FINDING SUITABLE STATISTICAL MODEL FOR GROSS DOMESTIC PRODUCT AND AGRICULTURE RELATED FACTORS

Aboobacker Jahufer

Department of Mathematical Sciences
Faculty of Applied Sciences
South Eastern University of Sri Lanka
jahufer@seu.ac.lk

ABSTRACT
Gross domestic product (GDP) is a measure of a country's economic performance. It is the market value of all officially recognized final goods and services produced within a country in a year, or over a given period of time. GDP per capita is often used as an indicator of a country's material standard of living. The economy of Sri Lanka is highly depending on agriculture, industry and services. The Objective of this research study is to find a suitable statistical model for GDP and agricultural related factors such as: Tea, Rubber, Coconut, Paddy, Fishing and Others. For this purpose annual data were collected form the central bank report from 1974 to 2014. Multiple regression models such as linear model, log linear model and first difference of log linear models are analyzed using the statistical tests: normality test, unit root test, residual test, autocorrelation and serial correlation test, heteroscedasticity, multicollinearity test and stability test. Among these three models the best model is selected based on the statistical significance of the models. The first difference of log linear model satisfied almost all the statistical tests. Hence, the first difference of log linear model is more appropriate for GDP and agriculture related factors.

Keywords: GDP, Agricultural Factors, Unit Root Test, Serial Correlation Test, Stability Test.

INTRODUCTION

Contribution of Agricultural Factors in Sri Lankan Economy

Sri Lankan economy has registered a strong growth in the first decade of 21st century, despite suffering civil war from 1983 to 2009s and some major natural disasters. The GDP grew at an average rate of 5% during the war period. One of the major factors behind development and economic growth is government spending. Agricultural production, tourism, textile, apparel and tea export are the biggest economic boosters. A lot of foreign exchange also comes through overseas employment especially in the Middle East. The country's economy has traditionally been based on agriculture, which now contributes less than 20% to the GDP and employs about a third of the work force. The emphasis is on export crops such as tea, rubber, and coconuts (all plantation-grown). Cinnamon, cardamom, pepper, cloves, nutmeg, citronella, tobacco, cocoa, and coffee are also exported. Rice, sugarcane, grains, pulses, oilseed, fruit and vegetables are grown for local use and consumption. Petroleum refining is important and amorphous graphite, precious and semiprecious gems, mineral sands, clay, and limestone are mined. Port construction, telecommunications, and offshore insurance and banking are also important industries. Remittances from Sri Lankans working abroad, mainly in the Middle East, contribute significantly to the economy. The island's swift rivers have considerable hydroelectric potential (See: http://www.economywatch.com/world_economy/sri-lanka).

Background to the Problem

The Sri Lanka's economy mainly depends on 3 sectors which are Agriculture, Industry & Services that are highly contributing to the Gross Domestic Product (GDP). According to Piana (2001), the Gross Domestic Product (GDP) is most important

measure of economic activity in a country. GDP is the market value of all officially recognized final goods and services produced within a country in a year, or over a given period of time. GDP per capita is often used as an indicator of a country's material standard of living.

The Study which analyzes the agriculture sector contributions to total GDP is pertinent in understanding the underpinnings of the economy. Therefore, this particular study which seeks to add to this understanding with regard to the contribution of agriculture sector to the GDP. This measurement is important for bringing the realistic data in evaluating of contribution of agriculture sector to GDP. The study will forecast the agriculture sector growth in view of predicting GDP growth in Sri Lanka. Furthermore this study will add more knowledge for economics in analyzing the growth of GDP through the established model.

Objectives of the Research Study

The main objective of this research study is to find a suitable statistical model for GDP and agricultural related factors based on the appropriate statistical analysis.

This paper is composed into five section. Section two illustrates review of the literature, section three explains methodology of the research study, in section four data analysis and presentations are given and conclusions are given in last section.

REVIEW OF LITERATURE

Udith K. Jayasinghe-Mudalig (2006) examines empirically the economic problem of whether the food and agricultural sector in Sri Lanka. Both descriptive and Multiple Regression Analysis were carried out to analyze the secondary data covering the period of 1970 to 2004.

Niringiye Aggrey (2007) studied simple stylized facts on relationships between the share of agriculture and its subsectors in GDP in Uganda. The study utilized both the trend and regression analyze is confirm this stylized fact.

In a comparative study by Papola (2005) agriculture, expectedly, registered a decline in its share in GDP in Thailand, Korea, India, China and Malaysia during 1960-2002.

METHODOLOGY

This study intends to investigate the effect of agriculture sector factors such as tea, rubber, coconut, paddy, fishing and others on GDP. The following methodological approach is adopted in this study for establishing the relationship between agriculture sector and GDP.

Data Collection

The research study uses the existing secondary data of GDP and agriculture sector data from 1974 to 2014 from the Annual Reports of the Central Bank of Sri Lanka (Reports: 2010, 2012 and 2014)

Research Design and Appropriate Statistical Analysis

This study is find a suitable statistical model for GDP and agriculture related factors tea, rubber, coconut, paddy, fishing and others from the time series data from 1974 to 2014 yearly data. Before to analyze the data, normality of the variables have to checked by Kolmogorov-Smirnov and Shapiro-Wilk approach for each of the agriculture factor variables and GDP variable. The parameters are estimated by ordinary least square (OLS) technique. Linear multiple regression model, log linear multiple regression model and first difference of log linear models are analyzed. Among these three models the best model is selected by checking the statistical

properties and conditions of appropriate statistical tools such as: coefficient of determination (R^2), normality of the residuals, stationary and non-stationary test, auto correlation and serial correlation test of the residual, heteroscedasticity test of the residual, stability test.

The Empirical Model of the Research Study

The data of this research study is time series. Hence, the following empirical models are studied:

Linear model:

$$ASPI = \beta_0 + \beta_1 Tea + \beta_2 Rubber + \beta_3 Coconut + \beta_4 Padddy + \beta_5 Fishing + \beta_6 Others + \varepsilon$$

Log linear model:

$$Ln(ASPI) = \beta_0 + \beta_1 ln(Tea) + \beta_2 ln(Rubber) + \beta_3 ln(Coconut) + \beta_4 ln(Padddy) + \beta_5 ln(Fishing) + \beta_6 ln(Others) + \varepsilon$$

First difference of log linear model:

$$D(Ln(ASPI)) = \beta_0 + \beta_1 D(ln(Tea)) + \beta_2 D(ln(Rubber)) + \beta_3 D(ln(Coconut)) + \beta_4 D(ln(Padddy)) + \beta_5 D(ln(Fishing)) + \beta_6 D(ln(Others)) + \varepsilon$$

DATA ANALYSIS AND PRESENTATION

Linear Model Test Results and Interpretation

The model tests statistical results are given in below table 1.

Table 1: Statistical test results for models

Description of the Statistical Test	Linear Model Test Results	Log Linear Model Test Results	First Difference of Log Linear Model Test Results
Model	Model is significant (P = 0.000)	Model is significant (P = 0.000)	Model is significant (P = 0.000)
β_0	Significant (P = 0.003)	Significant (P = 0.002)	Significant (P = 0.000)
β_1	Significant (P = 0.000)	Significant (P = 0.012)	Significant (P = 0.000)
β_2	Not Significant (P = 0.871)	Significant (P = 0.001)	Significant (P = 0.045)
β_3	Significant (P = 0.042)	Significant (P = 0.000)	Not Significant (P = 0.317)
β_4	Significant (P = 0.000)	Not Significant (P = 0.498)	Not Significant (P = 0.075)
β_5	Significant (P = 0.000)	Not Significant (P = 0.625)	Not Significant (P = 0.672)
β_6	Not Significant (P = 0.728)	Significant (P = 0.000)	Not Significant (P = 0.169)
R^2	0.998	0.997	0.498
Normality of the Variables	(Kolmogorov-Smirnov Test)	(Kolmogorov-Smirnov Test)	(Kolmogorov-Smirnov Test)
GDP	Not Normal (P = 0.000)	Normal (P = 0.200)	Normal (P = 0.300)
Tea	Not Normal (P = 0.000)	Normal (P = 0.200)	Normal (P = 0.200)
Rubber	Not Normal (P = 0.000)	Normal (P = 0.091)	Normal (P = 0.291)
Coconut	Not Normal (P = 0.000)	Normal (P = 0.200)	Normal (P = 0.200)
Paddy	Not Normal (P = 0.000)	Normal (P = 0.200)	Normal (P = 0.073)
Fishing	Not Normal (P = 0.000)	Normal (P = 0.200)	Not Normal (P = 0.006)
Others	Not Normal (P = 0.000)	Normal (P = 0.200)	Not Normal (P = 0.006)
Unit Root Test	(Augmented Dickey-Fuller Test)	(Augmented Dickey-Fuller Test)	(Augmented Dickey-Fuller Test)
GDP	Stationary (P = 0.0017)	Non Stationary (P = 0.0956)	Stationary (P = 0.0086)
Tea	Non Stationary (P = 0.1427)	Non Stationary (P = 0.4835)	Stationary (P = 0.0007)
Rubber	Non Stationary (P = 0.1705)	Non Stationary (P = 0.2723)	Stationary (P = 0.0028)
Coconut	Non Stationary (P = 0.0718)	Stationary (P = 0.0295)	Stationary (P = 0.0226)
Paddy	Stationary (P = 0.0226)	Non Stationary (P = 0.4330)	Stationary (P = 0.0024)
Fishing	Non Stationary (P = 0.2842)	Non Stationary (P = 0.4537)	Stationary (P = 0.0007)
Others	Non Stationary (P = 0.7794)	Non Stationary (P = 0.7975)	Stationary (P = 0.0280)
Residual Test	Not Normal (Jarque-Bera Test P = 0.000972)	Normal (Jarque-Bera Test P = 0.613301)	Normal (Jarque-Bera Test P = 0.260483)

Autocorrelation Test	Significant Positive Auto-Correlation Exist (Durbin-Watson Stat = 1.288255)	Positive Auto-Correlation exists (Durbin-Watson Stat =1.034396)	No Auto-Correlation Problems (Durbin-Watson Stat = 2.015919)
Serial Correlation Test	No Serial Correlation (Breusch-Godfrey Serial Correlation LM Test P = 0.1000)	Serial Correlation exists (Breusch-Godfrey Serial Correlation LM Test P = 0.0021)	No Serial Correlation Problem (Breusch-Godfrey Serial Correlation LM Test P = 0.0682)
Heteroscedasticity	Errors have not constant variance (Heteroscedasticity Test: White P = 0.0001)	Errors have constant variance (Heteroscedasticity Test: White P = 0.1119)	Errors have constant variance (Heteroscedasticity Test: White P = 0.6042)
Multicollinearity Test	Multicollinearity exists among independent Variables. (Highest VIF= 64.762)	Multicollinearity exists among independent Variables. Highest VIF= 71.316	No Multicollinearity among independent Variables. Highest VIF= 1.627
Stability Test (Recursive Estimation-CUSUM of Squares Test)	Stable	Not Stable	Stable

The linear model does not satisfied most of the statistical conditions. The linear variables are not normally distributed and except variables GDP and Paddy others are non-stationary variables. Errors are not normally distributed, errors have auto correlation, serial correlation and heteroscedasticity problems, furthermore not stable. In addition to these problem independent variables have multicollinearity problems. According to these overall statistical analysis linear multiple regression model is not suitable for this data.

The log linear model does not satisfied most of the statistical conditions. The linear variables are not normally distributed and except variable coconut others are non-stationary variables. Errors have auto correlation and serial correlation, furthermore not stable. In addition to these problem independent variables have multicollinearity problems. According to the overall statistical analysis above log linear multiple regression model is not suitable for this data.

The first difference of log model satisfied all most all appropriate statistical conditions. Hence, the first difference of log linear model is appropriate for this data.

CONCLUSION

This study is intended to find a suitable multiple regression model for GDP and agriculture related important factors contributed to GDP. Based on the appropriate statistical analysis the first difference of log linear model is more appropriate for this data.

REFERENCES

Annual Central Bank Report, (2010, 2012, 2014). Central bank of Sri Lanka.

http://www.economywatch.com/world_economy/sri-lanka

Niringiye Aggrey. (2009). Patterns of Agricultural Growth and Overall growth of Ugandan Economy. Presented at GLOBELICS 2009, 7th International Conference, Dakar, Senegal.

Papola, T. S. (2005). Structural Changes in the Indian Economy: Some Implications of the Emerging Pattern, *Artha Beekshan*, Vol. 13, No. 4.

Piana, V. (2001) Economics Web Institute. http://www.economicwebinsititute.org/

Udith K. Jayasinghe-Mudalige (2006). Role of Food and Agriculture Sector in Economic development of Sri Lanka: Wayamba University of Sri Lanka.