Day-of-the-week effects in stock market data

C H E U N G M E I T I N G
S U X U N

Master of Science Thesis
Stockholm, Sweden 2012

Day-of-the-week effects in stock market data

C H E U N G   M E I   T I N G
S U   X U N

Master's Thesis in Mathematical Statistics (30 ECTS credits)
Master Programme in Mathematics 120 credits
Royal Institute of Technology year 2012
Supervisor at KTH was Filip Lindskog
Examiner was Boualem Djehiche

TRITA-MAT-E 2012:08
ISRN-KTH/MAT/E--12/08--SE

Royal Institute of Technology
School of Engineering Sciences

KTH SCI
SE-100 44 Stockholm, Sweden

URL: www.kth.se/sci

Abstract

The purpose of this thesis is to investigate day-of-the-week effects for stock index returns. The investigations include analysis of means and variances as well as return-distribution properties such as skewness and tail behavior. Moreover, the existences of conditional day-of-the-week effects, depending on the outcome of returns from the previous week, are analyzed. Particular emphasis is put on determining useful testing procedures for differences in variance in return data from different weekdays. Two time series models, AR and GARCH(1,1), are used to find out if any weekday's mean return is different from other days. The investigations are repeated for two-day returns and for returns of diversified portfolios made up of several stock index returns.

Keywords: Day-of-the-week effect, Levene's test, Brown-Forsythe test, *GARCH*, *AR*, variance test, mean test

Acknowledgements

We would like to thank our supervisor Filip Lindskog at KTH Royal Institute of Technology for his guidance and helpful comments, Anders Modin at Industrivarden for the mentoring and the opportunity to work on this project. We would also like to thank our friends for their support during the course of this work.

Stockholm, October 2012

Cheung Mei Ting
Su Xun

Contents

Chapter 1

Introduction

The standard assumption in financial theory is that the distribution of stock returns are identical for all weekdays, however, stock market shuts down during Saturday and Sunday. This break provides the possibility of day-of-the-week effect, i.e. some weekday is different from other weekdays in the stock market. If the day-of-the-week effect exists, some investors can take an advantage from it to make arbitrage.

In the last three decades of financial researches, lots of work has been carried out on the study of day-of-the-week effects. The work of Cross [7], French [9], Harris [11] shows that returns of equity assets appear to be lower on Monday comparing to other days of the week in the 1970s. And based on the study of Lakoniskhok and Maberly [15], the average Monday return of stocks is negative in the US and some emerging stock markets. Similarly, Gibbons and Hess [10] document the day-of-the-week effects in treasury bill returns. Besides Yamori and Kurihara [19] find that the day-of-the-week effect exists in the 1980s for some currencies, but disappears for almost all currencies in the 1990s in the New York foreign exchange market.

This thesis mainly study the day-of-the-week effects on the one-day return of the Swedish index OMX which have not been studied by other researchers before. The other countries' indexes will also be examined to see if there is any special finding from them. Then it will extend to two days holding period and one-day return of a portfolio. The investigation will focus the day-of-the-week effects on variance and mean return. The traditional way for testing the equality of variance is F-test which bases on the assumption that the data are normally distributed. However, the daily stock returns are in fact non-normal. Levene [16] proposes another statistic used to assess the equality of variances in different samples and does not require normality of the underlying data. The Levene's statistic is then modified by Brown and Forsythe [4] and their alternative formulation shows more robust depending on the distribution of the underlying data. In this thesis, a comparison of different tests for variance will be carried out and applied to the return data.

To test the equivalent of mean, two common time series are used: AR and $GARCH$, see Al-Longhani and Chappell. [2]. How to decide a relatively small yet accurate order is vital when applying the time series models and it is analyzed and compared in details. When autocorrelation is found out in the data, AR model always come first. Then, if $ARCH$ effect is also found out in the data, an improved time series model $GARCH(1,1)$ is needed to get a more accurate result. If the $ARCH$ effect is too large, the results obtained from two models may be a big difference. This thesis will apply both AR and $GARCH$ models, and comparing the results between them. The relationship of the day-of-the-week effects with the previous week return is also studied, e.g whether the Monday effect shows up when the previous market has risen. Cross [7], Keim-Stambaugh [14], and Jaffe-Westerfield [13] point out that Monday return is positively correlated with the previous week return. This study presents the correlations between the daily week return and the weekly return. Unfortunately, as with the weekend effect, the twist is unable to be explained.

Chapter 2 describes the data, how it is modified and also theories about parametric models. Chapter 3 and 4 present the used methods and compare them with the traditional t-test and F-test. Chapter 5 states the theory of two time series models-AR and $GARCH(1,1)$ and also compares two common methods, AIC and $PACF$, used for order determination in time series model. The data analysis part will be presented in Chapter 6. Parametric models fitting, equality of variance testing, and also times series models for testing the mean difference can be found in this chapter. Then the conditional daily effect is examined in Chapter 7, i.e, if any day returns is related to the previous week return. All the methods mentioned can be used in many other situations, like the two more complicated examples presented in Chapter 8 and 9: the two-day effect and one-day portfolio. Chapter 10 is the conclusion of the findings. More studies using other indexes can be found in appendix.

Chapter 2

Financial data

2.1 Data resource

Six different indexes are examined for the existing of the day-of-the-week effect or conditional week effect. All the data are provided by Bloomberg. The Swedish market is primarily studied, and the sample is the OMX Index (OMX stockholm 30 Index) from Jan 1st 2001 to Mar 13th 2012.

Five other indexes are also analyzed, covering markets in Europe, Japan, United States and HongKong. They are SAX Index (OMX Stockholm All-Share Index), SX5E Index (EURO STOXX 50 (Price) Index), SPX Index (Standard and Poor's 500 Index), NKY Index (Nikkei-225 Stock Average) and HSI Index (Hang Seng Index), which are all daily closing prices from Jan 1st 2001 to Mar 13th 2012. The results are presented in the appendix. All the data are provided by Bloomberg and are dividend adjusted.

2.2 Return data

In finance, return is the ratio of money gained or lost (whether realized or unrealized) on an investment relative to the amount of money invested. Most of the financial studies focus more on return rather than prices of the assets. Campbell et al [6] gives out two reasons for that. First, for the average investor, financial markets may be considered close to perfectly competitive, so that the size of the investment does not affect price change. Therefore, since the investment "technology" is constant-returns-to-scale, i.e. output increases by that same proportional change in all inputs, the return is a complete and scale-free summary of the investment opportunity. Second, returns have more attractive statistical properties than prices which make them easier to be handled.

The rate of return can be calculated over a single period, or expressed as an average over multiple periods. Denoting the asset value on the t-th day by P_t, the single period arithmetic return, r_t, for day t is given by

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}. \tag{2.1}$$

When the frequency of measured asset prices is high, Eq. 2.1 can be approximated by

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right). \tag{2.2}$$

In the limit of infinite asset price frequency and thus time between P_t and P_{t-1} tending to 0, this approximation gets exact and is referred to as continuous compounding. The main advantage of continuously compounded returns is that they are symmetric, while the arithmetic returns are not: positive and negative percent arithmetic returns are not equal.

Eq. 2.2 assumes that the stock pays no dividends and does not take into account splits. However, in financial databases, reported price series are usually adjusted for splits and all cash dividends are reinvested into the asset. This procedure facilitates the calculation of return series as the adjusted price series already contain all distributions and Eq. 2.1 or 2.2 can be directly applied to produce accurate return series.

2.3 Volatility

The precise definition of the *volatility* of an asset is an annualized measure of dispersion in the stochastic process that is used to model the log returns. The most common measure of dispersion of a distribution is the standard deviation σ, which is a sufficient risk metric if returns are normally distributed. When dispersion increases as the holding period of returns increases, the standard deviation should be transformed into annualized terms to make comparisons possible.

Under the standard assumption that successive returns are independent from each other and are identically distributed, i.e. *i.i.d.* returns, risk and return can be converted into annualized terms by multiplication by an annualizing factor A. Formally:

$$\begin{aligned}
\text{Annualized mean} &= A\mu\ , \\
\text{Annualized variance} &= A\sigma^2\ , \\
\text{Annualized standard deviation} &= \sqrt{A}\sigma.
\end{aligned}$$

Since 252 risk days per year is commonly assumed, annualizing factor $A = 252$ is used for daily log returns.

2.4 Parametric model

Return data are random variables with unknown distributions. The standard assumption is that returns are normally distributed.

A parametric family of distribution functions is a set $\{F_\theta : \theta \in \Theta\}$ of distribution functions, where θ is the parameter and $\Theta \subset \mathbb{R}^k$ is the parameter space. For the family of Normal distribution functions, it has parameters (μ, σ^2), parameter space $\mathbb{R} \times (0, \infty) \subset \mathbb{R}^2$, and distribution function $\Phi((x - \mu)/\sigma)$.

However, when the return data are fitted with the Normal distributions, one usually observes that the return data have heavier tails. Therefore, it is commonly suggested to fit it with the Student's t-distribution. The Student's t location-scale family has parameters (μ, σ, ν), parameter space $(0, \infty) \times \mathbb{R} \times (0, \infty) \subset \mathbb{R}^3$, and distribution function $t_\nu((x-\mu)/\sigma)$ where $t_\nu(x)$ is the distribution function of a standard Student's t distributed random variables with ν degrees of freedom, see Hult et al [12].

After a parametric family is chosen, parameters can be estimated by maximum likelihood estimation (MLE). Consider the observations (historical return data) $z_1, ..., z_n$ of independent and identically distributed random variables $Z_1, ..., Z_n$ with the density function f_{θ_0}, where the parameter θ_0 is unknown. In MLE the unknown parameter θ_0 is estimated as the parameter value θ maximizing the probability of the observed data. The maximum likelihood estimator $\hat{\theta}$:

$$\hat{\theta} = argmax_\theta \prod_{k=1}^{n} f_\theta(z_k).$$

Since logarithm is strictly increasing, MLE is the same regardless of maximizing the likelihood or the log-likelihood function, so

$$\hat{\theta} = argmax_\theta \sum_{k=1}^{n} \log f_\theta(z_k).$$

In order to test whether it is reasonable to assume that the observations (historical return data) form a sample follows a suggested reference distribution F, a graphical test quantile-quantile plot (qq-plot) can be used.
A qq-plot is the plot of the points

$$\left\{ \left(F^{-1} \left(\frac{n-k+1}{n+1} \right), z_{k,n} \right) : k = 1, \ldots, n \right\}, \tag{2.3}$$

where $z_{1,n} \geq \cdots \geq z_{n,n}$ are the ordered sample data.

This qq-plot (Eq. 2.3) is the plot of the empirical quantiles against the quantiles of the reference distribution. If the data are generated by a probability distribution similar to the reference distribution then the qq-plot is approximately linear.

Chapter 3

Tests for Mean

Mean of the data is an important quantity to investigate. There are different ways to test whether the weekday data have different means. In this chapter, two kind of common used tests are introduced for various data situations. Methods presented in this chapter will be used in Chapter 7 to test the difference of mean.

3.1 The t-test

The t-test refers to the hypothesis test in which the test statistic follows a Student's t-distribution if the null hypothesis is true. The t-statistic is

$$t = \frac{Z}{S}, \tag{3.1}$$

where Z is a measure parameter shows the sensitive of the alternative hypothesis, and S is a scaling parameter that allows the distribution of t (Eq. 3.1) to be determined. In this thesis, one-sample t-test and unpaired t-test are used. For one-sample t-test, Z and S in Eq. 3.1 are

$$Z = \frac{X}{\sigma/\sqrt{n}}, S = \frac{\hat{\sigma}}{\sqrt{n}},$$

where X is the sample mean of the data, n is the sample size, and $\hat{\sigma}$ is the sample standard deviation while σ is the population standard deviation of the data.

According to the law of large numbers, the Student's t-distribution tends to be normally distributed for large degrees of freedom, thus t-tests can be applied here when the data are non-normally distributed but with large sample size. For comparing two groups with different sample sizes, the unpaired t-test is used in this thesis and the t-statistic is calculated as

$$t = \frac{X_1 - X_2}{S_{1,2}\sqrt{1/n_1 + 1/n_2}}, \tag{3.2}$$

where

$$S_{1,2} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2},$$

and S_i is the estimator of the unbiased standard deviation of the i-th group samples.

3.2 Welch's t-test

There are two types of error in hypothesis testing. Type I error occurs when the null hypothesis H_0 is true, but is rejected. A type II error occurs when the null hypothesis is false, but it is accepted as true.

Significance tests based on Normal theory, including the two-sample t-test, assume homogeneity of variance of treatment groups. Failure to satisfy this assumption, especially when sample sizes are unequal, alters Type I error rates and power. When a larger variance is associated with a larger sample size, the probability of a Type I error declines below the significance level. In contrast, the probability will increase, sometimes far above the significance level when the sample size is small. In practice, investigators do not often know the values of population variances, see Zimmerman [21].

Welch's t-test is an adaptation of typical t-test intending for using with two samples having possibly unequal variances. According to Welch [18], unlike the two-sample t-test, Welch's t-test do not pool variances in computation of an error term, thus it is insensitive to equality of variances regardless of whether the sample sizes are similar. It restores Type I error probabilities close to the significance level and also eliminates spurious increases or decreases of Type II error rates and power. Welch's t-test is better than typical t-test when population variances and sample sizes are unequal, see Albers et al. [1], Zimmerman [21]. The Welch's t-test is as follows:

$$t = \frac{X_1 - X_2}{\sqrt{s_1^2/N_1 + s_2^2/N_2}}, v = \frac{(S_1^2/N_1 + S_2^2/N2)^2}{S_1^4/N_1^2(N_1 - 1) + S_2^4/N_2^2(N_2 - 1)}, \tag{3.3}$$

where X_i is the mean of the i-th sample, s_i^2 is the unbiased estimator of the variance of the i-th sample, N_i is the size of the i-th sample, $i = 1, 2$.

The t-statistic from Eq. 3.3 is calculated, and then the p-value of the null hypothesis will be based on the degrees of freedom ν and the t-statistic.

Chapter 4

Tests for Variance

Besides mean value, it is also important to test for the equality of variance. Higher mean return does not mean better for investment. If there is a higher variance on Monday return, higher fluctuation it is, then it means there is a chance to get a higher positive return but also negative return. Sharpe ratio is a main index to measure the risk premium per unit of deviation in an investment asset.

$$S = \frac{E[R_a - R_b]}{\sqrt{var[R_a - R_b]}}, \tag{4.1}$$

Where R_a is the asset return, R_b is the risk free rate of return. The bigger value of S, the better performance it is. Besides mean, variance is also an important part of sharpe ratio, which requires more attention.

Suggesting a null hypothesis that all weekdays have same variances, one can check the p-value to see if the null hypothesis can be accepted or not. When sample sizes are identical, the null hypothesis is homoscedasticity, but when sample sizes are different then the null hypothesis is equality of the adjusted population variances,

$$H_0 : (1 - 1/n_1)\,\sigma_1^2 = (1 - 1/n_2)\,\sigma_2^2 = ... = (1 - 1/n_k)\,\sigma_k^2.$$

We use p-value to check if the hypothesis is to be believed or not in certain percent of probability. The p-value here is defined as the probability of obtaining a value of the test statistic as extreme as, or more extreme than, the actual value obtained when the null hypothesis is true. Thus, the p-value is the smallest significance level at which a null hypothesis can be rejected, given the observed sample statistic. The methods presented here will be used in Chapter 6 to test whether there is difference in the variance of weekday data.

4.1 Tests for variance

The daily returns of a stock index are defined as y_{ij}, for $i = 1, ..., k$, $j = 1, ..., n_i$ where k is the number of groups and n_i is the sample size of the i-th group. When the variances are compared among the weekdays, k will be equal to 5 and i will stands for which weekday return the data belong to.

4.1.1 The F-test

The F-test in one-way analysis of variance is used to assess whether the expected values of a quantitative variable within several pre-defined groups are different from each other.

$$F = \frac{\text{between group variability}}{\text{within group variability}}. \tag{4.2}$$

In this case,

$$F = \frac{(N-K)\sum_{i=1}^{k}(\bar{y}_i - \bar{y})^2 n_i}{(K-1)\sum_{ij}(y_{ij} - \bar{y}_i)^2}, \tag{4.3}$$

where y_{ij} is the j-th observation in the i-th out of K groups, and N is the overall sample size.

$$N = \sum_{i=1}^{k} n_i . y_i .$$

This F-statistic follows the F-distribution with $(K-1, N-K)$ degrees of freedom under the null hypothesis. The F-statistic is then applied to the F-distribution and the p-value for the null hypothesis can be calculated.

4.1.2 Levenes' test

Levene's test attempts to detect significant differences among the means of the absolute residuals, which are functions of the population variances and sample sizes. The Levene's test does not require normality of the underlying data, and is used for symmetric, moderate-tailed distributions. Levene's test is defined as the one-way analysis of variance on the absolute residual $z_{ij} = |y_{ij} - \bar{y}_i|$, where $\bar{y}_i$ stands for the mean value of y_{ij} in i-th group. The L-statistic is calculated as Eq. 4.3 with replacing y by z as following :

$$L = \frac{(N-K)\sum_{i=1}^{k}(\bar{z}_i - \bar{z})^2 n_i}{(K-1)\sum_{ij}(z_{ij} - \bar{z}_i)^2}, \tag{4.4}$$

where $z_{ij} = |y_{ij} - \bar{y}_i|$ with $\bar{y}_i$ is the mean of y_{ij} in i-th group, $\bar{z}_i$ is the mean of the z_{ij} for group i, $\bar{z}$ is the mean of all z_{ij}, K is the number of different groups to which the samples belong, n_i is the number of samples in the i-th group and N is the overall sample size.

When the L-statistic is calculated, it is applied into F-distribution as F-statistic and the p-value for testing the null hypothesis is extracted.

4.1.3 Brown-Forsythe tests

Brown and Forsythe [4] provide two alternative versions of Levene's test, named Brown-Forsythe tests. The equations of test statistic are similar with Levene's test as in Eq. 4.4 but different in forming z_{ij}.

Brown-Forsythe test I: the mean $\bar{y}_i$ in Eq. 4.4 is replaced by the median y'_i in forming z_{ij}, i.e. $z_{ij} = |y_{ij} - y'_i|$.

Brown-Forsythe test II: the mean $\bar{y}_i$ in Eq. 4.4 is replaced by the 10% trimmed mean $\tilde{y}_i$ to form $z_{ij} = |y_{ij} - \tilde{y}_i|$. To calculate $\tilde{y}_i$ of the values is to sort all the values and then discard 10% of the smallest and 10% of the largest values, then compute the mean of the remaining values.

Moreover, Brown and Forsythe perform a Monte Carlo studies before which shows that 10% trimmed mean performs best when the tested data followed a heavy-tailed distribution and the median performs best when the tested data followed a asymmetric distribution, see Brown et al [4].

4.2 Comparison of different tests

The F-test is commonly used for testing the equality of variances. However, it is sensitive to non-normality. There are many paper already show that the return data are far from Normal distribution but close to Student's t-distribution. Therefore, some alternative tests such as Levene's test and Brown-Forsythe test should be used.

To investigate whether these alternative tests show more robust when the data are not normal, two sets of random data are generated from Student's t-distribution with different degrees of freedom and non-centrality parameters, and Normal distributions with different means and variances. F-test, Levene's test and two types of Brown-Forsythe tests are applied individually on testing the variances. The results below show how many pairs which have different variances are found out in 1000 times tests. Since the data used in this thesis is around 500 pairs, here 500 data are tested.

As stated in Chapter 3, there are two types of errors, and here we will investigate how the four tests react to the Type I and Type II errors. Tab. 4.1 – 4.3 show the type II errors the four tests make, while Tab. 4.4 and 4.5 shows the type I errors.

In the tables below, $t(\nu)$ indicates the data being tested are generated from Student's t-distribution with ν degrees of freedom. $\chi^2(\nu)$ stands for χ^2 distribution with ν degrees of freedom.

Tab. 4.1 shows that when two data sets one follows Student's t-distribution and another one follows Normal distribution, Levene's test and Brown-Forsythe tests always perform better than F-test. And the closer Student's t-distribution becomes to standard Normal distribution, the better Levene's and Brown-Forsythe tests perform than F-test. Tab. 4.2 shows that when the two data sets follows Student's t-distribution but with different degrees of freedom, Levene's test and Brown-Forsythe tests always perform better than F-test.

Concluding from the above results, Levene's test and Brown-Forsythe test are more robust than F-test when data are non-normal, especially when the variances are close. Therefore, Levene's and Brown-Forsythe tests are more suitable when studying non-normally distributed financial return series.

Distributions	F-test	Levene's	Brown-Forsythe I	Brown-Forsythe II
$t(10)$,$N(0,1)$	690	880	879	879
$t(15)$,$N(0,1)$	359	585	586	585
$t(30)$,$N(0,1)$	130	272	270	271
$t(100)$,$N(0,1)$	50	137	133	137

Table 4.1: Number of the unequal variance pairs found in 1000 tests by different tests where 500 data are generated from t-distribution and from Normal distribution

Distributions	F-test	Levene's	Brown-Forsythe I	Brown-Forsythe II
$t(4)$,$t(5)$	497	632	631	633
$t(10)$,$t(11)$	110	226	232	229
$t(16)$,$t(17)$	87	172	174	170
$t(22)$,$t(23)$	61	149	140	154

Table 4.2: Number of the unequal variance pairs found in 1000 tests where 500 data are generated from t-distribution with different degrees of freedom

Distributions	F-test	Levene's	Brown-Forsythe I	Brown-Forsythe II
$\chi^2(2)$,$\chi^2(3)$	916	788	703	739
$\chi^2(4)$,$\chi^2(5)$	611	562	523	545
$t(2)$,$\chi^2(4)$	748	781	746	766
$t(3)$,$\chi^2(3)$	982	880	955	955

Table 4.3: Number of the unequal variance pairs found in 1000 tests where 500 pairs of data sets are generated from skewed distribution with different degrees of freedom or one is skewed while another one is not

Since the data used in this thesis shows a little skewed, χ^2 distribution is tested here as an example of skewed distribution. Tab. 4.3 shows that when both of the data are skewed, F-test performs better than the other three tests (more pairs of data are found). And when one of the set is skewed while another is not, there is no significant difference between all the four tests.

Distributions	F-test	Levene's	Brown-Forsythe I	Brown-Forsythe II
$N(1{,}2)$,$N(1{,}2)$	50	90	91	86
$t(4)$,$t(4)$	353	517	516	516
$t(8)$,$t(8)$	123	217	216	216
$\chi^2(4)$,$\chi^2(4)$	220	315	315	319

Table 4.4: Number of the unequal variance pairs found in 1000 tests by different tests where data are generated from same distributions

Distributions	F-test	Levene's	Brown-Forsythe I	Brown-Forsythe II
$t(6)$,$N(0{,}3/2)$	971	583	578	571
$t(10)$,$N(0{,}5/4)$	700	357	344	351
$t(12)$,$N(0{,}6/5)$	540	297	294	291

Table 4.5: Number of the unequal variance pairs found in 1000 tests by different tests where data are generated from different distributions with same variance

For Student's t-distribution $t(\nu)$, variance equals to $\nu/(\nu-2)$ for $\nu > 2$, where ν is the degree of freedom, while for $N(\alpha, \beta)$, variance is β. 1000 tests are carried out on the 500 pairs of data generated from Student's t-distribution and Normal Distribution with same variance. Tab. 4.4 shows how many pairs of data would be found to have different variance in 1000 tests when they follow the exactly same distribution. Tab. 4.5 shows how many pairs of data are found to be different in variance when there are the same variance. It shows that when the data are from the same distribution, F-test performs better than the other three, while it performs worse when there are Student's t distribution and Normal distributions. And the larger variance are, the smaller type I error are.

Size of data	F-test	Levene's	Brown-Forsythe I	Brown-Forsythe II
50,50	238	368	360	374
400,700	490	639	639	639
500,500	646	751	754	751
1000,1000	678	785	785	788

Table 4.6: Number of the unequal variance pairs found in 1000 tests by different tests where data are generated from $t(2)$ and $t(3)$

In the Tab. 4.6 – 4.8, the results indicate that the bigger the sizes are, the better performance of the tests when they are testing distributions with different variance, and the worse performance when there are same variance, and when data sizes are different, all four tests perform worse than compared

Size of data	F-test	Levene's	Brown-Forsythe I	Brown-Forsythe II
50,50	221	365	390	382
500,700	480	614	599	639
500,500	983	867	945	918
1000,1000	991	922	975	955

Table 4.7: Number of the unequal variance pairs found in 1000 tests by different tests where data are generated from $t(3)$ and $\chi^2(3)$

Size of data	F-test	Levene's	Brown-Forsythe I	Brown-Forsythe II
50,50	161	185	187	190
500,700	730	364	359	359
500,500	700	357	340	351
1000,1000	911	492	475	455

Table 4.8: Number of the unequal variance pairs found in 1000 tests by different tests where data are generated from $t(10)$ and $N(0, 5/4)$

with when the sizes are same, yet F-test still performs worse in most cases than the other three. Tab. 4.6 and 4.7 indicate the type II error while Tab. 4.8 shows the type I error.

Concluding from all the simulation tests above, none of all the tests could provide 100% correct result, but when comparing them, Levene's test and Brown-Forsythe I and II tests can often provide better results than F-test in the case of non-Normal distribution. Therefore, due to the unknown distribution of the data, the commonly used variance test F-test and also the other three tests are both used in Chapter 6 in this thesis to reach a more accurate result.

Chapter 5

Time Series Model

5.1 Stationarity

The theory introduced below follows Tsay [17] closely. A time series $\{r_t, t \in \mathbb{Z}\}$ is said to be *strictly stationary* if the distribution of $(r_{t_1}, ..., r_{t_k})$ and $(r_{t_1+t}, ..., r_{t_k+t})$ are the same for all t, where k is an arbitrary positive integer and $(t_1, ..., t_k)$ is a collection of k positive integers.

A time series $\{r_t, t \in \mathbb{Z}\}$ is said to be *weakly stationary* if
(a) The mean function is constant

$$E(r_t) = \mu \text{ for all } t \in \mathbb{Z},$$

(b) The covariance function

$$Cov(r_t, r_{t-l}) = Cov(r_0, r_{-l}) = \gamma_l \text{ , which only depends on } l,$$

and the covariance $\gamma_l = Cov(r_t, r_{t-l})$ is called the lag-l autocovariance of r_t which has two properties : (a) $\gamma_0 = Var(r_t)$ and (b) $\gamma_{-l} = \gamma_l$.

5.2 Autocorrelation function (ACF)

Assume r_t is a weakly stationary return serie. When the linear dependence between r_t and its past values r_{t-i} is of interest, the concept of correlation is generalized to autocorrelation. The correlation coefficient between r_t and r_{t-l} is called the lag-l autocorrelation of r_t and is commonly denoted by ρ_l, which under the weak stationarity assumption is a function of l only. More details can be found in Tsay [17].

$$\rho_l = \frac{Cov(r_t, r_{t-l})}{\sqrt{Var(r_t)Var(r_{t-l})}} = \frac{Cov(r_t, r_{t-l})}{Var(r_t)} = \frac{\gamma_l}{\gamma_0}. \tag{5.1}$$

5.3 Autoregressive model

When a return has a statistically significant lag-p autocorrelation, the lagged returns $r_{t-i}(i = 1, ..., p)$ are useful to predict the value of r_t.
The $AR(p)$ model is as follows:

$$r_t = \beta_0 + \beta_1 r_{t-1} + \cdots + \beta_k r_{t-p} + a_t, \tag{5.2}$$

where p is a non-negative integer and a_t is assumed to be a white noise series with mean zero and variance σ_a^2.

The order p of an AR time series can be determined by two approaches. The first one is using the partial autocorrelation function ($PACF$). The second one is utilizing the Akaike information criterion (AIC). Different approaches may result in different choices of the order p. However, one cannot determine that one approach always outperforms the other in application. They both play an important role in choosing an AR model for a given time series.

5.4 Partial autocorrelation function ($PACF$)

The $PACF$ of a stationary time series is a function of its ACF. Consider the following AR models in consecutive order:

$$\begin{aligned}
&r_t = \phi_{0,1} + \phi_{1,1} r_{t-1} + e_{1t},\\
&r_t = \phi_{0,2} + \phi_{1,2} r_{t-1} + \phi_{2,2} r_{t-2} + e_{2t},\\
&r_t = \phi_{0,3} + \phi_{1,3} r_{t-1} + \phi_{2,3} r_{t-2} + \phi_{3,3} r_{t-3} + e_{3t},\\
&r_t = \phi_{0,4} + \phi_{1,4} r_{t-1} + \phi_{2,4} r_{t-2} + \phi_{3,4} r_{t-3} + \phi_{4,4} r_{t-4} + e_{4t},\\
&\vdots
\end{aligned}$$

where $\phi_{0,j}, \phi_{i,j}$, and $\{e_{jt}\}$ are the constant terms, the coefficient of r_{t-i} and the error term of an $AR(j)$ model respectively.

These models are in the form of a multiple linear regression and can be estimated by the least squares method. The estimate $\hat{\phi_{1,1}}$ of the first equation is called the lag-1 sample $PACF$ of r_t. The estimate $\hat{\phi_{2,2}}$ of the second equation is the lag-2 sample $PACF$ of r_t, and so on.

From the definition, the lag-2 $PACF$ $\hat{\phi_{2,2}}$ shows the added contribution of r_{t-2} to r_t over the $AR(1)$ model $r_t = \phi_0 + \phi_2 r_{t-1} + e_{1t}$. The lag-3 $PACF$ shows the added contribution of r_{t-3} to r_t over an $AR(2)$ model, and so on. Therefore, for an $AR(p)$ model, the lag-p sample $PACF$ should not be zero, but $\hat{\phi_{j,j}}$ should be close to zero for all $j > p$. This property can be used to determine the order p. More details can be found in Tsay [17]

If one wants to test if the given partial correlation is zero at for example 5% significance level, it can be done by comparing the sample $PACF$ against the critical region with the approximate upper and lower confidence bounds,

$\pm 1.96/\sqrt{n}$, where n is the number of observations. This approximation relies on the assumption that the number of observation is large (say $n > 30$) and that the underlying process has a multivariate Normal distribution.

5.5 Akaike information criterion(AIC)

Akaike infomation criterion (AIC) is well-known to be used for determining the order p of the AR process. Burnham and Anderson [5] give the following expression of AIC. Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value. Hence AIC not only rewards goodness of fit, but also includes a penalty that is an increasing function of the number of estimated parameters. This penalty discourages overfitting (increasing the number of free parameters in the model improves the goodness of the fit, regardless of the number of free parameters in the data-generating process).

$$AIC = -2 \times \ln(\text{max. likelihood}) + 2 \times (\text{number of parameters}) \tag{5.3}$$

Let X be a random variable with a continuous probability distribution with density function f depending on a parameter θ. Then the likelihood function is

$$L(\theta|x) = f_\theta(x) \tag{5.4}$$

In the case of least squares (LS) estimation, if all the models in the sets assumed to have normally distributed errors with a constant variance, then Eq. 5.3 can be reduced to

$$AIC(k) = n \times \ln(\tilde{\sigma}^2) + 2 \times (\text{number of parameters}) \tag{5.5}$$

where

$$\tilde{\sigma}^2 = \frac{RSS}{n} = \frac{\sum \hat{\epsilon}_i^2}{n} = \text{the } MLE \text{ of } \sigma^2$$

and RSS stands for residual sum of squares, n is the number of observations and $\hat{\epsilon}_i$ are the estimated residuals for a particular candidate model. Then a model with the minimum AIC should be chosen. More details can be found in the book written by Burnham and Anderson [5]. More specifically, when AIC is used to determine the order of AR model, one computes the $AIC(l)$ for $l = 0, ..., m$, where m is a positive integer, then selects the order p that has the minimum AIC value.

5.6 Comparison of AIC and $PACF$

As mentioned before, different approaches may result in different choices of the order p. In order to clarify which one should be used here, a simple test is carried out to test the reliability of AIC and $PACF$ in different situations.

Two $AR(3)$ models E.q. 5.6, 5.7 and two $AR(3)$ models with $GARCH(1,1)$ as conditional variance model E.q. 5.8, 5.9 are created. Model 1 and 3 are chosen based on the coefficient gained from the time series model fitting with our data in Section 6.5. Model 2 and 4 are doubled the coefficients in model 1 and 3 respectively for checking the reliability of both methods if the coefficients are larger. The simulation are done by generating 500 or 2500 data points from each model, then find the $PACF$ values and also do the linear regression and calculate the AIC values by Eq. 5.5.

$$\text{Model 1}: \; r_t = -0.0134r_{t-1} - 0.0437r_{t-2} - 0.0679r_{t-3} + a_t, \tag{5.6}$$

$$\text{Model 2}: \; r_t = -0.0268r_{t-1} - 0.0874r_{t-2} - 0.1358r_{t-3} + a_t, \tag{5.7}$$

where a_t is assumed to be a white noise series with mean zero and variance equal to one.

$$\text{Model 3}: \; r_t = 0.0014 - 0.0210r_{t-1} - 0.0232r_{t-2} - 0.0434r_{t-3} + \epsilon_t, \tag{5.8}$$

$$\text{Model 4}: \; r_t = 0.0028 - 0.0420r_{t-1} - 0.0464r_{t-2} - 0.0868r_{t-3} + \epsilon_t, \tag{5.9}$$

where

$$\epsilon_t = \sigma_t z_t, \{z_t\} \sim \text{i.i.d.}N(0,1),$$
$$\sigma_t^2 = 0.000002 + 0.905\epsilon_{t-1}^2 + 0.0882\sigma_{t-1}^2,$$

For each model, we repeat the simulation for 100 times. For each simulation, we record if AIC or $PACF$ can correctly show us the correct order of the data.

	500 data		2500 data	
Model	AIC	$PACF$	AIC	$PACF$
1	26	32	49	36
2	51	59	72	50
3	1	7	12	3
4	6	31	65	13

Table 5.1: The number of times that AIC and $PACF$ can find out the correct order from 500 or 2500 simulated data out of 100 simulations

From Tab. 5.1, when the sample size is small, $PACF$ performs better than AIC. When the sample size is large, AIC performs better. And we can see that AIC can perform better when the sample size is larger. Therefore, we cannot say that one method is always outperforming than another, it depends on the sample size. By comparing the models, we can also observe that AIC and $PACF$ both find out more correct order in model 2 over model 1 and model 4 over model 3. It means that they show more reliable when the coefficients of the model are larger. However, we can also notice that the number of times that two methods find out the correct order are quite small especially when the sample size is small. Sometimes, it may happen that $PACF$ shows the correct order but AIC not in a large sample size or in the reverse way. Therefore, when we do the order determination, we cannot just trust either one of them totally. Instead we should put either one of them in the first priority for the consideration depends on the sample size and then we should check for another one if it also agrees with the result. For example, if the sample size is large, AIC will be considered first, and then we get the order p. If the $PACF$ of order p is within the confidence bounds and very small, we should not still use this order and we should reconsider maybe by $PACF$ values to find a more correct and reasonable one.

5.7 Conditional heteroscedastic models

Volatility is an important factor in option trading and also many other financial applications. Conditional heteroscedastic models are used for modeling the volatility of an asset return. Two of those models, $ARCH$ and $GARCH$, will be presented here .

Before setting up a model, a test for conditional heteroscedasticity is needed. Let $\epsilon_t = r_t - \mu_t$ be the residuals of the mean equation, then the squared series ϵ_t^2 is used for checking the conditional heteroscedasticity ($ARCH$ effect). The Lagrange multiplier test of Engle [8] can be used for checking the $ARCH$ effect. It is done by testing $\alpha_i = 0$ for $i = 1, ...m$, like the usual F-statistic. The null hypothesis $\alpha_1 = ... = \alpha_m = 0$, in the linear regression:

$$\epsilon_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + ... + \alpha_m \epsilon_{t-m}^2 + e_t, \; t = m+1, ..., T,$$

where e_t denotes the error term, m is a prespecified positive integer, and T is the sample size.

5.7.1 Autoregressive conditional heteroskedastic model

The autoregressive conditional heteroskedastic model ($ARCH$) is introduced by Engle [8]. This model includes the consideration weather the variance depends on the past. The basic idea is that (a) the shock (or innovation) of an asset return is serially uncorrelated, but dependent, and (b) the dependence of the shock can be described by a simple quadratic function of its lagged values. More details can be found in Tsay [17].
The $ARCH(q)$ model is given by:

$$\epsilon_t = \sigma_t z_t, \{z_t\} \sim \text{i.i.d.} N(0,1), \sigma_t^2 = a_0 + \sum_{i=1}^{q} (a_i \epsilon_{t-i}^2), \tag{5.10}$$

where $a_0 > 0$ and $a_i \geq 0$ for $i > 0$.

5.7.2 Generalized $ARCH$ ($GARCH$) model

In empirical applications of the $ARCH$ model, it often calls for a relatively long lag in the conditional variance equation. Therefore, an extension of the $ARCH$ class model to allow for both a longer memory and a more flexible lag structure is needed. The generalized $ARCH$ model is introduced by Bollerslev [3]. Besides the past sample variances, $GARCH(p,q)$ process allows lagged conditional variances to enter the linear function of the conditional variance too.

The process $\{\epsilon_t, t \in \mathbb{Z}\}$ is said to be a $GARCH(p,q)$ process if it is stationary and if

$$\epsilon_t = \sigma_t z_t, \{z_t\} \sim \text{i.i.d.} N(0,1),$$

where

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q}(a_i \epsilon_{t-i}^2) + \sum_{j=1}^{p}(b_j \sigma_{t-j}^2),$$

and $a_0 > 0, a_i \geq 0, b_j \geq 0$.

Chapter 6

Data analysis - Day-of-the-week effect

6.1 Description of the data source

All the data are extracted from Bloomberg. Six different indexes are chosen to examine for the existence of the day-of-the-week effect. They are OMX Index (OMX Stockholm 30 Index), SAX Index (OMX Stockholm All-Share Index), SX5E Index (EURO STOXX 50 (Price) Index), SPX Index (Standard and Poor's 500 Index), NKY Index (Nikkei-225 Stock Average) and HSI Index (Hang Seng Index). Daily closing prices of all indexes from 1/1/2001(Mon) to 13/3/2012 (Fri) are used. All data are adjusted for corporate actions such as dividends to assure fair return series.

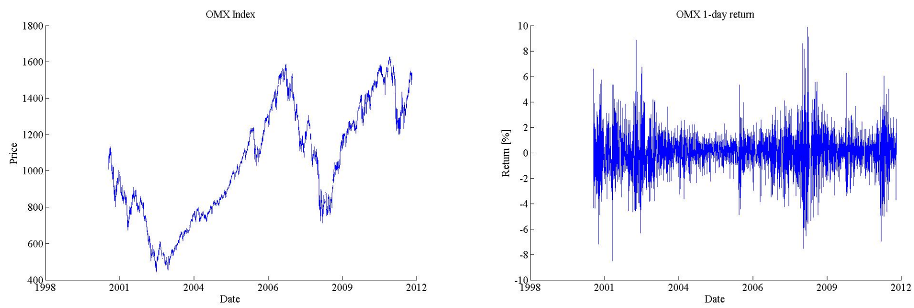
Index	Median	Mean	Variance	Lowest	Highest
OMX	1047.90	1045.47	96623	442.39	1627.19
SAX	306.92	311.03	9878	126.50	492.51
SX5E	3786.15	3774.44	608994	1958.94	5592.50
SPX	1307.35	1313.48	52705	795.83	1778.40
NKY	11550.89	12437.82	8291729	7697.85	19470.91
HSI	19326.43	21212.34	65341037	9264.55	40904.02

Table 6.1: Summary statistics for the daily price of the studied indexes

6.2 One-day weekday return

In order to examine the day-of-the-week effect, all prices are changed to one-day log returns.

$$R_t = \log\left(\frac{P_{t+1}}{P_t}\right) \tag{6.1}$$



(a) (b)

Figure 6.1: Time plots of (a) price for OMX index (b) one-day log returns for OMX index

The return data of

$$\log\left(\frac{\text{Closing price on Monday}}{\text{Closing price on Friday}}\right)$$

will be considered as Monday return. As some of the weekdays are holidays, the return data which include holidays in the formula for calculating the return value are removed.

Index	Mean $[10^{-4}]$	Variance $[10^{-4}]$	Kurtosis	Skewness
OMX All	1.08	2.72	6.35	0.10
OMX Mon	1.71	3.43	8.16	0.33
OMX Tue	-5.53	2.56	6.18	0.04
OMX Wed	5.19	2.72	4.46	-0.07
OMX Thu	6.38	2.72	4.85	0.00
OMX Fri	-2.62	2.19	6.60	0.07

Table 6.2: Summary statistics for one-day returns of OMX

Tab. 6.2 shows a first study about the one-day return of OMX data. Tuesday and Friday have negative mean returns while the others have positive mean returns and Thursday is the highest. The results also indicate that Monday returns have slightly higher variance. Moreover, it is known that the kurtosis of a Normal distribution is 3. By considering the kurtosis and the histogram graphs of the weekdays (Fig. 6.2), one can observe that all weekdays have relatively high peaks and heavy tails, especially on Monday. It also shows that their distributions are far from the Normal distribution. As for the value of skewness, most of return distributions have positive value, the histogram graphs however seem a bit left skewed, especially on the Monday returns. The positive skewness may be due to some relatively large positive values. By removing the 1-3 largest and smallest values of Monday returns, the value of skewness decreases to 0.20, 0.09, -0.08, respectively. This result is consistent with the assumption above.

6.3 Parametric models

All the weekday return data are fitted with different parametric families, Normal distribution, location-scale Student's t-distribution and third-degree polynomial of standard normal model. Maximum likelihood estimation (MLE) is used to estimate the parameters for the Normal and Student's t while least squares estimation is used for the last one.

6.3.1 Normal distribution

All the return data are first fitted to a Normal distribution and qq-plots Fig. 6.3 are obtained. From the qq-plots, one can observe that all the return data, especially Monday return data, are far from the Normal distribution with heavy tails. Therefore, location-scale Student's t-distribution is then fitted.

Weekday Return	$\mu[10^{-4}]$	$\sigma[10^{-2}]$
Monday	1.71	1.85
Tuesday	-5.53	1.60
Wednesday	5.19	1.65
Thursday	6.38	1.65
Friday	-2.62	1.48

Table 6.3: Estimated parameters by MLE for fitting OMX one-day return data with Normal distribution for different weekdays

6.3.2 Location-scale Student's t-distribution

Then location-scale Student's t-distribution is used as reference distribution, with results are presented in Tab. 6.4 and the qq-plots in Fig. 6.4. The estimated standard deviation is calculated by the formula $\sigma \times \sqrt{\nu/(\nu-2)}$. Tab. 6.5 shows the confidence intervals at level 95% for the estimated parameters from MLE by using bootstrap method which is described in Hult et al 's book [12].

From the qq-plots Fig. 6.4, one can observe that all the return data fit reasonable well with Student's t-distribution. Moreover, Monday has the lowest estimated degrees of freedom among all the weekday returns, so it has the heaviest tail. Comparing the estimated standard deviation and standard deviation obtained from the MLE with the Normal distribution, Monday returns show a significant difference while other weekdays have similar values.

Weekday Return	$\mu[10^{-4}]$	$\sigma[10^{-2}]$	ν	SD $[10^{-2}]$
Monday	8.65	1.07	2.54	2.31
Tuesday	-5.31	1.13	3.65	1.67
Wednesday	6.24	1.27	4.48	1.70
Thursday	5.98	1.24	4.28	1.70
Friday	1.09	1.04	3.62	1.55

Table 6.4: Estimated parameters by MLE for fitting OMX one-day return data with a location-scale Student's t-distribution $t(\mu, \sigma^2, \nu)$ for different weekdays.

Weekday Return	95% CI of $\mu[10^{-4}]$	95% CI of $\sigma[10^{-2}]$	95% CI of ν
Monday	-2.94 – 20.50	0.95 – 1.19	1.70 – 3.05
Tuesday	-16.04 – 5.97	0.99 – 1.26	1.91 – 4.49
Wednesday	-7.14 – 19.72	1.11 – 1.41	1.86 – 5.69
Thursday	-6.06 – 17.89	1.09 – 1.36	2.01 – 5.37
Friday	-9.61 – 12.13	0.92 – 1.15	2.08 – 4.44

Table 6.5: Confidence intervals at level 95% of the estimated parameters by using bootstrap

6.3.3 Third-degree polynomial of standard normal model

Since the qq-plots Fig. 6.3 against the Normal distribution turn out like a graph of a third-degree polynomial, it is reasonable to believe that the one-day log return samples can be seen as outcomes of the random variable $g(Y;\theta)$, where Y is standard normally distributed, $\theta = (\theta_0, \theta_1, \theta_2, \theta_3)$, and

$$g(y;\theta) = \theta_0 + \theta_1 y + \theta_2 y^2 + \theta_3 y^3,$$

which is also called third-degree polynomial of standard normal model. The quantile function of $g(Y;\theta)$ is given by

$$F^{-1}_{g(Y;\theta)}(p) = \theta_0 + \theta_1 \Phi^{-1}(p) + \theta_2 \Phi^{-1}(p)^2 + \theta_3 \Phi^{-1}(p)^3.$$

The return data are then tried to fit with this model and obtain the results in Tab. 6.6. By checking the value θ_3 across different weekdays in Tab. 6.6, one can notice that Monday has a relatively higher value than others. It seems that Monday has a relatively heavier tail comparing to other weekdays. It is consistent with the finding in Student's t-distribution where Monday has the lowest degrees of freedom.

The qq-plots of the third-degree polynomial of different weekdays are plotted in Fig. 6.5. Comparing them to the qq-plot fitted with location-scale Student's t-distribution(Fig. 6.4), they both give a good fit and it is difficult to see the difference between two models. Then three different models

Weekday Return	$\theta_0[10^{-4}]$	$\theta_1[10^{-4}]$	$\theta_2[10^{-4}]$	$\theta_3[10^{-4}]$
Monday	1.60	106.82	0.11	25.56
Tuesday	-6.40	113.21	0.89	15.92
Wednesday	6.69	132.86	-1.53	11.30
Thursday	5.90	128.62	0.49	12.69
Friday	-0.49	104.51	-2.17	14.69

Table 6.6: Estimated parameters by least squares estimation for fitting OMX one-day return data to the third-degree polynomial of standard normal for different weekdays

are plotted on the histograms of empirical distribution of different weekday returns for comparison. From Fig. 6.6, one can observe that Normal distribution has the poor fit. Location-scale Student's t-distribution and third-degree polynomial of the standard normal model both show better fit since they both have more parameters.

6.4 Test for homogeneity in variance

In order to test for the homogeneity in variance of the one-day return data among different weekdays, F-test is carried out with the null hypothesis that the population variances are equal and Tab. 6.7 is obtained. When 5% or even 1% significance level is used in F-test, the null hypothesis that Monday has the same variance as all the other weekday returns can be rejected. Also, the null hypothesis that Friday has the same variance with Wednesday and Thursday are also rejected under 5% significance level.

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.07	0.66	0.68	0.00
Tue	NaN	100.00	47.81	48.13	6.74
Wed	NaN	NaN	100.00	99.79	1.11
Thu	NaN	NaN	NaN	100.00	1.14
Fri	NaN	NaN	NaN	NaN	100.00

Table 6.7: The p-values of the F-test for OMX one-day returns

As already shown, the return data are far from normally distributed, the results from F-test may not be reliable. Therefore, the testing of the homogeneity in variance is also performed with Levene's and Brown-Forsythe tests. Due to the unknown return data distribution, Levene's test and Brown-Forsythe test with median and 10% trimmed mean are used here instead of just using either one of them. Again, those test are all used for testing the null hypothesis that the population variances are equal.

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.00	0.00	0.00	0.00
Tue	NaN	100.00	99.88	86.45	6.90
Wed	NaN	NaN	100.00	86.49	6.72
Thur	NaN	NaN	NaN	100.00	4.64
Fri	NaN	NaN	NaN	NaN	100.00

Table 6.8: The p-values of the Levene's test for OMX one-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.00	0.00	0.00	0.00
Tue	NaN	100.00	99.91	88.08	7.72
Wed	NaN	NaN	100.00	87.91	7.52
Thur	NaN	NaN	NaN	100.00	5.49
Fri	NaN	NaN	NaN	NaN	100.00

Table 6.9: The p-values of the Brown-Forsythe test (median) for OMX one-day returns

Referring to Tab. 6.8 – 6.10, and at 5% significance level, the null hypothesis that Monday has the same variance as all the other weekday returns can be rejected in all tests. It agreed with the result from the F-test. However, these new tests show the null hypothesis is not rejected for Friday and Wednesday return. Also, the null hypothesis for Friday and Thursday return is only rejected in Levene's test and Brown-Forsythe test with 10% trimmed mean but not in the Brown-Forsythe test with median. In addition, if one check for the p-values between Tuesday and Wednesday, one can see that they are all very close to 100% in Levene's and Brown-Forsythe tests. It seems that Tuesday and Wednesday have almost the same variance.

Concluding from all these tsets, the null hypothesis that the variances are all the same for all weekdays is rejected. Moreover, the p-values for Monday returns are very close to zero so one can strongly believe that Monday has a very different variance than the other weekday returns.

The same tests are repeated with other index data. SAX, SX5E and HSI indexes all show similar results as OMX that Monday return has a different variance with all the other weekday returns. While the American index SPX index not only Monday but also Friday has different variance with all the others. For the index in Japan, NKY, the result is different. It shows only Tuesday and Thursday have different variances.

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.00	0.00	0.00	0.00
Tue	NaN	100.00	99.85	86.74	7.07
Wed	NaN	NaN	100.00	86.50	6.94
Thu	NaN	NaN	NaN	100.00	4.80
Fri	NaN	NaN	NaN	NaN	100.00

Table 6.10: The p-values of the Brown-Forsythe test (10% trimmed mean) for OMX one-day returns

6.5 Time series model

6.5.1 *AR* model

From the autocorrelation graph Fig. 6.7, one can observe that there exists autocorrelation in the return data. Therefore, a mean equation is set up as an autoregressive model with dummy variables for different weekdays to examine for the day-of-the-week effect.

$$R_t = \alpha_0 + \alpha_1 D_T + \alpha_2 D_W + \alpha_3 D_H + \alpha_4 D_F + \sum_{i=1}^{k} \beta_i R_{t-i} \qquad (6.2)$$

where D_T, D_W, D_H and D_F are dummy variables for Tuesday, Wednesday, Thursday and Friday respectively.

Order	2	3	5	6	7
AIC [10^4]	-2.2491	-2.2492	-2.2481	-2.2479	-2.2475
$PACF$	-0.0431	-0.0682	-0.0531	-0.0534	0.0450

Order	10	17	18	19
AIC [10^4]	-2.2449	-2.2388	-2.2389	-2.2385
$PACF$	-0.0316	0.0401	-0.0440	-0.0529

Table 6.11: The values of AIC and $PACF$ of OMX one-day returns with different orders

From Tab. 6.11, one can observe that order 3 has the smallest AIC value -22 494 and the largest absolute $PACF$ value 0.0682 among all the orders. Therefore, $AR(3)$ model will be appropriate here. Then the built in function "regstats" in Matlab is used to get all the regression statistics.

From Tab. 6.12, all the p-values of the coefficient of the dummies are larger than 5% which mean the null hypothesis of the coefficients equal to zero in all the dummies are accepted. Therefore, it seems that there is no day-of-the-week effect in the mean equation.

parameter	α_0	α_1	α_2	α_3	α_4	β_1	β_2	β_3
value $[10^{-4}]$	2.67	-7.71	3.00	2.29	-5.59	-134.47	-437.13	-679.41
se$[10^{-4}]$	7.10	9.97	9.91	9.95	10.06	191.02	190.36	190.30
t-statistic	0.38	-0.77	0.30	0.23	-0.56	-0.70	-2.30	-3.57
p-value(%)	70.76	43.92	76.21	81.77	57.81	48.15	2.17	0.04

Table 6.12: Summary Statistics of regression for OMX one-day returns $AR(3)$ model by using Matlab

The same procedures are repeated with other indexes. SAX, SX5E, SPX and HSI index all show that they do not have any day-of-the-week effect. However, NKY index shows a different result. An AR model with order 6 is used as the mean equation for NKY. Even the smallest AIC value of NKY appears in order 1, the $PACF$ value of order 1 is within the confidence bounds, so order 6 which has the largest absolute $PACF$ value 0.049 among all the orders is chosen to get a more correct result. When 5% significance level is used, from Tab. 6.13, the null hypothesis for the coefficient of the Thursday dummy can be rejected. It means the day-of-the-week effect may exist on Thursday. However, if 4% or even lower significance level is used, the null hypothesis for Thursday cannot be rejected. Therefore, this special finding is not strong enough to be concluded here.

parameter	α_0	α_1	α_2	α_3	α_4
value $[10^{-4}]$	-12.32	1.47	11.89	19.63	8.05
se $[10^{-4}]$	7.08	9.97	9.74	9.73	9.73
t-statistic	-1.74	0.15	1.22	2.02	0.83
p-value(%)	8.17	88.26	22.20	4.37	40.83

β_1	β_2	β_3	β_4	β_5	β_6
-211.61	178.88	-406.93	-292.27	184.51	493.38
195.99	196.03	196.05	195.99	195.92	195.86
-1.08	0.91	-2.08	-1.49	0.94	2.52
28.04	36.16	3.80	13.60	34.64	1.18

Table 6.13: Summary Statistics of regression for NKY one-day returns $AR(6)$ model by using Matlab

6.5.2 $GARCH(1,1)$

By observing Fig. 6.7, it seems that the series is serially uncorrelated but dependent. The Lagrange multiplier test by Engle [8] is performed to test if there is evidence for conditional heteroscedasticity ($ARCH$ effect) in the series. It rejects the null hypothesis that no $ARCH$ effect in the residual of $AR(3)$ at the 5% significance level. Thus, indicating that there is time varying conditional heteroscedasticity in the return data. Therefore, a $GARCH(1,1)$ model with the mean equation Eq. 6.3 is used to examine if there is any day-of-the-week effect.

$$R_t = \alpha_0 + \alpha_1 D_T + \alpha_2 D_W + \alpha_3 D_H + \alpha_4 D_F + \sum_{i=1}^{3} \beta_i R_{t-i} + \epsilon_t, \tag{6.3}$$

$$\epsilon_t = \sigma_t z_t \ , \ \sigma_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2, \tag{6.4}$$

where D_T, D_W, D_H and D_F are dummy variables for Tuesday, Wednesday, Thursday and Friday respectively. $\{z_t\}$ is assumed to be a sequence of $i.i.d$ random variables with standard Normal distribution.

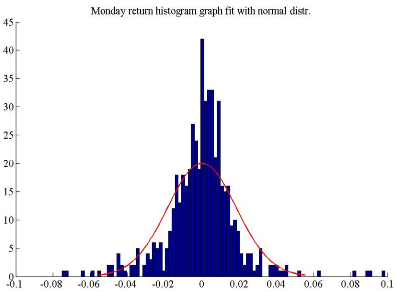
parameter	value $[10^{-4}]$	se $[10^{-4}]$	t-statistic	p-value(%)
α_0	14.02	5.12	2.74	0.62
α_1	-17.44	7.04	-2.48	1.33
α_2	-1.35	6.90	-0.20	84.49
α_3	-4.64	6.95	-0.67	50.48
α_4	-2.67	7.27	-0.37	71.37
β_1	-209.90	205.97	-1.02	30.82
β_2	-232.33	196.17	-1.18	23.64
β_3	-433.57	196.61	-2.21	2.75
a_0	0.02	0.00	4.50	0.00
a_1	9049.60	85.00	106.46	0.00
b_1	882.11	81.37	10.84	0.00

Table 6.14: Summary Statistics of $GARCH(1,1)$ model for OMX one-day returns by using Matlab

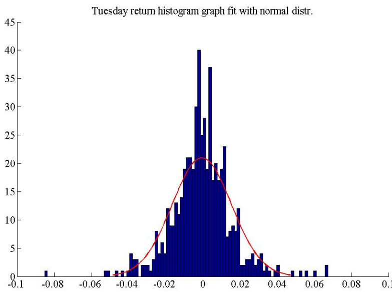
The built in function "garchfit" in Matlab is used and the statistics are obtained. From Tab. 6.14, the p-value of α_1 is just 1.33%. Therefore, the null hypothesis of no day-of-the-week effect on Tuesday is rejected at the 5% or even 2% significance level. While the null hypothesis of other weekdays still cannot be rejected. Moreover, the values in all the coefficients of dummies are negative and Tuesday is the most lowest one which has -0.001744. Recalling the result from AR model, which shows that there is no any day-of-the-week effect, the result from $GARCH$ model is inconsistent with the AR result. Yang and Chang [20]

The same procedures are repeated with other indexes. The null hypothesis of no day-of-the-week effect on Tuesday is also rejected on SAX and HSI indexes which mean day-of-the-week effect on Tuesday also exist for these two indexes. For SPX, SX5E and NKY indexes, no day-of-the-week effect is evident. Comparing the results from AR model, the results from

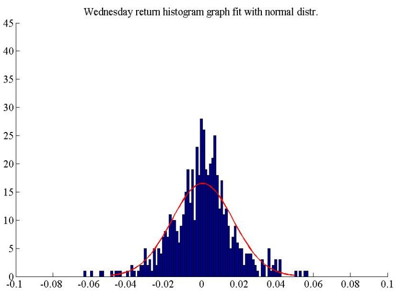
$GARCH(1,1)$ model are very different. In the AR model, SAX and HSI do not show any day-of-the-week effect on any weekday but they show day-of-the-week effect on Tuesday in $GARCH(1,1)$. For NKY index, it shows that it may have a little bit day-of-the-week effect on Thursday, but it no longer exists when $GARCH(1,1)$ model is used. It seems that $ARCH$ effect makes a very large impact there.



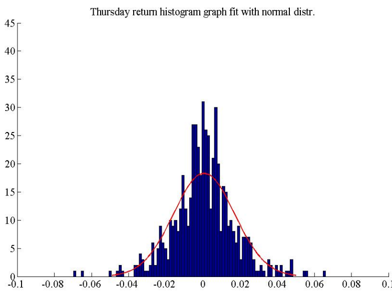
(a)



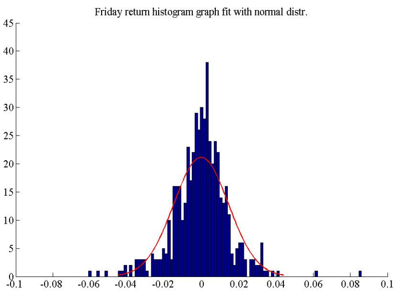
(b)



(c)

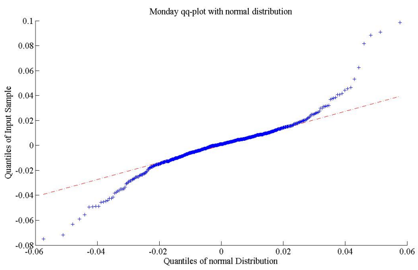


(d)

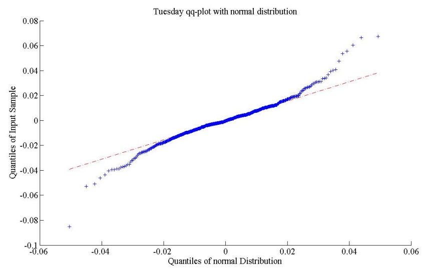


(e)

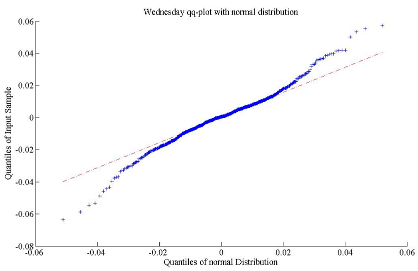
Figure 6.2: Histogram graphs fit with Normal distribution for (a) Monday (b) Tuesday (c) Wednesday (d) Thursday (e) Friday one-day returns of OMX.



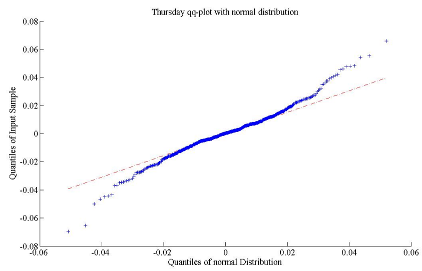
(a)



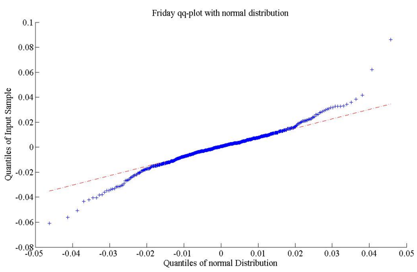
(b)



(c)

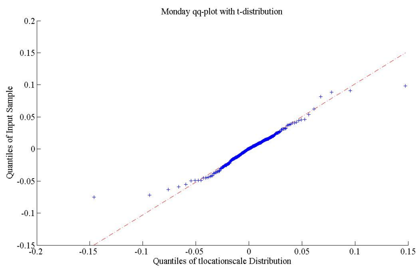


(d)

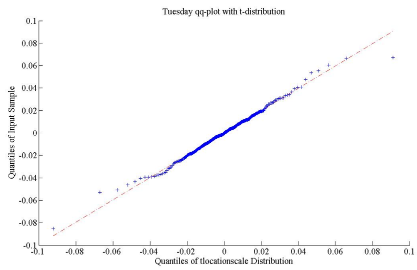


(e)

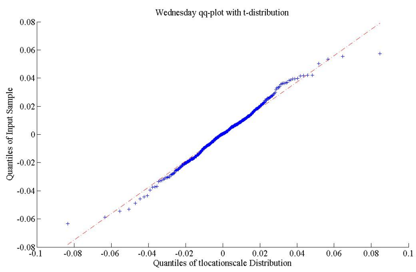
Figure 6.3: The *qq*-plots with Normal distribution for (a) Monday (b) Tuesday (c) Wednesday (d) Thursday (e) Friday one-day returns of OMX



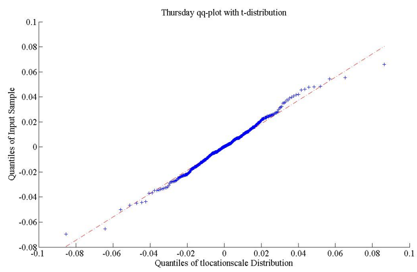
(a)



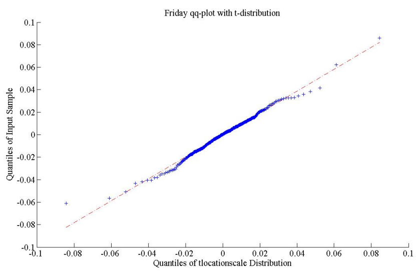
(b)



(c)

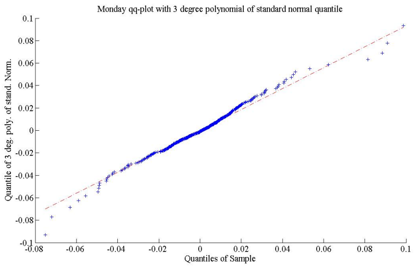


(d)

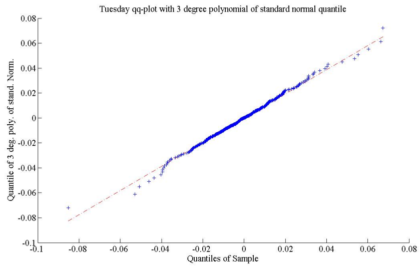


(e)

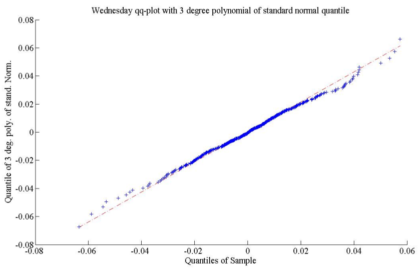
Figure 6.4: The qq-plots with Student's t-distribution for (a) Monday (b) Tuesday (c) Wednesday (d) Thursday (e) Friday one-day returns of OMX



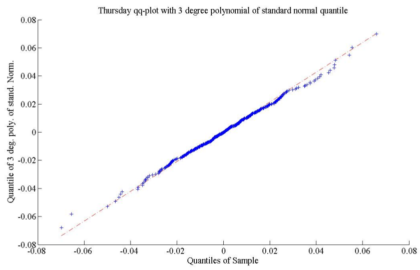
(a)



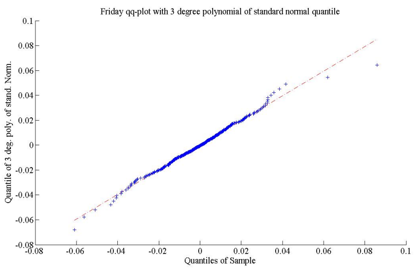
(b)



(c)

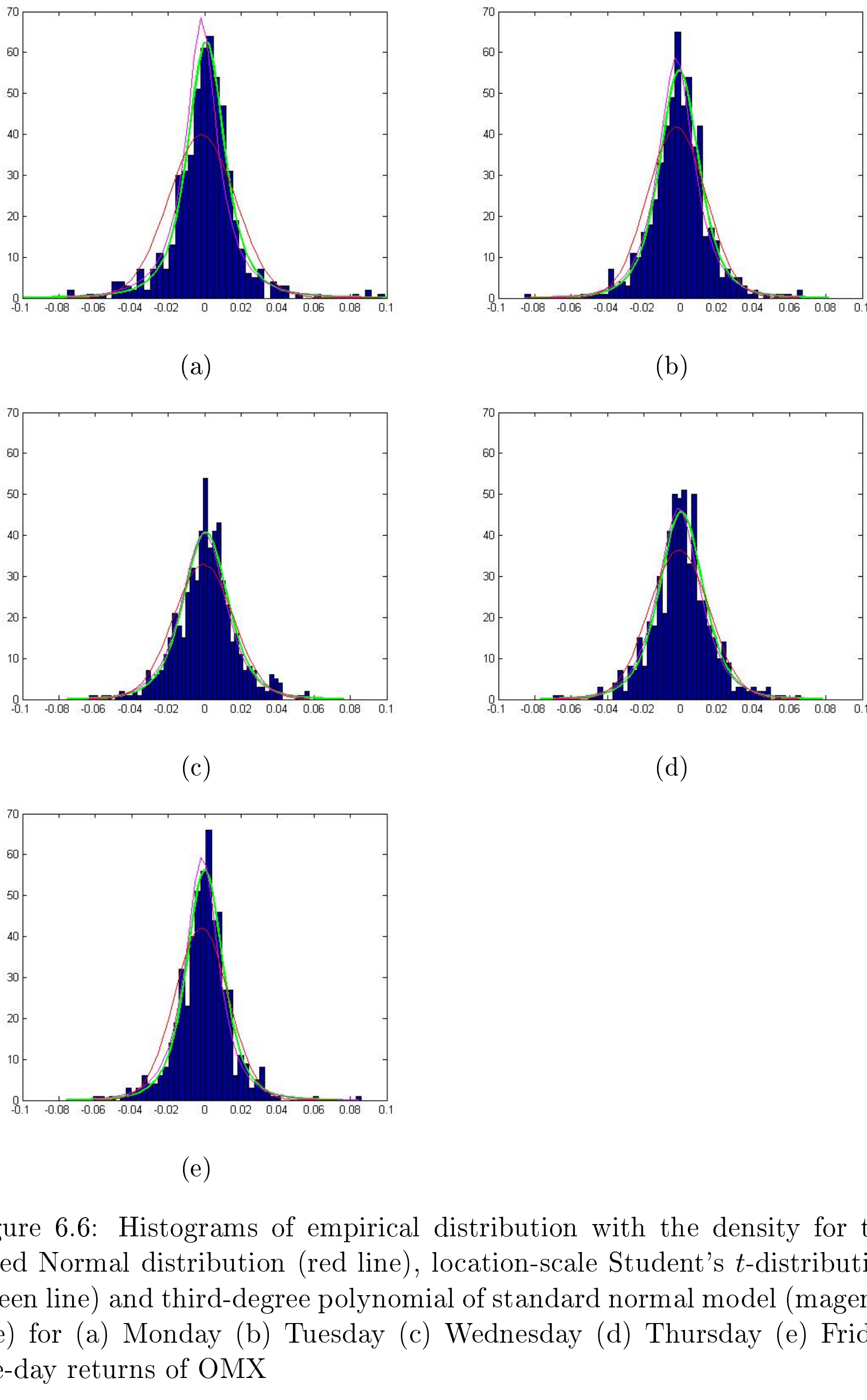


(d)



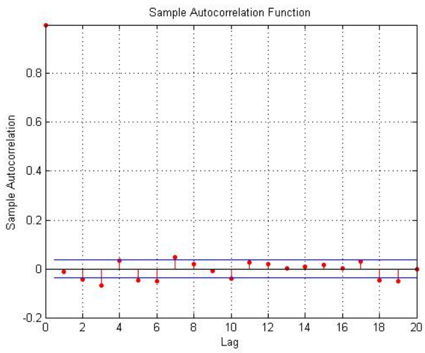
(e)

Figure 6.5: The qq-plots with third-degree polynomial of standard normal model for (a) Monday (b) Tuesday (c) Wednesday (d) Thursday (e) Friday one-day returns of OMX

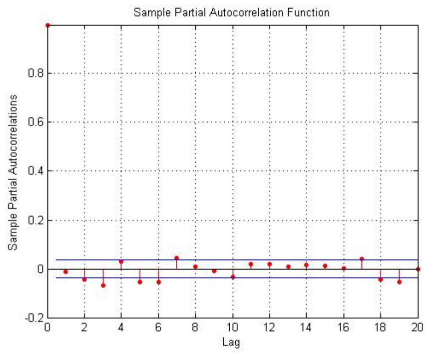


(a) (b) (c) (d) (e)

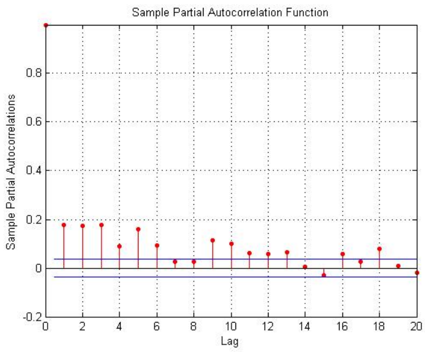
Figure 6.6: Histograms of empirical distribution with the density for the fitted Normal distribution (red line), location-scale Student's t-distribution (green line) and third-degree polynomial of standard normal model (magenta line) for (a) Monday (b) Tuesday (c) Wednesday (d) Thursday (e) Friday one-day returns of OMX



(a)



(b)



(c)

Figure 6.7: (a) Sample autocorrelation function of the OMX one-day return series, (b) Sample partial autocorrelation function of the OMX one-day return series (c) Sample partial autocorrelation function of the OMX squared one-day returns

Chapter 7

Conditional Analysis

Here the daily return to the return over the previous week return are studied. To define previous week's return, the following example introduces the methodology: Assuming today is Monday, July 18, and there is no holiday over the previous week, the previous week's return will be measured from the close of Friday, July 8 to the close of Friday, July 15 if no trading occurs on the weekend. Holidays are handled in the following simple manner. With no Saturday trading and a holiday on Friday, July 15, the previous week's return will be measured from the close of Friday, July 8 to the close of Thursday, July 14. As for a holiday on Friday, July 8, it will necessitate an interval from the close of Thursday, July 7 to the close of Friday, July 15.

7.1 Analysis based on previous week return

Daily returns are classified as being either above or below zero according to the previous week returns. Average returns across these two subsamples are presented in Tab. 7.1, 7.2. The one sample t-test is used to determine whether the daily return is 0 or not and the results are presented in those table also.

	Monday	Tuesday	Wednesday	Thursday
mean[10^{-3}]	-0.79	-1.85	-0.13	0.31
variance[10^{-4}]	2.41	2.23	2.32	2.02
p-value(%)	39.94	4.09	88.90	72.02

Table 7.1: Daily return summary when the last week return is positive

From Tab. 7.1, 7.2, under a significance of 5%, there is no Monday effect, and Monday's return is not influenced by previous week return. One can observe that only Tuesday's return is influenced by the previous week return but the Tuesday effect only exists when the previous week return is

	Monday	Tuesday	Wednesday	Thursday
mean[10^{-3}]	1.19	1.11	1.2	0.35
variance[10^{-4}]	4.21	2.80	2.97	3.26
p-value(%)	35.35	28.85	25.25	75.62

Table 7.2: Daily return summary when the last week return is negative

positive not in the case of negative. Also, it shows that the Tuesday return is significantly below zero when previous week return is positive while it is no difference than zero when the previous week return is negative.

Since the two subsamples of every weekday have different size from each other, normal t-test cannot be used. Welch's t-test is used here to test whether the two subsamples in every weekdays have same mean values or not.

	Monday	Tuesday	Wednesday	Thursday
t-statistic	-1.25	-2.15	-0.96	-0.03
ν	478.18	515.09	513.19	487.92
p-value(%)	21.24	3.20	33.64	97.62

Table 7.3: Welch t-test for the subsamples of everyday return

Tab. 7.3 shows that for Tuesday, there is difference between the two subsamples (based on whether previous week return is positive or negative), while all the others have no difference.

Chapter 8

Further study in longer holding period

8.1 Two-day return

After the day-of-the-week effects in one-day returns of the indexes are investigated, it is interesting to study longer holding periods using the same methodology to investigate existence of other time anomalies. Instead of one-day returns, two days holding periods are considered here. The same models and the same procedures are repeated but the studied series are changed to two-day log returns instead of one-day in this part.

$$R_t = \log\left(\frac{P_{t+2}}{P_t}\right)$$

The return data of

$$\log\left(\frac{\text{Closing price on Wedesday}}{\text{Closing price on Monday}}\right)$$

will be considered as Wednesday two-day return.

Index	Mean $[10^{-4}]$	Variance $[10^{-4}]$	Kurtosis	Skewness
OMX All	1.87	5.37	5.47	-0.05
OMX Mon	-4.58	5.39	4.84	-0.56
OMX Tue	-3.36	5.96	7.11	0.41
OMX Wed	-0.06	4.96	4.47	-0.09
OMX Thu	11.60	5.74	5.25	-0.07
OMX Fri	5.50	4.83	4.94	-0.04

Table 8.1: Summary statistics for two-day returns of OMX

Tab. 8.1 shows some summary statistics for the two-day returns of OMX data. Thursday and Friday have positive mean returns while the others have negative mean returns and Thursday is the highest. The results also indicate that Tuesday return has slightly higher variance. Moreover, it is known that the kurtosis of a Normal distribution is 3. By considering the kurtosis and the histogram graphs of the weekdays in Fig. 8.1, one can observe that all weekdays have relatively high peaks and heavy tails. It also shows that their distributions are far from normally distributed and have a better fit with location-scale Student's t-distribution and third-degree polynomial of standard normal model.

8.2 Tests for homogeneity in variance

The F-test, Levene's test and two type of Brown-Forsythe tests are used to test for homogeneity in variance of the two-day return data among different weekdays.

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	24.65	34.09	46.39	20.05
Tue	NaN	100.00	3.43	66.42	1.47
Wed	NaN	NaN	100.00	9.03	73.94
Thu	NaN	NaN	NaN	100.00	4.33
Fri	NaN	NaN	NaN	NaN	100.00

Table 8.2: The p-values of F-test for OMX two-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	2.19	8.79	42.27	9.79
Tue	NaN	100.00	0.01	13.19	0.01
Wed	NaN	NaN	100.00	1.17	96.33
Thu	NaN	NaN	NaN	100.00	1.37
Fri	NaN	NaN	NaN	NaN	100.00

Table 8.3: The p-values of the Levene's test for OMX two-day returns

Tab. 8.2 - 8.5 are the results from different tests. The null hypothesis is again the variance of two weekdays are equal. When 5% significance level is used, the null hypothesis for Monday and Tuesday are rejected in Levene's test and Brown-Forsythe test with 10% trimmed mean. While the null hypothesis for Monday and Wednesday and also Monday and Friday are just rejected in Brown-Forsythe test with median. As for Wednesday and Thursday, the null hypothesis cannot be rejected in F-test, but can be rejected for the other three tests. At last, the null hypothesis of Tuesday with Wednesday, Tuesday with Friday and Thursday with Friday are all rejected in all tests. Concluding from all these tests, the null hypothesis that the

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	7.84	2.56	78.68	3.33
Tue	NaN	100.00	0.01	13.32	0.01
Wed	NaN	NaN	100.00	1.18	92.30
Thu	NaN	NaN	NaN	100.00	1.59
Fri	NaN	NaN	NaN	NaN	100.00

Table 8.4: The p-values of the Brown-Forsythe test (median) for OMX two-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	3.01	6.79	49.00	7.68
Tue	NaN	100.00	0.01	13.53	0.01
Wed	NaN	NaN	100.00	1.14	95.97
Thu	NaN	NaN	NaN	100.00	1.35
Fri	NaN	NaN	NaN	NaN	100.00

Table 8.5: The p-values of the Brown-Forsythe test (10% trimmed mean) for OMX two-day returns

variances are all the same for all weekdays is rejected and the p-values for Tuesday and Friday two-day returns are very close to zero in all tests, so one can strongly believe that they have a very different variance. Moreover, the rejection of the null hypothesis for two consecutive weekday return, e.g. Tuesday and Wednesday return, maybe due to they both have one common day inside as two days holding period is considered here.

8.3 Time series model

8.3.1 AR model

From the autocorrelation graph Fig. 8.2, we can observe that there exists autocorrelation in the return data. Again, a mean equation as an autoregressive model with dummy variables for different weekdays is set up to examine for the day-of-the-week effect.

$$R_t = \alpha_0 + \alpha_1 D_T + \alpha_2 D_W + \alpha_3 D_H + \alpha_4 D_F + \sum_{i=1}^{k} \beta_i R_{t-i}, \tag{8.1}$$

where D_T, D_W, D_H and D_F are dummy variables for Tuesday, Wednesday, Thursday and Friday respectively.

From Tab. 8.6, order 8 has the smallest AIC value -21 684 with $PACF$ value -0.068. Therefore, $AR(8)$ model will be appropriate. Then the built-in function "regstats" in Matlab is used to get all the regression statistics.

Order	1	2	3	4	5	6
AIC [10^2]	-210.70	-214.58	-215.81	-216.65	-216.65	-216.69
$PACF$ [10^{-2}]	45.46	-36.50	21.60	-18.47	5.66	-6.48

Order	7	8	9	12	18	19
AIC [10^2]	-216.81	-216.84	-216.78	-216.63	-216.26	-216.16
$PACF$ [10^{-2}]	9.26	-6.80	2.78	-1.33	-7.84	-1.37

Table 8.6: The values of AIC and $PACF$ of the OMX two-day returns with different orders

parameter	α_0	α_1	α_2	α_3	α_4
value [10^{-4}]	0.51	-3.06	2.41	6.37	-3.60
se [10^{-4}]	7.89	11.15	11.10	11.08	11.15
t-statistic	0.07	-0.27	0.22	0.57	-0.32
p-value(%)	94.81	78.35	82.81	56.55	74.70

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8
value [10^{-2}]	76.79	-64.69	44.38	-32.74	19.20	-17.83	14.39	-6.73
se [10^{-2}]	1.92	2.41	2.69	2.80	2.80	2.69	2.41	1.92
t-statistic	39.96	-26.85	16.49	-11.69	6.85	-6.62	5.98	-3.51
p-value(%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05

Table 8.7: Summary Statistics of regression for OMX two-day return $AR(8)$ model by using Matlab

From Tab. 8.7, all the p-values of the coefficient of those dummies are much larger than 5% which means the null hypothesis that the coefficients of all the dummies are equal to zero cannot be rejected. Therefore, it seems that there are no day-of-the-week effect in the mean equation.

8.3.2 $GARCH(1, 1)$

The Lagrange multiplier test of Engle [8] is performed and it rejects the null hypothesis that no $ARCH$ effect in the residual of $AR(8)$ at the 5% significance level. It indicates that there is time varying conditional heteroscedasticity in the two-day return data. Therefore, a $GARCH(1, 1)$ model with the mean equation Eq. 8.2 is used to examine if there is any day-of-the-week effect in two-day return data.

$$R_t = \alpha_0 + \alpha_1 D_T + \alpha_2 D_W + \alpha_3 D_H + \alpha_4 D_F + \sum_{i=1}^{8} \beta_i R_{t-i} + \epsilon_t \tag{8.2}$$

$$\epsilon_t = \sigma_t z_t \ , \ \sigma_t^2 = a_0 + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2 \tag{8.3}$$

where D_T, D_W, D_H and D_F are dummy variables for Tuesday, Wednesday, Thursday and Friday respectively. $\{z_t\}$ is assumed to be a sequence of $i.i.d$ random variables with standard Normal distribution.

parameter	value $[10^{-3}]$	se $[10^{-3}]$	t-statistic	p-value(%)
α_0	1.55	0.54	2.84	0.46
α_1	-1.59	0.79	-2.02	4.37
α_2	-0.25	0.79	-0.31	75.35
α_3	-0.28	0.77	-0.37	71.22
α_4	-0.38	0.78	-0.49	62.57
β_1	741.70	20.13	36.84	0.00
β_2	-602.35	25.67	-23.47	0.00
β_3	419.49	28.44	14.75	0.00
β_4	-316.41	29.17	-10.85	0.00
β_5	197.79	30.23	6.54	0.00
β_6	-165.06	28.17	-5.86	0.00
β_7	110.80	25.52	4.34	0.00
β_8	-65.57	19.91	-3.29	0.10
a_0	0.00	0.00	4.87	0.00
a_1	883.64	10.47	84.39	0.00
b_1	107.20	10.37	10.34	0.00

Table 8.8: Summary Statistics of GARCH(1,1) for OMX two-day returns by using Matlab

The built in function "garchfit" in Matlab is used and then get the statistics in Tab. 8.8. When 5% significance level is used, the null hypothesis of α_1 will be rejected. It may indicate that there is day-of-the-week effect on Tuesday for the two-day returns of OMX. However, the p-value of the test on Tuesday two-day return is 4.37% which is close to 5%, so the result of existing Tuesday effect is not particularly strong.

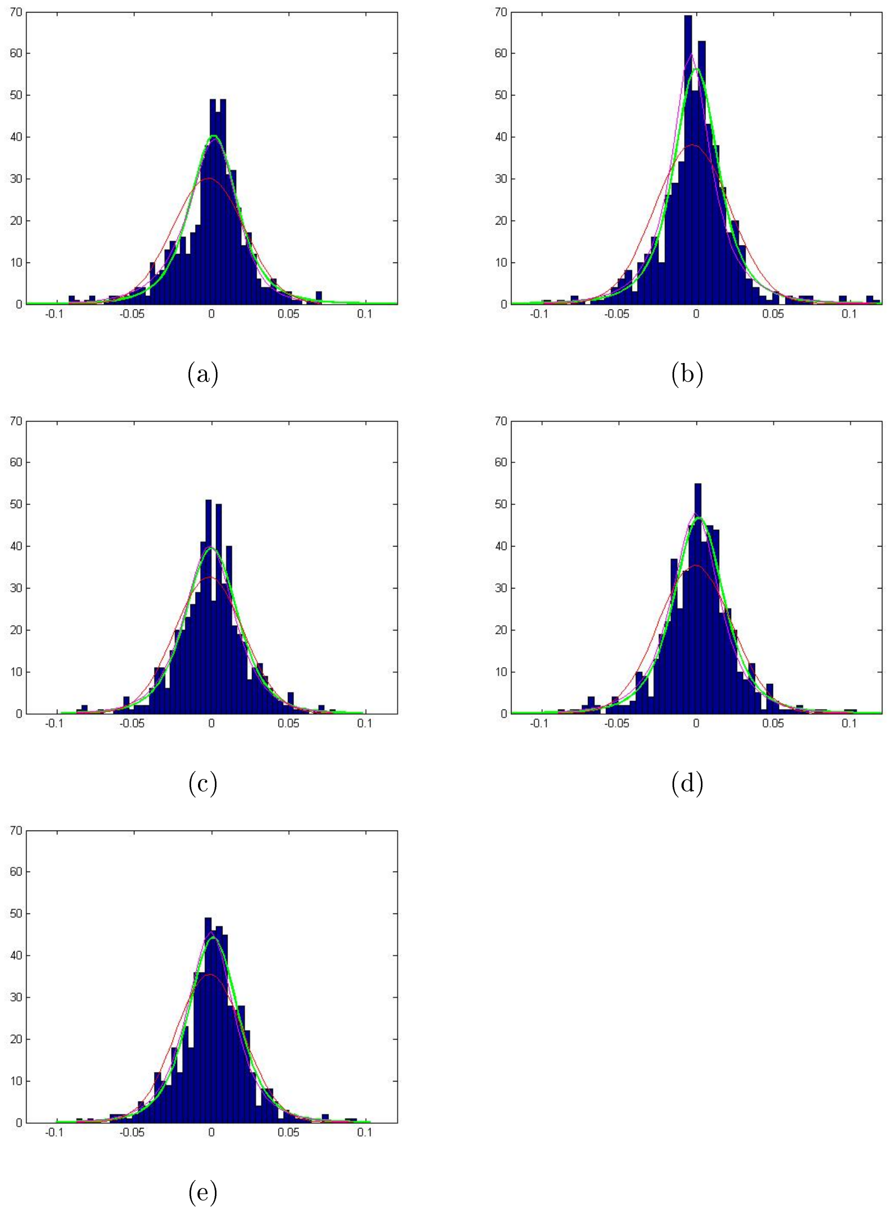
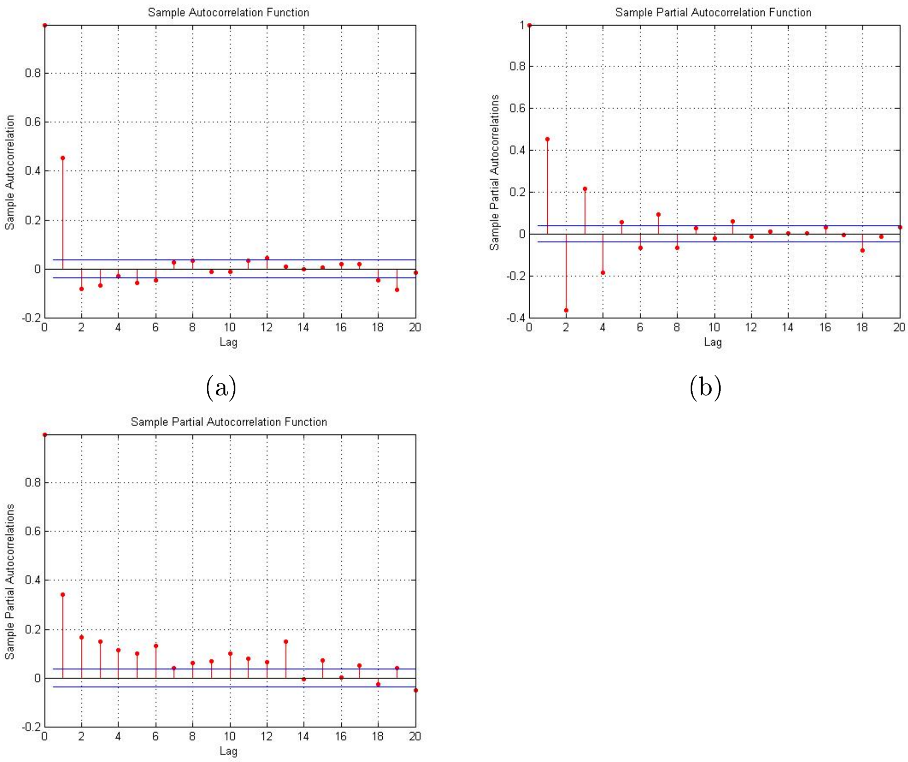


Figure 8.1: Histograms of empirical distribution with the density for the fitted Normal distribution (red line), location-scale Student's t-distribution (green line) and third-degree polynomial of standard normal model (magenta line) for (a) Monday (b) Tuesday (c) Wednesday (d) Thursday (e) Friday two-day returns of OMX index



(a) (b) (c)

Figure 8.2: (a) Sample autocorrelation function of the OMX two-day return series, (b) Sample partial autocorrelation function of the OMX two-day return series (c) Sample partial autocorrelation function of the OMX squared two-day returns

Chapter 9

Further study in day-of-the-week effect in portfolio

In the financial world, instead of investing in just one index, some will prefer to invest in a portfolio. The investigation is extended to find out if there are any day-of-the-week effects in a portfolio. There are uncountable numbers of combinations of portfolio. Here, a simple portfolio is used as an example to show how the tests and models can be applied in the case of using portfolio.

9.1 One-day return portfolio

A simple portfolio is set up by investing three indexes : OMX, SAX and SX5E with equal weight on each index. These indexes are in the same timezone so the time difference problem will not be considered when we are doing the analysis. Also, one day holding period is chosen to be investigated.

Then the one-day log return of the portfolio will be

$$R_t = \frac{1}{3}\left(\log \frac{P_{OMX,t+1}}{P_{OMX,t}} + \log \frac{P_{SAX,t+1}}{P_{SAX,t}} + \log \frac{P_{SX5E,t+1}}{P_{SX5E,t}}\right)$$

where $P_{OMX,t}$, $P_{SAX,t}$ and $P_{SX5E,t}$ are the prices of OMX, SAX and SX5E at time t respectively.

Tab. 9.1 shows some summary statistics for one-day returns of the portfolio. Tuesday and Friday have a negative mean returns while the others have positive mean returns and Thursday is the highest. The results also indicate that Monday return has higher variance. Moreover, by considering the kurtosis which is far from 3 and the histogram graphs of the weekdays in (Fig. 9.1), one can observe that all weekdays, especially on Monday, have relatively high peaks and heavy tails. Similar to those investigation in the previous sessions, the histogram graphs show again that their distributions are

	Mean [10^{-4}]	Variance [10^{-4}]	Kurtosis	Skewness
All	0.52	2.37	6.75	0.06
Mon	0.45	3.15	8.47	0.41
Tue	-4.07	2.18	6.35	-0.03
Wed	2.82	2.29	4.73	-0.18
Thu	4.89	2.33	4.70	-0.15
Fri	-1.65	1.91	7.32	0.01

Table 9.1: Summary statistics for one-day returns of the portfolio

far from normally distributed and have a better fit with location-scale Student's t-distribution and third-degree polynomial of standard normal model.

9.2 Test for homogeneity in variance

The F-test, Levene's test and two types of Brown-Forsythe tests are used to test the homogeneity in variance of the one-day return data among different weekdays. Tab. 9.2 – 9.5 are the p-values resulted from those tests with the null hypothesis of having the same variance. When 5% significance level is used, the null hypothesis for Monday with all the other weekdays can be rejected in all tests. Also, the p-values of its are all almost or equal to zero. It means all the tests support that Monday has a very different variance than the others. Moreover, this result is very similar to the result in OMX one-day return, but the null hypothesis for Thursday and Friday is not rejected here in Levene's test and Brown-Forsythe test.

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.00	0.02	0.05	0.00
Tue	NaN	100.00	58.07	42.32	12.12
Wed	NaN	NaN	100.00	80.09	3.52
Thu	NaN	NaN	NaN	100.00	1.89
Fri	NaN	NaN	NaN	NaN	100.00

Table 9.2: The p-values of the F-test for the portfolio's one-day returns

9.3 Time series model

9.3.1 AR model

The autocorrelation graph Fig. 9.2 shows that there exists autocorrelation in the return data which indicates an autoregressive model is needed. Therefore, a mean equation as an AR model with dummy variables for different

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.00	0.00	0.00	0.00
Tue	NaN	100.00	81.31	97.75	19.22
Wed	NaN	NaN	100.00	83.46	28.09
Thu	NaN	NaN	NaN	100.00	20.05
Fri	NaN	NaN	NaN	NaN	100.00

Table 9.3: The p-values of the Levene's test for the portfolio's one-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.00	0.00	0.00	0.00
Tue	NaN	100.00	82.55	97.63	19.56
Wed	NaN	NaN	100.00	80.17	27.83
Thu	NaN	NaN	NaN	100.00	18.43
Fri	NaN	NaN	NaN	NaN	100.00

Table 9.4: The p-values of the Brown-Forsythe test (median) for the portfolio's one-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.00	0.00	0.00	0.00
Tue	NaN	100.00	81.56	98.29	19.39
Wed	NaN	NaN	100.00	83.18	28.17
Thu	NaN	NaN	NaN	100.00	19.99
Fri	NaN	NaN	NaN	NaN	100.00

Table 9.5: The p-values of the Brown-Forsythe test (10% trimmed mean) for the portfolio's one-day returns

weekdays (Eq. 6.2) is set up to examine for the day-of-the-week effect.

From Tab. 9.6, order 3 has the smallest AIC value -22 853 and its $PACF$ value is much out of the confidence bounds. Therefore, an $AR(3)$ model will be an appropriate choice. Then the built in function "regstats" in Matlab is used to get all the regression statistics.

From Tab. 9.7, all the p-values of the coefficients of the dummies are much larger than 5% which means the null hypothesis that the coefficient equal to zero in all the dummies cannot be rejected. Therefore, it seems that there is no day-of-the-week effect in the mean equation.

Order	1	2	3	4	5
AIC $[10^2]$	-228.52	-228.50	-228.53	-228.47	-228.49
$PACF$ $[10^{-2}]$	-0.21	-4.45	-6.83	4.64	-6.46

Order	6	7	8	9	10
AIC $[10^2]$	-228.44	-228.39	-228.30	-228.19	-228.13
$PACF$ $[10^{-2}]$	-4.72	4.30	1.49	-1.61	-3.00

Table 9.6: The values of AIC and $PACF$ of the portfolio's one-day returns with different orders

parameter	α_0	α_1	α_2	α_3	α_4
value $[10^{-4}]$	1.10	-4.78	2.08	2.95	-3.35
se $[10^{-4}]$	6.64	9.32	9.26	9.29	9.40
t-statistic	0.17	-0.51	0.22	0.32	-0.36
p-value(%)	86.79	60.80	82.20	75.09	72.11

parameter	β_1	β_2	β_3
value $[10^{-4}]$	-45.59	-437.82	-680.47
se $[10^{-4}]$	191.01	190.72	190.79
t-statistic	-0.24	-2.30	-3.57
p-value(%)	81.14	2.18	0.04

Table 9.7: Summary Statistic of regression for the portfolio's one-day return $AR(3)$ model by using Matlab.

9.3.2 $GARCH(1,1)$

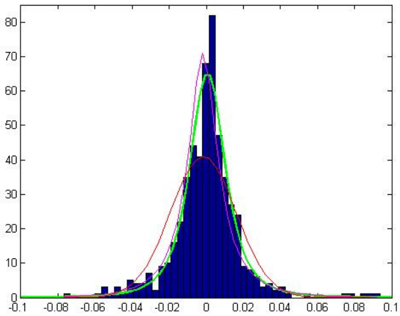
The Lagrange multiplier test of Engle [8] is performed and it rejects the null hypothesis that no $ARCH$ effect in the residual of $AR(3)$ at the 5% significance level. It indicates that there is time varying conditional heteroscedasticity in the one-day return data. Therefore, a $GARCH(1,1)$ model with the mean equation Eq. 6.3 and volatility equation Eq. 6.4 is used to examine if there is any day-of-the-week effect in the portfolio's one-day return data.

The built in function "garchfit" in Matlab is used and get the statistics in Tab. 9.8. When 5% significance level is used, the null hypothesis for the coefficient of Tuesday dummy equal to zero can be rejected while the other dummies cannot. It means that there may exist Tuesday effect for the one-day return in the portfolio. However, it is not consistent with the result in the AR model which accept the null hypothesis for all dummies. The $ARCH$ effect seems having a large impact here. Comparing with the result in the one-day return of OMX which shows that there exists Tuesday effect in the

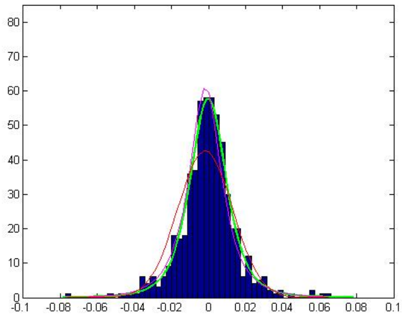
parameter	value $[10^{-3}]$	se $[10^{-3}]$	t-statistic	p-value(%)
α_0	1.09	0.47	2.35	1.89
α_1	-1.33	0.65	-2.05	4.00
α_2	0.03	0.63	0.05	96.40
α_3	-0.15	0.63	-0.25	80.59
α_4	-0.13	0.64	-0.20	84.43
β_1	-16.25	21.42	-0.76	44.80
β_2	-23.52	19.84	-1.19	23.59
β_3	-41.09	19.62	-2.09	3.63
a_0	0.00	0.00	4.64	0.00
a_1	895.36	9.27	96.55	0.00
b_1	97.71	8.99	10.87	0.00

Table 9.8: Summary Statistics of $GARCH(1,1)$ for the portfolio's one-day returns by using Matlab

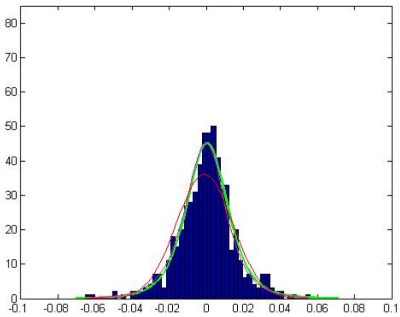
$GARCH(1,1)$ model, the Tuesday effect still exists when a portfolio where there are on two Swedish and one European indexes with equal weight.



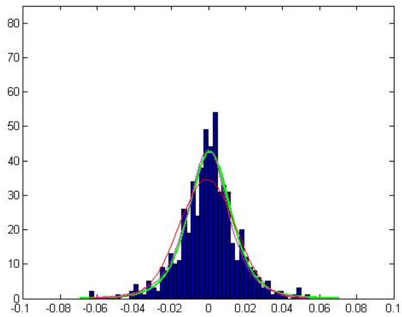
(a)



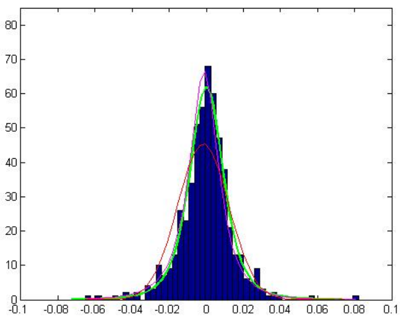
(b)



(c)

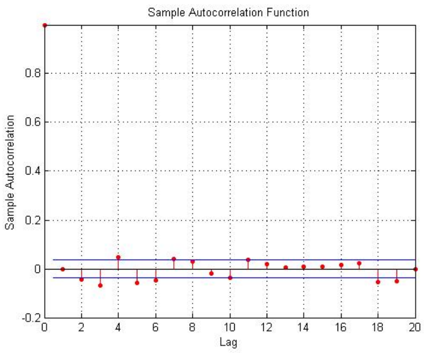


(d)

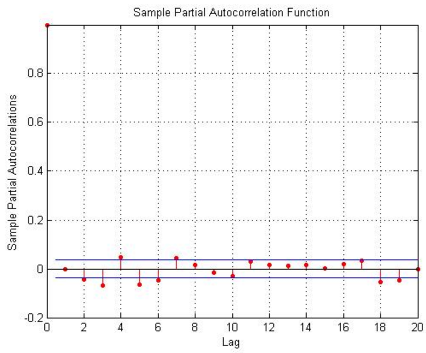


(e)

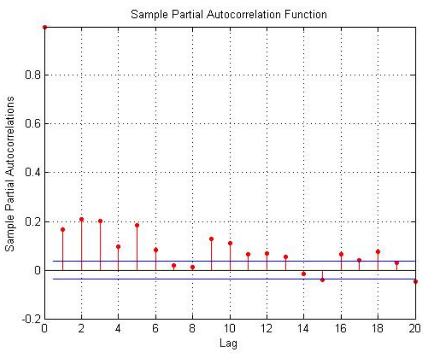
Figure 9.1: Histograms of empirical distribution with the density for the fitted Normal distribution (red line), location-scale Student's t-distribution (green line) and third-degree polynomial of standard normal model (magenta line) for (a) Monday (b) Tuesday (c) Wednesday (d) Thursday (e) Friday one-day returns of the portfolio



(a)



(b)



(c)

Figure 9.2: (a) Sample autocorrelation function of the portfolio's one-day return series, (b) Sample partial autocorrelation function of the portfolio's one-day return series (c) Sample partial autocorrelation function of the portfolio's squared one-day returns

Chapter 10

Conclusion

This thesis mainly study the day-of-the-week effects on the one-day return of the Swedish index OMX. The other countries' indexes are also being examined and the special findings are presented in Appendix.

To investigate the day-of-the-week effect among the stock index, two main quantities of the data are studied: mean and variance. To test the difference of variance, other than the commonly used F-test, three tests: Levene's test and Brown-Forsythe I and II tests are also presented in Chapter 4. And the comparison in Chapter 4 shows that in most cases when the data are non-normal, or centralized, the Levene's test and Brown-Forsythe I and II tests perform better than F-test no matter for type I or type II error.

In Chapter 6, by fitting OMX daily returns with some parametric models, return data are found to be non-normal with heavy tails, especially on Monday. Also, they are fitted well with Student's t-distribution and third-degree polynomial. Since the data showed non-normal, Levene's test and Brown-Forsythe tests are used along with F-test in Section 6.4 to test the homogeneity in variance of one-day returns. The results are consistent at 5% significance level that OMX Monday return variance is different from other weekdays, while Friday and Thursday return variances are considered different in Levene's test and Brown-Forsythe tests with 10% trimmed mean but not in other two tests. Results for other stock indexes, except NKY(see appendix), they show similar Monday effect in variance.

Other than variance, the difference of mean is also tested. Section 6.5 presents two time series models: AR and $GARCH(1,1)$. AR model is simpler and it shows that there are no mean difference in OMX one-day return, while $GARCH(1,1)$ shows significant Tuesday effect. As the $ARCH$ effect is proved existing in the return data, $GARCH(1,1)$ is more reliable than AR. The difference of the results also indicates that $ARCH$ effects make a large impact in the data. There are two ways for the order determination used in the time series models, AIC and $PACF$. The comparison part in Chapter 5 shows that when the sample size is small, $PACF$ is preferred, while AIC is more reliable when there is a large sample size in general.

Chapter 7 examines data in a different way that whether daily returns are affected by the previous week performance. Since the data sizes and variance are different, Welch's t-test as a modified t-test is introduced and used to test the means. And the results show that Tuesday return is negative in 5% significance level when the previous week return is positive, meaning that some daily return can be affected by the weekly return, indicating the correlation between the data.

All the methods used in Chapter 6 and Chapter 7 can be applied to other stock indexes, as stated in the appendix, as well as other investment horizons, like holding two-day instead of one-day (Chapter 8), or investing in a portfolio instead of single stock (Chapter 9). Chapter 8 shows that in two-day return analysis, the variances between two consecutive weekdays' two-day return are different at 5% significance level. It may due to the mutual day in them. Besides, Tuesday and Friday also show significance difference in variance. Also, there is no mean difference exists in AR model while a weak Tuesday effect exists in $GARCH(1,1)$. Comparing with one-day return, the data of two-day return for different weekdays are more similar to each other and there is obvious correlation within the data. In Chapter 9, a selected portfolio (OMX, SAX and SX5E) is analyzed. Monday one-day return still shows difference in variances as in the single stock index and the Tuesday effect in means still exists in the $GARCH(1,1)$ model for the portfolio.

This thesis not only study the day-of-the-week effect, but also presents methods that can be used in further studies. There are further problems which can be investigated, like the correlation within the data and between indexes, and how it affects the variance and mean. Also, one can check whether the difference in time zones will affect the portfolio performance. Besides, the weekly returns and yearly returns are all worth looking into.

Bibliography

[1] Albers W, Boon PC, Kallenberg WCM. *The asymptotic behavior of tests for normal means based on a variance pre-test*. Journal of Statistical Planning and Inference. 2000 7/1/;88(1):47-57.

[2] Al-Loughani N, Chappell D. *Modelling the day-of-the-week effect in the Kuwait Stock Exchange: a nonlinear GARCH representation*. Applied Financial Economics. 2001;11(4):353-9.

[3] Bollerslev T. *Generalized autoregressive conditional heteroskedasticity*. Journal of Econometrics. 1986;31(3):307-27.

[4] Brown MB, Forsythe AB. *Robust Tests for the Equality of Variances*. Journal of the American Statistical Association. 1974;69(346):364-7.

[5] Burnham KP, Burnham KP, Anderson DR. *Model selection and multimodel inference : a practical information-theoretic approach*. 2nd ed ed. New York, London: Springer; 2002. xxvi, 488 p. p.

[6] Campbell JY, Lo AW, MacKinlay AC. *The econometrics of financial markets*. Princeton, NJ: Princeton University Press; 2007. 611 S. p.

[7] Cross F. *The Behavior of Stock Prices on Fridays and Mondays*. Financial Analysts Journal. 1973;29(6):67-9.

[8] Engle RF. *Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation*. Econometrica. 1982;50(4):987-1007.

[9] French KR. *Stock returns and the weekend effect*. Journal of Financial Economics. 1980;8(1):55-69.

[10] Gibbons MR, Hess P. *Day of the Week Effects and Asset Returns*. The Journal of Business. 1981;54(4).

[11] Harris L. *A transaction data study of weekly and intradaily patterns in stock returns*. Journal of Financial Economics. 1986;16(1):99-117.

[12] Hult H, Lindskog F, Hammarlid O, Rehn CJ. *Risk and Portfolio Analysis.* 2012.

[13] Jaffe JF, Westerfield R. *The Week-End Effect in Common Stock Returns: The International Evidence.* Journal of Finance. 1985;40(2):433.

[14] Keim DB, Stambaugh RF. *A Further Investigation of the Weekend Effect in Stock Returns.* Journal of Finance. 1984;39(3):819.

[15] Lakonishok J, Maberly E. *The Weekend Effect: Trading Patterns of Individual and Institutional Investors.* Journal of Finance. 1990;45(1):231.

[16] Levene H.*Robust tests for equality of variances.* Stanford University Press, pp.278-292

[17] Tsay RS. *Analysis of Financial Time Series* John Wiley and Sons; 2005.

[18] Welch BL. *The Generalization of 'Student's' Problem when Several Different Population Variances are Involved.* Biometrika Trust. 1947;34:28-35.

[19] Yamori N, Kurihara Y. *The day-of-the-week effect in foreign exchange markets: multi-currency evidence.* Research in International Business and Finance. 2004;18(1):51-7.

[20] Yang Y-L, Chang C-L. *A double-threshold GARCH model of stock market and currency shocks on stock returns.* Mathematics and Computers in Simulation. 2008 12/1/;79(3):458-74.

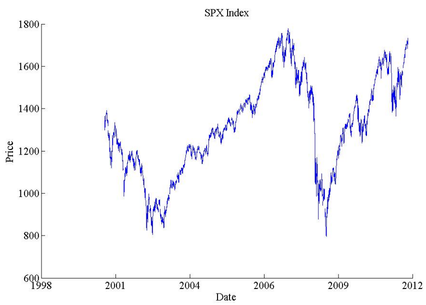
[21] Zimmerman DW. *Conditional Probabilities of Rejecting H_0 by Pooled and Separate-Variances t Tests Given Heterogeneity of Sample Variances.* Communications in Statistics - Simulation and Computation. 2004 2004/01/02;33(1):69-81.

Chapter 11

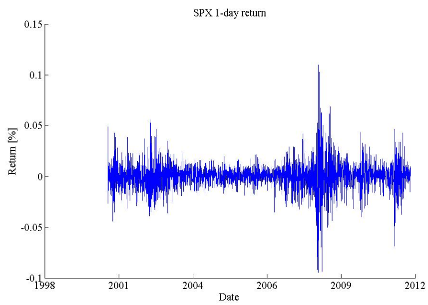
Appendix

This appendix will provide some special finding from United States index, SPX (Standard and Poor's 500 Index), and a Japan index, NKY (Nikkei-225 Stock Average).

11.1 Standard and Poor's 500 Index (SPX Index)



(a)



(b)

Figure 11.1: Time plots of (a) price for SPX index (b) one-day log returns for SPX index

11.1.1 Test for Homogeneity in variance

From Tab 11.1 – 11.4, we can observe that the null hypothesis for both Monday and Friday have the same variance as all the other weekday returns are rejected in all tests at 5% significance level. As this result is consistent in all tests and the p-values is very close to or equal to zero, we can strongly believe that they have a very different variances.

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.93	0.11	0.36	0.00
Tue	NaN	100.00	55.30	79.60	0.00
Wed	NaN	NaN	100.00	73.37	0.00
Thu	NaN	NaN	NaN	100.00	0.00
Fri	NaN	NaN	NaN	NaN	100.00

Table 11.1: p-values of the F-test for SPX one-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.01	0.00	0.00	0.00
Tue	NaN	100.00	52.77	71.42	0.00
Wed	NaN	NaN	100.00	78.81	0.00
Thu	NaN	NaN	NaN	100.00	0.00
Fri	NaN	NaN	NaN	NaN	100.00

Table 11.2: p-values of the Levene's Test for SPX one-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.01	0.00	0.00	0.00
Tue	NaN	100.00	58.33	86.45	0.00
Wed	NaN	NaN	100.00	70.06	0.00
Thu	NaN	NaN	NaN	100.00	0.00
Fri	NaN	NaN	NaN	NaN	100.00

Table 11.3: p-values of the Brown-Forsythe Test (median) for SPX one-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.01	0.00	0.00	0.00
Tue	NaN	100.00	53.14	74.40	0.00
Wed	NaN	NaN	100.00	76.14	0.00
Thu	NaN	NaN	NaN	100.00	0.00
Fri	NaN	NaN	NaN	NaN	100.00

Table 11.4: p-values of the Brown-Forsythe Test (10% trimmed mean) for SPX one-day returns

11.1.2 Time Series Model

As there exists autocorrelation in the one-day return data of SPX, a $AR(7)$ model with dummy variables for different weekdays is chosen and the regression is carried out and get Tab. 11.6. When 5% significance level is used, the

Order	1	2	3	4	5	6	7	8	9	10	11	12
AIC $[10^2]$	-233.05	-233.04	-232.96	-232.87	-232.76	-232.65	-232.64	-232.61	-232.49	-232.39	-232.29	-232.22
$PACF$ $[10^{-2}]$	-9.67	-6.28	1.23	-2.85	-1.90	0.13	-6.04	5.51	0.87	-0.99	2.53	4.07

Table 11.5: The values of AIC and $PACF$ of SPX one-day returns with different orders

parameter	α_0	α_1	α_2	α_3	α_4	β_1	β_2	β_3	β_4	β_5	β_6	β_7
value $[10^{-4}]$	-2.54	9.19	6.47	6.39	-2.90	-1025.42	-614.39	81.34	-293.37	-244.21	-64.36	-598.00
se $[10^{-4}]$	6.03	8.48	8.30	8.33	8.38	192.21	193.21	193.51	193.49	193.44	193.16	191.73
t-statistic	-0.42	1.08	0.78	0.77	-0.35	-5.33	-3.18	0.42	-1.52	-1.26	-0.33	-3.12
p-value(%)	67.38	27.85	43.57	44.31	72.94	0.00	0.15	67.43	12.96	20.69	73.90	0.18

Table 11.6: Summary Statistics of regression of SPX one-day returns by using Matlab

null hypothesis for all the coefficients in the dummies are accepted. Therefore, it seems that there is no day-of-the-week effect in the mean equation. Then, as we check that there is $ARCH$ effect in the series, $GARCH(1,1)$ is then used and get Tab.11.7. The result of $GARCH$ model is consistent with the AR model that no any day-of-the-week effect in the mean equation since the null hypothesis for all dummies are also cannot be rejected here. As we also observe that the null hypothesis for the coefficients for lag-2 to lag-7 ($\beta_2 - \beta_7$) can all be accepted under 5% significance level, we improve our model to a $GARCH(1,1)$ model with a AR(1) mean equation. Then the p-values for the dummies just change a little bit, but the result of no day-of-the-week effect in the mean equation is still can be concluded from this model.

parameter	value [10^{-4}]	se [10^{-4}]	t-statistic	p-value(%)
α_0	9.14	4.19	2.18	2.94
α_1	-4.60	5.49	-0.84	40.26
α_2	2.79	5.57	0.50	61.68
α_3	-3.19	5.69	-0.56	57.53
α_4	-9.07	5.66	-1.60	10.94
β_1	-652.12	227.69	-2.86	0.42
β_2	-389.99	200.35	-1.95	5.17
β_3	-55.25	206.56	-0.27	78.91
β_4	-266.72	192.01	-1.39	16.49
β_5	-329.12	201.58	-1.63	10.26
β_6	-271.24	195.53	-1.39	16.55
β_7	-169.17	186.78	-0.91	36.52
a_0	0.02	0.00	5.89	0.00
a_1	8967.60	90.39	99.21	0.00
b_1	930.88	84.30	11.04	0.00

Table 11.7: Summary Statistics of $GARCH(1,1)$ model for SPX one-day returns by using Matlab

11.2 Nikkei-225 Stock Average (NKY Index)

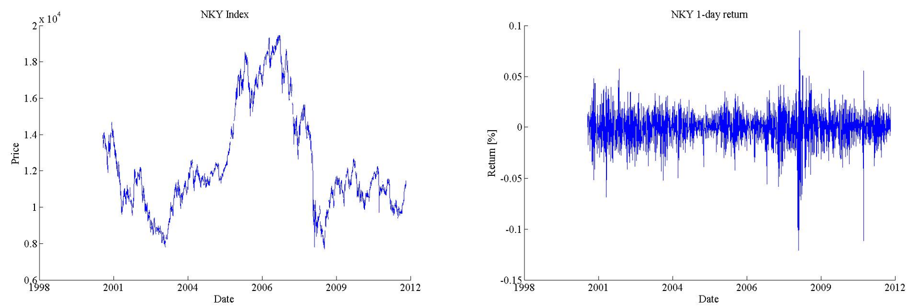
11.2.1 Test for Homogeneity in variance

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	0.58	7.91	71.71	12.45
Tue	NaN	100.00	28.05	1.35	19.48
Wed	NaN	NaN	100.00	15.16	82.23
Thu	NaN	NaN	NaN	100.00	22.70
Fri	NaN	NaN	NaN	NaN	100.00

Table 11.8: p-values of the F-test for NKY one-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	29.82	58.33	26.50	56.58
Tue	NaN	100.00	60.29	2.89	62.16
Wed	NaN	NaN	100.00	8.65	97.86
Thu	NaN	NaN	NaN	100.00	8.19
Fri	NaN	NaN	NaN	NaN	100.00

Table 11.9: p-values of the Levene's Test for NKY one-day returns



(a) (b)

Figure 11.2: Time plots of (a) price for NKY index (b) one-day log returns for NKY index

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	22.77	49.00	36.33	47.90
Tue	NaN	100.00	58.30	3.16	59.59
Wed	NaN	NaN	100.00	9.94	98.51
Thu	NaN	NaN	NaN	100.00	9.59
Fri	NaN	NaN	NaN	NaN	100.00

Table 11.10: p-values of the Brown-Forsythe Test (median) for NKY one-day returns

%	Mon	Tue	Wed	Thur	Fri
Mon	100.00	28.04	54.54	29.44	56.57
Tue	NaN	100.00	61.42	3.09	59.37
Wed	NaN	NaN	100.00	8.84	97.56
Thu	NaN	NaN	NaN	100.00	9.46
Fri	NaN	NaN	NaN	NaN	100.00

Table 11.11: p-values of the Brown-Forsythe Test (10% trimmed mean) for NKY one-day returns

From Tab. 11.8 - 11.11, we can observe that NKY has a very different results with other indexes. The null hypothesis for Monday having a different variance with all the other weekdays' return is not rejected here. Even null hypothesis for Monday and Tuesday having different variance is rejected in F-test in 5% significance level but not in other three tests. As we already know that the return data are non-normally distributed, so this result from F-test is not so reliable here. Therefore, the result for Monday and Tuesday return in NKY having different variance is not so strong. As for the null hypothesis for Tuesday with Thursday, we get a consistent result from all tests that it can be rejected in 5% significance level. Therefore, we can believe that Tuesday return has a different variance as Thursday.

11.2.2 Time Series Model

Order	1	2	3	4	5	6
AIC [10^2]	-217.42	-217.32	-217.27	-217.18	-217.08	-217.03
$PACF$ [10^{-2}]	-2.00	1.34	-4.38	-2.78	1.89	4.90

	7	8	9	10	11	12
AIC [10^2]	-216.96	-216.86	-216.76	-216.66	-216.56	-216.47
$PACF$ [10^{-2}]	3.72	1.36	-1.14	2.37	2.09	-2.87

Table 11.12: The values of AIC and $PACF$ of NKY one-day returns with different orders

parameter	α_0	α_1	α_2	α_3	α_4
value [10^{-4}]	-12.32	1.47	11.89	19.63	8.05
se [10^{-4}]	7.08	9.97	9.74	9.73	9.73
t-statistic	-1.74	0.15	1.22	2.02	0.83
p-value(%)	8.17	88.26	22.20	4.37	40.83

β_1	β_2	β_3	β_4	β_5	β_6
-211.61	178.88	-406.93	-292.27	184.51	493.38
195.99	196.03	196.05	195.99	195.92	195.86
-1.08	0.91	-2.08	-1.49	0.94	2.52
28.04	36.16	3.80	13.60	34.64	1.18

Table 11.13: Summary Statistics of regression of NKY one-day returns by using Matlab

As there exists a little bit autocorrelation in the one-day return data of NKY, a $AR(6)$ model with dummy variables for different weekdays is chosen and the regression is carried out and get Tab. 11.13. When 5% significance level is used, the null hypothesis for the coefficient of Thursday (α_3) is rejected while the other weekdays cannot. However, if 4% or even lower significance level is used, then the null hypothesis for Thursday no longer can be rejected. Therefore, the Thursday effect in the mean equation is not so strong. Then, as we check that there is $ARCH$ effect in the series, $GARCH(1,1)$ is then used and get Tab.11.14. The result of $GARCH$ model is not consistent with the AR model. In the $GARCH$ model, it shows that the null hypothesis for all weekday dummies, even the Thursday dummy,

parameter	value $[10^{-4}]$	se $[10^{-4}]$	t-statistic	p-value(%)
α_0	0.39	5.01	0.08	93.79
α_1	-6.26	7.79	-0.80	42.15
α_2	3.64	7.10	0.51	60.84
α_3	10.69	7.14	1.50	13.46
α_4	2.95	6.84	0.43	66.58
β_1	-14.10	230.49	-0.06	95.12
β_2	-148.95	213.16	-0.70	48.48
β_3	8.34	202.64	0.04	96.72
β_4	-160.39	201.10	-0.80	42.52
β_5	160.36	199.94	0.80	42.26
β_6	39.24	201.06	0.20	84.53
a_0	0.04	0.01	4.36	0.00
a_1	8734.60	111.65	78.23	0.00
b_1	1118.20	93.04	12.02	0.00

Table 11.14: Summary Statistics of $GARCH(1,1)$ model for NKY one-day returns by using Matlab

cannot be rejected under 5% significance level. Therefore, it means that there is no any day-of-the-week effect in the mean equation. As we also observe that the null hypothesis for the coefficient for lag-1 to lag-6 ($\beta_1 - \beta_6$) can all be accepted under 5% significance level, we improve our model to a $GARCH(1,1)$ model with no autocorrelation in the mean equation. Then the p-values for the dummies just change a little bit, but the result of no day-of-the-week effect in the mean equation is still can be concluded from this model.

TRITA-MAT-E 2012:08
ISRN-KTH/MAT/E--12/08-SE