Predicting the All Share Price Index (ASPI) of Colombo Stock Exchange on Macro Economic Variables and Exogenous Factors

DPPN Peiris[1], GP Lakraj[2]

[1] *Department of Statistics, University of Colombo, Sri Lanka*
[2] *Department of Statistics University of Colombo, Sri Lanka*
[#] GP Lakraj; <pemantha@stat.cmb.ac.lk>

Abstract**—***This study mainly focuses on developing a statistical model to predict ASPI with a reasonable accuracy and identifying the relationships between ASPI and other important socio-economic variables, such as macroeconomic variables, elections and war. The key results of this study include that the absence of the war has a positive relationship with ASPI compared to the presence of the war and the presence of severe bomb blasts in Colombo. The presence of presidential and general elections and the crude oil price have a positive relationship with ASPI. Furthermore, presence of the provincial elections, Imports, exports and exchange rates have a negative relationship with ASPI. Moreover, it was found that during the post war period the conditional heteroscedasticity was absent in ASPI. In addition to that the model which was chosen to predict the ASPI has a Mean Absolute Percentage error (MAPE) of 4.05% on the test set*

Keywords**—* Colombo Stock Exchange, ASPI, Macro Economic Variables, Elections, War, Conditional Heteroscedasticity**

I. INTRODUCTION

The Colombo Stock Exchange is one of the most preferred establishment of investors. Investors have a keen interest on predicting the future behaviour of the stock market. Moreover, they have a high curiosity to find out what really drives the stock markets and how the macroeconomic variables are related with the stock market. The study mainly concentrates on the performance of the Colombo Stock Exchange based on the ASPI.

ASPI is the broad market index of the CSE, and is designed to measure the movements of the overall market. The index is calculated in real-time as a market capitalization weighted index, which constitutes all voting and non-voting ordinary shares listed on the CSE (Colombo-Stock-Exchange, 2018).

$$ASPI = \frac{Market\ capitalization\ of\ all\ listed\ companies}{Base\ market\ capitalization} \times 100$$

Macroeconomics is a branch of economics that concerns itself with market systems that operate on a large scale. It deals with the performance, structure and behaviour of the entire economy. When investors talk about macroeconomics, policy decisions like raising or lowering interest rates or changing tax rates are discussed. The macroeconomic variables which were used in this study can be brought forward as follows.

1. Foreign exchange rate of United States Dollar
2. Imports
3. Exports

Factors that used in the study which are not considered as economic factors are defined as exogenous factors. Even though crude oil price is related to the international economy, it is not controlled according to the Sri Lankan economic factors. Crude oil price has been inserted into this study in order to include international influence. Below mentioned are the exogenous variables that were used.

1. Elections
2. War
3. Crude Oil Price

The main objective of this study is to predict the ASPI value under the given conditions of the macro economic variables and exogenous factors with a reasonable accuracy. In addition to that the following secondary objectives are achieved as well.

1. Identify the interactional relationships among the macro-economic variables.
2. Identify the behaviour of the ASPI on the period that the war was present and on the post war period.

Since the stock market has been a popular way of investing for a long time, it has become a popular topic among the literature as well. P. Sadorsky in his study, has investigated the dynamic interaction between the oil price and other economic variables which also included the stock returns (Sadorsky, 1999). M. Chesney, G. Reshetar and M. Karaman have studied the impact of several significant terror attacks on international stock markets, bond markets and commodity markets (Chesney, et al., 2011).

II. METHODOLOGY AND EXPERIMENTAL DESIGN

A. Time Series and Time Series Regression.

According to G.E.P. Box, G.M. Jenkins and G.C. Reinsell (Box, et al., 2008) in their book Time Series Analysis: Forecasting and Control**,** time series is a sequence of observations taken sequentially in time. According to them, adjacent observations being dependent is an inherent feature of time series. Time series analysis is concerned with techniques for the analysis of the above mentioned dependencies (Box, et al., 2008). There are different kind of time series data models which can be used according to the characteristics of the response variables. Hence some of them can be named as, Auto Regressive, Moving Average as well as combinations of them such as Auto Regressive Moving Average (ARMA) and Auto Regressive Integrated Moving Average (ARIMA) models.

More often, response variables depends on other variables, in such a way that the variation of the response variable can be explained by the variation of those other variables. With such occurrence in a time dependent variable, the concept of combining time series analysis and regression analysis arises. The concept behind this scenario can be described as follows (Fuller, 1976).

1. As the initial step the response variable is described under the explanatory variables and the coefficients are calculated for the regression equation.
2. Then the residual series of the fitted regression model is extracted and then a time series model is fitted to the residual series. (In usual regression analysis the residual series is assumed to be a normally distributed random variable having zero mean and a constant variance. But in this concept the residual series is assumed to be auto-correlated.)
3. The final step is estimating parameters of regression model and time series model for errors simultaneously.
4. The mathematical representation of a time series regression model having an error term of ARIMA(1,1,1) is as follows

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n + m_t$$

$$(1 - \phi_1 B)(1 - B) m_t = (1 + \theta_1 B) e_t$$

$$where\ e_t\ is\ a\ white\ noise\ process$$

B. Volatility and Conditional Heteroscedasticity.

The term "Volatility" refers to the amount of uncertainty or risk related to the size of changes in a security's value. In more of a statistical background the term can be defined as a measure of the dispersion of the returns for a provided market index or a security. Volatility can be measured either by using Standard Deviation or the Variance between the returns of the provided market index or the security. As a common practice it is defined that, higher the volatility higher the risk of the security (Tsay, 2005).

According to Engle, traditional econometric models assume a constant one-period forecast variance. To generalize this implausible assumption, a new class of stochastic processes called autoregressive conditional heteroscedastic (ARCH) processes were introduced in Engle's paper. According to Engle these were mean zero, serially uncorrelated processes with non-constant variances conditional on the past. For such processes, the recent past gives information about the one-period forecast variance. A regression model is then introduced with disturbances following an ARCH process. This model was used by Engle to estimate the means and the variances of the inflation in the U.K (Engle, 1982).

There are two tests to identify this ARCH effect. The first method is to apply the Ljung-Box statistic to the squared residual series. The second method is to apply the Lagrange Multiplier test of Engle which is also known as the ARCH-LM test. This test is equivalent to the F test used in regression to test the significance of the parameters.

Even though the ARCH model is simple it has a trade-off between the number of parameters, which are needed to adequately describe the volatility process. Taking that into concern, Bollerslev (1986) proposed an extension for the ARCH model which is known as the Generalized ARCH model (GARCH Model). Here the dependent volatility is describes as simple quadratic functions of its lagged values as well as quadratic functions of the variance's lagged values.

$$r_t = \mathrm{ARIMA(p, d, q)} + e_t$$

$$e_t = z_t \, {\sigma^2}_t$$

$${\sigma^2}_t = \omega_0 + \sum_{i=1}^{m} \alpha_i {e^2}_{t-i} + \sum_{j=1}^{n} \beta_j {\sigma^2}_{t-j}$$

Here r_t denotes the response variable, ASPI considering this study. e_t denotes the residual series where ${\sigma^2}_t$ denotes the variance of the residual at the lag t and z_t is a sequence of independent and identically distributed random variables with zero mean and unity variance. In practice, z_t is often assumed to follow the standard normal or a standardized Student-t distribution or a generalized error distribution

C. Goodness of Fit Test and Prediction Accuracy

Adjusted Pearson chi-square test which is an adjusted extension of the chi-square goodness of fit test has been used as the goodness of fit test in the study. The test has been used in the paper published by Peter J.G. Vlaar and Franz C. Palm (Vlaar & Palm, 1993). The prediction accuracy of the fitted models throughout this study were measured using the Mean Absolute Percentage Error.

D. Experimental Design.

1) Variable Selection: Before stepping into the modelling process it is necessary to have proper justification on the influence of the selected variables on the response. For this "Best Subset Selection Algorithm" is used. Variables which provided the best adjusted R square value were used.

In this study, apart from the presence and the absence of war, bomb attacks that happened in the district of Colombo are considered as a special case. Since Colombo district is considered as the heart of the Sri Lankan economy, there are hesitations that, changes happen in Colombo will have an influence on the country's economy. As the limits of the selections were too broad, the following criteria were used as a selection frame.

1. The bomb blast should be within Colombo district.
2. It should be an attack which affected general public, not a specific person. Even though the attack is for an individual and if the public had been affected, then the blast is taken into consideration.
3. It should be within the time frame 1994-Jan to 2017-Dec.

Following the above constraints 11 bomb blasts were selected. The details of the blasts are mentioned in table 1. Even though Katunayaka International Airport bomb blast took place in Gampaha district, it was considered, as it was the only international airport at the time.

Table 1. Selected Bomb Blasts

Attack	Date	Location
Thotalanga Bombing (Assassination of Gamini Dissanayake)	1994-10-24	Thotalanga, Colomb
Attack to the oil storage complexes at Kolonnawa and	1995-10-20	Kolonnawa,
Colombo Central Bank bombing	1996-01-31	Central Bank of
1996 Dehiwala train bombing	1996-07-24	Dehiwala,
1997 Colombo World Trade Centre bombing	1997-10-15	Colombo World
Colombo Town Hall Bombing (Assassination attempt on Chandrika Bandaranaike	1999-12-18	Colombo Town Hall
Bandaranaike Airport attack	2001-07-24	Katunayake,
Nugegoda shopping mall bombing	2007-11-28	Nugegoda
2008 Fort Railway Station bombing	2008-02-03	Fort Railway
Piliyandala bus bombing	2008-04-26	Piliyandala,
2008 Dehiwala train bombing	2008-05-26	Dehiwala,

Study was split into two parts as the concern is on predicting as well as identifying the influences of the factors. One path considers the best subset of the variable which will give the best prediction while the other path considers all the variables and identify their influence over ASPI.

To identify the influence of the war separately, the time window of the response variable was divided in to two parts where one window represent the movement of ASPI during the period of war and the other window represent the movement of the ASPI after the war. Then it was followed by a comparison of the two movements in order to identify the influence of war (presence and absence).

III. Results

ASPI from 1994 to 2017 which is the response variable of this study has ARCH effect, which eventually implies that the variance of the series is conditional.

As it was mentioned in the previous section this study consists of two parts where one path considers all the variables to fit a model and the other path consider a subset of a variables to fit a model. The subset of variables that were used in the aforesaid path are mentioned below.

1. ElectionsProv – Dummy variable for the Provincial Elections
2. TerrNo_Terr – Dummy variable for the absence of terrorism
3. CrudeOil_LKR – Crude Oil Price in Sri Lankan Rupees
4. ExchRate – Exchange Rate of United States Dollar
5. Imports_LKR_mill – Imports in Sri Lankan Rupees millions
6. Exports_LKR_mill - Exports in Sri Lankan Rupees millions

Table 2 depicts the performance measures of the fitted GARCH models from GARCH order from (1,1) to (2,3) with all the variables.

Table 2. Performance measure of the models with all the variables

Model	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)
RMSE	299.65	299.63	364.12	392.42	392.42	384.36

MAPE	4.38	4.38	5.30	5.73	5.73	5.61

GARCH (1,2) model is the selected model with all the variables which was used to describe the influence of the variables towards ASPI. The following equation represent the model while Table 3 shows the parameters of the selected model with all the variables.

$$X_t = 1425.2653 + 10.4761 * \text{TerrTerr\_North} + 100 * \text{TerrNo\_Terr} + 26.4535 * \text{Election\_Presi} + 18.2216 * \text{Election\_Gen} - 11.3087 * \text{Election\_Prov} + 0.0277 * \text{CrudeOil} - 0.0004 * \text{Exports} - 0.0016 * \text{Imports} - 2.1182 * \text{Exchange} + 0.7886 * X_{t-1} + 0.2134 * X_{t-2} + 0.2822 * \varepsilon_{t-1} + \varepsilon_t$$

$$\sigma^2_t = 113.4334 + 0.2209 * \varepsilon_{t-1}^2 + 0.8109 * \sigma^2_{t-1} + 0 * \sigma^2_{t-2}$$

Table 3. Parameters of the selected model with all the variables

Parameter	Estimate	Standard Error
Mu	1425.2652	259.0400
Ar1	0.7886	0.0244
Ar2	0.2134	0.0256
Ma1	0.2822	0.0809
TerrTerr_North	10.4761	14.7039
TerrNo_Terr	100.00	142.6266
Eelction_Presi	26.4535	21.5743
Election_Gen	18.2216	15.8806
Election_Prov	-11.308	13.3437
CrudeOil	0.028	0.0186
Exports	-0.0004	0.006
Imports	-0.0016	0.007
Exch	-2.1181	4.7821
Omega	113.433	100.1956
Alpha1	0.2209	0.1435
Beta1	0.8109	0.8625
Beta2	0.0000	0.7470

Table 4 depicts the performance measures of the fitted GARCH models from GARCH order (1,1) to (2,3) with the selected subset of the variables. The MAPE of the selected model is 4.05%.

Table 4. Performance measure of the models with the selected variables

Model	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)
RMSE	307.77	277.46	277.47	379.42	379.42	334.74
MAPE	4.49	4.05	4.05	5.54	5.54	4.86

The model with the GARCH order of (1,2) was selected as the predictive model with the selected subset of variables as it has the lowest MAPE score. The following equation represents the model while Table 5 shows the parameter values of the selected predictive model.

$$X_t = 1432.7179 + 100 * \text{TerrNo\_Terr} - 9.5713 * \text{Elections\_Prov} + 0.0238 * \text{CrudeOil} - 0.0005 * \text{Export} - 0.0017 * \text{Import} - 2.4891 * \text{Exchange} + 0.6972 * X_{t-1} + 0.3043 * X_{t-2} + 0.3918 * \varepsilon_{t-1} + 0.0692 * \varepsilon_{t-2} + 0.0096 * \varepsilon_{t-3} + \varepsilon_t$$

$$\sigma^2_t = 123.345 + 0.2188 * \varepsilon_{t-1}^2 + 0.8112 * \sigma^2_{t-1} + 0 * \sigma^2_{t-2}$$

Table 5. Parameters of the predictive model

Parameter	Estimation	Standard Error
Mu	1425.2653	280.4193
Ar1	0.6972	0.0321
Ar2	0.3043	0.0429
Ma1	0.3918	0.1919
Ma2	0.0691	0.0653
Ma3	0.0096	0.1748
TerrNo_Terr	100.00	262.0947
Elections_Prov	-9.5712	10.6067
CrudeOil	0.0238	0.0335
Exports	-0.0005	0.0009
Imports	-0.0017	0.0010
Exch	-2.4891	5.0961
Omega	123.3450	266.8363
Alpha1	0.2188	0.5473
Beta1	0.8112	3.1762
Beta2	0.0000	2.7255

Influence of the terrorism towards the ASPI was analysed using two time windows as mentioned in the previous section.

According to the results of the two tests that has been carried out to test the ARCH effect of the ASPI which was on the presence of war, concluded that there was ARCH effect in the response variable. Tables 6 and 7 respectively represent the results of Ljung-Box test and ARCH-LM test.

Table 6. Result of the Ljung-Box test when the war is present

Order	P-Value
4	7.87e-13
8	2.40e-11
12	1.60e-11
16	8.13e-14
20	1.11e16
24	0.00e+0

Table 7. Result of the ARCH LM test when the war is present

Order	P-Value
4	1.90e-09
8	1.14e-02
12	4.57e-01
16	9.73e-01
20	1.00e+00

24	1.00e+00

Even though the presence of war resulted a conditional heteroscedasticity of ASPI, the ASPI series during the absence of war does not show the property of conditional heteroscedasticity. The following Tables 8 and 9 presents the results of the two test for the ASPI series when the war is not present.

Table 8. Result of the Ljung-Box test when the war is not present

Order	P-Value
4	0.597
8	0.891
12	0.740
16	0.792
20	0.927
24	0.839

Table 9. Result of the ARCH LM test when the war is not present

Order	P-Value
4	3.55e-12
8	1.65 e-03
12	5.11e-01
16	9.67e-01
20	1.00e+00
24	1.00e+00

IV. DISCUSSION AND CONCLUSIONS

The first step of the predictive model building methodology of this study was to provide proper evidence that the selected variables are really related with the response variable. Hence the best subset selection algorithm was used.

Surprisingly, the algorithm left out some of the variables which was identified as important variables from previous studies (Jayasinghe, 2014). The dummy variables for the presence of the general and presidential elections as well as the dummy variables for the presence of war and bomb blasts were left out from the model.

However, (Jayasinghe, 2014) clearly concludes in his study that that there are abnormal returns and change in returns relative to the previous day's returns during election periods. Which indirectly implies the influence of the elections towards the performance of the stock market.

If this fact is discussed more deeply it can be found that there are some major differences between this study and the study carried out by him. The first and the most important difference is the frequency of the data. In his study (Jayasinghe, 2014) has used daily values, while in this study the data is in monthly basis. When this fact considered deeply, it can be argued that the influence that can be felt upon the ASPI for several days might be significant. But when the gap between two data points is a month, which is a relatively higher gap considered to daily data points, might miss the significant movements that take place between several days. Thus, the argument of this contradiction between the results can be converted into a conclusion that the influence of the elections towards the performance of the stock is only for a short period of time but not for months or larger periods of time. The parameter estimations for the absence of war is positive with a relatively high value. In other words the absence of the war has a positive influence with a relatively high magnitude on ASPI compared to the presence of severe bomb blasts in the Colombo district. The relatively high positive influence of the absence of war upon ASPI was observed as soon as the war ended. According to Colombo Stock Exchange, as an emerging stock market CSE was the best performing stock exchange in the world in 2009 as it jumped 125.2 % during that year.

Going in the same direction with (Jayasinghe, 2014) the presence of the presidential and general elections has a positive influence on the ASPI compared to the absence of elections. But with addition to (Jayasinghe, 2014) this study considers the provincial elections as well. And surprisingly presence of provincial elections has a negative influence on the ASPI with relative to the absence of elections. When such an election comes, the governing party will provide concessionaries toward the public as well as the business community most of the time. Which will result a good performance in the market.

Oil price has a positive effect towards the ASPI. It is a quite interesting result when it is compared to (Billmeier & Massa, 2009), since in their study they indicated that the increase in the oil prices had positive impact on the stock yields of the oil exporter countries but negative impact on the stocks of the oil importer countries. Contrary to the results of them, Sri Lanka being an oil importer country ended up with a positive influence upon the increase of the crude oil price.

Exchange rate also has a negative influence on the ASPI. It is a known fact that for a long time Sri Lanka has been importing goods and services more than exporting. Which will result of a higher demand in the foreign currency. When the country as a whole depends on imported goods, a reasonable amount of the revenue of the companies flows to the foreign markets. Moreover, when the currency rate increases the expenditures of companies are increased and will reduce the cash flow of the companies which will eventually decreases the dividends and the performance of the stock market.

Imports and exports also exhibit a negative influence upon the ASPI. And imports have a higher magnitude relative to the exports. Increase of the imports of a country implies either lack of local production or higher demand to the imported product. Both scenarios affect the local companies in a negative manner. Similar to the imports, increased number of exports implies either a production with a higher demand outside the country or a lower demand inside the country compared to the demand outside the country. Both situations lead the producer to rely on the international market rather than the local market. Which will eventually create a false demand for the product in the local market. This situation will lead the other companies to rely on the international market due to the shortage in the local market. This will eventually reduce the income of the companies and will decrease the performance of the stock market as well.

The war period was a crucial period for the Sri Lankan economy. The period is considered as crucial since the expenses of the war was very high at the time. In addition to that, foreign interventions in political perspective and economical perspective could be seen at the time. Moreover, investments were relatively low in the country due to the instability and the lack of peace in the island. These fact were the highly possible reasons behind the economic and political instability of the country. Given these facts, there was always a risk that act of terrorism to occur unnoticed. These facts being considered as risks in Sri Lankan economy, the performance of the stock market was also subjected to sudden changes, which was the main reason behind the conditional heteroscedasticity.

A list of the conclusions drawn from the study can be summarized as follows.

The predictive model that can predict the ASPI is the ARIMA (2,0,3)-GARCH(1,2) model with the selected features with an RMSE value of 277.46 and an MAPE of 4.05%. The selected features are,

1. Dummy variable for the presence of the war
2. Dummy variable for the presence of Provincial Elections
3. Crude Oil Price
4. Exports
5. Imports
6. Exchange Rate

Presidential Elections and General Elections were left out from the feature selection algorithm. But the full model with all the variables revealed that both election types has a positive influence on the ASPI with general elections having a slight lower magnitude than the presidential elections, when compared to the absence of the elections. Provincial elections has a negative influence but with a slightly lower magnitude than presidential and general elections towards the ASPI when compared to the absence of elections.

Absence of the war has a higher influence towards ASPI when compared to the presence of the war and presence of severe bomb attacks in the Colombo District. The absence of war has stabilized the performance of the stock market by removing the conditional heteroscedasticity which was present when the war is present.

Crude oil price has a positive influence toward the ASPI whereas the Export, Imports and Exchange Rates have a negative influence upon the ASPI.

ACKNOWLEDGEMENT

Conducting this study and composing this paper has been one of the most important and challenging tasks that I have been gone through. This small appreciation is for the numerous helping hands who were around me throughout this task. First and foremost I must convey my heartfelt gratitude to my supervisor who has been guiding me throughout this journey. I would like to extend my gratitude to Department of Statistics University of Colombo for their enormous support they have given me from the very first day of my undergraduate life. Finally I would like to thank each and every person that had not been mentioned in this text, who had supported me even in a single word throughout my life.

REFRENCES

Billmeier, A. & Massa, I., 2009. What drives stock market development in emerging markets-institutions,remittances or natural resources?. *emerging Markets Review,* 10(1), pp. 23-35.

Box, G. E. P., Reinsell, G. M. & Jenkins, G. M., 2008. *Time Series Analysis: Forecasting and Control, Fourth Edition.* s.l.:Wiley & Sons, Inc..

Chesney, M., Reshetar, G. & Karaman, M., 2011. The Impact of Terrorism on FinancialMarkets: An Empirical Study. *Journal of Banking and Finance,* Volume 35, pp. 253-267.

Colombo-Stock-Exchange, 2018. *Colombo SStock Exchange.* [Online]
Available at: https://www.cse.lk/home/market
[Accessed 2018].

Engle, R., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflations. *Econometrica,* 50(4), pp. 987-1007.

Fuller, W., 1976. *Introduction to Statistical Time series.* s.l.:John wiley & Sons Inc..

Jayasinghe, P., 2014. The Behavior of Stock Returns and Volatility around Elections: Evidence from Colombo Stock Exchange. *Journal of Multidisciplinary Research,* 1(1).

Sadorsky, P., 1999. Oil price shocks and stock market activity. *Energy Economics,* pp. 449-469.

Tsay, R. S., 2005. *Analysis of Financial Time Series.* s.l.:John Wiley & Sons Inc..

Vlaar, J. G. & Palm, F. C., 1993. The Message in Weekly Exchange Rates in the Europen Monetary System: Mean Reversion, Conditional Heteroscedasticity and Jumps. *Journal of Business & Economic Statistics,* 11(3), pp. 351-360.