Title: Correlation between Straits Times Industrial Index (STI) with Other Key Financial Stock Indexes

Executive Summary

This report studies the correlation between the percentage changes in Singapore's Straits Times Industrial Index (STI) with the percentage changes in other key financial stock indexes in Dow Jones Industrial Composite (New York, United States), Financial Times Stock Exchange 100 (London, United Kingdom), Nikkei 225 (Tokyo, Japan) and Hang Seng Index (Hong Kong SAR, China) from the period of Jan 2004 to Aug 2007. The regression model indicates a percentage change in STI index is directly correlated with the selected stock indexes. In the regression model, a small positive correlation between the percentage moves in STI index and Dow Jones Industrial (Composite) index is observed, while there are larger correlation relationships with other stock indexes. However, it is noted that there are other possible interaction of other economic factors (such as Singapore's macroeconomic performance and bilateral/multilateral trade relations) which could influence the changes in STI index.

Introduction

1. Often, we would notice from financial bulletins that the world's key stock indexes generally move in tandem with each other, in particular in the event of a major crisis where major stock indexes plunge drastically or a major economic breakthrough where stock markets boomed. Possible reasons contributing to this trend include (i) the strong interlink between the world's major financial markets where buyers and sellers are operating across the globe like in a single market, (ii) the wide implications to the entire world economy due to a major event such as a terrorist attack or a financial credit crunch and (iii) the general knee-jerk reaction of speculators.

2. While there is a general trend of stock markets moving together in major events, there are exceptions of stock indexes contradicting the trend. These then could possibly due to internal economic factors such as internal good/bad economic performance and bilateral economic relationships with closely linked economies.

3. To analyze this inter-relationship, this report looks specifically at Singapore's Straits Times Industrial Index (STI), the key stock index in Singapore and how the moves in STI is correlated with other major stock indexes and the economic relationship variables.

Areas of Analysis

STI and Other Stock Indexes

4. Essentially, the first question we would like to analyze is the correlation between STI and the major stock indexes.

5. Dow Jones Industrial Composite (New York, United States), Financial Times Stock Exchange 100 (London, United Kingdom), Nikkei 225 (Tokyo, Japan) and Hang Seng Index (Hong Kong SAR, China) are selected to compare against the STI index. As world's major financial hubs, the selected stock indexes listed in New York, London, Tokyo and Hong Kong are the main stock indicators and are good representations of the individual cities' general financial landscape. In addition, the selected countries/regions

are of similar economic development stage as Singapore (i.e. developed financial markets as opposed to emerging financial markets).

6. The next parameter to set is the time period of analysis. Specifically, there are factors such as availability of data, updated and sufficient data to reflect latest current trend, stability of economy (as opposed to choosing a period of turbulence which may over-skew the data and reflect wrong correlation due to a rare occurrence or low probability). Taking into account the above factors, the selected period is from Jan 2004 to Aug 2007. There are sufficient data points between this selected period as stock markets operate on all working days and that this period does not contain extreme world events such as 911 terrorist attacks. Also, this is also the period which all the selected economies are generally enjoying good growth with minimal economic disruptions or imbalances.

7. To measure the impact, the percentage changes in stock indices are being put into the regression model as opposed to having the absolute value of stock indices which would not reflect the changes. During the selected period between 2004 and 2007, there are more than 900 data points. As different countries have different public holidays, the days of operations vary, on such days, the rate of index is then computed as taking the value of the previous opening day[1]. In this analysis, only days which all stock exchanges were in operation are being used as the purpose of the report is to look at the correlation between STI and all other major stock indices.

Assumptions of Regression Model between STI Index and other Stock Indexes

8. Before running the regression model, it is assumed that the model of the correlation between STI index and other stock indexes (namely DJ Composite, FTSE100, Nikkei and Hang Seng), the model in the population is estimated to be as

Equation (1)

[1] This allows the measurement of changes in the market index which is usually based on the previous opening day, ie On day 1, the market opens, on Day 2, the market is closed and an event occurred which would have impact on the market performance. On Day 3, the market will respond.

$$STI\_percent = \beta_0 + \beta_1 DJ\_percent + \beta_2 FTSE\_percent$$
$$+ \beta_3 Hs\_percent + \beta_4 Nikkei\_percent + u$$

Where

STI_percent= % change in STI index of day 2 compared to the previous day

DJ_percent= % change in DJ Composite index

FTSE_percent= % change in FTSE 100 index

Hs_percent= % change in Hang Seng index

Nikkei_percent= % change in Nikkei index

9. This time series regression model takes into the linear time trend, commonly found in time series analysis where $y_t = \alpha_0 + \alpha_1 t + e_t$. It is to expect that stock markets are likely to contain time trend, ie stock index will rise over time along with general factors such as economic growth and development. The common method to include time trend in regression model is the use of $y^* = y_t - y_{t-1}$ the difference between 2 consecutive periods. From the equation 1, the STI_percent measures the % of the index of 2 consecutive opening days, factoring the time trend into the regression model.

Testing of the Regression Model

10. Table 1 below summarizes the regression results of the first model.

Table 1: Regression Results of model:
STI_percent= $\beta_0 + \beta_1$DJ_percent + β_2FTSE_percent + β_3Hs_percent + β_4Nikkei_percent + u

STI_percent	**Coefficient**	**Standard Error**	**t value**	**P>|t|**	**95% Confidence Interval**
DJ_percent	0.0128334	0.0291284	0.44	0.660	-0.0186621-0.0699958
FTSE_percent	0.1834248	0.0310964	5.90	0	0.1224004-0.2444493
Hs_percent	0.4481361	0.0246256	18.20	0	0.39981-0.4964622
Nikkei_percent	0.1881972	0.0221314	8.50	0	0.1447659-0.2316286
Constant	-0.0244687	0.0203138	-1.20	0.229	-0.0643332-0.0153958

R-squared = 0.5088
Adjusted R-squared = 0.5068
MSE = 0.62948
Number of observations = 963.

11. In analyzing the robustness of the coefficient estimates, there are statistical tests needed to ensure the unbiasedness in estimating the regression model. These are serial correlation and heterskedasticity.

Unbiasedness OLS estimator Assumptions

12. Under the Gauss Markov Theorem, there are assumptions made to the time series regression, which only when satisfied, the OLS estimators are BLUE (Best Linear Unbiased Estimator).

a. Assumption 1: Linear in Parameters

   The stochastic process $\{(x_{t,1}, x_{t,2}, \ldots x_{t,k}, y_t): t = 1,2,3 \ldots n\}$ follows the linear equation $y_t = \beta_o + \beta_1 x_{t,1} + \beta_2 x_{t,2} + \ldots + \beta_k x_{t,k} + u$ where $u_t : t = 1,2,3 \ldots n$ is the sequence of errors or disturbance and n is the number of observations (time periods).

b. Assumption 2: No Perfect Collinearity

   In the sample, no independent variable is constant nor a perfect linear combination of the others.

c. Assumption 3: Zero Conditional Mean

   For each t, the expected value of the error u_t, given the explanatory variables for all time periods is zero, $E(u \mid x_{t,1}, x_{t,2} \ldots x_{t,k}) = 0, t = 1,2,3 \ldots. n$

   In the sample, no independent variable is constant nor a perfect linear combination of the others.

d. Assumption 4: Homoskedasticity

   The variance of u_t, is the same for all t, $Var(u_t \mid x_{t,1}, x_{t,2}, \ldots x_{t,k}) = Var(u_t) = \sigma^2$ for $t = 1,2, \ldots n$

e. Assumption 5: No Serial Correlation

   The errors in two different time periods are uncorrelated, $Corr(u_t, u_s) = 0$ for all $t \neq s$

13. In the presence of serial correlation, the OLS estimator is no longer BLUE and the usual OLS standard errors and test statistics are not valid even asymptotically. If there is autocorrelation, the expression is as follows: $u_t = \hat{\rho} u_{t-1} + e_t$, $t = 1,2, \ldots. n$ where $|\hat{\rho}| < 1$

Testing for Serial Correlation

14. The model in equation (1) is being tested for serial correlation with the results being tabulated in Table 2. It shows that there is indeed serial correlation with H_0: $\hat{\rho} = 0$ being rejected. $\hat{\rho} = -0.1466321$, satisfies the inequality $|\hat{\rho}| < 1$. The model is therefore a Autocorrelation order 1 serial correlation (AR(1))

Table 2: Serial Correlation Test

$u_t = \rho u_{t-1} + e_t$, t=1,2,…n

u_t	Coefficient	Standard Error	t value	P>|t|	95% Confidence Interval
u_{t-1}	-0.1466321	0.0318804	-4.60	0.000	-0.209195 - -0.840693
constant	-0.000726	0.019983	-0.04	0.971	-0.0399411- -0.0384891

Correcting for Serial Correlation

15. In correcting the serial correlation, we continue to assume the Gauss Markov Model Assumptions 1 to 4 but we relax the Assumption 5 and takes into account the error term model, $u_t = \hat{\rho} u_{t-1} + e_t$. The equation (1) is being transformed to take into account (t-1) period to eliminate serial correlation. The transformed equation, known as the generalised least squares estimator (GLS) takes in the form of $\hat{y}_t = (1-\hat{\rho})\beta_0 + \beta_1(x_{t,1} - \hat{\rho} x_{t-1,1}) + \beta_2(x_{t,2} - \hat{\rho} x_{t-1,2}) + \ldots \beta_k(x_{t,k} - \hat{\rho} x_{t-1,k}) + e_t$.

Equation (2):

$$sti\_\hat{p}ercent_t = (1-\hat{\rho})\beta_0 + \beta_1(dj\_percent_t - \hat{\rho} dj\_percent_{t-1}) +$$
$$\beta_2(f\,tse\_percent_t - \hat{\rho} f\,tse\_percent_{t-1}) + .\beta_3(hs\_percent_t - \hat{\rho} hs\_percent_{t-1}) +$$
$$\beta_4(nikkei\_percent_t - \hat{\rho} nikkei\_percent_{t-1}) + e_t$$

16. The OLS on the transformed data is being tabulated in Table 3. This GLS estimator is BLUE and the errors in this transformed equation are serially uncorrelated, of which t-statistics and F-statistics are valid asymptotically.

Table 3: GLS estimators on transformed data for serial correlation correction:

sti_percent = $(1 - \rho)\beta_0 + \beta_1(dj\_percent_t - \rho\ dj\_percent_{t-1}) + \beta_2(ftse\_percent_t - \rho\ ftse\_percent_{t-1}) + \beta_3(hs\_percent_t - \rho\ hs\_percent_{t-1}) + \beta_4(nikkei\_percent_t - \rho\ nikkei\_percent_{t-1}) + e_t$

where serial_dj = $(dj\_percent_t - \rho dj\_percent_{t-1})$, …

STI_percent	Coefficient	Standard Error	t value	P>|t|	95% Confidence Interval
serial_dj	0.0523696	0.029651	1.77	0.078	-0.00058185 -0.058184

STI_percent	**Coefficient**	**Standard Error**	**t value**	**P>|t|**	**95% Confidence Interval**
serial_ftse	0.1656522	0.0323719	5.12	0	0.1021246-0.2291799
serial_hs	0.4497749	0.0247867	18.15	0	0.4011327-0.4984171
serial_nikkei	0.1934657	0.0222924	8.68	0	0.1497184-0.237213
Constant	-0.0216921	0.0204073	-1.06	0.288	-0.06174-0.0183558

R-squared = 0.5054
Adjusted R-squared = 0.5033
MSE = 0.63164
Number of observations = 962.

Testing for Heteroskedasticity

17. Similarly in cross sectional applications, heteroskedasticity in time series regression models while not causing bias or inconsistency in the coefficient estimate ß, invalidates the usual standard errors, t-statistics and F-statistics. In the test for heteroskedasticity, one key assumption is u_t should not be serially correlated; any serial correalation will generally invalidate a test for heteroskedasticity. After the serial correlation is corrected, we will apply the test for heteroskedasticity.

18. The test for heteroskedasticity for time series regression is the same as the cross sectional analysis which u_t^2 (in the equation (2)) is regressed with the variables, serial_dj, serial_ftse, serial_hs, serial_nikkei where

$u_t^2 = \delta_0 + \delta_1 serial\_dj + \delta_2 serial\_ftse + \delta_3 serial\_hs + \delta_4 serial\_nikkei + v$. The null hypothesis is being tested $H_0 : \delta_0 = \delta_1 = \delta_2 = \delta_3 = 0$.

19. Based on the results on F-statistics (Table 4), we fail to reject H_0 for significance level of 10%. Therefore, we can conclude that the transformed equation is not heteroskedastic.

Table 4: Heteroskedasticity Test:
$u_t^2 = \delta_0 + \delta_1$serial_dj + $\delta_2$serial_ftse+ $\delta_3$serial_hs + $\delta_4$serial_nikkei + v
where serial_dj = (dj_percent$_t$ – dj_percent$_{t-1}$), …

u_t^2	**Coefficient**	**Standard Error**	**t value**	**P>|t|**	**95% Confidence Interval**

serial_dj	0.050215	0.037517	-1.34	0.181	-0.0234097 -0.0534096
serial_ftse	0.103422	0.0409598	2.52	0.012	0.0230412 -0.1838028
serial_hs	0.0079294	0.0313623	-0.25	0.800	-0.0694758 -0.053617
serial_nikkei	0.0053556	0.0282063	0.19	0.849	-0.0499974 -0.0607086
Constant	-0.982814	0.0258211	15.42	0.000	0.3476092-0.4489535

R-squared = 0.0074
Adjusted R-squared = 0.0033
MSE = 0.7992
Number of observations = 962.
F-statistics (4, 962) = 1.80
Prob>F= 0.1269

Results and Interpretation

20. The regression model based on the results shown in Table 2 is estimated to be

sti_percent = -0.0216921 +0.0523696 dj_percent$_t$ + 0.1656522 ftse_percent$_t$ + 0.4497749 hs_percent$_t$ + 0.1934657nikkei_percent$_t$

(based on the p-values, we reject $H_0 : \beta_{dj\_percent} = \beta_{ftse\_percent} = \beta_{hs\_percent} = \beta_{nikkei\_percent} = 0$ at 5% significance level)

21. The model illustrates that there is correlation between the percentage changes in STI index and Dow Jones Industrial, FTSE100, Hang Seng Index and Nikkei Index.

22. With regard to the correlation between the stock indexes and the STI index, there are a few observations. First, it is in line with the common notion that there is some correlation between the markets in Singapore and the financial stock indexes in US, UK, HK and Japan (i.e. when Dow Jones suffers a loss, there may be a ripple effect on STI index to react on the following trading day). From the magnitude of the coefficients, we can observe that 1% change in Dow Jones has a return of 0.05% change in STI index. This is relatively small and reasonable correlation as a large correlation is not expected given the differences between the comparing indexes in terms of structure, composition and economic performance and fundamentals. The correlation with the Asian stock indexes is stronger, possibly due to closer proximity in terms of economic structure. The recent economic diversification efforts initiated by the Singapore government to rely less on the major world markets such US have also reduced this correlation relationship. This could also be attributed to the particular time period selected for analysis. This regression model is based on the time period from 2004 to Oct 2007, where there are no significant

turbulence economic shocks and even if during certain economic shocks i.e. sub-prime housing loan, the duration is not sufficiently long enough to capture the relationship effectively.

Possible Refinements

23. While efforts are being made to make data sets for analysis as robust as possible, nonetheless, there are some weaknesses in this selection. Only 4 other stock indices are being selected to conduct the regression model with STI index. These stock indexes are computed using different approaches (ie price weighted or scale weighted versus STI which is weighted-value stock index) and are therefore not an exact like-for-like comparison of stock indexes. There may be missing variables such as neighboring stock indexes (e.g. Kuala Lumpur Stock Exchange Index, KLSE) and emerging stock indexes with growing influence (e.g. China's Shanghai Stock Exchange).

24. The model can also be improved by separating into 2 periods where one of which is a stable period while the other could be a turbulent period, such that we can compare the correlation under different situations. Using the same analogy, the model can be expanded by separating into developed stock markets versus emerging stock markets.

Reference

1. Wooldridge "Introductory Econometrics: A Modern Approach" Chapters 10 to 12
2. International Enterprise Singapore website http://www.iesingapore.gov.sg
3. Singapore Stock Exchange Limited http://www.sgx.com.sg