CHAPTER 3

RESEARCH METHODOLOGY

3.1 DATA COLLECTION TECHNIQUE

The data utilized in this study is limited to the companies whose stocks are presently component stocks of the KLSE Composite Index (with 100 component stocks) and have been in existence for the past 12 years from 1984 to 1995. There are 32 companies identified for this purpose. The list of these companies and their market capitalization as at 31 October 1996 is given in Appendix A. Secondary data from the monthly closing prices for the 32 stocks were collected over the stipulated period. These prices were then adjusted for capital changes due to rights issue, bonus issue, stock split and stock dividends (if any). No adjustment is made for the Employees Stock Option Scheme (ESOS) as its effect is negligible.

3.2 SOURCES OF DATA

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

a) The KLSE Daily Dairy

b) Investors Digest

c) Stock Market Investment In Malaysia And Singapore - Neoh Soon Kean

d) Stock Market Gazette

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

$$R_t = (P_t - P_{t-1}) / P_{t-1}$$

where

R_t = Return for month t

P_t = Monthly closing price for month t

P_{t-1} = Monthly closing price for previous month t -1

Adjustments for capital changes are made and the computation for such changes are presented in Appendix I.

The objective of this study is to :

1. test the equality of mean returns for the months in the year. Rejection of this null hypothesis would indicate the month-of-the-year effect.

2. test whether there is significant difference in the mean returns across the 12 trading months.

3.3 MODEL OF STOCK RETURNS

Stock returns are assumed to follow Linear model. In this model, the constituents of the returns of a stock are the grand mean, effect specific to the month-of-the-year and random error.

$$R_{tm} = \mu + \alpha_m + \varepsilon_{tm}$$

where

R_{tm} = return in month m of the year

μ = the grand mean

α_m = the effect specific to month m

ε_{tm} = random error term associated with that month

It is also assumed that the error is randomly and normally distributed with zero mean and common variance.

3.4 HYPOTHESES

This study will use the parametric One-way ANOVA, F test and the non-parametric Kruskal-Wallis test to test the hypotheses :

$$H_0 = \mu_1 = \mu_2 = \mu_3 = \dots\dots\dots\dots = \mu_{12}$$

against

$$H_1 = \text{at least two } \mu \text{ are not equal}$$

where μ_i is the effect specific to month i of the year.

Rejection of the null hypothesis will infer that the stock returns R_m exhibit seasonality according to the month-of-the-year. If the F test rejects the null hypothesis, the Tukey's HSD test is than used for pairwise comparison of monthly returns for significance. Besides testing for seasonality, the study will also employ the t-test and the non-parametric Mann-Whitney test for a monthly effect, or a January effect. The hypothesis to be tested is :

$$H_0 = \mu_1 = \mu_m$$

against

$$H_1 = \mu_1 \neq \mu_m$$

$m \quad = \quad$ February to December

where μ_1 is the effect specific to January and μ_m is the average effect for the month of February to December. Rejection of the null hypothesis will indicate that there is a January effect.

3.5 STATISTICAL TESTS

Analysis of variance is a procedure used for comparing sample means of the monthly returns of a particular share to see if there is sufficient evidence to infer that the means of the corresponding population distribution of the monthly returns of the same share also differ. The null hypothesis for the test is :

$$H_0 = \mu_1 = \mu_2 = \mu_3 = \ldots\ldots\ldots\ldots = \mu_{12}$$

The assumptions made in One-way ANOVA are as follows :

(a) Population is normally distributed

(b) Variances are homogeneous

The test statistic is:

$$F = MS_{bg} \;/\; MS_{wg}$$

where

$F \quad = \quad$ Ratio of the variance between groups to the variance within groups

$MS_{bg} \quad = \quad$ Variance between groups

$MS_{wg} \quad = \quad$ Variance within groups

If Sig. $F < 0.05$ then reject the null hypothesis.

3.6 BARTLETT BOX TEST

This test is used to measure whether the group variances differ significantly from each other (heteroscedasticity). The null hypothesis of the tests is :

$H_0 \; : \; \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \ldots\ldots\ldots\ldots = \sigma_{12}^2$

H_1 : at least one pair of the σ^2 are not equal

The test statistic is χ^2 with 11 degrees of freedom.

The computation of χ^2 is as follows :

$$\chi^2 = \left(\left[\sum_{m-1}^{k} (n_m - 1) \right] \ln s^2 - \left[\sum_{m-1}^{k} (n_m - 1) \right] \ln s_m^2 \right) / C$$

where

$(n_m - 1)$ = degree of freedom of month m, m = 1,2,........,12

s_m^2 = variance of month m

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

$$s^2 = \frac{\sum_{m-1}^{k} (n_m - 1) s_m^2}{\sum_{m-1}^{k} (n_m - 1)}$$

$$c = 1 - 1/3(k - 1) \left[\sum_{m-1}^{k} 1/(n_m - 1) - \frac{1}{\sum_{m-1}^{k} (n_m - 1)} \right]$$

$k = 12$

The null hypothesis is rejected if the probability value (p) is less than 0.05.

3.7 T-TEST

This test is used to assess the statistical significance of the difference between two sample means X_1 and X_2 . The null hypothesis is :

$$H_0 = \mu_1 = \mu_m$$

$$H_1 = \mu_1 \neq \mu_m$$

where

μ_1 = the effect specific to a particular month (say January)

μ_m = the effect specific for the other months (say from February to december)

The test statistic is given by :

$$t = \frac{(X_1 - X_2)}{\dfrac{[\sqrt{(n_1 - 1) S_1^2 + (n_2 - 1) S_2^2}\,]\sqrt{(1/n_1 + 1/n_2)}}{\sqrt{(n_1 + n_2 - 2)}}}$$

which follows a t distribution with $(n_1 + n_2 - 2)$ degrees of freedom and

where X_1 and X_2 = sample means for group 1 and 2

n_1 = size of sample 1

n_2 = size of sample 2

S_1^2 and S_2^2 = unbiased estimates of the common population variance, σ^2 .

3.8 TUKEY HSD TEST

Although the univariable tests of ANOVA allow one to reject the null hypothesis that group means are equal, it does not pinpoint where the significant difference lie. For further investigation of specific group mean difference of interest, a post hoc method will be used i.e. Tukey Honestly Significant Difference (HSD). A pair of means is deemed to be significantly different at $\alpha = 0.05$ if their difference is equal to or greater than the critical difference MSD_{ij} i.e. if $| \bar{Y}_i - \bar{Y}_j | \geq MSD_{ij}$ (Sokal and Rohlf (1969)).

$$MSD_{ij} = Q_{\alpha(k,v)} * \frac{\sqrt{\{ MS\ within\ (1/n_j + 1/n_i) \}}}{\sqrt{2}}$$

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 = 12)

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

3.9 KRUSKAL-WALLIS TESTS

The non parametric test commonly used for testing the existence of month-of-the-year effect is the Kruskal-Wallis Test. This test is the non-parametric equivalence of the Oneway ANOVA F statistic test. Unlike the One-way ANOVA test, the Kruskal-Wallis test is not as stringent in its assumptions. This test is essentially a distribution-free test. 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 monthly returns of the 12 trading months of the year are equal. Rejection of the null hypothesis will indicate the month-of-the-year effect exists in the monthly returns.

The Kruskal- Wallis test assigns ranks to all the N observations (N= the total number of independent observations in the k samples) from the smallest to the largest. The statistic H used in the Kruskal-Wallis tests is defined below :

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

where

k = 12, number of trading months in a year

n_d = number of cases in month d of the year

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

R_d = sum of ranks in the month d of the year

If H_o 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 :

Reject Ho if $H > \chi^2 (11, \alpha)$

where $\chi^2 (11, \alpha)$ is the upper percentile point of a chi square distribution with 11 degrees of freedom. In this study, H_o 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, corrections must be made for tied observations.

The test statistic H corrected for tied observation is :

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

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

3.10 MANN-WHITNEY U TEST

The Mann-Whitney U test accomplishs essentially what a t-test does when the distributions of the two independent samples deviate significantly from normal. It does not require assumptions about the shape of the underlying distributions. It tests the hypothesis that two independent samples come from populations having the same distribution. The null hypothesis :

H_0 : Distribution A = Distribution B

is tested against

H_1 : Distribution A $\neq$ Distribution B

Reject the null hypothesis when the probability value (p) is less than 0.05. In other words, the rejection of H_0 suggests that the two independent samples come from populations having different distributions.