EMPIRICAL FINANCE
Theory, Econometrics, and Evidence

A Project Report

submitted by

P.RAAJ ROHAN

for the award of the degree

of

DUAL DEGREE

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

May 3, 13

THESIS CERTIFICATE

This is to certify that the thesis titled **EMPIRICAL FINANCE SUBMITTED TO IITM**, submitted by **RAAJ ROHAN**, to the Indian Institute of Technology Madras, Chennai for the award of the degree of **Masters in Electrical Engineering**, is a bona fide record of the research work done by him under our supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr.BHARATH BHIKKAJI
Professor
Dept. of Electrical Engineering
IIT-Madras, 600 036

Place: Chennai

Date: 3rd MAY 2013

ACKNOWLEDGEMENTS

This project thesis is the outcome of my one-year of work during which I have been accompanied, guided and supported by many people. It is my great pleasure to thank Indian Institute of Technology, Madras for giving me this opportunity to work in such a beautiful and cooperative environment. I express my earnest gratitude to my guide, Dr. Bharath, for his personal guidance, constant encouragement, inspiration and feedback during the entire period of my project work.

I also thank Srini, Abhishek and Sujith for giving me the necessary background to delve into this highly interesting field. I am also grateful to my labmates Nishidh, Vamsi, Immanuel, Mithun, Shyam for their support throughout the project. I also thank Venkat, my hostel senior for providing me with some very helpful inputs regarding the project.

I would like to thank all my batch mates, wing mates who either directly or indirectly helped me in my work. And finally, I am thankful to my parents and my sister for their constant motivation and support without which this research would not have been possible.

ABSTRACT

KEYWORDS: GARCH, ARCH, heteroskedasticity, VaR, CAPM.

Abstract--- The variance of a portfolio can be forecasted using index models. Conditional variance of a portfolio helps us in evaluating the Value at Risk of a portfolio or an individual stock. Conditional volatility models estimate conditional variance of portfolio returns using multivariate or univariate volatility models. The evaluated variance and the Value at Risk for different heteroskedastic models are compared to best fit a model. In this thesis we compare the performance of S&P index as a single index and as a portfolio. The S&P 500 is regarded as a gauge of the large cap U.S. equities market. The index incudes 500 leading companies in leading industries of the U.S. economy, which are publicly held on either the NYSE or NASDAQ, and covers 75% of U.S. equities. There are different methods of comparing the forecast performance of the various models. Here, we consider the deibold mariano test. Engle (2000) proposed a Dynamic Conditional Correlation (DCC) multivariate GARCH model which models the conditional variances and correlations using a single step procedure and parameterizes the conditional correlations directly in a bivariate GARCH model. In this approach, a univariate GARCH model is fitted to a product of two return series. Parameters or model coefficients of GARCH model can be estimated by log likelihood estimation. We evaluate the forecast performance of the various volatility models on the S&P 500 data from 1990 to recent times. We later build the theoretical background necessary for constructing the capital asset pricing model. The Capital Asset Pricing (CAPM) model is a factor model that helps in Portfolio Selection. CAPM relates expected return to a measure of risk. This measure, now known as beta, used the theoretical result that diversification allows investors to escape the company specific risk. Thus, they only get rewarded for their portfolio's sensitivity to the level of economic activity. We apply this model on the discussed S&P index to test for the mean-variance efficiency of the proxy for the market portfolio and the tangency portfolio.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

ARCH	Auto Regressive Conditional Heteroskedasticity.
GARCH	Generalized Auto Regressive Conditional Heteroskedasticity.
EGARCH	Exponential-Generalized Auto Regressive Conditional Heteroskedasticity.
VaR	Value at Risk.
S&P	Standard & Poor’s.
CAPM	Capital Asset Pricing Model.
CML	Capital Market Line.

NOTATIONS

r_t	Return at time *t*.
P_t	Price at time *t*.
μ	Mean.
σ	Variance.
u_t	Residuals at time *t*.
δ	Adjusted residuals at time *t*.
Z_t	Risk free returns.
R_f	Risk free rate.
α	Intercept.
β	Beta.

CHAPTER 1

INTRODUCTION

This thesis evaluates the performance of single index and portfolio models in forecasting value at risk (VaR) by using GARCH, Risk metric and rolling window-type models. The performances of each forecast model on data from S&P index of the U.S. stock exchange are compared using deboild mariano tests. Then, the theoretical background for the Capital Asset Pricing model is established to check the performance of portfolios that constitute the index. We check the hypothesis on monthly data from 1926 to 2011 and 1963 to 2011 and also test for the tangency portfolio.

1.1.STOCHASTIC PROCESS

Economic data such as daily stock prices, foreign exchange, GDP etc. are examples of discrete stochastic processes.

1.1.1. STATIONARY TIME SERIES

A time series $\{r_t\}$ is strictly stationary if the joint distribution of $(r_{t_1}, r_{t_2}, \ldots, r_{t_k})$ is identical to that of $(r_{t_1+t}, r_{t_2+t}, \ldots, r_{t_k+t})$ for all *t*, where *k* is an arbitrary positive integer and $(t_1, t_2, ..., t_k)$ is a collection of random integers. A time series is weakly stationary if:

$$\begin{aligned} &E(P_t) = \mu \\ &Var(P_t) = E(P_t - \mu) = \sigma^2 \\ &Cov(P_t, P_{t+k}) = E[(P_t - \mu)(P_{t+k} - \mu)] = \gamma_k \end{aligned} \tag{1. 1}$$

1.1.2. NON-STATIONARY TIME SERIES

Any time series that does not follow the above properties is a non-stationary time series. A non-stationary process can have a varying mean and/or variance. Asset prices like the daily stock prices follow a Random walk, which is non-stationary. There are two types of these random walks, with and without drift.

- **RANDOM WALK WITHOUT DRIFT**

Let r_t be white noise with zero mean and a variance of σ^2. This series can be called a random walk if:

$$P_t = P_{t-1} + r_t \tag{1. 2}$$

In a random walk model the value at t is equal to the value at t-1 plus a random shock i.e. an AR(1) process. In general, if the process started at some time 0 with the value Y_0, we have

$$\begin{aligned} P_t &= P_0 + \sum_{t=1}^{t} r_t \\ E(P_t) &= E\left(P_0 + \sum_{t=1}^{t} r_t\right) = P_0 \\ V(P_t) &= t\sigma^2 \end{aligned} \tag{1. 3}$$

- **RANDOM WALK WITH DRIFT**

Let

$$P_t = \mu + P_{t-1} + r_t \tag{1. 4}$$

Where μ is known as the drift parameter.
We can also write as:

$$\begin{aligned} P_t - P_{t-1} &= \mu + r_t \\ E(P_t) &= P_0 + t\mu \\ V(P_t) &= t\sigma^2 \end{aligned} \tag{1. 5}$$

This is an AR(1) process too. Random Walk Model with drift has its mean and variance increasing over time and so violates the weak stationary conditions.

1.2. VALUE-AT-RISK (VaR)

The Value at Risk gives an estimate of the probability with which a loss of not more than *V* is made in the next *N* days. Here, *V* is the VaR of the portfolio. It depends on the time duration *N*, the confidence level needed. It is the loss level over *N* days that the analyst is *P*% certain will not exceed. *P* here is he confidence level. VaR gives an estimate of how bad, things can get in the next N days. A conditional VaR is the expected loss during an N-day period that we are in the left tail of the distribution. In practice we shrink the N-day interval to 1 day, to get a comparative study. The assumption thus made is:

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

The formula is exact when change in the portfolio valuation on successive days follows an independent identical normal distribution with zero mean. In other cases it is an approximation.

1.3. CONDITIONAL VOLATILITY MODELS

The basic model to forecast variances and covariance (volatility and correlations) of a multivariate normal distribution is through the rolling window forecast. For a zero mean return r_t this means:

$$\sigma_t^2 = \frac{1}{q}\sum_{s=1}^{q} r_{t-s}^2 = (r_{t-1}^2 + r_{t-2}^2 + \ldots + r_{t-q}^2)/q \qquad (1.6)$$

Where the last q observations are used. Here σ^2 depends on lagged information and can be thought of as a prediction (made at t-1) of volatility in t.

J.P Morgan Risk Metrics (1995) introduced an alternative to the rolling window forecast called the Exponentially Weighted Moving Average (EWMA). This estimator of volatility uses all data points since the beginning of the sample but with recent observations carrying larger weights. The weight for a lag s = $(1-\lambda)\lambda^s$, where $0 < \lambda < 1$, so

$$\sigma_t^2 = (1-\lambda)\sum_{s=1}^{\infty} \lambda^{s-1} r_{t-s} = (1-\lambda)(r_{t-1} + \lambda r_{t-2} + \lambda^2 r_{t-3} + \ldots) \qquad (1.7)$$

For the EWMA model, the conditional variance or r_t is proportional to the time horizon k. EWMA is a special form of GARCH(1,1) without drift or zero intercept. Time series realization of returns often shows time-dependent volatility. Engle's (1982) ARCH (Auto Regressive Conditional Heteroskedasticity) model was the first to include a time dependent volatility process. Volatility, thus computed, is a two-step regression on asset returns. GARH is an advanced form of ARCH. Other models include GJR, EGARCH, (G)ARCH-M, and etc.

1.4. FACTOR MODELS

One of the cornerstones of investment science is the theory of Portfolio Selection. Many major hedge funds and investment banks offer different portfolios to their clients based on their outlook of the market with a considerable amount of money riding on the market. The capital asset pricing model is considered to check if the market portfolio is the tangency portfolio and various other tests to check the robustness of the model and other discrepancies.

Various aspects of the CAPM model are investigated in this thesis. The first aspect is the theoretical background of the model. Here, mean-variance analysis (MVA) is thoroughly examined. First, a mathematical argument from utility theory that can motivate the implementation of MVA is presented. Then, efficient portfolios are examined in a mean-standard deviation space assuming there is no risk- free asset. We show the incentive to diversify ones portfolio and derive the efficient frontier consisting of the portfolios with the maximum expected return for a given variance. Using mathematical and economic arguments, it is shown that the market portfolio consisting of all risky assets is mean-variance efficient. A riskless asset in then included in the analysis to get the Capital Asset Market Line (CML) in a mean-standard deviation space. We argue that this is the efficient frontier when a risk-free rate exists. Based on the CML, the Capital Asset Pricing Model (CAPM) is derived. The CAPM relates the expected return on any asset to its beta. As beta is based on the covariance of returns between an asset and the market portfolio, it follows that CAPM only rewards investors for their portfolios responsiveness to swings in the overall economic activity.

With this theoretical background, the econometric methods for testing the CAPM are developed. First, the traditional model is rewritten in order to work with excess returns. Then, the market portfolio is tested for its mean-variance efficiency. Later, the assumption that returns are independent and identically distributed and jointly multivariate normal is introduced. Based on this assumption, the joint probability density function (pdf) of excess returns conditional on the market risk premium is derived. Using this pdf derive maximum likelihood estimators of the market model parameters. A number of different test statistics (for CAPM) are derived based on these estimators. The first is an asymptotic Wald type test, which is then transformed into an exact F test.

1.5.METHODOLOGY

Risk is often measured as a change in stock price. Most processes measure change in value of a portfolio in terms of log price changes. Continuously compounded returns are:

$$r_{t+1} = \ln(P_{t+1}) - \ln(P_t) \quad (1.8)$$

Where P_{t+1} and P_t are the known prices of an asset at time *t+1* and *t*. We first check for the autocorrelation among the returns and later check for ARCH and other characteristics using the Ljung-Box-Perice and Engle's ARCH Test applied on daily squared returns. After which we estimate the parameters of each of the given models. We then evaluate the forecast performance of these models using the Deibold Mariano test statistic. On the other hand, back testing on each model and confirm the number of VaR violations. Suppose that the financial position is a long position so that loss occurs when there is a big price drop. If the probability is set to 5%, then RiskMeterics uses $1.65\sigma_{t+1}$ to measure the risk of the portfolio-that is, it uses the one sided 5% quantile of a normal distribution with mean and variance 0 and σ_{t+1}. The actual 5% quantile is $-1.65\sigma_{t+1}$, but the negative sign is ignored with the understanding that it signifies a loss. Consequently, if the standard deviation is measured in percentage, the daily VaR of the portfolio under RiskMeterics is:

$$\textbf{VaR} = \textbf{(Amount of Position)}\ (1.65\sigma_{t+1})$$

In addition, that of a k-day horizon is

$$\textbf{VaR} = \textbf{(Amount of Position)}\ (1.65\sigma_{t+1})$$

Where the argument (k) of VaR is used to denote the time horizon. We have

$$\textbf{VaR} = \sqrt{}\ \textbf{k x VaR}$$

This is referred to as the square root of time rule in VaR calculation under RiskMetrics.

The capital asset pricing model is introduced with the necessary background assumptions. We later test these assumptions with the relevant test statistics developed. We discuss the choice of proxies, construction of dependent variables and various other aspects. Then, the tests are carried out on a 85 year sample of American stock. And for this period we conduct mean variance analysis and test to confirm the hypothesis using the *F* statistic.

CHAPTER 2

CONDITIONAL HETEROSCEDASTIC MODELS

2.1. INTRODUCTION

Time-variation in volatility, also known as heteroskedasticity is a common feature of macroeconomic and financial data.

2.2. ASSUMPTIONS FOR OLS REGRESSION

Ordinary least square (OLS) regression analysis relies on four assumptions to give the Best Unbiased Linear Estimate (BLUE). We also consider the assumption of homoscedasticity. We assume a linear relation between dependent and explanatory variables i.e.

$$Y = \beta_0 + \beta_1 X + u \tag{2. 1}$$

Where β_0 is the intercept, β_1 the slope and u is the error term with factors that affect Y and are not the specified independent variable (linearity). Second, it is intuitively a necessity that the sample to be analyzed must consist of a random sample of the relevant population to yield an unbiased result (random sampling). Third, a zero conditional mean is assumed. This means that the average error term of the function should always be zero. This can be refined into the assumption that the average value of $\boldsymbol{u}$ does not depend on the value of X as for any value of X the average value of $\boldsymbol{u}$ will be equal to the average value of $\boldsymbol{u}$ in the entire population, which is zero (zero conditional mean). Increasingly however, econometricians are being asked to forecast and analyze the size of the errors of the model. In this case, the questions are about volatility, and the standard tools have become the ARCH/GARCH models be expressed as:

$$E(u \mid x) = E(u) = 0 \tag{2. 2}$$

Fourth, a sample variation in the independent variables X_i. This means that X is not a constant and is fixed in repeated samples (sample variation). This is however is not an assumption that is likely to fail in statistical analysis, as a completely homogenous population is not the typical target for statistical analysis. These assumptions assure an unbiased result where the sample β_n is equal to the population β_n.

Finally, homoscedasticity is assumed to obtain a consistent result. This assumption states that the value of the variance of error term u conditional on the explanatory variable *X* is constant. In other words, the pattern of distribution of error terms at any given value of *X* will show the same distribution with a mean around the sample *βnX*. This is expressed as:

$$Var(u \mid x) = \sigma^2 \qquad (2.\ 3)$$

2.3. HETEROSKEDASTICITY

A time variation in volatility is termed as heteroskedasticity [1]. The effect of violation of homoscedasticity assumption in the BLUE model is that the test statistic (regression output) is no longer reliable. The distribution of the error term is no longer a constant variance distribution. It is often observed that there exists a relationship between volatilities from one period of volatility cluster to another.

2.4. HETEROSKEDASTIC MODELS AND ITS SPECIFICATIONS

The rolling window uses:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(r_t - \bar{r})^2} \qquad (2.\ 4)$$

The assumption that all past prices have equal relevance is a very crude one. A better way to estimate is to use *q* recent observations and ignore the previous (under the assumption that prices, several years ago do not affect current day prices. An even more sophisticated model would be the exponentially weighted volatility model (EWMA).

Models typically estimate the conditional variance of a portfolio by modeling historic stock/ returns or by modeling conditional variance of each asset and the conditional correlation of each pair of assets.

2.4.1. ARCH MODEL

Variance as we know changes over time i.e. it is not homoscedastic but rather a heteroskedastic process. More over it is also not beneficial to apply equal weights. ARCH overcomes these assumptions by letting the weights of the parameters to be estimated thereby determining the most appropriate weights to forecast the variance.

Consider the regression:

$$y_t = x'_t b + \delta_t, \text{ where } E\left[\delta_t\right] = 0 \text{ and } Cov\left(x_{it}, \delta_t\right) = 0 \tag{2. 5}$$

$$\delta_t = \varepsilon_t \sigma_t, \text{ where } \varepsilon_t \sim iid \text{ with } E_{t-1}\left[\varepsilon_t\right] = 0 \text{ and } E_{t-1}\left[\varepsilon_t^2\right] = 1 \tag{2. 6}$$

The ARCH(1) model states:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \delta_{t-1}^2 \tag{2. 7}$$

By increasing the lags we capture more of the dependence on previous prices. The ARCH(p) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \delta_{t-1}^2 + \alpha_2 \delta_{t-2}^2 + \ldots + \alpha_p \delta_{t-p}^2 \tag{2. 8}$$

2.4.2. GARCH MODEL

Bollerslev (1986) introduced an extension calling it the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. For a log return r_t, let δ_t be the residuals in the mean adjusted log returns. Under the assumption that δ_t follows a GARCH model, we have:

$$\delta_t = \sigma_t \varepsilon_t; \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^{u} \alpha_i \delta_{t-i}^2 + \sum_{j=1}^{v} \beta_j \sigma_{t-j}^2 \tag{2. 9}$$

Since, we know that r_t is log return of daily stock prices with mean μ and δ_t is the mean corrected log return:

$$\delta_t = r_t - \mu \tag{2. 10}$$

Where δ_t is a random variable with mean 0 and variance 1, and assumed to be normally distributed. If $v = 0$, then the GARCH(u,v) reduces to an ARCH(u) process. The restrictions $0 \le \alpha, \beta \le 1$, $\alpha + \beta < 1$ are needed to ensure that $\sigma_t^2 > 0$ and to make σ_t^2 stationary and therefore the unconditional variance finite. In period t, we can forecast the future conditional variance (σ_{t+s}^2) as (since σ_{t+1}^2 is known in t)

$$E_t\sigma_{t+s}^2 = \bar{\sigma}^2 + (\alpha+\beta)^{s-1}(\sigma_{t+1}^2 - \bar{\sigma}^2), \text{ with } \bar{\sigma}^2 = \frac{\alpha_0}{1-\alpha-\beta} \quad (2.\ 11)$$

where $\bar{\sigma}^2$ is the unconditional variance.

2.4.3. EGARCH MODEL

an asymmetric GARCH model can be constructed as :

$$\sigma_t^2 = \alpha_0 + \alpha r_{t-1}^2 + \beta\sigma_{t-1}^2 + L_i G(r_{t-1} > 0)\delta_{t-1}^2, \text{ where } G(q) = 1 \text{ if q is true else } G(q){=}0 \quad (2.\ 12)$$

This means that the effect of shock r_{t-1}^2 is α if the shock was negative and is $(\alpha + G)$ if the shock is positive. With *G<0*, volatility increases to a negative r_{t-1} (bad news) than to a positive return.

The exponential GARCH (EGARCH) sets:

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{u}(\beta_i \ln\sigma_{t-1}^2) + \sum_{j=1}^{v}\alpha_j\left[\frac{|\delta_{t-j}|}{\sigma_{t-j}} - E\left(\frac{|\delta_{t-j}|}{\sigma_{t-j}}\right)\right] + \sum_{j=1}^{u}L_j\left(\frac{\delta_{t-j}}{\sigma_{t-j}}\right) \quad (2.\ 13)$$

Where $\delta_t = \sigma_t\varepsilon_t$, Being written in terms of log makes $\sigma_t^2 > 0$ hold without any restrictions on the parameters. This being an asymmetric model both negative and positive return affects the volatility in the same way, the linear term makes the effect asymmetric.

2.4.4. GJR MODEL

GJR is an extension to GARCH model, it captures the asymmetries between effects of positive and negative shocks of similar magnitude. The model is given by:

$$\sigma_t = \alpha_0 + \sum_{j=1}^{u}\alpha_j\delta_{t-j}^2 + \gamma D(\psi_{t-1})\delta_{t-1}^2 + \sum_{i=1}^{v}\beta_i\sigma_{t-i} \quad (2.\ 14)$$

Where $\delta_t = \sigma_t\varepsilon_t$ and $D(\psi_t)$ is defined as:

$$D(\psi_t) = \begin{cases} 1 & \text{if } \delta_t \le 0 \\ 0 & \text{if } \delta_t > 0 \end{cases}$$

for u=1, $\alpha_0 > 0,\ \alpha_1 > 0,\ \alpha_1 + \gamma_1 > 0, \beta_1 \ge 0$, to ensure positive variance. The indicator variable D distinguishes between shocks of short run persistence ($(\alpha_1 + \gamma_1)/2$) and long run persistence ($(\alpha_1 + \beta_1 + \gamma_1)/2$)

2.5. PARAMETER ESTIMATION

Historical observations are used to estimate the parameters with the help of Maximum likelihood functions. With the assumption that the probability distribution of ε_i conditional on the variance is normal, we maximize:

$$\prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_t^2}} e^{\left(\frac{-\delta_i^2}{2\sigma_i^2}\right)} \tag{2. 15}$$

Where m is the number of observations. On taking logarithm and ignoring constants, the equation boils down to maximizing:

$$\sum_{i=1}^{m} \left(-\ln \sigma_t^2 - \frac{\delta_i^2}{\sigma_t^2} \right) \qquad 2. 16)$$

An iterative search helps estimate the parameters [1]. First we guess values of parameter b and α_0 and α. The guess of b can be taken from the LS estimation of (2.5), and guess, the guess of α_0 and α from an LS estimation of $\hat{\delta}_t^2 = \omega + \alpha\hat{\delta}_{t-1}^2 + \varepsilon_t$ where $\hat{\delta}_t$ are the fitted residuals from the LS estimation of (2.5). Now loop over the sample (for t=1, t=2,…) and calculate $\hat{u}_t$ from (2.5) and σ_t^2 from (2.6). Plug these numbers in (2.15) to find the likelihood value. Third make better guesses of the parameters and do the second step again. Repeat until the likelihood value converges (at a maximum).

The Value at Risk thus evaluated, say at 0.95 confidence level is $VaR_\alpha = -\text{cdf}^{-1}(1-\alpha)$, i.e. for a normal distribution we have $VaR_{95\%} = -(\mu - 1.64\sigma)$.

CHAPTER 3

DATA

For this study, S&P 500 index data from dec 1998 to dec 2010 is considered.

3.1. PRE-ESTIMATION ANALYSIS

The presence of GARCH is tested using the Ljung-Box Pierce test and Engle' ARCH test. The Ljung-Box Pierce Q-Test (LBQ Test) performs a lack-of-fit hypothesis test for model misspecification, which is based on the Q-statistic:

$$Q = N(N+2)\sum_{k=1}^{L}\frac{r_k^2}{(N-k)} \tag{3. 1}$$

Where N = sample size, L = number of autocorrelation lags included in the statistic, and r_k^2 *is* the squared sample autocorrelation at lag k. Once you fit a univariate model to an observed time series, you can use the Q-statistic as a lack-of-fit test for a departure from randomness [1]. The Q-test is most often used as a post estimation lack-of-fit test applied to the fitted innovations (i.e., residuals). In this case, however, you can also use it as part of the pre-fit analysis because the default model assumes that returns are just a simple constant plus a pure innovations process. Under the null hypothesis of no serial correlation, the Q-test statistic is asymptotically Chi-Square distributed.

As for Engle's ARCH Test, the ARCH test also tests the presence of significant evidence in support of GARCH effects (i.e. heteroskedasticity). It tests the null hypothesis that a time series of sample residuals consists of independent identically distributed (i.i.d.) Gaussian disturbances, i.e., that no ARCH effects exist. Given sample residuals obtained from a curve fit (e.g., a regression model), this test tests for the presence of u^{th} order ARCH effects by regressing the squared residuals on a constant and the lagged values of the previous M squared residuals. Under the null hypothesis, the asymptotic test statistic, $T(R^2)$, where T is the number of squared residuals included in the regression and R^2 is the sample multiple correlation coefficients, is asymptotically chi square distributed with M degrees of freedom. When testing for ARCH effects, a GARCH(u,v) process is locally equivalent to an ARCH($u+v$) process.

All the analysis about pre estimation analysis were did in MATLAB 7.0, Q-statistics or Engle's ARCH-statistics, p-value and critical values at 95% confidence level for 10, 15 and 20 lags are generated in MATLAB. Both functions return identical outputs. The first output, H, is a Boolean decision flag. H = 0 implies that no significant correlation exists (i.e., do not reject the null hypothesis). H = 1 means that significant correlation exists (i.e., reject the null hypothesis). The remaining outputs are the P-value (p-Value), the test statistic (Stat), and the critical value of the Chi-Square distribution (Critical Value).

Table 3.1: The results for the LBQ-test:

Lags	H	P-Value	Stat	Critical Value
10	1	0	48.47	18.31
15	1	0	71.16	25.00
20	1	0	98.33	31.41

Table 3.2: The results for ARCH-test

Lags	H	P-Value	Stat	Critical Value
10	1	0	38.48	18.31
15	1	0	39.03	25.00
20	1	0	46.15	31.41

Using LBQ-test, It is observed that no significant correlation is present in the raw returns when tested for up to 10, 15, and 20 lags of the ACF at 0.05 level of significance. Under the null hypothesis H_0: No correlation and H_a: correlation is present.

H is a Boolean decision flag, is equal to 1, which indicates that significant exist (i.e. rejects the null hypothesis that correlation is present). However, P-value at any lags of 10, 15 and 20 are less than .05 and all *Q*- statistics exceeds its corresponding critical values, suggesting that correlation is present in log returns of S&P-500 at 5% level of significance. In the above output H=1, p-value less 0.05 and equal to zero, ARCH test statistics exceeds its critical value. Therefore, ARCH test strongly rejects the null hypothesis that there is no ARCH/GARCH effect in given return of S&P-500. Finally, log returns have an ARCH effect at significance level of 5% and given time series has no random sequence of Gaussian disturbance.

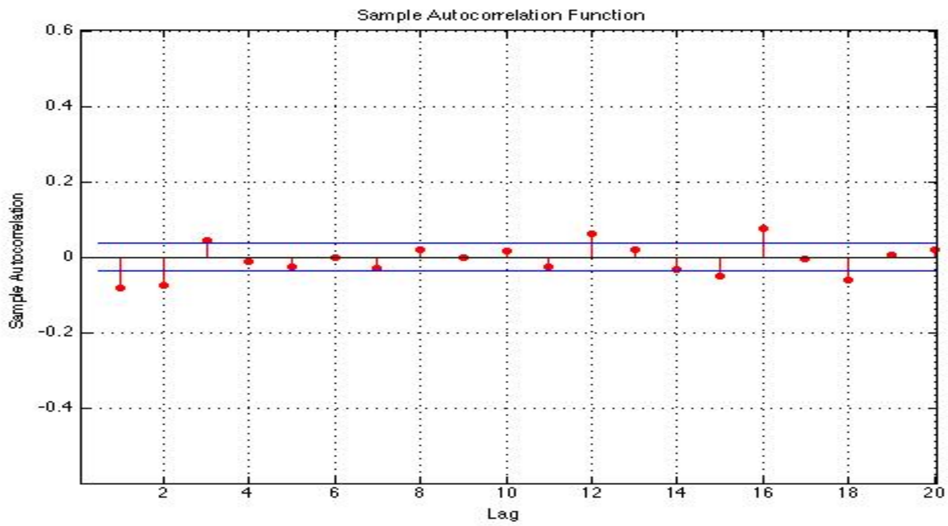


Figure 3.1:Autocorrelation of returns.

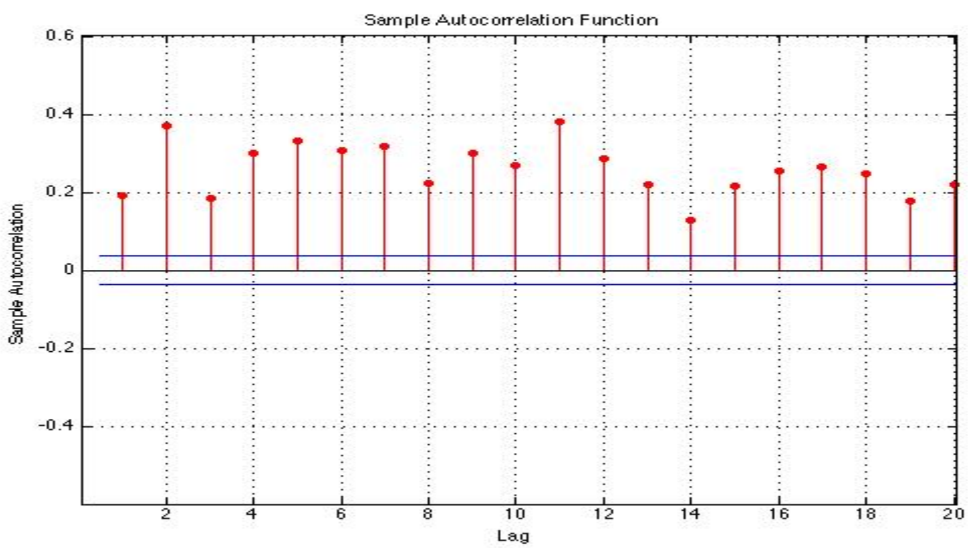


Figure 3.2:Autocorrelation of squared returns.

3.2. APPLYING MODELS

ESTIMATION OF PARAMETERS

Using MATLAB, the estimates of the parameters of GARCH, EGARCH, GJR models are:

GARCH: >> garchfit(spec,r)

Mean: ARMAX(0,0,0); Variance: GARCH(1,1)

Conditional Probability Distribution: Gaussian

Number of Model Parameters Estimated: 4

Parameter	Value	Standard Error	T Statistic
C	0.0012041	0.00054956	2.1909
K	2.6269e-06	1.4529e-06	1.8081
GARCH(1)	0.85767	0.031521	27.2094
ARCH(1)	0.12624	0.033811	3.7336

Log Likelihood Value: 863.123

EGARCH: >> garchfit(spec1,r)

Mean: ARMAX(0,0,0); Variance: EGARCH(1,1)

Conditional Probability Distribution: Gaussian

Number of Model Parameters Estimated: 5

Parameter	Value	Standard Error	T Statistic
C	0.000368	0.00059546	0.6180

K -0.81735 0.12281 -6.6554

GARCH(1) 0.91197 0.014528 62.7746

ARCH(1) 0.00032102 0.073434 0.0044

Leverage(1) -0.33169 0.058887 -5.6326

Log Likelihood Value: 875.409

GJR: >> garchfit(spec2,r)

Boundary constraints active: standard errors may be inaccurate.

Mean: ARMAX(0,0,0); Variance: GJR(1,1)

Conditional Probability Distribution: Gaussian

Number of Model Parameters Estimated: 5

		Standard	T
Parameter	Value	Error	Statistic
-----------	-----------	------------	-----------
C	0.00074968	0.00055576	1.3489
K	3.8638e-06	1.3381e-06	2.8875
GARCH(1)	0.85825	0.039838	21.5434
ARCH(1)	0	0.05589	0.0000
Leverage(1)	0.21119	0.071504	2.9535

Log Likelihood Value: 870.114

Based on the estimation the GARCH(1,1) conditional variance model that best fits the data:

$$\sigma_t^2 = 0.0000026269 + 0.12624\delta_{t-1}^2 + 0.\ 85767\sigma_{t-1}^2 \quad (3.\ 2)$$

It follows that the long-run average variance per implied by this model is 0.00016326 corresponding to a $\sqrt{0.00016326} = 0.01277$ or 1.27% per day volatility.

For the EGARCH(1,1) the conditional volatility model that best fits the data:

$$\ln(\sigma_t^2) = \text{-}0.81735 + 0.91197 \ln \sigma_{t-1}^2 + 0.00032102\left[\frac{|\delta_{t-1}|}{\sigma_{t-1}}\right] - 0.33169\left(\frac{\delta_{t-1}}{\sigma_{t-1}}\right) \quad (3.\ 3)$$

For the GJR(1,1) the conditional volatility model that best fits the data:

$$\sigma_t = 0.0000038638 + 0 * \delta_{t-1}^2 + 0.21119 D(\psi_{t-1})\delta_{t-1}^2 + 0.85825\sigma_{t-i} \quad (3.\ 4)$$

The forecasts from the models are plotted below

.

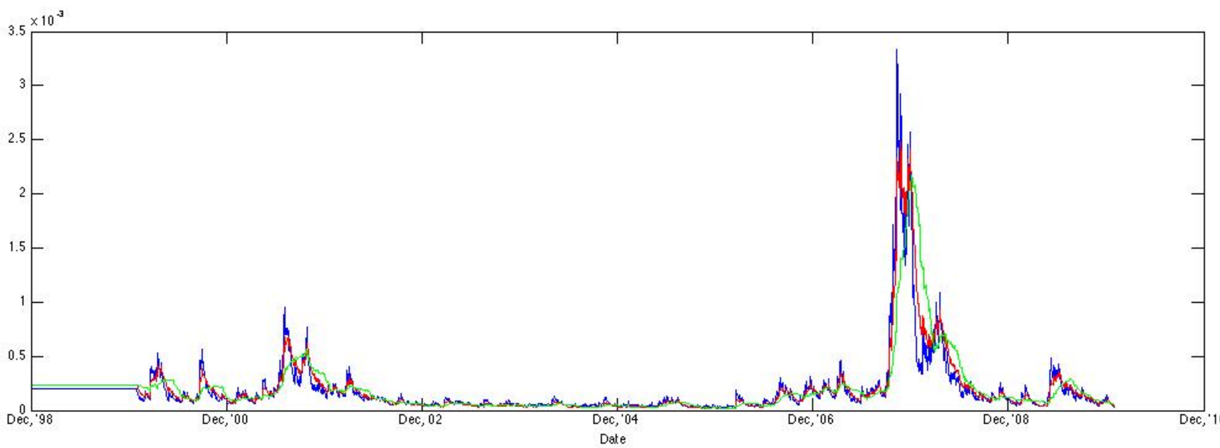


Figure 3.3: The above figure is a plot of volatility forecast of the rolling window forecast (green), risk metric (red), GARCH (blue). For Dec 1998 to Dec 2010.

CHAPTER 4

TESTING MODELS

There are different approaches to check which model better explains our stock exchange data. Each model can be applied with different parameters, i.e. u and v, but in limited way, each parameter cannot be greater than three for ARCH and simple GARCH models and two for EGARCH and GJR models. The evaluation of these forecasts is done using the Diebold Mariano test.

Evaluation is done using a sample of forecast history and resulting forecast errors:

$$e_t = y_t - \hat{y}_t \tag{4. 1}$$

$\hat{y}_t =$ forecast and $y_t =$ actual outcome Most statistic methods use the idea of minimizing the sum of squared forecast errors. The least squares method picks regression coefficients in $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$ to minimize the sum of squared residuals. This will give a zero mean of fitted residuals and also a zero correlation between the fitted residuals and the regressor. Evaluating forecast performance involves studying the following:

1) The forecast error has zero mean.
2) The forecast error is uncorrelated to the variables used (information) used in the construction of the forecast.
3) To compare the sum (or mean) of the squared forecasting error of different forecast approaches.
Zero mean implies bias, non zero correlation implies information not used efficiently (a forecast error should not be predictable)

An efficient h-step ahead forecast error must have a zero correlation with the forecast error (h-1) and the periods earlier.

The Diebold and Mariano test tests for forecasting superiority [5]: For instance to compare the MSE of two methods (a and b), we first define $g_t = (e_t^a)^2 - (e_t^b)^2$, for mean absolute errors, use $g_t = |e_t^a| - |e_t^b|$ where e_t^i is the forecast error of model i . Treating this as a GMM problem, we test for $Eg_t = 0$. By applying t-test on the same:

$$\frac{\overline{g}}{Std(\overline{g})} \sim N(0,1), \text{ where } \overline{\mathrm{d}}=\sum_{t=1}^{T}\frac{d_t}{T} \tag{4. 2}$$

Here the standard error is typically estimated using Newey-West approach.

$$Std(\overline{g}) = \left(\sum_{s=-\infty}^{\infty}\frac{Cov(g_t, g_{t-s})}{T}\right)^{\frac{1}{2}} \tag{4. 3}$$

Table 4.1: The Diebold Mariano test statistic for our forecasts is:

	Rolling Window	Risk Metric	GARCH
Rolling Window	0.071	2.31	1.97
Risk Metric	-2.34	0.065	0.52
GARCH	-1.96	-0.46	0.071

The test statistic is constructed by considering the three-implemented models pairwise and it can be observed that GARCH is a better estimate for the data set considered, Risk Metric ranks second and rolling window forecasts give the largest error. The error function considered here is the MSE function, which is the most widely used unbiased loss function.

The above tests are done on 3000 data points. The first 252 data points are used to estimate the parameters and the remaining are for the forecasts. The error in parameter estimation is considerably high as the ratio of data points used for estimation (r) far exceeds the data points used for the forecasts (t). Parameters are frequently updated to keep the ratio of (*r/t*) a constant. We notice here that the GARCH model out performs other forecast models even for crude data and with errors in parameter estimation.

CHAPTER 5

CAPITAL ASSET PRICING MODEL

One of the cornerstones of investment science is the theory of Portfolio Selection. The first foray into this field was in 1952 by Harry Markowitz, who made a historical contribution to financial mathematics with his classic article "Portfolio Selection". In the article, he incorporated the quantification of risk in the portfolio choice problem. He developed a framework where investors who like wealth and dislike risk would hold mean-variance efficient portfolios. Building on his work, Sharpe (1964) and Lintner (1965 b) almost simultaneously developed a model to price capital assets. The equation they derived has later been christened as the Capital asset Pricing Model.

In the following part of the thesis, the theoretical foundation for the capital asset pricing model will be formed and economic techniques will be applied to test the implications of CAPM.

This chapter will first lay the foundation by introducing mean-variance analysis. Then the efficient frontier is derived with and without a riskless asset. With this we can derive the CAPM. Finally the security market line is defined and a more intuitive interpretation of the CAPM is given [1].

With the following assumptions:

1) There exists a risk-free asset that investors may borrow or lend any amount of at the same rate.
2) Investors assign the same probability distributions to end of period values of a given asset. i.e. investors have homogenous expectations regarding means, variances and covariances.

It is also assumed that investors are wealth loving risk averters and that the market is purely competitive, frictionless (no transaction costs) and without taxes. The assumption that all assets are marketable and infinitely divisible is also assumed.

5.1. MEAN-VARIANCE ANALYSIS

Mean-Variance Analysis (MVA) is the process of selecting a portfolio that provides maximum expected return for a given variance and minimum variance for a given expected return [1]. It is based on the assumption that investors like wealth and dislike risk. i.e. an investor prefers portfolio 1 over portfolio 2 if the following mean variance criterion are fulfilled.

$$E[R_1] > E[R_2] \tag{5. 1}$$

and

$$\sigma^2(R_1) \leq \sigma^2(R_2) \tag{5. 2}$$

Consider a von Neumann-Morgenstern type utility function of an investor's end of period wealth:

$$u(W) = \sum_{n=0}^{\infty} \frac{u^{(n)}(a)}{n!}(W-a)^n, \tag{5. 3}$$

where *u(W)* is evaluated at a=E[W].

$$u(W) = u(E[W]) + u'(E[W])(W - E[W]) + \frac{1}{2}u''(E[W])(W - E[W])^2 + R_3, \tag{5. 4}$$

where

$$R_3 = \sum_{n=3}^{\infty} \frac{1}{n!} u^{(n)}(E[W])(W - E[W])^n. \tag{5. 5}$$

$$E\left[u(W)\right] = u(E[W]) + \frac{1}{2}u''(E[W])\sigma^2(W) + E\left[R_3\right], \tag{5. 6}$$

where

$$E[R_3] = \sum_{n=3}^{\infty} \frac{1}{n!} u^{(n)}(E[W]) m^n(W), \tag{5. 7}$$

Note that $m^n(W)$ is the *n*'th central moment of the end of period wealth. With the assumption that the investors prefer wealth and are variance aversive, it is assumed that investors have strict concave utility functions. With the assumptions of MVA, the straightforward implication of the above is that the utility function is quadratic. The problem with assuming a quadratic utility function is that the function involves a point of saturation ($u'(W) = 0$) where investors begin having preference against more wealth. The implication of the concave nature is that $E[u(W)]$ is a decreasing function. An alternate assumption is that wealth or equivalently returns are normally distributed. In normal distribution the moments of higher order than two can be expressed as functions of first and second moments. Thus in effect investors can only care about mean and variance under the normality assumption. The distribution assumptions will be dealt more thoroughly in chapter 6. From now on, it shall be assumed that investors' utility functions are concave and strictly increasing with no point of saturation. Moreover, we also assume that wealth is multivariate normally distributed.

5.2. MINIMUM VARIANCE FRONTIER

Consider the above analysis to include every feasible portfolio. The ones that satisfy MVC when compared to every other are said to be efficient. Efficient meaning that you cannot find a better portfolio in the mean-variance space. The set of feasible portfolios that meet the MVC constitute the efficient frontier [12].

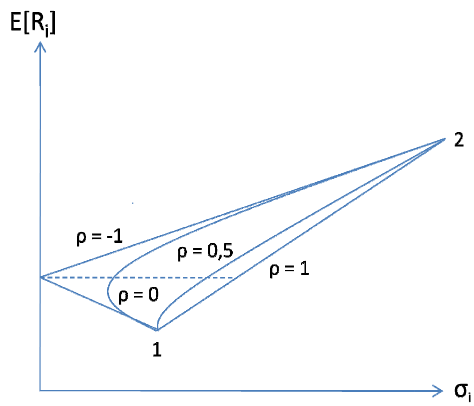
Consider combining two risky assets 1 and 2 in a portfolio, the expected return would be:

$$E_p = X_1 E\left[R_1\right] + X_2 E\left[R_2\right], \tag{5. 8}$$

where $X_1 + X_2 = 1$. The variance for the same can be found as:

$$\sigma_p = X_1^2 \sigma_1^2 + X_2^2 \sigma_2^2 + 2 X_1 X_2 \rho_{1,2} \sigma_1 \sigma_2, \tag{5. 9}$$

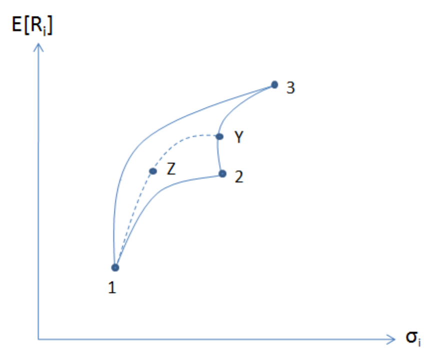
where $-1 \le \rho_{1,2} \le 1$ is the correlation coefficient or the covariance σ_{12}



Source: (Cuthbertson & Nitzsche, 2004)

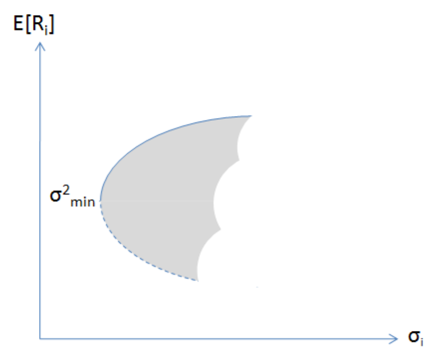
Figure 5.1:Risk-return frontiers in the two-asset case.

The above figure demonstrates the various possible portfolios for the two-asset case. To understand the concept of efficient frontier, we consider the case with *N* risky assets. The boundary of the feasible region is called the investment opportunity line or the minimum variance frontier and it contains the portfolios with the lowest variance for every level of expected return.



Source: Sharpe, 1970

Figure 5.2:Risk return frontier in the 3-asset case.



Source: Sharpe, 1970

Figure 5.3:Risk-return frontier with *N* assets.

The above figures depict the possible portfolios for three and multiple asset cases. The curves connecting points 1, 2 and 3 represent the available combinations when working with two assets at a time. Consider the combination of 2 and 3 marked by Y. When combined with portfolio 1 would give point Z. Continuing this, it is easy to see that the entire feasible region is covered. Figure 5.3 considers the case with multiple risky assets. It is the feasible region resulting from the infinite combinations available, and the blue line is the minimum variance frontier.

The portfolios satisfying the mean variance criteria can be interpreted as the portfolios farthest to the northwest. With MVC it is clearly observed that the solid line on the upper part of the minimum variance curve consists of the efficient portfolios. This line is the efficient frontier. Every other portfolio gives less return for the same levels of variance.

In order to construct the line portfolio weights of individual assets are needed. These can be found by minimizing the variance for every level of expected return. This involves finding the optimal distribution of wealth between individual stocks with the following minimization problem:

$$\min \sigma_p^2 = \sum_{i=1}^{N}(X_i^2\sigma_i^2) + \sum_{i=1}^{N}\sum_{\substack{j=1 \\ j\neq i}}^{N}(X_i X_j \sigma_{ij}), \tag{5. 10}$$

Consider the matrix notation for the portfolio variance:

$$\sigma_p^2 = \omega'\Omega\omega. \tag{5. 11}$$

Let Ω be the N×N variance-covariance matrix of the N assets in the portfolio. Let ω be the N×1-vector of portfolio weights on risky assets in P.

$$\min_{\omega} \frac{1}{2}\omega'\Omega\omega, \tag{5. 12}$$

subject to

$$\omega'\mu = \mu_p,$$

and

$$\omega'\iota = 1,$$

here μ is the N×1-vector of expected returns of the risky assets and ι is the N×1-vector of ones. Applying the Lagrange multiplier:

$$\min_{\omega,\delta^1,\delta^2} L = \frac{1}{2}\omega'\Omega\omega + \delta^1(\mu_p - \omega^1\mu) + \delta^2(1-\omega'\iota), \tag{5. 13}$$

where δ^1 and δ^2 are the Lagrange multiplier. The first order conditions can be formed as

$$\frac{\partial L}{\partial \omega} = \Omega\omega + \delta^1(-\mu) + \delta^2(-\iota) = 0,$$
$$\frac{\partial L}{\partial \delta^1} = \mu_p - \omega'\mu = 0, \tag{5. 14}$$

and

$$\frac{\partial L}{\partial \delta^2} = 1 - \omega'\iota = 0 \tag{5. 15}$$

Ω must be positive definite as $\omega'\Omega\omega$ is the variance of a risky portfolio. With L being convex, he first order conditions are both necessary and sufficient to find a global solution to the minimization problem. Solving for portfolio weights we have:

$$\begin{aligned}\omega_p &= -\delta^1\Omega^{-1}(-\mu) - \delta^2\Omega^{-1}(-\iota)\\ &= \delta^1(\Omega^{-1}\mu) + \delta^2(\Omega^{-1}\iota).\end{aligned} \tag{5. 16}$$

$$\begin{aligned}\mu'\omega_p &= \delta^1(\mu'\Omega^{-1}\mu) + \delta^2(\mu'\Omega^{-1}\iota).\\ \mu_p &= \delta^1(\mu'\Omega^{-1}\mu) + \delta^2(\mu'\Omega^{-1}\iota) \qquad \text{since } \mu'\omega_p = \omega'\mu\end{aligned} \tag{5. 17}$$

multiplying by ι'

$$1 = \delta^1\left(\iota'\Omega^{-1} - \mu\right) + \delta^2\left(\iota'\Omega^{-1}\iota\right). \tag{5. 18}$$

With two unknowns and two equations, we have, as shown in the appendix A.2

$$\begin{aligned}\delta^1 &= \frac{C\mu_p - A}{D}\\ \delta^2 &= \frac{B - A\mu_p}{D},\end{aligned} \tag{5. 19}$$

where $A = \iota'\Omega^{-1}\mu$, $B = \mu'\Omega^{-1}\mu$, $C = \iota'\Omega^{-1}\iota$ and $D = BC - A^2$. Substituting in (5.16):

$$\omega_p = g + h\mu_p,$$
$$g = \frac{1}{D}\left[B\left(\Omega^{-1}\iota\right) - A\left(\Omega^{-1}\mu\right)\right] \quad (5.20)$$
$$h = \frac{1}{D}\left[C\left(\Omega^{-1}\mu\right) - A\left(\Omega^{-1}\iota\right)\right].$$

The weights given by the vector ω_p result in a portfolio on the minimum variance frontier with expected return equal to μ_p. Using the above equations we can find every frontier portfolio. In fact every frontier portfolio can be constructed using two random frontier portfolios with different level of expected return. Let q be any portfolio on the minimum variance frontier. Since these two random frontier portfolios have different levels of expected return, there exists a combination such that

$$\mu_p = X\mu_1 + \left(1 - X\right)\mu_2, \quad (5.21)$$

Assigning these weights to the weight-vectors of portfolio 1 and 2,

$$\begin{aligned} X\omega_1 + \left(1 - X\right)\omega_2 &= X\left[g + h\mu_1\right] + \left(1 - X\right)\left[g + h\mu_2\right] \\ &= g + h\left[X\mu_1 + \left(1 - X\right)\mu_2\right] \\ &= g + h\mu_q \\ &= \omega_q \end{aligned} \quad (5.22)$$

5.3. CAPITAL MARKET LINE

Now we consider an asset with a riskless rate also known as a pure interest rate. Let us consider a combination of any risky portfolio and the risk-free asset. The expected return on the combination would be [1]:

$$E\left[R_p\right] = \left(1 - X\right)R_f + XE\left[R_T\right]$$
$$E\left[R_p\right] = R_f + X\left(E\left[R_T\right] - R_f\right),$$

Here X=1 implies that the investor holds all his wealth in the risky portfolio. Conversely, X<1 means that the investor lends some of his money at the risk-free rate. Finally, X>1 means a leveraged position where he borrows money at the risk-free rate and invests the proceeding from the loan plus his wealth in stocks and other risky assets. The variance of the above portfolio will be:

$$\sigma_p^2 = X^2\sigma_T^2$$
$$\sigma_p = X\sigma_T, \tag{5. 23}$$

We rewrite equation (5.1) as

$$E\left[R_p\right] = R_f + \theta\sigma_p, \tag{5. 24}$$

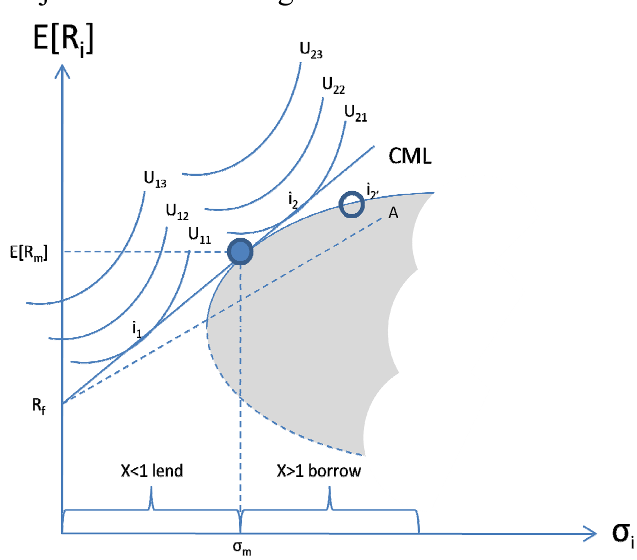
where

$$\theta = \frac{\left(E\left[R_T\right] - R_f\right)}{\sigma_T}.$$

The expression (5.2) says that the net expected rate of return on the total investment is linearly related to the risk of the net investment. This expression can be used for any portfolio of risky assets and is called the market opportunity line. In figure 5.4 this is the line R_fA for some random risky portfolio. As opposed to the minimum variance frontier in figure 5.3, the market opportunity line expresses a constant trade off between risk and expected return. Therefore, the price of risk reduction is independent of the risk level of a given position.

θ, the slope of this line is known as Sharpe ratio [16] or the price of risk for efficient portfolios. The intercept R_f is referred to as the price of time.

According to (5.24), no matter which risky portfolio an investor chooses to combine with the risk free asset, he can reach any desired net expected return of his total investment. He just has to leverage his investment in the risky asset enough by borrowing mon



Source: Lintner, 1965

Figure 5.4:The optimal portfolio choices in a mean-standard deviation space.

Yet, (5.23) reveals a serious drawback to leveraging a random portfolio: the risk of the total investment increases proportionately with the magnitude of the leverage X. Regardless of his utility function (subjective preferences), a risk averse investor will always choose the risky portfolio with the highest θ. As illustrated in figure 5.4, the only difference between investors is the proportion of their wealth placed in the tangency portfolio of risky assets.

Consider investor 1 with indifferences curves U_{11}, U_{12}, and U_{13}. Investor 1 is relatively risk averse, so he maximizes his utility where U_{11} is tangent to the CML in point i_1. As a consequence, investor 1 chooses to lend some of his money resulting in less risk and expected return than that of a pure risk asset portfolio. As another example, we have investor 2 with indifference curves U_{21}, U_{22}, and U_{23}. He is a lot less risk averse than investor 1. Therefore, he borrows money in order to invest more than his wealth in the fixed mix of risk assets. As long as investors like wealth and dislike risk, some combination of the tangency portfolio and the risk-free asset will always maximize their utility.

The implication of the separation theorem is important when we want to describe the tangency portfolio. Recall the assumptions set up in the beginning of the chapter. Two of them are of particular interest when we want to make conclusions about the tangency portfolio. The first is homogeneity of expectations, and the second is unlimited lending and borrowing at the risk-free rate. When everyone have the same distributional expectations, and everyone can use the same risk- free rate, every investor must face the same minimum variance frontier and opportunity market line as illustrated in figure 5.4. Thus, every investor faces the same tangency portfolio. Then, as the separation theorem suggests, the optimal portfolio is independent of individual preferences. When all investors assign the same weight in their risk portfolios to any given security, they must all hold the market portfolio (Sharpe, 1970).

We have made the assumption that investors' utility curves were strictly increasing. It can be seen in figure 5.4 that they consequently always choose efficient portfolios. As the market portfolio is a convex combination of every individual portfolio, the market portfolio must also be mean-variance efficient when involving a risk-free asset. Alternatively, one could argue that if the market portfolio is the tangency portfolio, it must lie on the CML and thus be efficient. Then, the slope in (5.24) can be written as:

$$\theta = \frac{\left(E\left[R_m\right] - R_f\right)}{\sigma_m}. \qquad (5.25)$$

5.4. DERIVATION OF THE CAPM

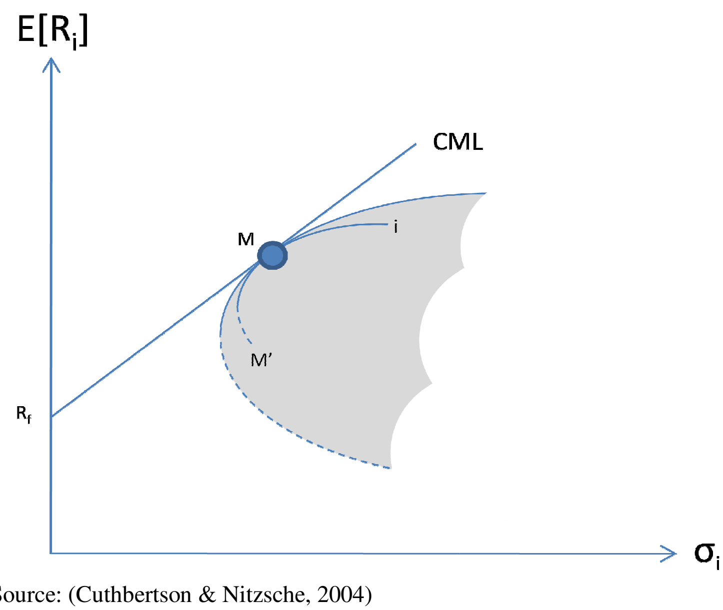
Sharpe, in 1964, and Lintner, in 1965, independently derived this model.

5.4.1. SHARPE'S DERIVATION

Consider a combination of the market portfolio and any other portfolio *i*. Let this be portfolio *Z* [16]. It follows that the expected return and the standard deviation on this combination will be

$$E[R_Z] = (1-X)R_m + XE[R_i]$$
$$\sigma_Z = \left[(1-X)^2\sigma_m^2 + X^2\sigma_i^2 + 2X(1-X)\rho_{m,i}\sigma_m\sigma_i\right]^{\frac{1}{2}} \quad (5.26)$$

At X = 1 only asset *i* is held while X = 0 implies that only the market portfolio is held. Points *i* and *m* in figure 5.5 illustrates the two cases. Note however, that at X = 0 some wealth is still placed in asset *i* as it is a part of the market. Thus, at X = 0 the weight placed on asset *i* corresponds to its values proportion of the total market wealth. *M'* indicates the case where *i* is not held at all making X < 0 (Sharpe, 1964). It should be marked that this combination curve does not intersect the minimum variance frontier. By definition, when there is no riskless asset all efficient portfolios lie on the curved minimum variance frontier, so the combinations between *m* and *i* cannot dominate them



Source: (Cuthbertson & Nitzsche, 2004)

Figure 5.5:Combining the market portfolio with some random risk asset

To find the slope of the curve we differentiate [R_f] with respect to σ_z

$$\frac{\frac{\partial[R_Z]}{\partial X}}{\frac{\partial\sigma_Z}{\partial X}}=\frac{\partial[R_Z]}{\partial\sigma_Z}.$$

$$\begin{aligned}\frac{\partial\sigma_Z}{\partial X}&=\frac{1}{2}\left[(1-X)^2\sigma_m^2+X^2\sigma_i^2+2X(1-X)\rho_{mi}\sigma_m\sigma_i\right]^{\frac{1}{2}}\\&\quad\times\left[-2(1-X)\sigma_m^2+2X\sigma_i^2+2(1-2X)\rho_{mi}\sigma_m\sigma_i\right]\\&=\frac{1}{\sigma_Z}\left[-(1-X)\sigma_m^2+X\sigma_i^2+(1-2X)\rho_{mi}\sigma_i\sigma_m\right].\end{aligned} \qquad (5.\ 27)$$

$$\frac{\partial[R_Z]}{\partial X}=-E[R_m]+E[R_i].$$

At point M we have X=0. At this point $\sigma_z = \sigma_m$ and $\sigma_{mi} = \rho_{mi}\sigma_m\sigma_i$. So:

$$\begin{aligned}\frac{\partial\sigma_Z}{\partial X}&=\frac{1}{\sigma_m}\left[-\sigma_m^2+\sigma_{mi}\right].\\\frac{\partial[R_Z]}{\partial\sigma_Z}&=\frac{\left(E[R_i]-E[R_m]\right)\sigma_m}{\sigma_{mi}-\sigma_m^2}=S_Z.\end{aligned} \qquad (5.\ 28)$$

In equilibrium the curve at point *m* must be a tangent to the CML. Equating slopes at X=0 we have:

$$\begin{aligned}\frac{\left(E[R_i]-E[R_m]\right)\sigma_m}{\sigma_{mi}-\sigma_m^2}&=S_Z=\theta=\frac{\left(E[R_m]-R_f\right)}{\sigma_m}\\E[R_i]&=\frac{\left(E[R_m]-R_f\right)\left(\sigma_{mi}-\sigma_m^2\right)}{\sigma_m^2}+E[R_m]\\&=\mathrm{R}_f+\beta_i\left(E[R_m]-R_f\right),\\\beta_i&=\frac{\sigma_{mi}}{\sigma_m^2}\end{aligned} \qquad (5.\ 29)$$

This is the traditional CAPM model derived by Sharpe in 1964. It says that the expected return of asset *i* equals the risk-free rate plus a reward for bearing risk.

CHAPTER 6

ECONOMETRIC TECHNIQUES FOR TESTING THE CAPM

For the purpose of statistical testing it is convenient to express the Sharpe-Lintner model in terms of excess returns. That is, in terms of the risk premium you receive in addition to the risk-free rate. Let Z_i and Z_m denote the expected excess return of asset *i* and the market portfolio respectively. Then, the traditional CAPM model can be written as:

$$E\left[Z_i\right]=\beta_{im}E\left[Z_m\right], \tag{6. 1}$$

where

$$\begin{aligned} Z_i &= R_i - R_f, \\ \beta_{im} &= \frac{Cov\left[Z_i, Z_m\right]}{Var\left[Z_m\right]}. \end{aligned} \tag{6. 2}$$

The empirical research that tests the CAPM primarily concentrates on three aspects of (6.1):

1) There is no intercept.
2) Beta can explain the variation in the cross-sections of expected excess returns.
3) The risk premium in the market is positive.

6.1. THE MARKET MODEL

In the traditional Sharpe-Lintner version, as well as the excess return model in (6.1), there is no time dimension as the CAPM is a one-period model [1]. However, when conducting empirical tests of the model, we use panel data in order to apply the econometric methods. Thus, there is a time dimension in the sample data. Therefore, asset returns are assumed to be independently and identically distributed over time (IID) and jointly multivariate normal. Then, it makes sense to make single estimates of the model parameters using data collected over time. Intuitively, when the properties of the data do not change over time, the CAPM can theoretically hold period by period. The assumptions that returns are IID and jointly normal are strong, and there is vast empirical evidence indicating that they are too strong. Therefore, we will also derive a test robust to this notion. This can serve as a check that the inferences are not biased by the distributional assumptions. In our analysis N will denote the number of assets or portfolios under consideration, and T will denote the number of periods under consideration. Consider the market model:

$$Z_t = \alpha + \beta Z_{mt} + \varepsilon_t \qquad (6.3)$$

where

$$E[\varepsilon_t] = 0,$$
$$E[\varepsilon_t \varepsilon_t'] = \sum,$$
$$Cov[Z_{mt}, \varepsilon_t] = 0.$$

Here Z_t denotes the $N \times 1$-vector of expected excess returns at time t, α is the $N \times 1$-vector of intercepts, β the $N \times 1$-vector of betas, the scalar Z_{mt} is the market risk premium at time t, and ε_t is the $N \times 1$-vector of error terms. Note that the notion of temporal IID returns does not imply anything about the cross-sections of data. In fact, there should be some correlation between returns for the CAPM, and in particular beta, to make sense. Thus, (6.3) implies that there can be heteroskedasticity (different variances in the diagonal of the matrix) and cross-correlation (other terms in the matrix above zero) in the *cross-sections* of data. Accordingly, the $N \times N$-matrix Σ contains the variances and covariances of the disturbances. From the above assumptions regarding the behavior of returns over time, we can say the following about the vector of error terms:

$$\varepsilon_t \sim N(0, \sum).$$

When comparing (6.3) with (6.1) it is obvious that if the CAPM holds, the intercepts in the vector α should all be zero. Testing whether the elements in α equals zero can be seen as a test of the exact linear relationship expressed in the CAPM. If this relationship does not exist, the market portfolio is not efficient and the CAPM is rejected. Also, if the market portfolio is not mean variance efficient, it cannot be the tangency portfolio. Hence, testing whether the intercepts are zero can also be interpreted as testing whether the market portfolio is the tangency portfolio. When the intercepts are all zero, every asset must lie on the SML just as the CAPM predicts (Roll, 1977). This gives the following null hypothesis:

$$H_0 : \alpha = 0$$

against

$$H_A : \alpha \neq 0$$

Notice that this is a joint hypothesis as we simultaneously test if every intercept α_i is zero. To test the H_0 against the alternative, we will use several test statistics. To implement these tests, we first have to estimate the parameters of the model *α,* *β* and *Σ*. Here we will apply Maximum Likelihood Estimation (MLE). As the single independent variable Z_{mt} is the same in all the equations expressed by the vectors of (6.3), the OLS-estimators of α and β are identical to the ML-estimators (Cuthbertson and Nitzsche, 2004).

For the assumption that the results are IID and jointly multivariate normal, we have the following multivariate pdf for the jointly normal excess returns in time *t* conditional on the observed conditional market risk premium (Campbell et al.1997).

$$f\left(Z_t \mid Z_{mt}\right) = \frac{1}{\left(2\pi\right)^{\frac{N}{2}} \left|\Sigma\right|^{\frac{1}{2}}} \times \exp\left[-\frac{1}{2}\left(Z_t - \left(\alpha + \beta Z_{mt}\right)\right)' \Sigma^{-1} \left(Z_t - \left(\alpha + \beta Z_{mt}\right)\right)\right] = \frac{1}{\left(2\pi\right)^{N/2} \left|\Sigma\right|^{1/2}} \times \exp\left[-\frac{1}{2}\left(Z_t - \alpha - \beta Z_{mt}\right)' \Sigma^{-1} \left(Z_t - \alpha - \beta Z_{mt}\right)\right]. \quad (6.4)$$

Results are assumed to be IID through time. Hence the conditional pdf of the excess return vectors $\mathbf{Z_1, Z_2, Z_3, Z_4, \ldots, Z_T}$ is:

$$f\left(Z_1, Z_2, \ldots, Z_T \mid Z_{m1}, Z_{m2}, \ldots, Z_{mT}\right) = \prod_{t=1}^{T} \frac{1}{\left(2\pi\right)^{\frac{N}{2}} \left|\Sigma\right|^{\frac{1}{2}}} \times \exp\left[-\frac{1}{2}\left(Z_t - \alpha - \beta Z_{mt}\right)' \Sigma^{-1} \left(Z_t - \alpha - \beta Z_{mt}\right)\right].$$

The log-likelihood function thus formed:

$$L\left(\alpha, \beta, \Sigma\right) = -\frac{NT}{2}\log(2\pi) - \frac{T}{2}\log\left|\Sigma\right| - \frac{1}{2}\sum_{t=1}^{T}\left[\left(Z_t - \alpha - \beta Z_{mt}\right)' \Sigma^{-1} \left(Z_t - \alpha - \beta Z_{mt}\right)\right]. \quad (6.5)$$

Since Σ^{-1} is symmetric we have $\partial\left[\left(x - s\right)' \Sigma^{-1} \left(x - s\right)\right] = -2\Sigma^{-1}\left(x - s\right)\partial s$, so,

$$\frac{\partial L}{\partial \alpha} = -\frac{1}{2}\sum_{t=1}^{T} -2\Sigma^{-1}\left(Z_t - \alpha - \beta Z_{mt}\right)$$

$$= \Sigma^{-1}\left[\sum_{t=1}^{T}\left(Z_t - \alpha - \beta Z_{mt}\right)\right].$$

Similarly for β, we have:

$$\frac{\partial L}{\partial \beta} = -\frac{1}{2}\sum_{t=1}^{T} -2\Sigma^{-1}\left(Z_t - \alpha - \beta Z_{mt}\right)Z_{mt}$$

$$= \Sigma^{-1}\left[\sum_{t=1}^{T}\left(Z_t - \alpha - \beta Z_{mt}\right)Z_{mt}\right].$$

And for Σ, we have:

$$\frac{\partial L}{\partial \Sigma} = -\frac{T}{2}\Sigma^{-1} - \frac{1}{2}\sum_{t=1}^{T} -\Sigma^{-1}\left(Z_t - \alpha - \beta Z_{mt}\right)\left(Z_t - \alpha - \beta Z_{mt}\right)'\Sigma^{-1}$$

$$= -\frac{T}{2}\Sigma^{-1} + \frac{1}{2}\Sigma^{-1}\left[\sum_{t=1}^{T}\left(Z_t - \alpha - \beta Z_{mt}\right)\left(Z_t - \alpha - \beta Z_{mt}\right)'\right]\Sigma^{-1}.$$

Since $\partial \ln\left(|X|\right) = \left(X'\right)^{-1}\partial X$ and $\partial \ln\left(a'X^{-1}b\right) = \left[X'\right]^{-1}ab'\left[X'\right]^{-1}\partial X.$

Setting the above partial differentials to zero, and solving for α, we have:

$$\hat{\alpha} = \hat{\mu} - \hat{\beta}\hat{\mu}_m, \quad (6.\ 6)$$

where

$$\hat{\mu} = \text{mean excess return} = \frac{1}{T}\sum_{t=1}^{T}\left(Z_t\right)$$

and

$$\hat{\mu}_m = \frac{1}{T}\sum_{t=1}^{T}\left(Z_{mt}\right)$$

Similarly

$$\hat{\beta} = \frac{\sum_{t=1}^{T}\left(Z_t - \hat{\mu}\right)\left(Z_{mt} - \hat{\mu}_m\right)}{\sum_{t=1}^{T}\left(Z_{mt} - \hat{\mu}_m\right)^2},$$

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\left(Z_t - \hat{\alpha} - \hat{\beta}Z_{mt}\right)\left(Z_t - \hat{\alpha} - \hat{\beta}Z_{mt}\right)'. \tag{6. 7}$$

The OLS estimators also give the same results as above.

In order to construct a test statistic for testing our hypothesis, we need to know the distribution of the estimators. We have assumed that excess returns are jointly normal and temporally IID. The estimators' distributions result from this supposition. Conditional on the market risk premium, $\hat{\alpha}$ and $\hat{\beta}$ follow a normal distribution (Campbell et al., 1997). Their expected values equal the true parameter values, so they are unbiased. This follows from the specification of the market model, and in particular the exogeneity of the explanatory variable combined with the IID assumption.

With $Cov\left(\hat{\mu}, \hat{\beta}\hat{\mu}_m\right) = 0$ we have:

$$Var\left(\hat{\alpha}\right) = Var\left(\hat{\mu} - \hat{\beta}\hat{\mu}_m\right) = Var\left(\hat{\mu}\right) - Var\left(\hat{\beta}\hat{\mu}_m\right)$$

$$Var\left(\hat{\alpha}\right) = Var\left(\frac{1}{T}\sum_{t=1}^{T}\left(Z_t\right)\right) + Var\left(\frac{\sum_{t=1}^{T}\left(Z_t - \hat{\mu}\right)\left(Z_{mt} - \hat{\mu}_m\right)}{\sum_{t=1}^{T}\left(Z_{mt} - \hat{\mu}_m\right)^2}\hat{\mu}_m\right). \tag{6. 8}$$

Consider each term independently:

$$\begin{aligned} Var\left[\frac{1}{T}\sum_{t=1}^{T}\left(Z_t\right)\right] &= \frac{1}{T^2}\sum_{t=1}^{T}Var\left(Z_t\right) \\ &= \frac{1}{T^2}\sum_{t=1}^{T}\Sigma \\ &= \frac{1}{T}\Sigma. \end{aligned}$$

$$\sum_{t=1}^{T}(Z_t-\hat{\mu})(Z_{mt}-\hat{\mu}_m)$$

$$=\sum_{t=1}^{T}\left(Z_t-\frac{1}{T}\sum_{t=1}^{T}(Z_t)\right)\left(Z_{mt}-\frac{1}{T}\sum_{t=1}^{T}(Z_{mt})\right)$$

$$=\sum_{t=1}^{T}Z_tZ_{mt}-\sum_{t=1}^{T}Z_t\left[\frac{1}{T}\sum_{t=1}^{T}(Z_{mt})\right]-\sum_{t=1}^{T}\left[\frac{1}{T}\sum_{t=1}^{T}(Z_t)\right]Z_{mt}+\sum_{t=1}^{T}\left[\frac{1}{T}\sum_{t=1}^{T}(Z_t)\frac{1}{T}\sum_{t=1}^{T}(Z_{mt})\right]$$

$$=\sum_{t=1}^{T}Z_tZ_{mt}-\frac{1}{T}\sum_{t=1}^{T}\left[\sum_{t=1}^{T}(Z_t)\right]Z_{mt}$$

$$=\sum_{t=1}^{T}Z_tZ_{mt}-\frac{1}{T}\left[\sum_{t=1}^{T}(Z_t)\right]\hat{\mu}_m$$

$$=\sum_{t=1}^{T}Z_t(Z_{mt}-\hat{\mu}_m).$$

Substituting back in the second term we have:

$$Var\left(\frac{\sum_{t=1}^{T}(Z_t-\hat{\mu})(Z_{mt}-\hat{\mu}_m)}{\sum_{t=1}^{T}(Z_{mt}-\hat{\mu}_m)^2}\hat{\mu}_m\right)$$

$$=Var\left(\frac{\sum_{t=1}^{T}(Z_t)(Z_{mt}-\hat{\mu}_m)}{\sum_{t=1}^{T}(Z_{mt}-\hat{\mu}_m)^2}\hat{\mu}_m\right)$$

$$=\sum_{t=1}^{T}Var\left([Z_t]\frac{(Z_{mt}-\hat{\mu}_m)}{\sum_{t=1}^{T}(Z_{mt}-\hat{\mu}_m)^2}\hat{\mu}_m\right) \tag{6. 9}$$

$$=\sum_{t=1}^{T}Var[Z_t]\left(\frac{(Z_{mt}-\hat{\mu}_m)}{\sum_{t=1}^{T}(Z_{mt}-\hat{\mu}_m)^2}\hat{\mu}_m\right)^2$$

$$=\sum\left(\frac{\sum_{t=1}^{T}(Z_{mt}-\hat{\mu}_m)}{\sum_{t=1}^{T}(Z_{mt}-\hat{\mu}_m)^2}\hat{\mu}_m\right)^2=\frac{1}{T}\sum\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2},$$

where

$$\hat{\sigma}_m^2 = \frac{1}{T}\sum_{t=1}^{T}\left(Z_{mt} - \hat{\mu}_m\right)^2.$$

hence:

$$\begin{aligned} Var(\hat{\alpha}) &= \frac{1}{T}\Sigma + \frac{1}{T}\Sigma\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2} \\ &= \frac{1}{T}\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]\Sigma. \end{aligned} \tag{6. 10}$$

and

$$Var(\hat{\beta}) = \frac{1}{T}\left[\frac{1}{\hat{\sigma}_m^2}\right]\Sigma. \tag{6. 11}$$

Using the following theorem (Uhlig, 1994):

Consider a random variable $Y_t \sim N(0,\Sigma)$ for t=1,2,…,T where Σ is an $N \times N$ positive definite matrix. Then,

$$X = \sum_{t=1}^{T} Y_t Y_t' \sim W_N(T,\Sigma).$$

This means that X follows a Wishart distribution with T degrees of freedom and variance-covariance matrix Σ. The Wishart distribution is the chi-square distribution generalized to a multivariate case.

Previously, we found out that $\varepsilon_t \sim N(0,\Sigma)$. Let $Y_t = \varepsilon_t$ and recall from (6.3) that $E[\varepsilon_t\, \varepsilon'_t] = \Sigma$. Then, it can be shown that $T\hat{\Sigma}$ follows a Wishart distribution with T-2 degrees of freedom. And so to summarize, we have:

$$\begin{aligned} &\hat{\alpha} \sim N\left(\alpha, \frac{1}{T}\left[1 + \frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]\Sigma\right) \\ &\hat{\beta} \sim N\left(\beta, \frac{1}{T}\left[\frac{1}{\hat{\sigma}_m^2}\right]\Sigma\right) \\ &T\hat{\Sigma} \sim W_N(T-2,\Sigma). \end{aligned} \tag{6. 12}$$

6.2. THE STANDARD TESTS

Now we have built the necessary foundation to construct test statistics for testing the null hypothesis. We wish to test simultaneously whether all the intercepts are zero [18]. Thus, there are in fact multiple hypotheses. Here, the hypotheses can be seen as exclusion restrictions, as they practically exclude the intercepts from the market model. There are different ways of testing multiple restrictions. One possibility is the Wald test

In a Wald test, we evaluate the difference between the estimate of the market model parameter and its value under the null hypothesis. Formally, the squared difference is benchmarked against the variance of the parameter. This result is then compared to the test statistic's distribution under the null hypothesis to see whether the above difference is significant. Using result (6.3) we formed the null hypothesis α = 0. Thus, the Wald test statistic is

$$
\begin{aligned}
J_0 &= \hat{\alpha}\left[Var(\hat{\alpha})\right]^{-1}\hat{\alpha}' \\
J_0 &= \hat{\alpha}\left[\frac{1}{T}\left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]\Sigma\right]^{-1}\hat{\alpha}' \\
J_0 &= T\left[1+\frac{\hat{\mu}_m^2}{\hat{\sigma}_m^2}\right]^{-1}\hat{\alpha}\Sigma^{-1}\hat{\alpha}'.
\end{aligned}
\qquad (6.\ 13)
$$

Under the null hypothesis *J0* has a chi-square distribution with *N* degrees of freedom. Here, *N* corresponds to the number of restrictions imposed, as there are *N* different intercepts.

Using the above results, we can also construct a test statistic in a *finite-sample* setting. To do this, we employ the following theorem:

Let **x** be a *m* × 1-vector and **A** be a *m* × *m* -matrix. Furthermore, let **x** and **A** be independent, **x**~*N*(**0**, Ω), and **A** ~W_N (*n*, Ω), where *m* ≤ *n*. Then

$$\frac{(n-m+1)}{m}\mathrm{x}'\mathrm{A}^{-1}\mathrm{x}\sim F_{m,n-m+1}. \qquad (6.\ 14)$$

Using the above, the central F-test statistic is:

$$\frac{(T-N+1)}{TN}J_0\sim F_{N,T-N-1}. \qquad (6.\ 15)$$

CHAPTER 7

DATA

For the empherical tests, we use S&P 500 index data from Dec, 1926 to Dec, 2011

We use monthly realizations of the excess return to the aggregate U.S. stock market and monthly returns to portfolios of U.S. stocks. Following the restrictions from sections on true size, power and portfolio construction the eligible stocks of the S&P are assigned to ten value-weighted portfolios based on their size. Specifically, at the end of each quarter, eligible companies in the U.S. exchange are sorted on market capitalization and then are divided into 10 deciles of equal populations. The companies with the largest capitalization in each decile serve as the breakpoints when assigning all the sampled companies. Portfolio 1 contains the largest companies, portfolio 2 the next largest, and so on. The returns recorded are total returns, so dividends are included as demanded by the CAPM. As indicated above, the returns are computed from a value-weighted portfolio of the securities in each decile.

We also choose the theoretical consistent value-weighted proxy of the market portfolio. Specifically, we use the CRSP value weighted basket of American stocks with dividends. To check the robustness of the inference to the market portfolio proxy used, we also employ the equivalent CRSP equal weighted portfolio. For the proxy of the risk free rate, we use the 30-day US Treasury Bill as recommended above. All data is collected from the CRSP tapes of Wharton Research Data Services (WRDS, 2011). We use 85 years of monthly data from December 1926 to December 2011. This way we can also test the model's performance on newer data, as the main part of the empirical literature deals with samples from before the 1990s.

7.1. ESTIMATION AND ANALYSIS

Using the aggregate U.S. stock market as the proxy for total wealth, we construct the individual statistics of the hypothesis that $\alpha_i = 0$ for the portfolios i =1,2,...,10. We do this over the full sample July 1926 through December 2011, then over the sample July 1963 through December 2011. Note that the five percent critical value for this test is 3.852 for 1024 degrees of freedom (the length of the full sample minus two) and 3.858 for 580 degrees of freedom (the length of the shorter sample minus two). The following are the test statistics:

CAPM results for 1963:12 through 2011:12

Portfolio	Full sample (26-11)	Shorter sample (63-11)
Port 1:	1.454	1.755
Port 2:	0.493	0.897
Port 3:	1.345	2.810
Port 4:	1.924	2.575
Port 5:	1.724	4.830
Port 6:	3.333	3.973
Port 7:	3.332	5.699
Port 8:	2.635	4.020
Port 9:	1.539	2.987
Port 10:	0.034	0.415

Joint test

F	pval
2.465	0.007

For the full sample, the monthly Sharpe ratio for the market portfolio is 0.113. The maximum monthly Sharpe ratio for the residual strategies is 0.109. We sum the squares, and then take the square root, to determine the Sharpe ratio of the tangency portfolio. It is 0.157. The corresponding numbers for the shorter sample are 0.096, 0.209, and 0.230.

For both sample periods, the MV-efficient portfolios have a decidedly odd structure. They consist of extremely large short positions in the aggregate market, combined with extremely long positions on the portfolio of largest-cap stocks. The weights for individual portfolios for the tangency portfolio are evaluated to construct the tangency portfolio.

Sharpe ratios

Market:	0.096	0.113
Max, 10 resid assets:	0.209	0.109
Max, all eleven assets:	0.230	0.157

Loading of tangency portfolio on market and 10 ME ports

Portfolio	Full sample	Shorter sample
Market:	−15.029	-28.016
Port 1:	0.807	2.570
Port 2:	−1.811	-3.110
Port 3:	0.216	1.461
Port 4:	0.588	-1.290
Port 5:	0.396	4.417
Port 6:	1.937	1.002
Port 7:	1.958	3.180
Port 8:	1.741	1.663
Port 9:	1.179	3.483
Port 10:	9.017	15.642

For the full sample, the F statistic is 1.195 with a p-value of 0.29, thus the model is not close to rejection. For the more recent sample, the F statistic is 2.465, which has a p-value of 0.007. Thus the model is rejected at the 1% level. The efficient frontier, market portfolio and the tangency portfolio are plotted below.

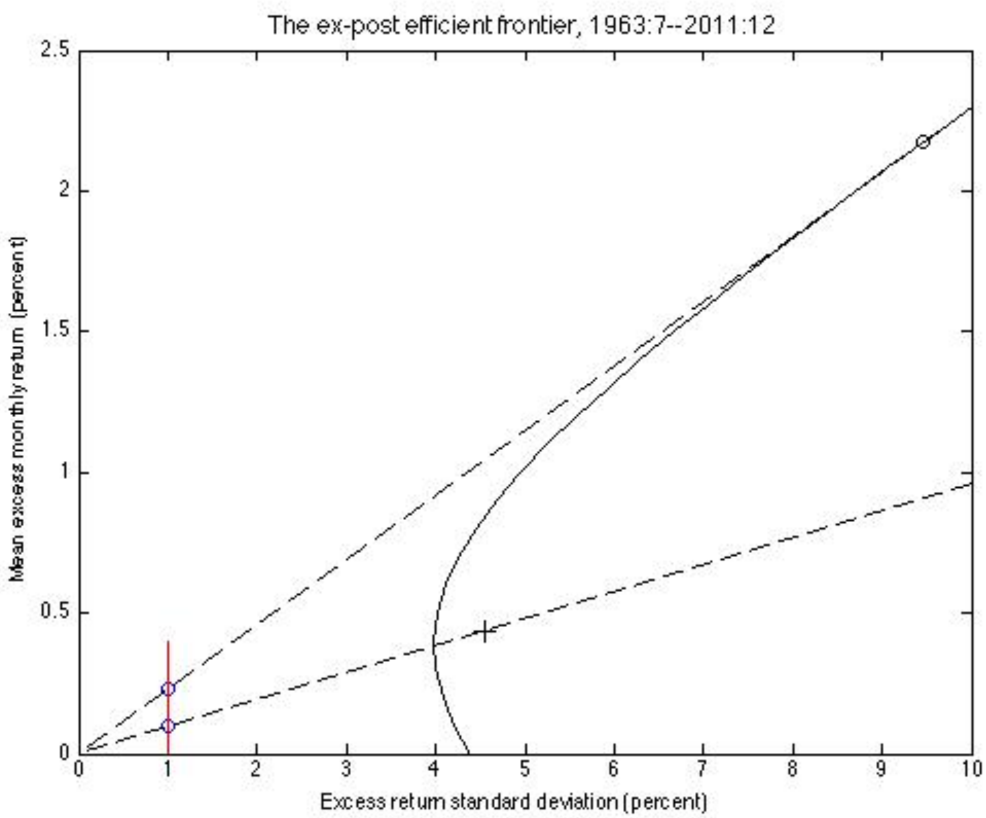


Figure 7.1:Efficient portfolios full sample.

CHAPTER 8

CONCLUSION AND FUTURE RESEARCH

The Endeavour to examine if GARCH-type models were the better models in describing return series for VaR was done statistically and empirically. The other companies including the S&P-500 index contained correlation in its returns or squared returns, which meant that modeling with GARCH was appropriate. After testing the dataset on the normality assumptions using variance ratio tests, the models were set up and run; the parameters were estimated for each of the model with their conditional volatility as the conditional volatility is the main ingredient for forecasting VaR and its depends on Conditional variance. ARCH test strongly rejects the null hypothesis that there is no ARCH/GARCH effect in given return of S&P-500. Then we check the quality of our estimated parameters and volatility. Finally, log returns have an ARCH effect at significance level of 5% and given time series has no random sequence of Gaussian disturbance. A comparative statistic for each of the forecast has been established to confirm the GARCH model's accuracy.

In CAPM, it is clear that the linear relationship in the model only holds when the market portfolio is efficient. Thus, a central, testable implication of the CAPM is the efficiency of the market portfolio. In empirical applications, this is equivalent to testing the intercept against zero in the excess return version of the traditional model. Testing N assets, the null hypothesis is that all the estimated N intercepts are jointly zero. Several statistics allow for joint tests. A number of them are only asymptotically distributed under the null-hypothesis. One such test statistic is a Wald type test. Another is the likelihood ratio test. In finite samples, these asymptotic approximations can behave rather perversely. They tend to reject too often. Fortunately, the Wald type test can be transformed into an exact F test, and the likelihood ratio test can be corrected to perform much better in finite samples. We have tested the CAPM on a sample of American stocks assigned to 10 size-sorted portfolios. Monthly, total returns were collected from January 1926 to December 2011. For the full sample, the model is not close to rejection. But for the more recent sample, the model is rejected at 1% level.

Further study with no assumption on third and fourth order moments can give better accuracies. Using multi-variate models can help better understand deviations from the regular assumptions. We also highlight the necessity of the use of higher order test statistics on QLIKE, MSE and other loss functions to better explain the deviations.

REFERENCES

1. Söderlind, Paul. "Lecture Notes in Empirical Finance (PhD)." (2013).
2. Lo, Andrew W., and Archie Craig MacKinlay. "Stock market prices do not follow random walks: Evidence from a simple specification test." *Review of financial studies* 1.1 (1988): 41-66.
3. Awartani, Basel, and Valentina Corradi. "Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries." *International Journal of Forecasting* 21.1 (2005): 167-183.
4. Hasanhodzic, Jasmina, Andrew Lo, and Emanuele Viola. "Is it real, or is it randomized?: A financial turing test." *A Financial Turing Test (February 23, 2010)* (2010).
5. Costantini, Mauro, and Robert M. Kunst. *On the usefulness of the Diebold-Mariano test in the selection of prediction models: some Monte Carlo evidence*. No. 276. Economics Series, Institute for Advanced Studies, 2011.
6. Conrad, Christian, Menelaos Karanasos, and Ning Zeng. "Multivariate fractionally integrated APARCH modeling of stock market volatility: A multi-country study." *Journal of Empirical Finance* 18.1 (2011): 147-159.
7. Bollerslev, Tim. "Generalized autoregressive conditional heteroskedasticity."*Journal of econometrics* 31.3 (1986): 307-327.
8. Lopez, Jose A. "Evaluating the predictive accuracy of volatility models."*Journal of Forecasting* 20.2 (2001): 87-109.
9. Baker, Malcolm, and Jeffrey Wurgler. "Investor sentiment in the stock market." (2007).
10. Leung, Mark T., Hazem Daouk, and An-Sing Chen. "Forecasting stock indices: a comparison of classification and level estimation models." *International Journal of Forecasting* 16.2 (2000): 173-190.
11. Engle, R.F. (2004), "Risk and Volatility: Econometric Models and Financial Practice," *American Economic Review*, Vol.94, p.405-420.
12. Patton, Andrew J., and Allan Timmermann. "Monotonicity in asset returns: New tests with applications to the term structure, the CAPM, and portfolio sorts."*Journal of Financial Economics* 98.3 (2010): 605-625.
13. Gibbons, Michael R., Stephen A. Ross, and Jay Shanken. "A test of the efficiency of a given portfolio." *Econometrica: Journal of the Econometric Society* (1989): 1121-1152.
14. Cuthbertson, Keith, and Dirk Nitzsche. *Quantitative financial economics: stocks, bonds and foreign exchange*. Wiley, 2005.

15. Bollerslev, Tim. "A conditionally heteroskedastic time series model for speculative prices and rates of return." *The review of economics and statistics*(1987): 542-547.
16. Sharpe, William F. "CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK*." *The journal of finance* 19.3 (1964): 425-442.
17. Sharpe, William F. "The sharpe ratio." *Journal of portfolio management* 21 (1994): 49-49.
18. Brooks, Chris. *Introductory econometrics for finance*. Cambridge university press, 2008.