Univariate and Multivariate Financial Time Series Analysis

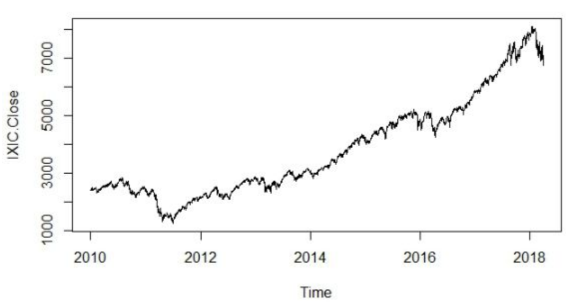
Jaechan Park

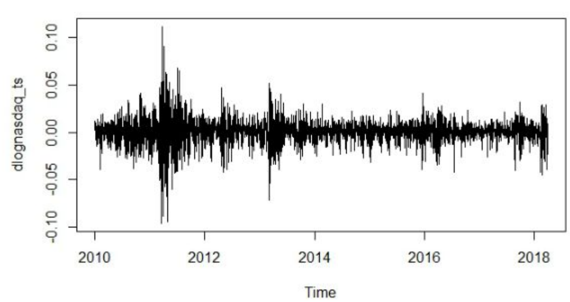
E-mail : figo721@hotmail.com

1 Univariate Time Series Analysis

1.1 Time series plot and Log-transformation

A fundamental approach for time series analysis is a creation of a time series plot which is a plot with observed data points (on the response axis) versus xed time order. In this study, the daily close index of the NASDAQ market was used to analyze the time series model and its forecasting. In Figure 1, the time series plot is for daily NASDAQ close index from the rst trading day in January 2018 to current time (as of 19.Dec.2018). From the gure, the time series shows that it does not have periodical behavior. As well, the time series plot has uctuation with an upward trend which is a strong indication for non-stationary. A major problem of a non-stationary time series has often violated either normality or homoscedasticity or both.





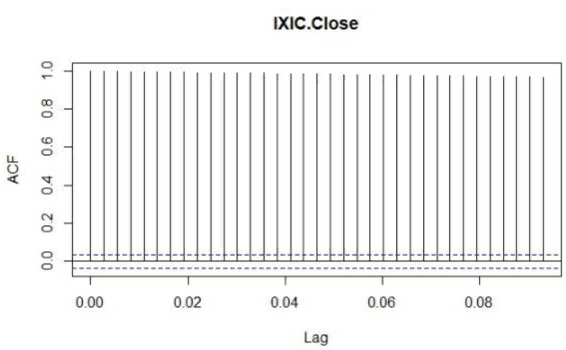
(a) Figure 1: daily Nasdaq close index

(b) Figure 2: di erence for log transformation

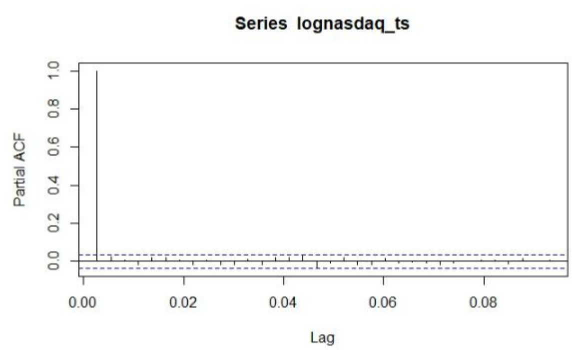
A log-transformation is commonly used to address non-stationary time series in statistical analysis. After log-transformation, the time series still represents a non-stationary. Hence, it is natural to consider making the rst order di erence in the log-transformed time series for stationary. As seen in gure 2, the time series plot, the rst order di erence in the log transformation, represents a stationary stochastic process which is an indication to build a time series model for forecasting. Indeed, it is a strong signal the log-transformed time series has the rst order di erence in the model to minimize the standard deviation by tting AR(I) MA model for its forecast.

1.2 AR(I)MA and Model validation

In time series analysis, the AR(I)MA model is a widely used statistical method for forecasting time series models. To an AR(I)MA model, it is essential to plot both autocorrelation function (ACF) and partial autocorrelation function (PACF). Based on the ACF plot in gure 3, it is not appropriate to model with moving average for the time series. Apart from the ACF plot. PACF plot represents the time series assumes ARIMA(1,1,0) seems to be a proper model for the from time series for now. In this section, ARIMA(1,1,) and ARIMA(2,1,0) models will be tested for the log-transformed time series and then and a proper model to be used in this study.



(c) Figure 3: ACF for log Nasdaq close index



(d) Figure 4: PACF for log Nasdaq close index

Table 1: Q-test results

Model	p-value
ARIMA(1,1,0)	0.03176
ARIMA(2,1,0)	0.3584

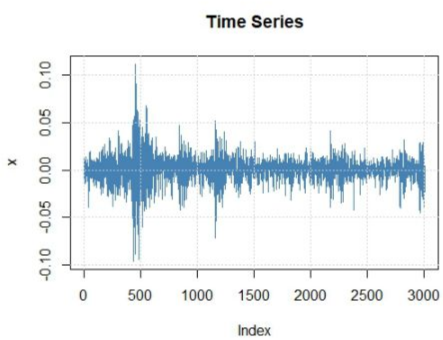
Table 2: Diebold-Mariano Test

Result	p-value
dm.test result	0.02816

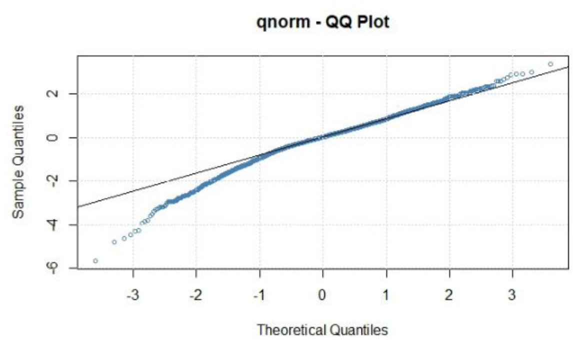
As see table 1, the p-value for unit root test of ARIMA(1,1,0) has 0.03176 which means rejecting the null hypothesis and the ARIMA(1,1,0) model is not valid for the log-transformed time series. While ARIMA(2,1,0) is valid as its p-value (0.3584) is greater than its type I error (=0.05). Therefore, it can conclude the ARIMA(2,1,0) model is a rather proper model than the ARIMA(1,1,0) model in this study. To make a comparison between the two models, Diebold-Mariano test statistics (table 2) has been conducted and the p-value (0.02816) is less than 5%. Hence, it is sufficient to determine that the two models are different.

1.3 GARCH Model Estimation

(G)ARCH is an acronym meaning Generalized Autoregressive Conditional Heteroskedastic model is commonly used for time series with volatility (conditional variance). GARCH model is also known as a model for time series model in the nancial data set. As the GARCH(1,1) model is often su cient model for time series, ARIMA(2,1,0)-GARCH(1,1) is a proposed model for this time series analysis. Figure 5 represents the time series plot for ARIMA(2,1,0)-GARCH(1,1) model. Also, normality validation could be con rmed with its Q-Q plot in gure 6. According to the Q-Q plot, its standardized residuals plot derailed from its theoretical Q-Q normal straight line at the edges of the line. Thereby, it could be concluded the time series model is non-normality which is non-normality.



(e) Figure 5:ARIMA(2,1,0)-GARCH(1,1)



(f) Figure 6: Q-Q Plot of standardized residuals

To address the non-normality issue, Quasi-Maximum likelihood estimation, QMLE, is used in this study. From the outputs in table 5, ar2 coefficient has its p-value (0.6147), of which the coefficient is insignificant. In table 4, Jarque-Wilk Test and Shapiro-Wilk Test have zero value for p-values. Hence, the null hypotheses are rejected and conclude the model is non-normality. For the Q-test, the p-values of the standardized residuals and the squared standardized residuals have all greater than the type I error (=0.05). Therefore, the null hypotheses for the Q-test do not reject and it could be concluded the model is valid.

Table 3: Error Analysis of ARIMA(2,1,0)-GARCH(1,1) with "QMLE"

	Estimate	Std. Error	t-value	P r(> jtj)
mu	8:444e^{04}	1:808e^{04}	4.670	< 3:01e^{06}
ar1	3:969e^{02}	1:906e^{02}	-2.082	0.0373
ar2	9:908e^{03}	1:968e^{02}	-0.503	0.6147
omega	3:479e^{06}	8:101e^{07}	4.295	1:75e^{05}
alpha1	1:0681e^{01}	1:509e^{02}	7.076	1:48e^{12}
beta1	8:702e^{01}	1:628e^{02}	53.468	< 2e^{16}

Table 4: Standardized Residuals Tests of ARIMA(2,1,0)-GARCH(1,1) with "QMLE"

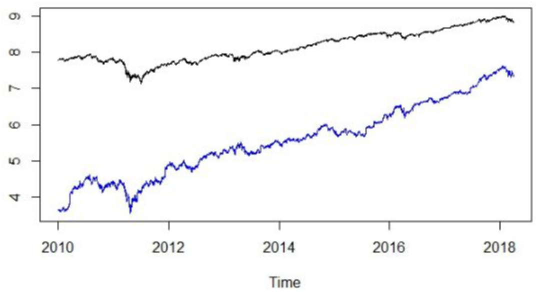
			Statistics	p-value
Jarque-Wilk Test	R	Chi^2	436.3476	0
Shapiro-Wilk Test	R	W	0.97796	0
Ljung-Box Test	R	Q(10)	12.8965	0.2295164
Ljung-Box Test	R	Q(15)	19.86186	0.17729632
Ljung-Box Test	R	Q(20)	25.50927	0.1826374
Ljung-Box Test	R^2	Q(10)	15.30524	0.12132344
Ljung-Box Test	R^2	Q(15)	23.06336	0.082803881
Ljung-Box Test	R^2	Q(20)	29.45516	0.079176248
LM Arch Test	R	T R^2	15.97094	0.1925705

2 Bivariate Time Series Analysis

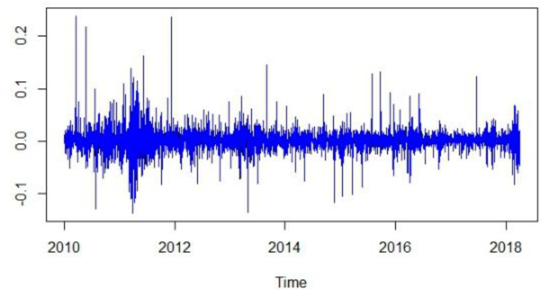
2.1 Model Estimation

Another study conducts for multivariate time series analysis to explain cointegration or to end any relationship between two time series models. In this study, the purpose of multivariate time series analysis is to con rm cointegration between two time series; NASDAQ close index (dependent variable) and Amazon daily close price (independent variable). In gure 7, the time series plot is log-transformed on the two time series models. Also, figure 8 is the first difference in the log-transformed two time series models. Based on the two previous time series plots, movement of the two time series models tend to resemble each other with the xed time interval. However, it is insufficient to confirm the two models are cointegrated.

To confirm the existence of cointegration between the two time series models, Engle-Granger Test is used in this study. The null hypothesis of the test is no cointegration and rejection of the null hypothesis is statistical evidence of cointegration in this study. Based on the output of p-value (0.0123) from Engle-Granger Test, the null hypothesis is rejected and conclude the two time series models are cointegrated each other.



(g) Figure 7: Nasdaq(black),Amazon(blue)



(h) Figure 8: di erence for log transformation of two data sets

2.2 Vector Error Correction Model (VECM)

Thanks to the earlier analysis, the two time series models are stationary when they are in the rst order di erence in log-transformation. In this case, the Vector Error Correction Model (VECM) is the most applicable model for multivariate time series analysis. To identify whether the two models are cointegrated, Johansen's test uses the following conditions; the lag order of 5 and containing a constant and the transitory e ects. The lag order of 5 is based on Schwartz Information Criterion.

Table 5: Trace test

	test	10pct	5pct	1pct
r <= 1	5.46	7.52	9.24	12.97
r = 0	20.95	17.85	19.96	24.60

Table 6: Maximum eigenvalue test

	test	10pct	5pct	1pct
r <= 1	5.46	7.52	9.24	12.97
r = 0	15.49	13.75	15.67	20.20

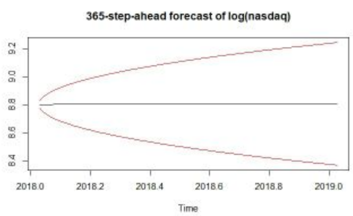
According to the output from Johansen's trace test statistic (table 5), the test statistic value (20.95) for r = 0 is greater than its critical value (19.96). As well, the test statistic value (5.46) for r = 1 is less than its critical value (9.24). The outcomes lead to conclude log(nasdaq) and log(amazon) have one cointegration relation. The cointegration equation is as below:

$$5{:}3339541 + \log(nasdaq) \quad 0{:}4835237\log(amazon) = \boldsymbol{\delta}(t) \tag{1}$$

While using Johansen's maximum eigenvalue test statistic, the null hypothesis (no cointe-gration) does not reject since the test statistic (15.49) for r = 0 is less than its critical value (15.67). For r = 1, the test statistic (5.46) is less than 9.25 (in 95%). Based on the result from the test with maximum eigenvalue test statistic, the two time series models are integrated, however, not cointegrated.

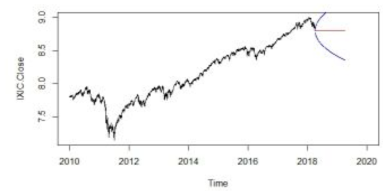
2.3 Forecasting

Figure 9 and gure 10 represent forecasting plots for the time series of log (Nasdaq) and log (amazon) with the VECM model, respectively. In addition, gure 11 is a forecast of the ARIMA (2,1,0) model of log (NASDAQ). All three groups have their prediction intervals are gradually increasing with time. Therefore, it could state the forecasts are not trustful. It is natural to conclude that the prediction of a specific stock price or a stock market index is one of the most challenging tasks in finance and econometric.



(i) Figure 9: forecast for log(nasdaq)

(j) Figure 10: forecast for log(amazon)



(k) Figure 11: ARIMA(2,1,0) forecast for log(amazon)

3 Conclusion

In this study, univariate time series analysis examined the time series model of the NAS-DAQ close index from January 2010. Due to high volatility and complexity of nancial data, single classical time series models have limitations to make an accurate interpretation and prediction. In the multivariate time series analysis, the study found cointegration relation between the NASDAQ index and Amazon stock price within a xed past time. In another word, the stock market index and Amazon stock price at the stock market, the trend of their uctuations are very similar to each other. To be simplified, when Amazon stock price closed higher, the stock market index also went up within a past time period. However, predicting future nancial price/index in multivariate case, it is one of the most difficult tasks even experts in this led.