Chapter 3

Data and Methodology

3.1 Data

3.1.1 Selection of Data

Daily seasonality anomalies thus far are believed to have a more significant effect on the returns of smaller capital stocks than that of large capital stocks as concluded by previous studies regarding firm size. This study aims to determine if the daily seasonality effect is also significant for the large capital stocks on the KLSE. As such, thirty large capital stocks on the Main Board are randomly selected, based on the market capital of the stock on 31st December 1995. The thirty stocks selected are :-

1. Boustead	*2. Faber Group*	*3. Genting*
4. Golden Hope	*5. Guiness Anchor*	*6. High & Low*
7. Hong Leong Indusries	*8. Kian Joo*	*9. Kuala Lumpur Kepong*
10. Magnum	*11. MayBank*	*12. Malayan Cement*
13. MMC	*14. Malaysia Oxygen*	*15. Multi-purpose Holding*
16. MUI	*17. Nestle*	*18. NSTP*
19. Oriental	*20. Perlis Plantation*	*21. Renong*
22. Rothman	*23. Sime Darby*	*24. Sungei Way*
25. Shell	*26. Tan Chong*	*27. Telekom*
28. Tenaga	*29. Time*	*30. UMW Holding*

The study covers the period from 2nd January 1992 to 31st December 1995.

3.1.2 Sources of Data

The data used in this study were obtained from the following sources :-

a. The KLSE Daily Dairy, and

b. Investors Digest.

The closing prices of shares, capital changes due to rights and bonus issues were obtained based on the information published from the KLSE Daily Dairy.

The daily stock returns are computed by comparing the daily price changes as shown by the following formula :-

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

where R_t = return for day t;

P_t = daily closing price for day t;

P_{t-1} = daily closing price for previous day t-1.

If there are any capital changes, adjustment will have to be made to the closing prices by the following formulae :-

$$R_t = \frac{P_t - P'_{t-1}}{P'_{t-1}}$$

for bonus issue, $$P'_{t-1} = \frac{n_1}{n_1 + n_2} P_{t-1}$$

for right issue, $$P'_{t-1} = \frac{n_1 P_{t-1} + n_2 E}{n_1 + n_2}$$

where P'_{t-1} = adjusted daily closing price for previous day t-1.

n_1 = original number of shares required to be entitled to n_2 new shares.

E = subscription price.

From the daily return, R_t, calculated, the following can be carried out :-

- to test for any significant differences among mean returns of the five trading days in the week;
- to test the existence of the weekend effect, that is, low and negative return on Monday;
- to test if there is any significant differences in the mean returns across the 5 trading days.

3.2 Research Model of Stock Returns

The statistical model used in this study is Analysis of Variance Model based on fixed effects where the many returns of any stock is composed of a simple sum of overall mean, effect specific to the *day-of-the-week* and the random error associated with that effect.

$$R_{td} = \mu + \alpha_d + \varepsilon_{td} \tag{3.1}$$

where R_{td} = return in day d of week

μ = the grand mean

α_d = the effect specific to day d

ε_{td} = random error term associated with that day

It is assumed that the error term is normally and independently distributed with mean zero and a common variance.

3.3 Statistical Test

The main hypothesis to be tested under this study using Oneway ANOVA (F statistics) and t-test for independent samples is :-

$H_o : \mu_{d1} = \mu_{d2} = \mu_{d3} = \mu_{d4} = \mu_{d5}$

H_1 : at least two μ not equal

μ = mean return

d1 = Monday, d2 = Tuesday, d3 = Wednesday, d4 = Thursday, d5 = Friday

where μ_d is the effect specific to *day-of-the-week*. Rejection of the null hypothesis will suggest that the daily mean returns R_t are a result of the actual differences in the mean that exhibit seasonality according to the *day-of-the-week* effect and not due to chance occurrence.

The weekend effect will also be tested in the study. The hypothesis to be tested is as follow :-

$H_o : \mu_{d1} = \mu_d$ where d1 = Monday, and d = Tuesday to Friday

$H_1 : \mu_{d1} \neq \mu_d$

where μ_{d1} is the effect specific to Monday and μ_d is the average effect for the days of Tuesday to Friday. Rejection of the null hypothesis will indicate that there is a Monday effect.

There are two major types of statistical tests used in this study, the parametric test and the non-parametric test. Parametric tests assume that the populations are normally distributed. Non-parametric tests do not require the presence of nomality of the distribution. They require fewer and less stringent conditions for their use and are more flexible. The non-parametric tests are also known as distribution free tests.

3.3.1 Parametric Tests

Test for day-of-the-week effect

To determine the existence of the *day-of-the-week* effect, the equality of means of the various daily average returns is being tested. The null hypothesis for the test is :-

$$H_o : \mu_{d1} = \mu_{d2} = \mu_{d3} = \mu_{d4} = \mu_{d5}$$

where μ = mean return

d1 = Monday, d2 = Tuesday, d3 = Wednesday, d4 = Thursday, d5 = Friday

Rejection to the hypothesis will show the existence of day-of-the-week effect.

The Statistical technique used to test whether the daily returns have equal means is the F statistic of the Oneway Analysis of Variance (ANOVA), used in conjunction with the Tukey's test for further testing the significance of differences between means of paired groups.

Oneway ANOVA

The Oneway ANOVA makes comparison of the means of k, (where $k \geq 2$) independent samples based on the fixed effect situation as in Equation 3.1 above. It is used to test whether the probability that differences in means across several groups are due solely to sampling error. The Oneway ANOVA assumes normal populations and homogeneity of population variances.

In the Oneway ANOVA test, comparison of the Mean Square$_{\text{(between groups)}}$ to the Mean Square$_{\text{(within groups)}}$ is carried out to derive the F ratio. The resulting F-ratio is mathematically wriiten as :

$$F = \frac{MS_{(between.\ groups)}}{MS_{(within.\ groups)}}$$

The larger the F ratio, the higher the possibility to reject the null hypothesis of equal means.

Bartlett's Test

This is a statistical test used to test the homogeneity of variances when there are more than two groups. The test statistic is χ^2, which is to be compared against the critical value of chi square.

The null hypothesis of the test is :-

$$H_o : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2$$

$$H_1 : \text{ at least two } \sigma^2 \text{ are not equal}$$

The test statistic is χ^2 with 4 degrees of freedom.

The computation of χ^2 is as follows:-

$$\chi^2 = \frac{1}{c}\left(\left[\sum_{d=1}^{k}(n_d - 1)\right]\ln s^2 - \sum_{d=1}^{k}(n_d - 1)\ln s_d^2\right)$$

where (n_d-1) = degree of freedom of day d, d = 1,2,3,4,5

s_d^2 = variance of day d returns

s^2 = weighted average of variance, computed as :-

$$s^2 = \frac{\sum_{d-1}^{k}(n_d - 1)s_d^2}{\sum_{d=1}^{k}(n_d - 1)}$$

$$c = 1 + \frac{1}{3(k-1)}\left[\left(\sum_{d=1}^{k}\frac{1}{n_d - 1}\right) - \frac{1}{\sum_{d=1}^{k}(n_d - 1)}\right]$$

and $k = 5$

Tukey Test

Tukey test is used to test the significance of differences between the means of any paired groups as part of the unplanned comparison of means technique. A pair of means is considered to be significantly different at a = 0.05 if their differences is equal or greater than the critical difference MSD_{ij}, that is, if $|\overline{Y}_j - \overline{Y}_i| \geq MSD_{ij}$ (Sokal and Rohlf (1969)).

$$MSD_{ij} = Q_{\alpha(k,v)}\sqrt{\frac{MS_{within}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}{2}}$$

where	$Q_{\alpha(k,v)}$	=	critical value of studentized range
	n_i	=	sample size of group *i*
	n_j	=	sample size of group *j*
	a	=	0.05, significance level
	k	=	number of groups (in this case, k=5)

v = degree of freedom, $\sum(n_i - 1)$

3.3.2 Non-Parametric Test

The Kruskal-Wallis test is the non-parametric test used for testing the existence of *day-of-the-week* effect. This test is the non-parametric equivalent to the Oneway ANOVA F statistical test. It is a very useful test in determining k independent samples are from different populations. The test only assumes that the variable under study has an underlying continuous distribution.

In the computation, each of the N observations is replaced by rank. All the stock returns from all the five days are ranked in a series. The smallest return is replaced by rank 1, the next smallest by rank 2, and the largest by rank N. Then the sum of the ranks in each of the days is found. The test statistic is given by H,

$$H = \frac{12}{N(N+1)} \sum_{d=1}^{k} \frac{R_d^2}{n_d} - 3(N+1)$$

where k = 5, number of trading days d of the week

n_d = number of cases in day d of the week

N = $\sum n_d$, the number of cases in all days combined

R_d = sum of ranks in the day of the week

If H_o is true, and the sample sizes of the k samples are not too small, then the H statistic has approximately a chi-square, χ^2, distribution under the hypothesis with degree of freedom of k-1. Therefore, the decision rule is as follow :-

Reject H_o if $H > \chi^2\,(4,a)$

where χ^2 $(4,a)$ is the upper a percentile point of a chi-square distribution with 4 degrees of freedom. In this study, H_o is tested at 5 percent significance level.

When two or more scores in a sample have the same value, ties occur. In this case, each score is given the mean of the ranks for which it is tied. Hence, correction must be made for tied observations in computing H. The H statistic corrected for tied observations is :-

$$H = \frac{\dfrac{12}{N(N+1)}\sum_{d=1}^{k}\dfrac{R_d^2}{n_d} - 3(N+1)}{1 - \dfrac{\sum T}{N^3 - N}}$$

where $T = t^3 - t$ (when t is the number of tied observations in a tied group of scores)

$\sum T$ = summation of all groups of ties.