAPM. However, their two-pass regression method fails to account for estimation errors among cross-sectional correlation in residuals. Bootstrap methods, on the other hand, can adapt to the sample's characteristics. In fact, standard errors and residuals from the CAPM model can be derived using a new procedure of simulation and re-sampling called independent bootstrapping or re-sampling of residuals. The general idea is to generate bootstrapped samples $\mathbf{Y}_i^b$ by re-sampling from the residuals $\mathbf{E}_i$. To simplify the mathematical notations, let $\mathbf{Y}_i$ be the vector of excess returns (dependent variable) on the i asset and $\mathbf{X}_i$ be the excess market return (independent variable). For i asset, the model (2.3) can be rewritten into a model of matrix form given data in a period of T months:

$$\mathbf{Y}_i = \alpha_i + \beta_i \mathbf{X}_i + \mathbf{E}_i \tag{3.19}$$

where, $\mathbf{Y}_i = (R_{i1} - R_{f1}, R_{i2} - R_{f2}, \cdots, R_{iT} - R_{ft})'$, $\mathbf{X}_i = (R_{m1} - R_{f1}, R_{m2} - R_{f2}, \cdots, R_{mT} - R_{fT})'$, and $\mathbf{E}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{iT})'$. The procedure for the independent bootstrap method contains the following steps:

- Step 1: Fit the model for all N assets. Obtain all estimates of α_i and β_i,and residuals for all N assets from the model (2.3). Let $\hat{\alpha}_i$ be the estimate of α_i, $\hat{\beta}_i$ be the estimate of β_i, and $\hat{\varepsilon}_t = (\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t}, \cdots, \hat{\varepsilon}_{Nt})$ where $\hat{\varepsilon}_{it}$ $=Y_{it} - \hat{\alpha}_i - \hat{\beta}_i X_{it}$ is the residual at t^{th} moment of the i^{th} asset. Then calculate the test statistic $\hat{S}_{obs} = \Sigma_{i=1}^{N} \hat{\alpha}_i^2$

- Step 2: Set the number of bootstrap replicates to B. Generate B bootstrap samples of moments from $1, 2, \cdots, T$. Let the b^{th} bootstrap sample of moments be $(i_1^b, i_2^b, \cdots, i_T^b)$

- Step 3: Generate b^{th} bootstrap sample data based on the CAPM with $\alpha_i = 0$ and $\beta_i = \hat{\beta}_i$ at the time t

$$\mathbf{Y}_t^b = \begin{pmatrix} Y_{1t}^b \\ Y_{2t}^b \\ \vdots \\ Y_{Nt}^b \end{pmatrix} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_N \end{pmatrix} (R_{mt} - R_{ft}) + \mathbf{E}_t^b \tag{3.20}$$

  where $\mathbf{E}_t^b$=$(\varepsilon_{i_1^b}, \varepsilon_{i_2^b}, \cdots, \varepsilon_{i_N^b})'$ are the re-sampled residuals for the b^{th} bootstrap.

- Step 4: Rewrite bootstrap sample at moment $t = 1, 2, \cdots, T$ to bootstrap sample for different assets by $\mathbf{Y}_i^b = (Y_{i1}^b, Y_{i2}^b, \cdots, Y_{iT}^b)'$

- Step 5: Regress the bootstrap $\mathbf{Y}_i^b$ values on the original *fixed* value $\mathbf{X}_i$ to obtain the bootstrap bootstrap test statistic S_b^* for $b = 1, 2, \cdots, B$

- Step 6: Repeat the Step 3 through 5 for B times

- Step 7: Using the p-value formula in (3.21) to determine the p-value

- Step 8: If p-value is less than the significance level, reject the null hypothesis.

Bootstrap hypothesis tests can be proceeded by constructing an empirical sampling distribution for the test statistic. The estimate of the p-value of a one-sided test for the null hypothesis is simply given by:

$$p - value = \frac{\text{number of times}[S_b^* > \hat{S}_{obs}]}{B} \tag{3.21}$$

where, p-value is the calculated probability under the null hypothesis testing, S_b^* and $\hat{S}_{obs}$ are the bootstrap test statistic and sample test statistic respectively and B is the number of replications. Reject the null hypothesis H_0 if p-value is less than α (significant level) and retain H_0 otherwise.

3.4 Cochrane- Orcutt Independent Bootstrap

One of the popular procedures that deal with serial autocorrelation in data is the Cochrane-Orcutt approach, which involves iterating in the calculation. The Cochrane-Orcutt method was introduced by Cochrane and Orcutt (1949). Their procedure allows the linear model to be adjusted for serial correlation in the error term [11]. We adopted the Cochrane-Orcutt (1949) procedure in estimating the coefficient $\hat{\beta}_i$ and incorporated their regression method into the block bootstrap procedure. Let $\hat{\rho}$ denote the estimated autocorrelation parameter of the model. The Cochrane-Orcutt two-step procedure is summarized as follows:

- Step 1: Run OLS regression on the original data using equation (2.3) to obtain coefficients and residuals ε_t^i, where ε_t^i are residuals at the moments t^{th} of the i^{th} asset

- Step 2: Run OLS regression on the residuals ε_t^i to obtain $\hat{\rho}$ using the equation below:

$$\varepsilon_t^i = \rho_i \varepsilon_{t-1}^i + \delta_t^i \tag{3.22}$$

  where δ_t^i are the error terms for (3.22) of the i^{th} asset.

- Step 3: Run regression on the new transformed variables X_{it}^c and Y_{it}^c to obtain the estimate of $\hat{\alpha}_i^c$. Then calculate the Cochrane-Orcutt test statistic $\hat{S}_{obs}^c = \Sigma_1^N(\hat{\alpha}_i^c)^2$

$$Y_{it}^c = \alpha_i^c + \beta_i^c X_{it}^c + \omega_{it} \quad \text{for} \quad i = 1, 2, \cdots, N \quad \text{and} \quad t = 1, 2, \cdots, T \tag{3.23}$$

where, $X_t^c = X_{it} - \hat{\rho} X_{i(t-1)}$, $Y_{it} = R_{it}$, and $Y_{it}^c = Y_{it} - \hat{\rho} Y_{i(t-1)}$ The Cochrane-Orcutt procedure above gives an estimate of $\hat{\beta}_i^c$ for all asset over the period of time T. Let $\mathbf{E}_i^c = (\hat{\omega}_{i1}, \hat{\omega}_{i2}, \cdots, \hat{\omega}_{iT})$ be the Cochrane-Orcutt residuals at t^{th} moment of the i^{th} asset. The Cochrane-Orcutt independent bootstrap procedure is described as follows:

- Step 1: Fit the model using Cochrane-Orcutt estimation method for all N assets. Obtain all estimates of α_i and β_i,and residuals for all N assets from the model (2.3). Let $\hat{\alpha}_i^c$ be the estimate of α_i^c, $\hat{\beta}_i^c$ be the estimate of β_i^c, and $\hat{\omega}_t = (\hat{\omega}_{1t}, \hat{\omega}_{2t}, \cdots, \hat{\omega}_{Nt})$ where $\hat{\omega}_{it} = Y_{it}^c - \hat{\alpha}_i^c - \hat{\beta}_i^c X_{it}^c$ is the residual at t^{th} moment of the i^{th} asset. Then calculate the test statistic $S_{obs} = \Sigma_{i=1}^N \hat{\alpha}_i^2$

- Step 2: Set the number of bootstrap replicates to B. Generate B bootstrap samples of moments from $1, 2, \cdots, T$. Let the b^{th} bootstrap sample of moments be $(i_1^{c^b}, i_2^{c^b}, \cdots, i_T^{c^b})$

- Step 3: Generate b^{th} bootstrap sample data based on the CAPM with $\alpha_i^c = 0$ and $\beta_i^c = \hat{\beta}_i^c$ at the time t

$$\mathbf{Y}_t^{b_c} = \begin{pmatrix} Y_{1t}^{c^b} \\ Y_{2t}^{c^b} \\ \vdots \\ Y_{Nt}^{c^b} \end{pmatrix} = \begin{pmatrix} \hat{\beta}_1^c \\ \hat{\beta}_1^c \\ \vdots \\ \hat{\beta}_N^c \end{pmatrix} (R_{mt} - R_{ft}) + \mathbf{E}_t^{c^b} \tag{3.24}$$

  where $\mathbf{E}_t^{c^b}$=$(\omega_{i_1^{c^b}}, \omega_{i_2^{c^b}}, \cdots, \omega_{i_T^{c^b}})'$ are the re-sampled residuals for the b^{th} bootstrap.

- Step 4: Rewrite bootstrap sample at moment $t = 1, 2, \cdots, T$ to bootstrap sample for different assets by $\mathbf{Y}_i^{c^b} = (Y_{i1}^{c^b}, Y_{i2}^{c^b}, \cdots, Y_{iT}^{c^b})'$

- Step 5: Regress the bootstrap $\mathbf{Y}_i^{c^b}$ values on the original *fixed* value $\mathbf{X}_i$ to obtain the bootstrap bootstrap test statistic $S_b^* = \sum_1^N (\hat{\alpha}_i^c)^2$ for $b = 1, 2, \cdots, B$

- Step 6: Repeat the Step 3 through 5 for B times

- Step 7: Using the p-value formula in (3.21) to determine the p-value

- Step 8: If the p-value is less than the significance level, reject the null hypothesis H_0

3.5 Circular Block Bootstrap

In this section, another bootstrap approach is applied as the the residuals may be correlated over time. This dependence is known as serial correlation meaning that residuals in one period ε_t are correlated with residuals in previous periods ε_{t-1}. Early in the bootstrap literature, Efron (1979) introduces a re-sampling method for independent data but fails to replicate the properties of dependent sequences. The solution for this problem is to use the block bootstrap method. Block re-sampling was introduced by Carlstein (1986). Liu and Singh (1992) apply the general results of block jackknife and bootstrap approaches to develop another bootstrap method that deals with weak dependence and general types of data sets that are not identical independently distributed.The performance of independent bootstrap may not be accurate when serial correlation is found in the data. In this case, the block bootstrap may improve the accuracy as it does relax the independent and identically distributed assumption for residuals of the data. By relying on the moving block bootstrap, it is expected to be robust to serial dependence and cross sectional correlation types of data even through the residuals are heavy-tailed [17].

The block bootstrap method, like independent bootstrap, attempts to capture the temporal dependence of the data. The general idea is to sample bootstrap blocks with length L and paste them together to form the bootstrap sample [6]. The bootstrap procedure can preserve the original properties of the data within a block. Selecting a block length is important in the block bootstrap procedure. Lahiri, Furukawa, and Lee (2007) review and compare the two data-based methods for selecting an optimal block size for the block bootstrap method [21]. The two approaches to the selection of the optimal block size are plug-in method and empirical criteria-based method. Data sample size also needs to be taken into account when determining the block length. Hall, Horowitz, and Jing (1995) suggest the possible optimal block lengths for the block bootstrap method being equal to $T^{1/3}$, $T^{1/4}$, and $T^{1/5}$ for the block bootstrap method[15].

The philosophy for the moving block bootstrap is quite simple. As the block length L goes large, the block's distribution should converge to the joint distribution of elements within a block. The blocks are assumed to be approximately independent for each asset. Similarly, the bootstrap draws are approximately close to the distribution of the entire data being re-sampling. In the presence of serial dependence in the residuals, the block bootstrap is used to draw samples of residuals in blocks. The residuals time series are divided into G blocks such that each block has L elements (block length), that is, $T = GL$. Intuitively, the value of G can be interpreted as the L^{th} lag that the block length captures dependence of the data.

The circular block bootstrap procedure to generate the bootstrap samples of the time indices is described as follows. Set the block length $L = T^{1/3}$. Let $B_j = \{i_j, i_{j+1}, \cdots, i_{j+L-1}\}$ denotes a j^{th} block of length L of the time indices. Then, each block starts from the time index j for $1 \le j \le T$. The bootstrap procedure will draw G blocks of B_j randomly with replacement from the set of blocks $\{B_1, B_2, \cdots, B_{T-L+1}\}$. Let $(B_1^*, \cdots, B_G^*)$ represents a bootstrap sample of size T with replacement from the set $\{B_1, B_2, \cdots, B_{T-L+1}\}$. Then, within the bootstrap block, B_k^* is defined as $(i_{(j-1)L+1}, \cdots, i_{jL})$ for $k = 1, \cdots, G$. The re-sampled blocks of vectorized observations $(i_1, \cdots, i_L)'$, $(i_{L+1}, \cdots, i_{2L})'$, $(i_{(G-1)L}, \cdots, i_{kL})'$. Let the b^{th} circular block bootstrap sample of the time indices be $(i_1^{b^*}, i_2^{b^*}, \cdots, i_T^{b^*})$ at the time t. Then the bootstrap $Y_i^{b^*}$ for $i = 1, 2, \cdots, N$ are generated by applying the steps given below:

- Step 1: Fit the model for all N assets to obtain all estimates of α_i and β_i,and residuals for all N assets from the model (2.3). Let $\hat{\alpha}_i$ be the estimate of α_i, $\hat{\beta}_i$ be the estimate of β_i and let $\hat{\varepsilon}_t = (\hat{\varepsilon}_{1t}, \hat{\varepsilon}_{2t}, \cdots, \hat{\varepsilon}_{Nt})$ where $\hat{\varepsilon}_t^i = Y_t^i - \hat{\alpha}_i - \hat{\beta}_i X_t$ is the residual at t^{th} moment of the i^{th} asset. Calculate the test statistic $\hat{S}_{obs} = \sum_{i=1}^{N} \hat{\alpha}_i^2$
- Step 2: Set the number of bootstrap replicates to B. Implement the circular block bootstrap

procedure described above to generate B bootstrap samples of time indices $(i_1^{b^*}, i_2^{b^*}, \cdots, i_T^{b^*})$

- Step 3: Generate b^{th} bootstrap sample data based on the CAPM with $\alpha_i = 0$ and $\beta_i = \hat{\beta}_i$ at the time t

$$\mathbf{Y}_t^{b^*} = \begin{pmatrix} Y_{1t}^b \\ Y_{2t}^b \\ \vdots \\ Y_{Nt}^b \end{pmatrix} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_N \end{pmatrix} (R_{mt} - R_{ft}) + \mathbf{E}_t^{b^*} \tag{3.25}$$

  where $\mathbf{E}_t^{b^*}$ =$(\varepsilon_{i_1^{b*}}, \varepsilon_{i_2^{b*}}, \cdots, \varepsilon_{i_N^{b*}})'$ are the re-sampled residuals for the b^{th} bootstrap

- Step 4: Rewrite bootstrap sample at moment $t = 1, 2, \cdots, T$ to bootstrap sample for different assets by $Y_i^{b^*} = (Y_{i1}^{b^*}, Y_{i2}^{b^*}, \cdots, Y_{iT}^{b^*})$

- Step 5: Regress the bootstrap $\mathbf{Y}_i^{b^*}$ values on the original *fixed* value $\mathbf{X}_i$ to obtain the bootstrap bootstrap test statistic S_b^* for $b = 1, 2, \cdots, B$

- Step 6: Repeat the Step 3 through 5 for B times

- Step 7: Using the p-value formula in (3.21) to determine the p-value.

3.6 Cochrane-Orcutt Circular Block Bootstrap

The last method we adopt to test the CAPM is the combination of Cochrane-Orcutt (1949) and the circular block bootstrap method. In this method, the estimators α_i and β_i are estimated using the Cochrane-Orcutt method, which allows the model adjust serial autocorrelation in the data. The autocorrelation test can be done by calculating the Durbin-Watson (DW) *d statistic* for lack of independence of residuals. The DW statistic is range from 0 to 4. Values of DW around 2 indicate no serial correlation in the residuals. In our method, we adopt the Cochrane-Orcutt two-step procedure to adjust the first order serial autocorrelation in residuals . The next stage is to implement the circular block bootstrap residuals to estimate the new bootstrap estimator. The Cochrane-Orcutt Circular Block Bootstrap procedure is simiarly to the circular block bootstrap described in the previous section except for the implement of Cochrane-Orcutt two-step process to obtain the estimates of α_i^c and β_i^c.

All bootstrap tests for $\Sigma\hat{\alpha}_i^2$ are conducted at the significance levels of 5% and 10%, which is related to the 95% and 90% confidence intervals of the one right-sided test. If the p-value is less than or equal to the significance levels (5% or 10%), then the null hypothesis H_0 is rejected.

CHAPTER IV: RESULTS

This section summarizes the results of bootstrap tests for the CAPM. The pricing errors α is estimated using OLS and Cochrane-Orcutt estimation method. Figure 4.4 provides the plots of the average log returns and estimated betas for both stock markets. Figure 4.5 shows the histograms of distribution for α_i for both OLS (left) and Cochrane-Orcutt estimation methods.

We conduct different bootstrap tests for both stock markets. For each bootstrap test, the histogram of distribution for the bootstrap test statistics S_b^* is plotted to visualize the shape of bootstrap distribution. Figures 4.6 and Figures 4.7 show the histograms of distribution of the bootstrap test statistics $\hat{S}_{obs}^*$ for both independent bootstrap and circular block bootstrap methods.

The numerical results are summarized in the tables at the significance level of 5% and 10% for both US and VN stock markets during the testing periods. The test statistics $\hat{S}_{obs} = \sum_{i=1}^{N} \hat{\alpha}_i^2$ obtained from the OLS and Cochrane-Orcutt estimation method are shown in the second column. Table 4.3 and 4.4 summarize the results of the bootstrap tests conducted for the US and VN stock markets. The test statistics $\hat{S}_{obs}$ obtained from the OLS and Cochrane-Orcutt estimation method are shown in the second column. The next columns show critical values and p-values for each bootstrap test at 95% and 90% CI , which is related to 5% and 10% significance levels of the right-sided test.

Table 4.3
Summary results of model-based bootstrap methods for the 30 selected U.S. stocks during the observed period (2007-2016).

MODEL-BASED BOOTSTRAP METHODS US STOCK MARKET	Bootstrap Test Significance Level			
	test statistic	p-value	$\alpha = 5\%$ critical value	$\alpha = 10\%$ critical value
Independent Bootstrap	0.00158	0.732	0.0047	0.004
Cochrane-Orcutt Independent Bootstrap	0.00159	0.720	0.004	0.0049
Circular Blocks Bootstrap	0.00158	0.584	0.0046	0.0036
Cochrane-Orcutt CBB	0.00198	0.602	0.0037	0.0043

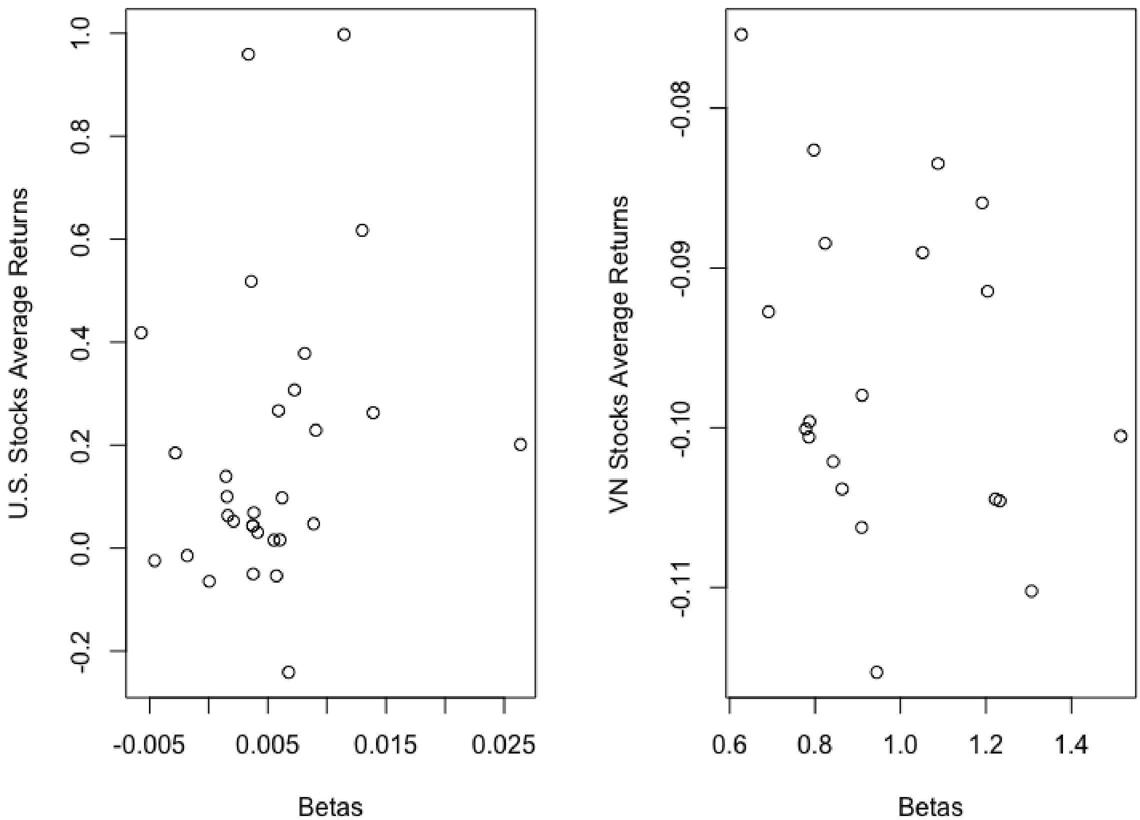

Figure 4.4
The plots of average returns and estimated betas for selected stocks for U.S. and VN stock markets during the observed periods

Table 4.4
Summary results of model-based bootstrap methods for the 20 selected VN stocks during the observed period (2008-2017)

MODEL-BASED BOOTSTRAP METHODS VN STOCK MARKET	Bootstrap Test Significance Level			
	test statistic	p-value	$\alpha = 5\%$ critical value	$\alpha = 10\%$ critical value
Independent Bootstrap	0.010	0.000**	0.006	0.0054
Cochrane-Orcutt Independent Bootstrap	0.011	0.000**	0.0064	0.0055
Circular Blocks Bootstrap	0.010	0.004*	0.0063	0.0053
Cochrane-Orcutt CBB	0.011	0.008*	0.0060	0.0053

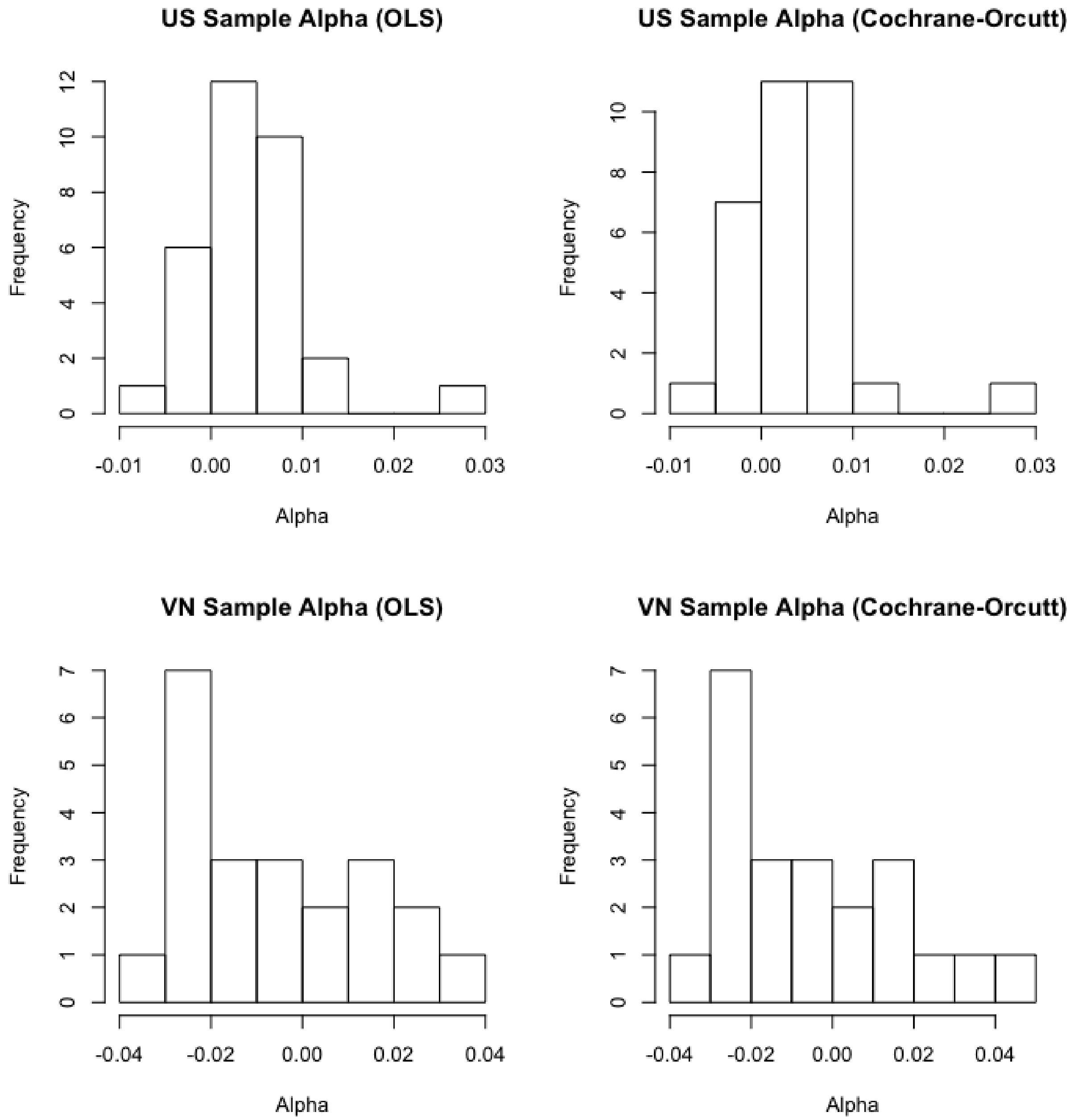

Figure 4.5
The figure shows the histograms of distribution of the sample $\hat{\alpha}_i$ obtained from using the OLS (left) and Cochrane-Orcutt (right) method of estimation for US and VN stock market during the testing periods.

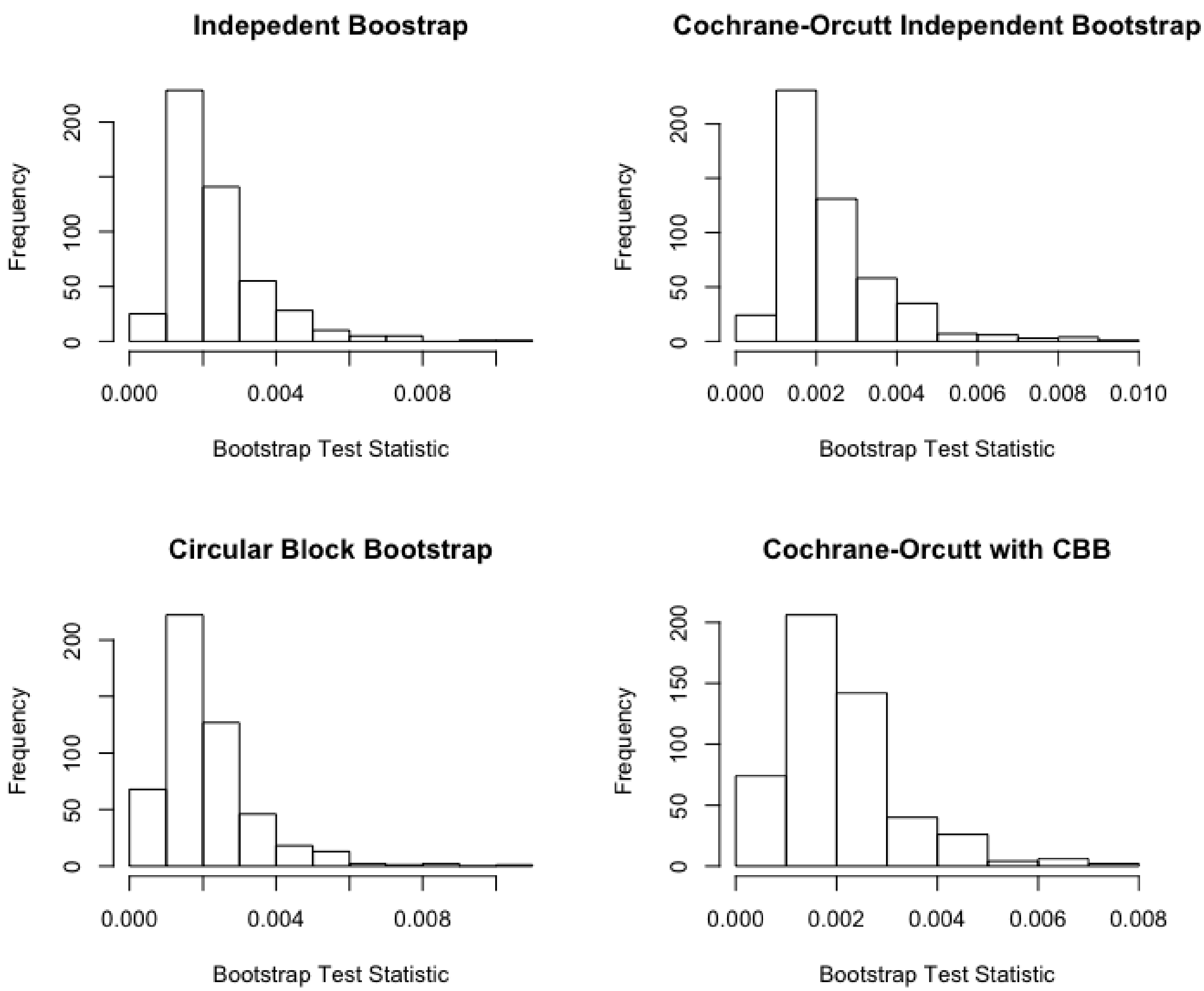

Figure 4.6
Histogram of distribution of the bootstrap test statistics S_b^* for independent bootstrap, Cochrane-Orcutt independent bootstrap, moving block bootstrap, and Cochrane-Orcutt moving blocks bootstrap for the US stocks ($B = 500$) times.

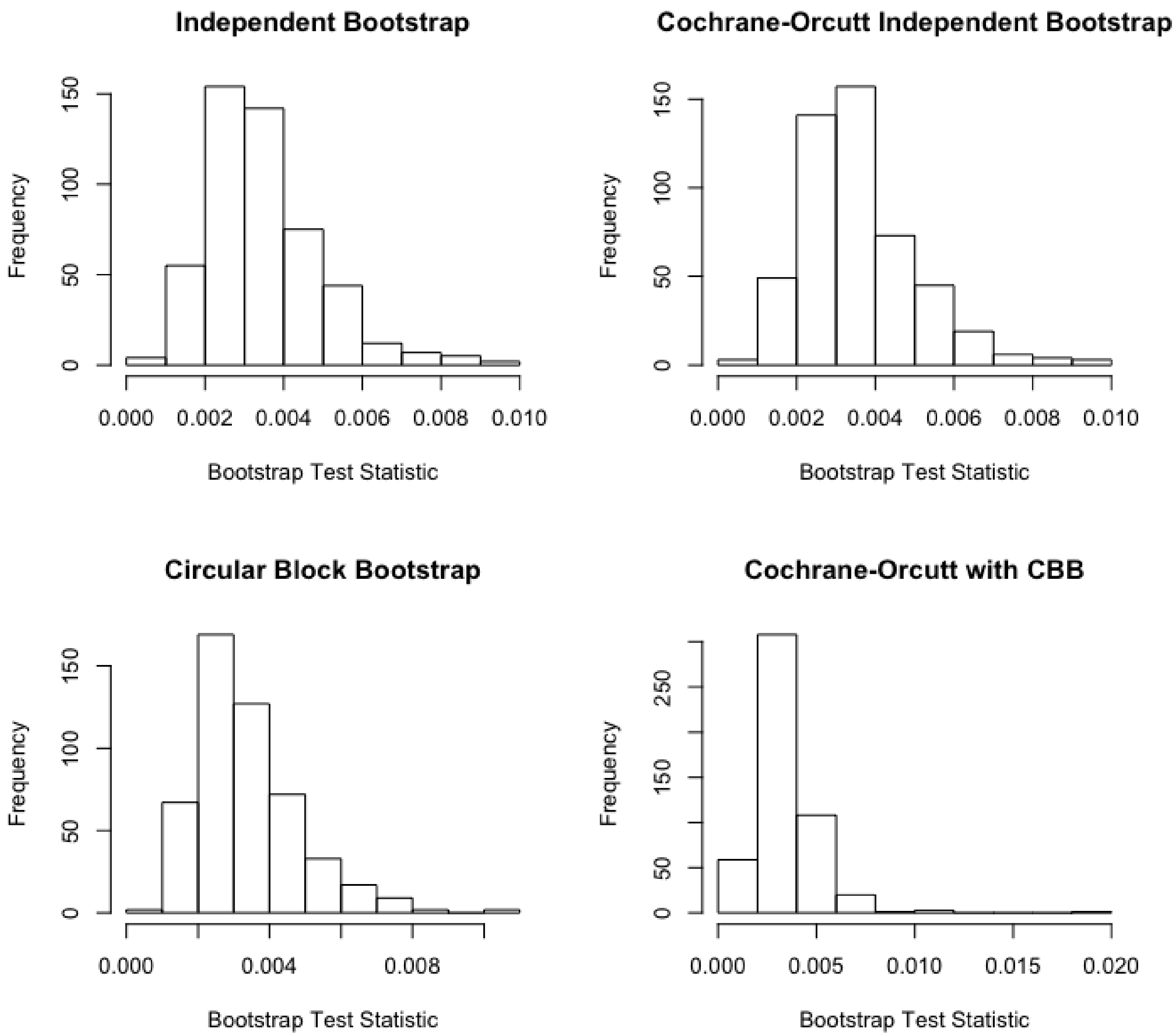

Figure 4.7
Histogram of distribution of the bootstrap test statistics S_b^* for independent bootstrap, Cochrane-Orcutt independent bootstrap, moving block bootstrap, and Cochrane-Orcutt moving blocks bootstrap for the 20 selected VN stocks ($B = 500$) times.

CHAPTER V: DISCUSSION

Figure 4.4 provides the plots of the monthly average returns and estimated betas for both the US and VN stock markets. The betas are estimated using OLS estimation method. According to the CAPM, there exists a positive linear relationship between stock returns and stock betas. For individual assets, the expected returns are found positively correlated with the market returns but statistically insignificant. Nevertheless, the joint tests are more efficient rather than looking at the betas individually.

The results are statistically significant for the VN stock market (5% significance level). That is, all pricing errors α_i are jointly different from zero at significance level of 5% over the testing period of November 2007 through October 2017 inclusively. The test statistics $\hat{S}_{obs} = \sum_{i=1}^{N} \hat{\alpha}_i^2$ are equal to 0.016 and 0.01 for the OLS and Cochrane-Orcutt regression method respectively (Table 4.3).

The null hypothesis is not rejected for the US stock market. However,the results are positive but not statistically significant over the testing period January 2007 through December 2016. Table 4.3 summarizes the results of model-based bootstrap test for the 30 selected stocks. The bootstrap test statistics $\hat{S}_{obs}$ are relatively small over the test period for both estimation methods. The average p-value obtained from the independent bootstrap is 0.726. When the serial dependence in the data are taken into account, the average p-value obtained from the block bootstrap procedure is 0.593. Both bootstrap tests failed to reject the null hypothesis H_0 over the test period.

We conduct different bootstrap methods to compare the results. The results are slightly different but agree with each other in the hypothesis testing. Our data are collected at a monthly frequency, so it may not be a big difference between the different independent bootstrap and block bootstrap. For future research in which the data are high frequency like daily and contain a large number of assets, bootstrap method may work better and yield an advantage over traditional methods.

CHAPTER VI: SUMMARY AND CONCLUSIONS

This study examines the CAPM model in the US and VN stock market. According to the CAPM, all α_i should equal zero to hold the CAPM. Under the assumption of no serial dependence in the data, we implement the independent bootstrap to capture the cross-sectional correlation between assets. Furthermore, we introduce the block bootstrap with blocks over time to capture the temporal dependence in the data. Our empirical study for the US and VN stock markets results in interesting findings. We find the pricing errors of the CAPM for the US stock market are not jointly different from zero at the significance level of 5%. On the other hand, the empirical findings on the VN stock market do not support the CAPM. The pricing errors are statistically different from zero at the significance level of 5%, which suggests the need for an improved pricing model. The Fama-French (1992) three-factor model is a possible model specification for the VN stock market.

The bootstrap method is simple and quick to implement. For this reason, the bootstrap is preferred to the traditional methods and can be used in place of asymptotic approximation methods in financial economics for a number of reasons (1) provides better approximation and higher level of refinement (2) avoid the complicated formula in calculation such as variance formula (3) correct biased estimators to less biased estimators by subtracting the bootstrap bias estimate from the original parameter estimator and (4) that when the asymptotic distribution of the estimator is not theoretically available or difficult to compute, the bootstrap methods sometimes may be used as an alternative for statistical inferences.

The use of the bootstrap test does not involve a lot of computational effort in computing the estimated p-values along with the confidence intervals. It works nicely for complicated statistical problems such as test statistics of nonparametric analysis. Specifically, for high temporal dependence data such as daily stock prices, we suggest the use of the bootstrap test for assessing a test statistic and constructing confidence intervals along with corresponding estimated p-value.

NOTES

REFERENCES

[1] Mohammed T Abusharbeh and Sana Jamil Sous. Testing the validity of capital assets pricing model: Evidence from palestinian exchange market. *Journal of Accounting, Finance and Economics*, 6(2):99–107, September 2016.

[2] Rolf W. Banz. The relationship between returns and market value of common stocks. *Journal of Financial Economics*, 9(1981):3–18, 1981.

[3] Arduino Cagnetti. Capital asset pricing model and aebitrage pricing theory in the italian stock market: an empirical study. 2017.

[4] John Y. Campvel, Andrew W. Lo, and A. Craig Mackinlay. *The Econometrics of Financial Markets*. Princeton University Press, 1996.

[5] Ke Chen and Anton Ravindran, editors. *Forging Connections between Computational Mathematics and Computational Geometry*, number 1. The 3rd Annual International Conference on Computational Mathematics and Computational Geometry (CMCGS 2014), Springer, 2016.

[6] Pin-Huang Chou and Guofu Zhou. Using bootstrap to test portfolio efficiency. *Annals of Economics and Finance*, 2:217–249, 2006.

[7] John Cochrane. *Asset Pricing*. Graduate School of Business University of Chicago, 2000.

[8] Bradley Efron. Bootstrap methods: Another look at the jacknife. *The Annals of Statistics*, 7(1):1–26, 1979.

[9] Eugene F. Fama and Kenneth R. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.

[10] Eugene F. Fama and Kenneth R. French. Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 51(1):55–84, 1996. An optional note.

[11] Eugene Famma. and James D. MacBeth. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy*, 81, 1973.

[12] John Fox and Georges Monette. *R and S Plus Companion to Applied Regression*. Sage Publications, 2002.

[13] H. Swint Friday, Shelly Howton, and Mitchell Conover. An analysis of the cross section of returns for ereits using a varying-risk beta model. *Real Estate Economics*, 28(1):141–163, 2000.

[14] James Guszcza. Introduction to bootstrapping. Technical report, Deloitte Consulting, 2005.

[15] Peter Hall, Joel Horowitz, and Bing-Yi Jing. On blocking rules for the bootstrap with dependent data. *Biometrika*, 82:561–574, 1995.

[16] Zobaer Hasan, Anton Abdulbasah Kamil, Adli Mustafa, and Azizul Baten. A validity test of capital asset pricing model for dhaka stock exchange. *Journal of Applied Sciences*, 11(20):3490–3496, 2011.

[17] Scott E. Hein and Peter Westfall. Improving tests of abnormal returns by bootstrapping the multivariate regression model with event parameters. *Journal of Financial Economics*, 2(3):451–471, 2004.

[18] Michael C. Jensen. The performance of mutual funds in the period 1945-1964. *Journal of Finance*, 23(2), 1967.

[19] Michael C. Jensen, Fischer Black, and Myron Scholes. The capital asset pricing model: Some empirical tests. *STUDIES IN THE THEORY OF CAPITAL MARKETS*, 1972.

[20] Sinem Derindere Koseoglu and Burcu Adiguzel Mercangoz. Testing the validity of standard and zero beta capital asset pricing model in istanbul stock exchange. *International Journal of Business, Humanities and Technology*, 3(7), September 2011.

[21] S.N. Lahiri, F Furukawa, and Y. Lee. A non-parametric plug-in rule for selecting optimal block lengths for block bootstrap methods. *Statistical Methodology*, 4:292–321, 2007.

[22] Cheng-Few Lee, Joseph Finnerty, John Lee, Alice C Lee, and Donald Fort. *Security Analysis, Portfolio Management, and Financial Derivaties*. World Scientific Publishing Company, 2012.

[23] John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1):13–37, 1965.

[24] Harry M Markowitz. *Portfolio Selection*. New York: Wiley, 6 edition, 1959.

[25] Andre F. Perold. The capital asset pricing model. *Journal of Economic Perspectives*, 18(3):3–24, 2004.

[26] Nobel Prizes and Laureates. The prize in economics 1990 - press release, October 1990.

[27] Boris Radovanov and Aleksandra Marcikic. A comparison of four different block bootstrap methods. *Croatian Operational Research Review*, 5:189–202, 2014.

[28] Richard Roll. A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of Financial Economics*, 4(2), 1977.

[29] Stephen Ross. The aribitrage theory of captial asset pricing. *Journal of Economic Theory*, 13(1):341–360, 1976.

[30] Esther Ruiz and Lorenzp Pascual. Bootstrapping financal time series. *Journal of Economic Surveys*, 16(3):271–300, 2002.

[31] Cosma Rohilla Shalizi. *Advanced Data Analysis from an Elementary Point of View*. Cambridge University Press, 2017.

[32] William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, September 1964.

APPENDIX A: R CODE

```
#Loading data
capm.data = read.csv("capm.csv", header=1)

#delete the missing data
capm.data.new = capm.data[!is.na(capm.data$R_SP500), 1:32]
Num.of.Assests = length(capm.data.new[1,])
T = length(capm.data.new[,1])
excess.return  = capm.data.new - capm.data.new$r_f
market.return = capm.data.new$R_SP500 - capm.data.new$r_f
Alpha.all = rep(0, Num.of.Assests)
Beta.all = rep(0, Num.of.Assests)
fitted.values.matrix = matrix(0, nrow= T, ncol=Num.of.Assests)
residual.matrix = matrix(0, nrow= T, ncol=Num.of.Assests)

#Find betas for all assests using OLS
for(a in 3:Num.of.Assests)   #the first two are risk free and market return
{
  fit = lm( excess.return[,a] ~ market.return)
  reg.coef = coef(fit)
  Alpha.all[a] = reg.coef[1]
  Beta.all[a]= reg.coef[2]
# simulating bootstrap data when alpha =0
  fitted.values.matrix[,a] = fit$fitted.values  -  Alpha.all[a]
  residual.matrix[,a] = fit$residuals
}

#Calculation of test statistic, sum square of all alpha
TestStatistic= sum(Alpha.all^2)

set.seed(1000243151)
Boot.Num = 500
Test.Stat.Boot.Vec = rep(0, Boot.Num)
Alpha.all.boot = rep(0, Num.of.Assests)

#Method 1: Independent Bootstrap
for(iboot in 1:Boot.Num)
{
  index.boot =sample(1:T, T, replace=1)
  Boot.data = fitted.values.matrix + residual.matrix[index.boot,]

  for(a in 3:Num.of.Assests)
  {
    fit = lm( Boot.data[,a] ~ market.return)
    reg.coef = coef(fit)
```

```
    Alpha.all.boot[a] = reg.coef[1]
  }
  Test.Stat.Boot.Vec[iboot]= sum(Alpha.all.boot^2)
}

hist(Test.Stat.Boot.Vec,main="Indepedent Boostrap",xlab="")
 #critical value at significance level of 0.05
cvalue = quantile(Test.Stat.Boot.Vec, 0.95)

#bootstrap p-value
pvalue = sum(Test.Stat.Boot.Vec>TestStatistic)/Boot.Num

######################################################

#Method 2: Cochrane-Orcutt Independent Bootstrap
#Using Cochrane-Orcutt in estimating coefficients
library(orcutt)
#Find Betas for all assests , The first two are risk free and market
for(a in 3:Num.of.Assests)
{
  fit.ols = lm( excess.return[,a] ~ market.return)
  fit = cochrane.orcutt(fit.ols)
  reg.coef = coef(fit)
  Alpha.all[a] = reg.coef[1]
  Beta.all[a]= reg.coef[2]
# simulating bootstrap data when alpha =0
  fitted.values.matrix[,a] = fit$fitted.values  -  Alpha.all[a]
  residual.matrix[,a] = fit$residuals
}

set.seed(1000243151)
Boot.Num = 500
Test.Stat.Boot.Vec = rep(0, Boot.Num)
Alpha.all.boot = rep(0, Num.of.Assests)

for(iboot in 1:Boot.Num)
{
  index.boot =sample(1:T, T, replace=1)
  Boot.data = fitted.values.matrix + residual.matrix[index.boot,]

  for(a in 3:Num.of.Assests)
  {
    fit.boot = lm( Boot.data[,a] ~ market.return)
    fit= cochrane.orcutt(fit.boot)
    reg.coef = coef(fit)
    Alpha.all.boot[a] = reg.coef[1]
  }
```

```
  Test.Stat.Boot.Vec[iboot]= sum(Alpha.all.boot^2)
}
TestStatistic= sum(Alpha.all^2)
hist(Test.Stat.Boot.Vec,main="Cochrane-Orcutt Independent Bootstrap")

pvalue = sum(Test.Stat.Boot.Vec>TestStatistic)/Boot.Num

cvalue = quantile(Test.Stat.Boot.Vec, 0.95)    #critical value at significance level of 0.05


######################################################
#Block bootstrap
b.size = 5
Num.of.blocks = ceiling(T/b.size)

#Block Boostrap  - Circular
for(iboot in 1:Boot.Num)
{
  index.boot =sample(1:T, Num.of.blocks, replace=1)
  boot.index = c()
  for(index.block in 1:Num.of.blocks)
  {
    boot.index = c(boot.index, index.boot[index.block]+(1:b.size)-1)
  }
  boot.index = boot.index[1:T]
  for( index.o in 1:T)
  {
    if(boot.index[index.o]>T)
      boot.index[index.o]=boot.index[index.o] - T
  }
  boot.index

  Boot.data = fitted.values.matrix + residual.matrix[boot.index,]

  for(a in 3:Num.of.Assests)
  {
    fit = lm( Boot.data[,a] ~ market.return)
    reg.coef = coef(fit)
    Alpha.all.boot[a] = reg.coef[1]
  }
  Test.Stat.Boot.Vec[iboot]= sum(Alpha.all.boot^2)
}
#critical value at significance level of 0.05
cvalue = quantile(Test.Stat.Boot.Vec, 0.95)

#Bootstrap p-value
pvalue = sum(Test.Stat.Boot.Vec>TestStatistic)/Boot.Num
```

```
########################################################
#Block bootstrap with cochrane.orcutt
b.size = 5
Num.of.blocks = ceiling(T/b.size)

#Block Boostrap  - Circular
for(iboot in 1:Boot.Num)
{
  index.boot =sample(1:T, Num.of.blocks, replace=1)
  boot.index = c()
  for(index.block in 1:Num.of.blocks)
  {
    boot.index = c(boot.index, index.boot[index.block]+(1:b.size)-1)
  }
  boot.index = boot.index[1:T]
  for( index.o in 1:T)
  {
    if(boot.index[index.o]>T)
      boot.index[index.o]=boot.index[index.o] - T
  }
  boot.index

Boot.data = fitted.values.matrix + residual.matrix[boot.index,]

  for(a in 3:Num.of.Assests)
  {
    fit.boot = lm( Boot.data[,a] ~ market.return)
    fit = cochrane.orcutt(fit.boot)
    reg.coef = coef(fit)
    Alpha.all.boot[a] = reg.coef[1]
  }
  Test.Stat.Boot.Vec[iboot]= sum(Alpha.all.boot^2)
}

#Bootstrap p-value
pvalue = sum(Test.Stat.Boot.Vec>TestStatistic)/Boot.Num

 #critical value at significance level of 0.05
cvalue = quantile(Test.Stat.Boot.Vec, 0.95)

end
```