

**A STATISTICAL MODEL TO FORECAST FOREIGN  
DIRECT INVESTMENT OF A COUNTRY: A CROSS  
COUNTRY ANALYSIS**

W.A.Kushan Sumedha

(148870P)

Dissertation submitted in partial fulfillment of the degree Master of Science in  
Business Statistics

Department of Mathematics

University of Moratuwa

Sri Lanka

December 2017

## **DECLARATION**

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:

Date:

The above candidate has carried out research for the Master's Dissertation under my supervision.

Name of the supervisor:

Signature of the supervisor:

Date:

## ABSTRACT

In this research it is aimed to study the effects of variety of different socio economic factors for foreign direct investment. The multiple linear regression model is developed to forecast the foreign direct investment. The dependent variable of the model is foreign direct investment of the country in calendar year 2015. (in USD) and the independent variables are the gross domestic product in USD, adult literacy rate, gross national income, gross domestic products annual growth and gross national income. Hypothesis tested was all socio economic factors affect equally to the attracting foreign direct investment. Further the extent of which various socio economic indexes affect a country's economic well-being was evaluated to determine which indexes have stronger effect on the foreign direct investment of a country when compared with the others. Having identified the stronger and more impact socio economic indexes, it was tried to improve the econometric model by using variable selection methods and possibly multiple regression methods to further understand the relationship between foreign direct investment and the socio economic factors that affected. Finally exploratory factor analysis was carried out to examine the common factors among socio-economic variables. For this study a pool of forty five countries were selected and ten models were developed using thirty five countries from the pool and the ten models illustrated that there is no equal influence from eight macro-economic variables to FDI amount of the country at 95% significant level. The result emphasis that the macro-economic variables differently affect in attracting foreign direct investment. Moreover the factor analysis asserts that there are similar patterns among macro-economic variables.

**Keywords:** Foreign Direct Investment, Macro-economic Indicators, Multiple Regression Analysis, Factor Analysis

## **ACKNOWLEDGEMENT**

Apart from my effort, the success of this project depends largely on the encouragement and guidelines of many others. I take this opportunity to express my gratitude to the people who have participated in the successful completion of this project. I would like to show my greatest appreciation to Professor T.S.G Peiris, the Professor in Applied Statistics and the course coordinator of Masters of Business Statistics, who gave me knowledge and support to carry out the concept of this project. I feel motivated and encouraged every time when I attend his meetings. Without his encouragement and guidance this research project would not have materialized. I am very much grateful to all my lecturers of the M.Sc. in Business Statistics for their knowledge and support shared with me. I would like to give my thanks to the Department of Mathematics in the Faculty of Engineering at the University of Moratuwa for providing me with this opportunity to present my knowledge through practical aspect of business environment. I would like to thank my work place, Micro Cars Ltd for the support given me to carry out this course work.

The help and support provided by my father and mother motivated me very much to complete this project successfully.

# TABLE OF CONTENT

<b>DECLARATION</b> .....	i
<b>ABSTRACT</b> .....	ii
<b>ACKNOWLEDGEMENT</b> .....	iii
<b>TABLE OF CONTENT</b> .....	iv
<b>LIST OF TABLES</b> .....	viii
<b>LIST OF FIGURES</b> .....	xi
<b>LIST OF EQUATIONS</b> .....	xii
<b>LIST OF ACRONYMS</b> .....	xiv
<b>CHAPTER 1</b> .....	1
<b>INTRODUCTION</b> .....	1
1.1 Introduction of Foreign Direct Investments (FDI).....	1
1.2 History of FDI and Trade across Boundaries.....	1
1.3 Foreign Direct Investment and Trades in Sri Lanka .....	2
1.4 Definitions for FDI.....	3
1.5 Forms of FDI.....	4
1.6 Importance of FDI to a Country .....	5
1.6.1 Benefits to Home Country .....	5
1.6.2 Benefits to Host Country .....	6
1.7 Advantages of Foreign Direct Investments .....	6
1.8 Disadvantages of Foreign Direct Investments.....	8
1.9 Role of FDI for the Economic Development of a Country .....	9
1.10 Factors Affecting Foreign Direct Investments .....	10
1.11 Objective of the Research.....	13
1.12 Hypothesis Tested on the Research.....	14
1.13 Significant of the Research .....	14
1.14 Summary of the Chapter .....	14
1.15 Outline of the Dissertation .....	15

<b>CHAPTER 2</b> .....	16
<b>LITERATURE REVIEW</b> .....	16
2.1 Introduction .....	16
2.2 Theoretical Framework of FDI .....	16
2.3 Empirical Works Based on the Foreign Direct Investment.....	17
2.4 Summary of the Chapter.....	23
<b>CHAPTER 3</b> .....	24
<b>MATERIAL AND METHODS</b> .....	24
3.1. Introduction .....	24
3.2 Conceptualization of Framework Followed in the Research .....	24
3.3 Secondary Data .....	25
3.4 Definition of the Indices.....	26
3.4.1 Foreign direct investment, net inflows (BOP, current US\$) - .....	26
3.4.2 GNI Growth (Annual %) - .....	27
3.4.3 GNI per capita, Atlas method (current US\$)-.....	27
3.4.4 Labor force with tertiary education (% of total) - .....	27
3.4.5 Wage and salaried workers, total (% of total employed) - .....	27
3.4.6 Total tax rate (% of profit) - .....	28
3.4.7 Central government debt, total (% of GDP) - .....	28
3.4.8 Procedures required to start a business (number) - .....	28
3.4.9 Time required to start a business (days)- .....	28
3.5 Statistical Methodology Used .....	29
3.5.1 Modeling via Multiple Regression .....	29
3.5.2 Hypothesis testing in multiple linear regression.....	30
3.5.3 Problems in multiple regression models.....	31
3.5.4 Model Selection Process.....	32
3.5.5 Model Diagnostics methods .....	33
3.6 Sampling Methods.....	36
3.6.1 Stratified sampling method.....	36
3.6.2 Mean and the Variance of the Estimator .....	38

3.6.3 Proportionate stratified random sample.....	39
3.7 Factor Analysis.....	39
3.7.1 Basic Theory of FA .....	40
3.7.2 Appropriateness of Data for FA .....	40
3.7.3 Extraction of Factors.....	41
3.7.4 Identification of Number of Factors .....	41
3.7.5 Rotating Factors.....	42
3.7.6 Factor Score Coefficients .....	43
3.8 Statistical Tools Used in Data Analysis. ....	43
<b>CHAPTER 4</b> .....	<b>45</b>
<b>RESULT AND DISCUSSION</b> .....	<b>45</b>
4.1 Introduction .....	45
4.1.1 Mapping of Variables. ....	45
4.2 Sampling Procedure and Sample Size.....	45
4.2.1 Sample Size .....	47
4.2.2 Distribution of Sample Size.....	47
4.3 Identification of Significant Factors.....	48
4.4 Diagnostic Tests for the selected Model .....	59
4.4.1 Randomness.....	59
4.4.2 Constant Variance.....	60
4.4.3 Normality.....	60
4.4.4. Multicollinearity .....	61
4.5 Detecting of Influential Points .....	61
4.5.1 Cook’s distance statistics.....	61
4.5.2 DFITS .....	62
4.5.3 Leverage .....	62
4.5.4 Externally Studentized residuals (RSTUDENT).....	62
4.6 Interpretation of the Model .....	63
4.7 Identification of Common Factors among Macro-Economic Variables .....	63
4.7.1 Association among variable.....	64
4.7.2 Number of Factors .....	65
4.7.3 Identification of Factor Loadings .....	66

4.7.4 Defining the Factors .....	70
4.7.5 Identification of Linear Relationship among Factors .....	71
4.8 Testing the Hypothesis. ....	72
4.9 Summary of the Chapter.....	72
<b>CHAPTER 5</b> .....	<b>74</b>
<b>CONCLUSIONS AND RECOMMENDATIONS</b> .....	<b>74</b>
5.1 Conclusions .....	74
5.2 Recommendation.....	76
<b>REFERENCES</b> .....	<b>77</b>
<b>ANNEXURE 1</b> .....	<b>88</b>
<b>ANNEXURE 2</b> .....	<b>89</b>
<b>ANNEXURE 3</b> .....	<b>97</b>



## LIST OF TABLES

Table: 3.1 Indices for identifying influential factors. ....	25
Table: 3.2 Sources for the Data.....	25
Table 3.3: Interpretation of VIF.....	36
Table 4.1 Descriptive statistics of FDI amount .....	46
Table 4.2 Range of FDI amount for the selected 5 strata .....	46
Table 4.3 Distribution of sample size .....	47
Table 4.4 Descriptive statistics of dependent and explanatory variables .....	48
Table 4.5 Output for Data Set 1 .....	49
Table 4.5.1 Model Summary.....	49
Table 4.5.2 ANOVA for the fitted model .....	49
Table 4.5.3 Significance test for the coefficients .....	49
Table 4.6 Output for Data Set 2 .....	50
4.6.1 Model Summary .....	50
4.6.2 ANOVA for the fitted model .....	50
4.6.3 Significance test for the coefficients .....	50
Table 4.7 Output for Data Set 3 .....	51
4.7.1 Model Summary .....	51
4.7.2 ANOVA for the fitted model .....	51
4.7.3 Significant test for the coefficients.....	51
Table 4.8 Output for Data Set 4 .....	52
4.8.1 Model Summary .....	52
4.8.2 ANOVA for the fitted model. ....	52
4.8.3 Significant test for the coefficients.....	52
Table 4.9 Output for Data Set 5 .....	52
4.9.1 Model Summary .....	52
4.9.2 ANOVA for the fitted model .....	53
4.9.3 Significant test for coefficients.....	53

Table 4.10 Output for Data Set 6 .....	53
4.10.1 Model Summary .....	53
4.10.2 ANOVA for the fitted model .....	53
4.10.3 Significant test for the coefficients.....	54
Table 4.11 Output for Data Set 7 .....	54
4.11.1 Model Summary .....	54
4.11.2 ANOVA for the fitted model .....	54
4.11.3 Significant test for the coefficient. ....	54
Table 4.12 Output for Data Set 8 .....	55
Table 4.12.1 Model Summary .....	55
Table 4.12.2 ANOVA for the fitted model .....	55
4.12.3 Significant test for the coefficient .....	55
Table 4.13 Output for Data Set 9 .....	56
4.13.1 Model Summary .....	56
4.13.2 ANOVA for the fitted model .....	57
4.13.3 Significant test for the coefficients.....	57
Table 4.14 Output for Data Set 10 .....	57
4.14.1 Model Summary .....	57
4.14.2 ANOVA for the fitted model .....	58
4.14.3 Significant test for the coefficients.....	58
Table 4.15 List of significance variable for the models.....	58
Table 4.16 Tests of Normality .....	60
Table4.17 VIF values of the explanatory variables in the model .....	61
Table 4.18 Secondary Model of the model building process.....	63
4.18.1 Model summary.....	63
4.18.2 ANOVA for the fitted model .....	63
4.18.3 Significant test for the coefficients.....	63
Table 4.19 Correlation Matrix .....	64
Table 4.20 - KMO and Bartlett's Test.....	65
Table 4.21 - Eigenvalues Analysis.....	65

Table 4.22 Factor loading of three factor model extracted using PCF and rotated via varimax. ....	66
Table 4.23 Factor loading of three factor model extracted using PCF and rotated via quartimax. ....	67
Table 4.24 Factor loading of three factor model extracted using PCF and rotated via equamax. ....	68
Table 4.25 Summary of the result obtained from three rotation types. ....	69
Table 4.26 Component Score Coefficient Matrix .....	70
Table 4.27 Factor Regression Model .....	71
4.27.1 Model summary.....	71
4.27.2 ANOVA for the fitted model .....	71
4.27.3 Significant test for the coefficients.....	71

## LIST OF FIGURES

Figure 3.1 Conceptual framework of the study.....	24
Figure 3.2 three dimensional interpretations for multiple linear regressions .....	30
Figure 3.3 Stratified sampling methodology.....	37
Figure 4.1 Scatter plot of Predicted value vs. standardized residuals.....	60
Figure 4.2 Scree plot .....	66

## LIST OF EQUATIONS

Equation 3.1: Common multiple regression model.	29
Equation 3.2 matrix form of the multiple regression model	30
Equation 3.3 Estimator value of Y	30
Equation 3.4 F statistics	31
Equation 3.5 test statics for MR model	31
Equation 3.6 Estimator of X	38
Equation 3.7 Variance estimator of X	38
Equation 3.8 Total Estimator of X	38
Equation 3.9 Formulas for stratified sampling	39
Equation 4.1 Formula for sample size	46
Equation 4.2 Equation of the first model	49
Equation 4.3 Equation of the second model	50
Equation 4.4 Equation of the third model	51
Equation 4.5 Equation of the fourth model	52
Equation 4.6 Equation of the fifth model	53
Equation 4.7 Equation of the sixth model	53
Equation 4.8 Equation of the seventh model	54
Equation 4.9 Equation of the eighth model	55
Equation 4.10 Equation of the ninth model	56
Equation 4.11 Equation of the tenth model	57
Equation 4.12 First factor model equation	70
Equation 4.13 second factor model equation	70

Equation 4.14 Third factor model equation	70
Equation 4.15 Composite factor model equation 1	70
Equation 4.16 Composite factor model equation 2	70
Equation 4.17 Composite factor model equation 3	70

## **LIST OF ACRONYMS**

ANOVA Table: Analysis of Variance Table.

BOP: Balance of Payment.

BRICS- Brazil, Russia, India, China and South Africa.

CGE: Computable General Equilibrium

DW Statistics: Durbin Watson Statistics.

ECOWAS: Economic Community of West African States

EU: European Union Countries

FA: Factor Analysis

FDI: Foreign Direct Investments.

GLS: Generalized least squares

GNI: Gross National Income.

KMO Test: Kaiser-Meyer-Olkin Test.

MANOVA: Multivariate analysis of variance.

MENA: Middle East and North African Countries

MNE: Multinational Enterprises.

MR: Multiple Regression.

OECD: The Organization for Economic Co-operation and Development.

OLS: Ordinary least squares

PCA: Principle Component Analysis.

SSR: Residual Sum of Squares.

SST: Total Sum of Squares.

UK: United Kingdom.

US: United States.

USD: United States Dollar.

VIF: Variance Inflation Factor.