# CHAPTER FOUR

# METHODOLOGY

## *4.0 Introduction*

Chapter four discusses the methodology, which can be used to measure the cointegration restriction in the simplified generalized version of Grossman's Health Capital Model employed in this study. The Singapore and Malaysia cointegration restrictions used in this model helps to determine that the log of the demand for medical care depends on the log of net consumption expenditure and the log of the relative price of medical care.

The annual medical care expenditures data, net consumption expenditures data and the relative price of medical care data (cited in Lee and Kong, 1999) are used in this study. The Singapore case consists of 38 observations while that for the Malaysia case consists of 24 observations. All the testing is carried out by using the E-views program version 3.0.

## *4.1 Data*

Data for medical care demand for Singapore and Malaysia were obtained from the breakdown components of national accounts as provided by the Department of Statistics of both countries. Data used are in line with the data classes employed at constant market price series based on the year 1990 for Singapore and 1987 for Malaysia.

The study period for Singapore was from the year 1964 until 2001, while the study period for Malaysia was from 1978 until 2001. The aspects considered for both countries were the medical care expenditure, the net consumption expenditure and also the relative price of medical care. The net consumption expenditure was calculated as total consumption expenditures less medical care expenditures. Implicit deflators were constructed by dividing nominal series through the constant market price series. The ratio of implicit deflators was respectively used as the relative price of medical care for both countries (cited in Lee and Kong, 1999; pp. 328).

All these series take the natural logarithmic ($\log_e$) transformation to have the advantage of stabilizing the variances of the series. For both countries, the log of medical care expenditure in period t is represented by $\ln M(t)$; while $\ln C(t)$ stands for the log of net consumption expenditure in period t. Whereas, the log of the relative price of medical care in period t is depicted as $\ln P_{M}(t)$. However, $\ln M(t)$ is used as the dependent variable, where $\ln C(t)$ and $\ln P_{M}(t)$ are considered as explanatory or independent variables.

### *4.2 Framework of Analysis*

The regression model employed in Singapore and Malaysia can be divided into ten categories and displayed in terms of Unit Root Test, Cointegration Test (only Johansen Cointegration Test is available here), Normalized Cointegrating Regression and Granger Causality Test. Finally, the volatility of the studied series as additional information will be discussed. Since the research is in two different countries analyses, so the explanation of the findings in next chapter will be separated into two parts. Part I is the analysis for Singapore case, while part II analyzes the Malaysian case.

#### *4.2.1 Unit Root Test*

Unit Root Test is an alternative test of stationarity. According to **Cuthbertson, Hall and Taylor (1992)**, a time series ($x_t$) is stationary if its mean, $E(x_t)$ is independent of t and its variance, $E[x_t-E(x_t)]^2$ is bounded by some finite number and does not vary systematically with time. Thus, it will tend to return to its mean and fluctuates around its mean and have broadly constant amplitude. A non-stationary time series would have a different mean at different points in time. One of the characteristics of a stationary series is that it tends to return to, or cross its mean values repeatedly and this property is the one, which is exploited by most stationarity test[41].

[41] Quoted from Keith Cuthbertson, Stephen G. Hall, Mark P. Taylor (1992): "Non-stationarity and Cointegration", Applied Econometric Techniques, Phillips Allan, Hertfordshire, U.K. in pp. 130.

To avoid obtaining results that may be spurious, the order of integration of the logarithmic terms of each series is determined by using Augmented Dickey-Fuller (1979) Test[42] and also Phillips-Perron (1988) Test. Unit Root Test is conducted on the level of the logarithms of the medical care demand, consumption differentials and the relative price of medical care. These three series are transformed into logarithms because often a series with a non-stationary variance will be stationary in the natural logarithms.

The existence of unit root in a time series indicates that the time series is non-stationary. If there is a unit root in the null hypothesis, time series maintains difference-stationary and the null hypothesis cannot be rejected unless there is overwhelming evidence to reject it.

The null hypothesis and alternative in Unit Root Tests for stationary and non-stationary can be adopted as:

$H_0$: $\delta = 1$ (The model is non-stationary)

$H_A$: $\delta < 1$ (The model is stationary)

As noted, if stationary, it is an I(0) stochastic process whereas if it is non-stationary, it is an I(1) time series. The requirement to reject the null hypothesis is when the series is stationary which means the $\tau$ (Tau) calculated values for the variables in absolute terms exceed the t-critical values. Vice versa when the t-calculated values in

[42] Dickey and Fuller (1979), Said and Dickey (1984), Phillips (1987), Phillips and Perron (1988), and others developed modifications of the Dickey-Fuller tests when $\varepsilon_t$ is not white noise. These tests, called the 'Augmented Dickey-Fuller (ADF) Tests'.

absolute terms are less than t-critical values, the null hypothesis cannot be rejected and the series is non-stationary.

### *4.2.1a Augmented Dickey-Fuller Test*

This test was introduced and developed by **Dickey and Fuller (1979)**[43]. The test for unit root indicates whether an individual series ($Y_t$), is stationary by running an OLS regression equation. For a time series $Y_t$, two forms of the "Augmented Dickey-Fuller" regression equations are:

$$\Delta Y_t = \mu + \delta Y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta Y_{t-i} + e_t \quad \text{(i)}$$

$$\Delta Y_t = \mu + \beta t + \delta Y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta Y_{t-i} + e_t \quad \text{(ii)}$$

Where

$\Delta Y_t = Y_t - Y_{t-1}$ (the first-differenced of the series)

$\delta, \beta$ = constant parameters

$e_t$ = white noise disturbance term

t = time or trend variable

p = the number of lagged terms

Equation (i) is with-constant, no trend and (ii) is with-constant, with trend. The number of lagged terms p is chosen to ensure the errors is uncorrelated.

---

[43] Dickey, D. A. and Fuller, W. A. (1979), Distribution of the Estimators for Autoregressive Time-series with a Unit Root, *Journal of the American Statistical Association*, 74, pp. 427-431.

If the autoregressive representation of $Y_t$ contains a unit root (i.e., integrated of order one), the t-ratio for $\delta$, should be consistent with the hypothesis $\delta = 0$. The conventional of t-tables are inappropriate for this hypothesis test, so we use the results of **Dickey and Fuller (1979)** and the tabulated distribution in **Fuller (1976)** to interpret the t-ratio.

### *4.2.1b Phillips-Perron Test*

**Phillips and Perron (1988)**[44] introduced the Phillips-Perron Test. They proposed non-parametric correction to allow for series correction as an alternative to the inclusion of lag terms. It accounts for non-independent and identically distributed processes using a non-parametric adjustment to the standard Dickey-Fuller procedure to ensure that the error terms are uncorrelated and have constant variance. In other words, **Phillips and Perron (1988)** developed a generalization of the Dickey-Fuller procedure that allows for fairly mild assumptions concerning the distribution of the errors.

The method is first calculated the Unit Root Test from regression equations with p = 0. The statistics were then transformed to remove the effects of serial correlation on the asymptotic distribution of the test statistic. The critical values are the same as those used for the Dickey-Fuller Test.

[44] Refer to Phillips and Perron (1988),"Testing for a Unit Root in Time Series Regression," Biometrika, 75, pp. 335-346.

Applying the test to in level then to first-differenced to measure the order of integration. Normally, it is to make sure that all the series are integrated of order one I(1), and is stationary after first-differenced. If each series is I(1), it is possible that common trends exist within them as a group, so that they could be cointegrated. Therefore, further analysis requires using the first-differenced of each series, instead of the levels.

### *4.2.1c Unit Root Test Application*

As previously discussed the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) Unit Root Test are used to examine the stationarity of the data series. It consists of running a regression of the level and the first-differences of the series against the series lagged once, lagged difference terms, and optionally, a constant and a time trend. This is expressed in equation (1) as follows:

$$\Delta \ln M_t = \mu + \beta t + \delta \ln M_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta \ln M_{t-i} + u_t \qquad (1)$$

The test for a unit root is on the coefficient of $\ln M_{t-1}$ in the regression. In other words, it is tested on the null hypothesis that $\delta = 0$. If $\delta$ is negative, it means a rejection of the null hypothesis implies stationary [$M_t \sim I(0)$], which do not exhibit a unit root in $M_t$. However, if $\delta = 0$, that is there exist a unit root, meaning that the log of medical care demand under consideration is non-stationary [$M_t \sim I(1)$]. Then, we need to proceed into first-differences to test for presence of unit root, such as:

$$\Delta^2 \ln M_t = \mu + \beta t + \delta \Delta \ln M_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta^2 \ln M_{t-i} + u_t \qquad (2)$$

From equation (1), rejection of unit root would imply that the log of the medical care demand is integrated of order 1, I(1). Therefore, it can be concluded that $\ln C(t)$, $\ln M(t)$ and $\ln P_M(t)$ are difference-stationary variables. Besides, the explanation for Phillips-Perron (PP) Test is also the same as the explanation for ADF Test. The unit root hypothesis can be rejected if the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) t-test statistics are smaller than the critical value for all tests at the 1%, 5% and 10% significance level.

After testing using Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) Unit Root Test, we proceed into cointegration test, such as Johansen Cointegration Test.

### *4.2.2 Cointegration Test*

Cointegration means that one or more linear combinations of non-stationary individually tend to be stationary. The theory of cointegration is developed by **Granger (1981)** and elaborated by **Engle and Granger (1987)**[45]. Granger addresses this issue of integrating short run dynamics with long run equilibrium. He stated that cointegration analysis could be used to discover a tendency for some linear relationships to hold

[45] Granger, C. W. J. (1981): "Some Properties of Time Series Data and Their Use in Econometric Model Specification," *Journal of Econometrics*, Vol. 16, No. 1., pp. 121-130; Engle, R. F. and Granger, C. W. J. (1987): " Cointegration and Error Correction: Representation, Estimation and Testing," *Econometrica*, Vol. 55, No. 2, pp. 251-276.

between a set of variables in the long term if it does exists. The basic idea of this explanation is if in the long run, two or more series move closely together, even though the series themselves are trended, the difference between them is constant. These series may be regarded as defining a long run equilibrium relationship and as the difference between them is stationary, the error term in a regression will have well-defined first and second moments.

A time series $M_t$ (in this research) is said to be integrated of order 1 or I(1) [$M_t \sim I(1)$] if $\Delta M_t$ is a stationary time series. A stationary time series is said to be I(0) [$M_t \sim I(0)$] if $M_t$ is a random walk, or a white noise process. Similarly, a time series is said to be integrated of order 2 or I(2) if $\Delta M_t$ is I(1) and so on. If $M_t \sim I(1)$ and $u_t \sim I(0)$, then their sum can be considered as $Z_t = M_t + u_t \sim I(1)$.

### *4.2.2a Engle and Granger Cointegration Test*

The Engle and Granger Cointegration Test is also called as Engle and Granger Two-Step Cointegration Test. Based on the statistical perspectives, the Engle and Granger Two-Step Cointegration Test is used to examine the long run equilibrium or long run relationships between the variables, which means that the variables moved together over a long period and short term disturbance. If these variables cointegrated among themselves, they should not move "too far" away from one another. In other words, they should not diverge, but trend up and down together. The Engle and Granger Cointegration Test is better to test for two cointegrating vectors.

However, since this test is only used to test two cointegrating vectors and furthermore, is more suitable to predict the normal endogenous as well as exogenous variables. Therefore, it is not appropriate to apply in the model, which consists of lagged variables (predetermined variables). As a result, the Engle and Granger Cointegration Test is not used in this study because the model comprises lagged variables. Thus, Johansen's Cointegration Test is more appropriate for this study.

### *4.2.2b Johansen Cointegration Test*

**Johansen-Juselius (1990)**[46] proposed cointegration test (with unrestricted intercept and no trends) to examine the long run relationship between two or more variables. To investigate the Johansen Cointegration Test, first, we test for the null hypothesis. If the result implies that the cointegrating vector is zero (r = 0), we cannot reject the null hypothesis. It shows that no cointegrating relationship exists between the two vectors. However, if the result shows we can reject $H_0$, it means we accept alternative hypothesis. This implies that at least one cointegrating vector exists. To test whether there is only one or more than one cointegrating vectors, we continue to conduct another test for the null hypothesis (r ≤ 1). If the result, again shown the existence of

[46] Johansen, S. and Juselius, K. (1990), Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics*, 52, No.2, pp. 305-9049

only one cointegrating vector, we cannot reject the null hypothesis. However, if there is more than one cointegrating vector, we reject null hypothesis and conclude that there exist two cointegrating vectors. In other words, the null hypothesis shows there is no cointegrating relationship among the variables is rejected at most for two cointegrating vectors. While, the maximum eigen value test confirms non-zero vectors among the variables.

The null hypothesis and the alternative hypothesis in cointegration test are represented as below:

$H_o$: r = 0 (No cointegrating relation exists)

$H_A$: r > 0 (At least one cointegrating relation exists)

$H_o$: r = 1 (At most one cointegrating relation exists)

$H_A$: r > 1 (More than one cointegrating relations exist)

$H_o$: r = 2 (At most two cointegrating relations exist)

$H_A$: r > 2 (More than two cointegrating relations exist)

The Johansen Cointegration Test is conducted to establish the relationships between these series within the study periods. The null hypothesis (r = 0) in this test means that no cointegration is found among the element of $\ln M(t)$, where r represents the number of cointegrating vectors. Rejection of null hypothesis indicates the presence of cointegrating relations and accepts the alternative of one or more cointegrating vectors (r > 0).

The procedure begins with the equation (3) and equation (4), which indicate in term of the least square estimating regressions for Singapore and Malaysia cases, as:

$$\hat{u}_t = M_t - \beta_0 - \beta_1 C_t - \beta_2 P_{M,t} \qquad (3)$$

$$\ln M_t = \alpha_0 + \alpha_1 \hat{u}_{t-1} + \sum_{i=1}^{p} \alpha_{2i} \ln M_{t-i} + \sum_{i=1}^{p} \alpha_{3i} \ln C_{t-i} + \sum_{i=1}^{p} \alpha_{4i} \ln P_{M,t-i} + \varepsilon_t \qquad (4)$$

Equation (4) defines the product moment matrices of the estimated residuals in level.

**Johansen (1988)**[47] also shows that the likelihood ratio test statistic ($Q_r$) is used for the hypothesis of at most equilibrium relationships, as shown by equation (5):

$$Q_r = -n \sum_{j=r+1}^{m=k} \ln(1 - \hat{\lambda}_j) \qquad (5)$$

Where $\lambda_1 > \lambda_2 > \ldots\ldots\lambda_p$ are the eigen values to solve the following equation:

$$\left| \lambda S_{11} - S_{10} S^{-1}{}_{10} S_{01} \right| = 0 \qquad (6)$$

Which is considered as equation (6). The eigen values are also called the squared canonical correlation of $u_{2t}$ with respect to $u_{1t}$. The limiting distribution of the $Q_r$, statistic is given in terms of a p-r dimensional Brownian motion process, and the quantiles of the distribution are tabulated in **Johansen and Juselius (1990)** for p-r+1,....,5 and in the **Osterwald-Lenum (1992)** for p-r+1,.....,10.

Equation (5), usually referred to as the trace test, may be rewritten as:

---

[47] Refer to Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, *Journal of Economic Dynamics and Control*, 12, pp. 231-54.

$$L_{trace} = -T\sum_{i=r+1}^{p}\ln(1-\lambda_i) \qquad (7)$$

This model is considered as equation (7). Where $\lambda_{r+1},......,\lambda_p$ are the p-r smallest squared canonical correlation or eigen values.

The other test for cointegration is the maximal eigen value test. Based on the following statistic:

$$L_{\max} = -T\ln(1-\lambda_{r+1}) \qquad (8)$$

Which considered as equation (8). While, the $\lambda_{r+1}$ is the $(r+1)^{th}$ largest squared canonical correlation or eigen value.

Equation (7), finds that the null hypothesis is at the most r-cointegrating vectors. Whereas equation (8) shows the null hypothesis is that there are r-cointegrating vectors against the alternative of r+1 cointegrating vectors.

**Johansen and Juselius (1990)** deduced that the trace test might lack power relative to the maximal eigen value test. Based on the power of the test, the maximal eigen value test statistic is often preferred. Long run elasticity is obtained from the normalized equation that is only conducted if non-zero vectors are confirmed by the cointegration test. This is obtained by setting each cointegrating vector equal to one.

### 4.2.3 *Cointegrating Regression*

Cointegrating regression is a regression (in level) of one series against a constant and another series. There can be as many cointegrating regression as the number of variables. The residuals of the cointegrating regression that provide information about the long run relationships among variables are then tested for stationarity.

#### 4.2.3a *Normalized Cointegrating Regression*

As previously mentioned, normalization is only conducted if non-zero vectors are confirmed by the cointegration test which setting each cointegrating vector equal to 1. In this study, the normalized cointegrating regression is demonstrated in least square estimating regression to reveal the long run elasticities obtained by Singapore and Malaysia. A linear time trend is included in below cointegrating regression since the model assumed that the preference shock in the health sector is trend-stationary. The cointegrating regression is written as:

$$\ln M_t = \alpha_0 + \beta t + \alpha_1 \hat{u}_{t-1} + \sum_{i=1}^{p} \alpha_{2i} \ln C_{t-i} + \sum_{i=1}^{p} \alpha_{3i} \ln P_{M,t-i} + \varepsilon_t \qquad (9)$$

Where $\alpha_0$ indicates as constant term; $\alpha_1, \alpha_{2i}$ and $\alpha_{3i}$ are used as coefficient of the series for both countries, while $\beta t$ indicates as coefficient of the trend variable. On the other hand, $\hat{u}_{t-1}$ represents the estimated one-period lagged value of the residual and $\varepsilon_t$ is the random error term for both countries. All the series, demand for medical care expenditure

$(\ln M(t))$, net consumption expenditure $(\ln C(t))$ and relative price of medical care $(\ln P_M(t))$ are written in natural logarithmic forms.

### *4.2.4 Granger Causality (GC) Test*

A useful test for causality in time series models is related to correlation among the variables and temporal asymmetry (i.e., time precedence) among phenomena. The causality test, based on the **Granger (1969)**[48] approach, is conducted to see if there is any influence between the variables. A dependent variable is said to be granger-cause by an independent variable if lagged values of the independent variable can help to improve the explanation of the current dependent variable apart from its past values Vice-versa explanation is also correct, an independent variable is said to be granger-cause by a dependent variable if it helps in the prediction of independent variable, or equivalently if the coefficients on the lagged values of independent variables are statistically significant.

The Granger Causality relations between a dependent and independent variables for both countries of restricted and unrestricted regression model are indicated as follows, for instance:

$H_0$: M –|> C (medical care demand does not granger-causes net consumption expenditure)

$H_A$: M —> C (medical care demand does granger-causes net consumption expenditure)

---

[48] Granger, C. W. J. (1969): "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", *Econometrica*, (July), pp. 424-438.

To test for causality when variables are cointegrated, the following Granger Causality equation is conducted. Lets assume two variables only for both countries, for example:

$$\Delta \ln C_t = a_{10} + \sum_{j=1}^{p} a_{11,j} \Delta \ln C_{t-j} + \sum_{j=1}^{p} a_{12,j} \Delta \ln M_{t-j} + e_{1t}$$

$$\Delta \ln M_t = a_{20} + \sum_{j=1}^{p} a_{21,j} \Delta \ln M_{t-j} + \sum_{j=1}^{p} a_{22,j} \Delta \ln C_{t-j} + e_{2t} \qquad (10)$$

M Granger Causes C if $a_{12,j} \neq 0$ for any j

C Granger Causes M if $a_{22,j} \neq 0$ for any j

Where independent and dependent variables have been identified as first-differenced stationary time series and $e_{1t}$ and $e_{2t}$ are disturbance terms for both countries, which are serially uncorrelated. Bilateral causality is suggested if both $a_{12}$ and $a_{22}$ is statistically different from zero, and inter-temporal lead-lag relationship does not exist if both of them are not significant.

### *4.2.5 Series Volatility*

**Gujarati (2003)** mentioned that a time series model often exhibits phenomenon of volatility clustering, that is, the error variance may be correlated over time. Knowledge of volatility is crucially important in many areas, including the area of health and medical care. Volatility is employed to model "varying variance[49]" in demand and consumption as well as in the relative price.

In short, modeling of volatility is useful to analyze the sensitivity of the series when other series have changed. In addition, it also helps to obtain more accurate confidence intervals for forecasting. Finally, it is used to obtain more efficient estimators by handling heteroscedasticity in error properly.

The volatility of the studied series for Singapore and Malaysia is conducted as additional information to test the sensitivity (volatile) and also the movement of the series. This test draws upon raw data for the entire series, whereas, the volatility of the given series refers to the coefficient of variation, which is written in percentages. It is obtained from the results to illustrate the trend of the series. Furthermore, this test helps to explain the situation of the series when the utility preferences of the market have changed.

## *4.3 Statistical Test*

Hypotheses are used to test the validity of the model and to improve the model specification. To test whether an individual slope coefficient ( $\beta_i$ ) is significantly different from zero, t-tests are carried out.

[49] Varying variance in (Gujarati, 2003; pg. 856) refers to the variance of time series varies over time.

Individual Regression Coefficient Test (T-test)[50]

$H_0$: $\beta_2 = 1$ $H_0$: $\beta_3 = -1$

$H_A$: $\beta_2 \neq 1$ $H_A$: $\beta_3 \neq -1$

In the language of statistics, the stated hypothesis, which is denoted by the symbol $H_0$, is known as the null hypothesis. The null hypothesis is usually tested against an alternative hypothesis (also known as maintained hypothesis[51]). Whereas, the hypothesis denoted by the symbol $H_A$, is recognized as the alternative hypothesis. For instance, alternative hypothesis may stated as, true $\beta_2$ is different from unity. The alternative hypothesis can be simple or composite[52], this is shown, for example, $H_A: \beta_2 = 1$ is a simple hypothesis, while $H_A: \beta_2 \neq 1$ is a composite hypothesis.

Test Statistic for t-value

From the computer output, $t_c$

$$t = \frac{\hat{\beta i} - \beta i}{s.e.(\hat{\beta i})} \sim t_{\alpha/2;n-k}$$

From the t-table, $t^{*}_{\alpha/2;n-k}$

Result: reject the null hypotheses if $t_c > t^{*}_{\alpha/2;n-k}$ or $t_c < -t^{*}_{\alpha/2;n-k}$

---

50 Refer to Gujarati (1995), "Hypothesis Testing: General Comments" from page 121-122 and "Hypothesis Testing: The Test-Of-Significance Approach" from page 124-128.

51 According to Damodar N. Gujarati (1995) in his "Hypothesis Testing: General Comments" in "Basic Econometrics" has mentioned that 'alternative hypothesis' also known as 'maintained hypothesis', pp. 121.

52 A statistical hypothesis is called a simple hypothesis if it specifies the precise value(s) of the parameter(s) of a probability density function; otherwise, it is called a composite hypothesis (Gujarati, 1995; pp. 122).

n = number of observations

k = number of parameters

$\alpha/2$ = level of significance

n-k = degrees of freedom

$\beta i$ = slope coefficient for individual variable

$\hat{\beta} i$ = estimated slope coefficient for individual variable

t = t-calculated value

$t_c$ = t-critical value

s.e. = standard error

$s.e.(\hat{\beta} i)$ = estimated standard error of estimator for individual variable

$-t_{\alpha/2}$ and $t_{\alpha/2}$ = values of the t-variable obtained from the t-distribution for $\alpha/2$- level of significance

The above test statistic defined follows the t-distribution with n-k degrees of freedom. This test statistic can also be written as:

$$t = \frac{\text{estimator(i)} - \text{parameter(i)}}{\text{estimated standard error of estimator(i)}}$$

Where the $s.e.(\hat{\beta}_i)$, refers to the estimated standard error for individual i. If the value of true $\beta_i$ is specified under the null hypothesis, the above t-value can serve as a test statistic, which can readily be computed from the available sample. While, $-t_{\alpha/2}$ and $t_{\alpha/2}$ are often called the critical t-value at $\alpha/2$-level of significance.

Test Statistic for F-test[53]

In order to determine whether the model is poor or a good one, the test for this purpose requires reformulation. As a result, a joint F-test, which is a measure of the overall significance of the estimated regression, is carried out to test for overall goodness of fit.

$H_0$: $\beta_2 = \beta_3 = 0$ (All slope coefficients are simultaneously zero)

$H_A$: At least one $\beta_k \neq 0$, k = 2,3 (Not all slope coefficients are simultaneously zero)

The above null hypothesis represented by $H_0$, is a joint hypothesis, which shows $\beta_2$ and $\beta_3$ are jointly or simultaneously equal to zero. Whereas, the alternative hypothesis denoted by the symbol $H_A$, reveals not all the slope coefficients are jointly or simultaneously equal to zero. Normally, the null hypothesis is tested against the alternative hypothesis. Here, if the null hypothesis is being rejected, that means the alternative hypothesis is accepted and the model is statistically significant since it lies in the critical region. Therefore, it can be concluded that the model obtains goodness of fit. In contrast, if the null hypothesis is being accepted, that means the alternative hypothesis is rejected. By the same token, the model is statistically insignificant because it lies in the acceptance region. In this case, the model does not obtain goodness of fit.

[53] Refer to Gujarati (2003), "Testing the Overall Significance of the Sample Regression" from page 253-264.

$$F_c = \frac{ESS/(k-1)}{RSS/(n-k)} \sim F_{\alpha;k-1,n-k}$$

From the F-table, $F^{*}_{\alpha,k-1,n-k}$

Result: reject the null hypotheses if $F_c > F^{*}_{\alpha;k-1,n-k}$ for a α-level test

n = number of observations

k = number of parameters

α = level of significance

ESS = Explained Sum of Squares

RSS = Residual Sum of Squares

k-1 = degrees of freedom for the numerator (not including constant)

n-k = degrees of freedom for the denominator (losing some of the degrees of freedom)

F = F-calculated value

$F_c$ = F-critical value

$F_\alpha$ = values of the F-variable obtained from the F-distribution for α-level of significance

From the regression viewpoint, the above F-statistic is known as the analysis of variance (ANOVA) and it follows the F-distribution. ESS is the explained sum of squares and RSS is the residual sum of squares. ESS is divided by k-1 degrees of freedom, whereas, RSS is divided by n-k degrees of freedom. $F_\alpha$ is usually called the critical F-value at $\alpha$-level of significance.

F-test for Linear Restrictions[54]

$H_0$: Restrictions are true

$H_A$: At least one restriction is not true

Above linear restrictions postulate that $H_0$, which indicates as null hypothesis represents the restrictions are true. Whereby $H_A$, considers as alternative hypothesis exhibits that at least one restriction is not true. As usual the null hypothesis is tested against the alternative hypothesis. As previously mentioned, if the null hypothesis is rejected and the alternative hypothesis is accepted, that means at least one restriction is not true. Otherwise, if the alternative hypothesis is rejected and the null hypothesis is accepted, this demonstrates that the restrictions are true.

F-test for Linear Restrictions

$$F_c = \frac{(RSS_r - RSS_u)/s}{RSS_u/(n-k)} \sim F_{\alpha;s,n-k}$$

From the F-table, $F^*_{\alpha;s,n-k}$

Result: reject the null hypotheses if $F_c > F^*_{\alpha;s,n-k}$ for a α-level test

$RSS_u$ = Total Residual Sum of Squares ($\sum \hat{u}i^2$) for the unrestricted model

$RSS_r$ = Total Residual Sum of Squares ($\sum \hat{u}i^2$) for the restricted model

n = number of observations

[54] Refer to Gujarati (2003), "Restricted Least Squares: Testing Linear Equality Restrictions -- The F-test Approach: Restricted Least Squares" from page 267-273.

k = number of parameters in the unrestricted model

s = number of restrictions for the numerator

α = level of significance

n-k = degrees of freedom for the denominator (losing some of the degrees of freedom)

F = F-calculated value

$F_c$ = F-critical value

$F_\alpha$ = values of the F-variable obtained from the F-distribution for α-level of significance

Above F-statistic also defined follows the F-distribution but it is divided by the number of restrictions. $RSS_u$ is considered as Total Residual Sum of Squares for the unrestricted model, while $RSS_r$ represents the Total Residual Sum of Squares for the restricted model. The outcome of $RSS_u$ minus $RSS_r$ is divided by the number of restrictions, whereas, the $RSS_u$ only divided by n-k degrees of freedom. $F_\alpha$ is the critical F-value at α-level of significance.

### *4.4 Lag Determination*

The idea of imposing a penalty for adding regressors to the model has been carried further in the Akaike Information Criterion (AIC) and Bayesian Schwarz Criterion (BIC) or Schwarz Information Criterion (SIC)[55]. Both tests are used to determine the optimum lag-length (p), which optimum lag-length is determined by

---

[55] Refer to Gujarati (2003), "Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC)" from page 537-538.

choosing the minimum values of these two tests. In other words, the lower values of Akaike and Schwarz are chosen to obtain the goodness of fit of the model. On this basis, the more parsimonious model is preferable. Parsimonious is a tradeoff of a model and the model is considered parsimonious when the number of explanatory variables is less. A better model comprises less explanatory variables.

From the aspect of number of parameters (k) and also number of observations in natural logarithm ($\ln n$), BIC is preferred because $\ln n > 2$, which is greater than the number of k. Therefore, BIC tends to choose models that are more parsimonious with less parameter. However, AIC tends to choose models that are less parsimonious with more parameters. Below is the explanation for both AIC and BIC. According to **Gujarati (2003)**:

Akaike Information Criterion (AIC)

$$AIC = n\sum \hat{u}_t^2 + 2k$$

Where n = number of observations

$\sum \hat{u}_t^2$ = Total Residual Sum of Squares in period t

k = number of parameters including the intercept

For mathematical convenience, above equation can also be written as:

$$\ln AIC = \left|\sum \hat{u}_t\right| + \frac{2k}{n}$$

Where $\ln AIC$ = natural log of AIC

$$\frac{2k}{n} = \text{penalty factor}$$

AIC imposes a harsher penalty than $\bar{R}^2$ for adding more regressors. In comparing two or more models, the model with the lowest value of AIC is preferred. One advantage of AIC is that it is useful for not only in-sample but also out-of-sample forecasting performance of a regression model. Also, it is useful for both nested and non-nested models.

Bayesian Schwarz Criterion (BIC) or Schwarz Information Criterion (SIC)

Similar in spirit to the AIC, the BIC or SIC criterion is defined as:

$$BIC \text{ or } SIC = n\sum \hat{u}_t^2 + k \ln n$$

Where n = number of observations

$\sum \hat{u}_t^2$ = Total Residual Sum of Squares in period t

k = number of parameters including the intercept

$\ln n$ = number of observations in logarithmic ($\log_e$) term

While in log-form:

$$\ln BIC \text{ or } \ln SIC = \ln\left|\sum \hat{u}_t^2\right| + k\frac{\ln n}{n}$$

Where $\frac{k}{n}\ln n$ is the penalty factor. BIC or SIC imposes a harsher penalty than AIC. Like AIC, the lower the value of BIC or SIC, the better the model. Again, like AIC, BIC or SIC can be used to compare in-sample or out-of-sample forecasting performance of a model.

## *4.5 Conclusion*

Current time-series techniques, namely Unit Root Test, Johansen Cointegration Test and Granger Causality Test applied in this study are used to investigate whether a stable long run relationship exists for medical care. These techniques are employed to examine the series for Singapore and Malaysia. The data is also studied to see whether there is any persistent pattern between them, which will cause a long run stable relationship.