Tse Chun Hei Vincent

Stock Index Forecast Report

1. Introduction

The forecast report focuses on forecasting of stock indexes in emerging markets by using time series analysis. In this report, Jakarta Composite Index and IBOVESPA index are chosen from Yahoo Finance for analysis. Emerging market economies like Indonesia and Brazil are rapidly growing as measured by GDP but they are also more volatile compared to mature markets. Emerging market economies are sensitive to economic and financial changes in macroeconomic environment such as interest rate change and exchange rate change. After the COVID-19 pandemic, the markets are facing capital outflows in financial markets. Hence, the report aims to study the pattern of leading stock indexes of Indonesia and Brazil between 2020 and 2021, in order to investigate the time dependence behind stock price movements after the pandemic.

2. Data and forecast horizon

For the forecast, the daily stock index data, Jakarta Composite Index and IBOVESPA from Yahoo Finance website are selected for analysis. The data are divided into two periods, the estimation period and the forecasting period. The length of the estimation period is one year, from 20th April 2020 to 16th April 2021. The forecasting period is 5 days, from 19th April 2021 to 23rd April 2021.

3. Methodology

In this report, 3 separate stock indices analysis, are made by using time series ARIMA model, which are the Jakarta Composite Index, IBOVESPA and an equally weighted portfolio with the two stock indexes. There are three stages in building the above 3 time series models, which include model identification, model estimation and model selection.

4. Model identification

The first stage involves the stationarity analysis through visual examination and Augmented Dickey Fuller test. It also involves the autocorrelation and partial autocorrelation analysis.

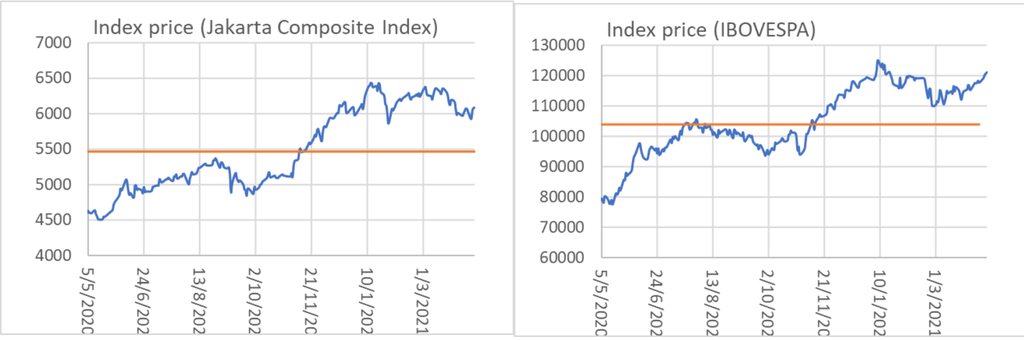
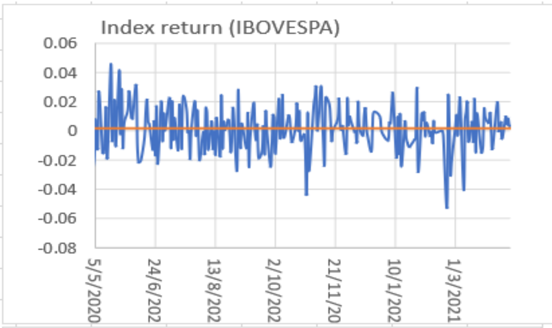
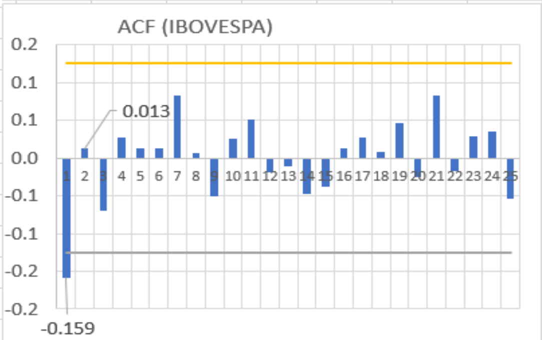


Figure 1: The Index daily prices

Through visual inspection, it is clear that Figure 1 is not covariance stationary. The data points do not have a constant mean and variances. In this case Augmented Dickey-Fuller test is not necessarily used. In order to secure the covariance stationarity condition, the data is transformed to stock return by using log differencing.



Descriptive Statistics (IBOVESPA)	
Number of Observations	243
Average	0.18%
Median	0.27%
Maximum	4.58%
Minimum	-5.60%
Standard Deviation	1.59%
Skewness	-0.313
Kurtosis - 3	0.769
Total Frequency Count	243



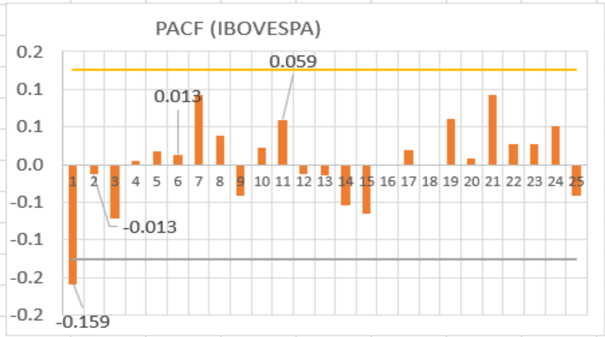
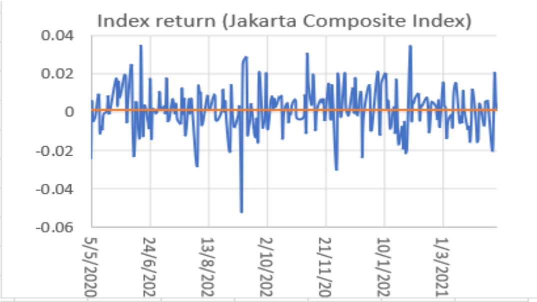
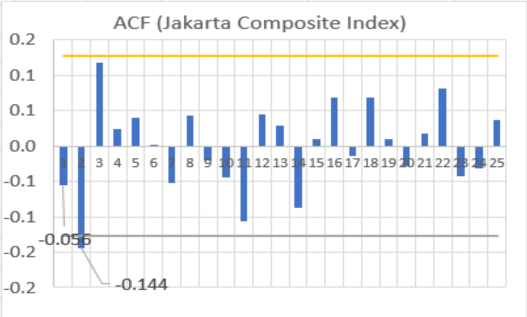


Figure 2: IBOVESPA plots

From the plots of IBOVESPA from Brazil, the natural log return looks covariance stationary from visual examination. The Augmented Dickey Fuller test is -5.295 with p-value = 0.01. The time series data is stationary, which means the series can be used for estimation the model and forecasting. There is one significant spike in both ACF and PACF plots. The ACF and PACF plots in Figure 3 suggest that MA(1), AR(1) or ARIMA(1,1) will be suitable candidate of the underlying time series process.



Descriptive Statistics (Jakarta Composite Index)	
Number of Observations	238
Average	0.12%
Median	0.14%
Maximum	3.47%
Minimum	-5.14%
Standard Deviation	1.17%
Skewness	-0.228
Kurtosis - 3	1.691
Total Frequency Count	238



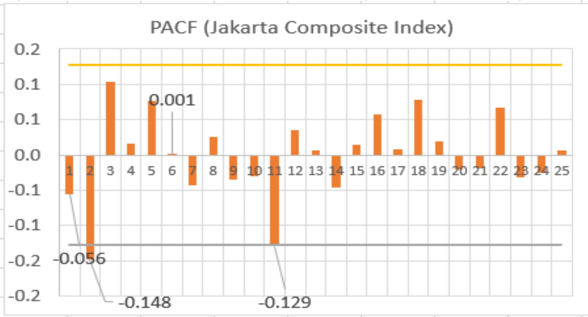


Figure 3: Jakarta Composite Index plots

From the plots of Jakarta Composite Index from Indonesia, the natural log return looks covariance stationary from visual examination. The Augmented Dickey Fuller test is -5.5531 with p-value = 0.01. The time series data is stationary, which means the series can be used for estimation the model and forecasting.

There is one significant spike in lag 2 in both ACF and PACF plots. The ACF and PACF plots in Figure 3 suggest that MA(2), AR(2) or ARIMA(2,2) will be suitable candidate of the underlying time series stochastic process.

5. Model estimation and selection

Note: below estimates are calculated using maximum likelihood estimation, Akaike info criterion and Augmented Dickey-Fuller Test in R but not in excel. Measurement of coefficients may be different.

For IBOVESPA,

	MA(1)	AR(1)	ARMA(1,1)
Constant	0.0017	0.002	0.002
Coefficient MA(1)	-0.1637		-0.2611
Coefficient AR(1)		-0.1591	0.099
Standard error MA(1)	0.0642		0.4841
Standard error AR(1)		0.0634	0.4971
Akaike Criteria	-1324.24	-1324.12	-1322.27
Log likelihood	665.12	665.06	665.13
Estimated σ^2	0.0002475	0.0002477	0.0002485

MA(1) is chosen for the model as it has the lowest Akaike information criteria. ARMA(1,1) does not have statistically significant coefficients for both MA(1) and AR(1).

In general, the chosen model can be expressed as:

$$Y(t) = 0.0017 - 0.1637\varepsilon_{t-1} + \varepsilon_t$$

For Jakarta Composite Index,

	MA(2)	AR(2)	ARMA(2,2)
Constant	0.0012	0.0015	0.0026
Coefficient MA(1)	-0.0285		0.5583
Coefficient AR(1)		-0.0632	-0.6054
Standard error MA(1)	0.0668		0.3600
Standard error AR(1)		0.0643	0.3478
Coefficient MA(2)	-0.1327		0.1065
Coefficient AR(2)		-0.1481	-0.2834
Standard error MA(2)	0.065		0.2668
Standard error AR(2)		0.0643	0.2561
Akaike Criteria	-1439.24	-1440.19	-1438.55
Log likelihood	723.62	724.1	725.27
Estimated σ^2	0.0001355	0.000135	0.0001348

AR(2) is chosen for the model as it has the lowest Akaike information criteria. ARMA(1,1) does not have statistically significant coefficients at all. Note that the AR(1) coefficient in model AR(2) is not statistically significant, so it is eliminated in the model by considering the coefficient to be zero.

In general, the chosen AR(2) model can be expressed as:

$$Y(t) = 0.0015 - 0.1481y_{t-2} + \varepsilon_t$$

6. Forecasting

This stage validates out of sample data in order to see whether the chosen model fits the data well and forecasting the stock prices for some future days and comparing them with the actual prices. This report uses fixed scheme to do the forecast as it is fast and convenient with only one estimation but it does not allow parameter updating.

For IBOVESPA,

Date	Point forecast	Lower 95% interval	Upper 95% interval	Actual log return	Absolute Error
19th April	0.001399	-0.02943646	0.0322363	-0.00148	0.002879
20th April	0.001747816	-0.02949907	0.0329947	-0.00723	0.008977
21st April	0.001747816	-0.02949907	0.0329947	No trading record	-
22nd April	0.001747816	-0.02949907	0.0329947	-0.00577	0.007517
23rd April	0.001747816	-0.02949907	0.0329947	0.00966	0.007913

From the forecast result, the point forecast converges quickly after 1 period. When using MA(1) in forecasting, the point forecast quickly drives into the unconditional mean and it is no longer powerful in forecasting the date after 19th April. The forecast error becomes larger after 19th April.

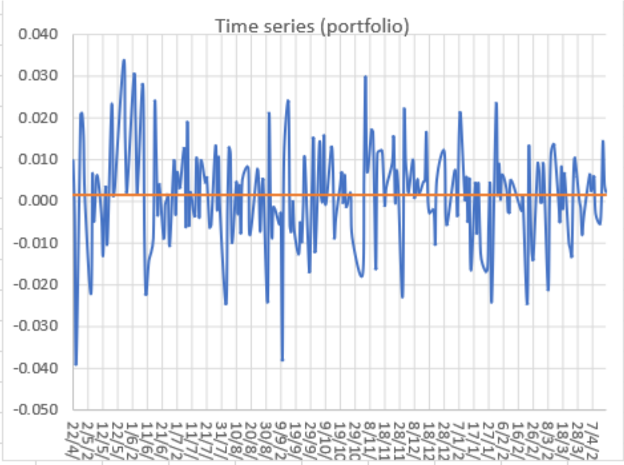
For Jakarta Composite Index,

Date	Point forecast	Lower 95% interval	Upper 95% interval	Actual log return	Absolute Error
19th April	0.0006735	-0.02209922	0.02344633	-0.005555	0.006229
20th April	0.001250	-0.02156797	0.02406858	-0.002352	0.003602
21st April	0.001279	-0.02177441	0.02433157	-0.007494	0.008773
22nd April	0.001191	-0.02186545	0.02424821	0.0001571	0.001034
23rd April	0.001193	-0.02186870	0.02425411	0.003776	0.002583

From the forecast result, the point forecast converges quickly after 2 period. In the AR(2) model, the first lag is not significant and the lag 2 is significant. It is clearly to see that the first forecast on 19th April is less accurate than that on 20th April. Although the model is less "converging" compared to the MA(1) model above, the forecasting power is diminishing along the time. The farther the time goes, the less the forecasting power it has by using the same information set. Also, the errors become harder to estimate as the errors in different time correlate with each other in AR(2).

7. Portfolio analysis

An equal weighted portfolio is created by mixing both the IBOVESPA and Jakarta Composite Index log return in the same date.



Descriptive Statistics (portfolio)	
Number of Observations	228
Average	0.16%
Median	0.19%
Maximum	3.37%
Minimum	-3.87%
Standard Deviation	1.12%
Skewness	-0.284
Kurtosis - 3	1.251
Total Frequency Count	228

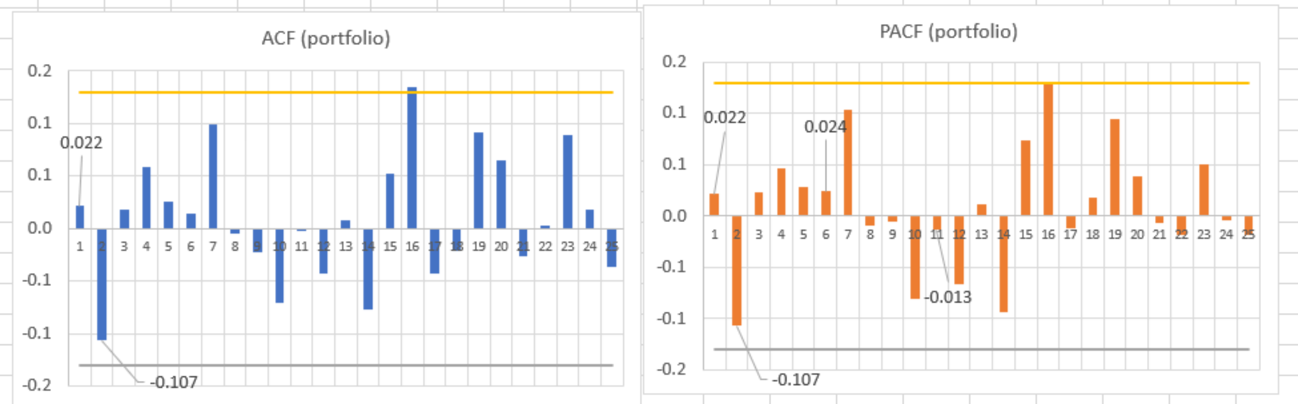


Figure 4: Equal weighted portfolio plots

From Figure 4, the equal weighted portfolio acts like white noise from visual examination. It is covariance stationary, but there are no significant spikes in both ACF and PACF plots. The combination of two markets is highly efficient, compared to the single market of IBOVESPA and Jakarta Composite Index. There are no methods to predict the white noise process using ARIMA models. Using the Box-Pierce test, the p-value is 0.9928 for this time series, which is extremely significant that the underlying process follows white noise process.

8. Conclusions

The forecast study aims to investigate the emerging financial markets by analyzing their underlying time series model. MA(1) and AR(2) models are selected for forecasting. It is not surprised that the markets are very efficient and act closely to white noise process. Obtaining highly time dependence evidence from the above experiments is difficult. Investors could not easily predict the market return by using time series in this case. For future implications, ARCH and GARCH models can be used as tools to investigate the volatility changes in time series, which could provide more market insights for the more “volatile” emerging economies.