Chapter 3

Data and Methodology

3.1 Data

3.1.1 Selection of Data

In view of the short history of the KLSE Second Board (it was only launched on 2 January 1991), therefore the data used in this study is limited by the number of companies that were among the first to be listed on the Second Board so that sufficient time series data are available for analysis. With this limitation, only 31 companies' daily transacted closing prices which were among the first to be listed on the KLSE Second Board since its launched were selected. Of these 31 companies, 2 companies namely, Polypulp Paper Industries and Corrugated Carton Paper were promoted to the Main Board on 8 August 1995 and 11 August 1995 respectively. The price data for these 2 companies utilised in this study was therefore confined to the period of their listing on the Second Board. Besides these 31 companies, the Second Board Index and KLSE Composite Index data for the same period were also collected for the study. The study covered the period from 2 January 1992 to 15 September 1995 with a total of 915 trading days. The list of selected stocks is given in Appendix I. The market capitalisation of the 31 stocks as at 15 September 1995 is RM6,054.5 million.

Besides the whole period of study which was from 2 January 1992 to 15 September 1995, the study also sub-divided the study period into two sub-periods, the first sub-period from 2

January 1992 to 31 December 1993 and the second sub-period from 3 January 1994 to 15 September 1995.

3.1.2 Sources of Data

The data for this study were obtained from the following sources:

a) The KLSE Daily Dairy.

b) Investors Digest

Adjustments for capital changes (due to Rights and Bonus Issues) and dividends were made based on information obtained from the Daily Dairy.

The daily stock return in this study is computed using the capital price change measure below:

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

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.

As stated earlier, the objective of this study is to:-

1. test the equality of mean returns for the days in the week. Rejection of the null hypothesis would indicate day-of-the-week effect:
2. test the existence of the weekend effect, i.e. low (and negative) return for Monday;
3. test whether there is significant difference in the mean returns across the 5 trading days.

3.2 Research Model of Stock Returns

The statistical model of stock returns assumed for this study is the Linear Model in Fixed Effects situation. This is called the Linear Model because the returns of any stock is composed of a simple sum of overall mean or grand 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 so that

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

It is also assumed that the error term is normally and independently distributed with mean zero and a common (equal) 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_0 : $\mu_{d1} = \mu_{d2} = \mu_{d3} = \mu_{d4} = \mu_{d5}$ against;

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 stock returns R_t exhibit seasonality according to the day of the week and not due to chance occurrence.

Besides testing for seasonality, the study will also test for a weekend effect, or a Monday effect. Thus, the hypothesis to be tested is as follows:

$$H_0 : \mu_1 = \mu_d \qquad d = \text{Tuesday to Friday}$$

$$H_1 : \mu_1 \neq \mu_d$$

where u_1 is the effect specific to Monday and u_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. Two types of statistical tests will be employed in this study, namely the parametric test and the non-parametric test. Parametric tests make a number of assumptions about the parameters e.g. normal population distribution, homogeneity of variances, etc while the non-parametric tests require fewer and less stringent conditions for their use (Siegel (1956)). The non-parametric tests are also known as *distribution free* tests.

3.3.1 Parametric Tests

Test for Day-of-the-Week Effect

The method used in this study to determine the existence of the day-of-the-week effect is to test the equality of means of the various daily average returns. The null hypothesis for the test is:

$$H_0 : \mu_{d1} = \mu_{d2} = \mu_{d3} = \mu_{d4} = \mu_{d5}$$

Rejection of the hypothesis will indicate the existence of day-of-the-week effect.

The equality of means test for testing the null hypothesis is the F statistic of the Oneway Analysis of Variance (ANOVA), used in conjunction with Tukey's test for pairs of groups which are significantly different.

Oneway ANOVA

The Oneway ANOVA makes comparison of the means of *k, (where* $k \geq 2$*)* independent samples based on the fixed effect situation of the Linear Model type in Equation 3.1 above. The assumptions applied in the Oneway ANOVA is normally distributed population and homogeneity of variances.

In the Oneway ANOVA test, comparison of the Mean $\text{Square}_{\text{(between groups)}}$ to the Mean $\text{Square}_{\text{(within groups)}}$ is carried out to derive the F ratio. The mathematical formula for computating F is:

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

The larger the F ratio of $MS_{\text{(between groups)}}$ to $MS_{\text{(within groups)}}$, the higher the possibility of rejection of the null hypothesis of equal means.

Bartlett's Test

This test is normally 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:

$$H_0: \ \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2$$

$$H1: \ \text{at least two } \sigma^2 \text{ are not equal}$$

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

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 difference between any pair of means as part of the *unplanned comparison* of means technique. The test will carry out comparison of all possible pairs of means to determine if there is significant difference. A pair of means is deemed to be significantly different at α (=0.05) if their difference is equal or greater that the critical difference MSD_{ij}, i.e. if $|\bar{Y}_j - \bar{Y}_i| \geq MSD_{ij}$ (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*
	α	=	0.05, significance level
	k	=	number of groups (in this case, k=5)
	v	=	degree of freedom, $\Sigma(n_i-1)$

3.3.2 Non-Parametric Test

The non-parametric test commonly used for testing the existence of day-of-the-week effect, is the Kruskal-Wallis Test. This test is the non-parametric equivalence of the Oneway ANOVA F statistic test. Unlike the Oneway ANOVA test, the Kruskal Wallis test is not as stringent in its assumptions.

This test is essentially a distribution free test which uses ranks based only on assumptions that the samples are continuous and rankable. It is analogous to the one-way ANOVA but is less restrictive in the treatment of assumptions. The Kruskal-Wallis test is used to test the null hypothesis that the *k* independent samples are drawn from the same population or from identical population (Siegel (1956)). In this study, the Kruskal-Wallis test is used to test the null hypothesis that daily mean returns of the 5 trading days of the week are equal. Rejection of the null hypothesis will indicate day-of-the-week effect exists in the daily returns. The test statistic is given by H where,

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

The Kruskal-Wallis test assigns ranks to all the N observations, (N=the total number of independent observations in the *k* samples) of R_{td} from the smallest to the largest. The statistic H used in the Kruskal Wallis test is defined below:(Siegel (1956))

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

If H_0 is true, and the sample sizes of the k samples are not too small, then the test statistic H is distributed as chi-square, χ^2, with degree of freedom, *df* = k-1. Therefore, the decision rule is as follows:

$$\text{Reject } H_0 \text{ if } H > \chi^2(4,\alpha)$$

where $\chi^2(4,\alpha)$ is the upper α percentile point of a chi square distribution with 4 degree of freedom. In this study, H_0 is tested at 5 percent significance level.

It is not uncommon for ties to occur i.e. two or more scores in a sample with the same score. In this case, each score is given the mean of the ranks for which it is tied. Therefore, in computing H, correction must be made for tied observations. The test statistic H corrected for tied observation 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)

ΣT = summation of all groups of ties