

## **CHAPTER 4**

### **STOCK RETURN AND TRADING VOLUME RELATIONSHIP**

#### **4.1 Introduction**

From Chapter 3, the results show that there are variations in the return and trading volume for the time-of-the-day and day-of-the-week. Therefore, it is interesting to examine the basic characteristics of the return-volume relationship. In this chapter, we examine the contemporaneous relationship and the causality between stock return and trading volume in the Malaysian securities market.

#### **4.2 Contemporaneous Relationship Between Return and Volume**

There are two different sets of trading data in the previous chapter; one starts from 1000 hour till 1600 hour and the other from 0930 hour till 1700 hour which arises from the extended trading hours. To provide data consistency in the analysis, the

earlier set of data which covered period before extended trading hours is dropped. Thus, the analysis concentrates on the period after extended trading hours which consists of 3740 intraday observations, made up of 170 trading days of 22 intraday sessions (15-minute interval) . The return and volume series used in this section refer to the index return and volume ratio, respectively.

In examining the contemporaneous relationship between the intraday return and volume, the model used by Jain and Joh (1988) and Ho et al (1993) is adopted. The model is as follows:

$$\begin{aligned}
 V_t = & a + b |R_t| + c [ D_t |R_t| ] \\
 & + \sum_{k=1}^n e_k DD_{kt} + \sum_{k=1}^n f_k [ DD_{kt} |R_t| ] \\
 & + \sum_{k=1}^n g_k [ DD_{kt} |R_t| D_t ] + \mu_t
 \end{aligned}$$

where  $V_t$  is the volume ratio used in earlier chapter;  $R_t$  denotes returns for period  $t$ ;  $D_t = 0$  if the period  $t$  return is positive and 1 otherwise;  $DD_{kt}$  is a vector of dummy variables,  $k=1,2,...,n$ ;  $n$  is 25 which includes 4 dummy variables for days of the week and 21 for times of the day;  $U_t$  is a random error term.

The three specifications are described as below:

- (a)  $DD_{kt}$  variables are omitted, i.e.  $e_k=f_k=g_k=0$ ;
- (b) The seasonal variables affect only the intercept but not the slope of the regression, i.e.  $f_k=g_k=0$ ;
- (c) Full model as in the above equation.

The regression results are presented in Table 4.1. It is found that all the three coefficients are highly significant in the first specification. The positive coefficient (0.0263) for the absolute return indicates positive relationship between absolute return and the volume; in other words, volume tends to move in the same direction as absolute return. When  $R_t$  is positive, the slope of the regression is **b**; while when  $R_t$  is nonpositive, i.e.  $D_t$  is 1, the slope is the sum of **b** and **c**, that is, 0.0108 (0.0263 - 0.0155). Hence, it is obvious that the slope for the positive return is greater than the value when the returns are nonpositive. The results here are consistent with the asymmetric price-volume hypothesis proposed by Karpoff (1987) and the empirical evidence found by Jain and Joh (1988) in the US market and Ho et al (1993) in the Hong Kong market. The highly statistically significant coefficient of  $|R_t|D_t$  implies that the asymmetrical response is stronger than the Hong Kong market but not as strong as the US market.

To allow the influence of the day-of-the-week and time-of-the-day

Table 4.1  
 Regression Results on Specification 1

LS // Dependent variable is V Included observations: 3740				
Variable	Coefficient	Std. Error	T-Statistic	Prob.
a	0.004422	9.92E-05	44.60143	0.000000
R	0.026321	0.002271	11.59024	0.000000
D  R	-0.015457	0.003067	-5.040149	0.000000
R-squared	0.034811	Mean dependent var		0.005032
Adjusted R-squared	0.034294	S.D. dependent var		0.004616
S.E. of regression	0.004536	Akaike info criterion		-10.790490
Sum squared resid	0.076899	Schwartz criterion		-10.785500
Log likelihood	14874.39	F-statistic		67.389630
Durbin-Watson stat	0.720624	Prob(F-statistic)		0.000000

Note:

1. D=0 if R greater than 0 and D=1 if R less than or equal to 0
2. R is the average return for the Composite Index.



effects, additional dummy variables are included in the second specification. The results are presented in Table 4.2. It is found that the adjusted  $R^2$  increases significantly from 3.43 per cent to 9.27 per cent. Most of the coefficients for dummy variables are statistically significant either at 1, 5 or 10 per cent level except the ones for Tuesday and Wednesday. Such findings are, however, different from those obtained by Ho et al (1993) and Jain and Joh (1988). Generally, this specification suggests that the seasonal effects affect only the intercept but not the slope of the equation. The analysis reveals that the time-of-the-day effect has greater influence on the intercept of the equation as compared to the day-of-the-week effect.

Table 4.3 presents the results of the full model as stated in the third specification. The specification allows intercepts and slopes to differ across the weekday and the 15-minute trading interval. The results show that the adjusted  $R^2$  increases further to 10.18 per cent. There are more significant coefficients for the time-of-the-day effect than the day-of-the-week. In fact, most of the coefficients (11 out of 12) for the dummy variables for the day-of-the-week are not significant. This seems to imply that time-of-the-day has greater influence on the intercept and slope coefficients as compared to the day-of-the-week.

Table 4.2  
Regression Results on Specification 2

LS // Dependent variable is V Included observations: 3740				
Variable	Coefficient	Std. Error	T-Statistic	Prob.
a	0.009242	0.000394	23.475160	0.000000
R	0.016793	0.002332	7.200962	0.000000
D  R	-0.007534	0.003057	-2.464736	0.013800
D(MON)	0.000603	0.000236	2.549789	0.010800
D(TUE)	0.000237	0.000228	1.039348	0.298700
D(WED)	0.000270	0.000228	1.186675	0.235400
D(THU)	0.000421	0.000229	1.834461	0.066700
D(0945)	-0.004405	0.000477	-9.227089	0.000000
D(1000)	-0.004745	0.000483	-9.821652	0.000000
D(1015)	-0.004931	0.000485	-10.165030	0.000000
D(1030)	-0.005390	0.000485	-11.103270	0.000000
D(1045)	-0.005361	0.000486	-11.020060	0.000000
D(1100)	-0.005177	0.000487	-10.623720	0.000000
D(1115)	-0.005441	0.000487	-11.184240	0.000000
D(1130)	-0.005844	0.000487	-11.990780	0.000000
D(1145)	-0.005677	0.000487	-11.658720	0.000000
D(1200)	-0.005737	0.000487	-11.785480	0.000000
D(1215)	-0.005667	0.000486	-11.655590	0.000000
D(1230)	-0.005757	0.000487	-11.820500	0.000000
D(1445)	-0.003732	0.000486	-7.684780	0.000000
D(1500)	-0.004707	0.000486	-9.685560	0.000000
D(1515)	-0.004973	0.000487	-10.218840	0.000000
D(1530)	-0.005390	0.000486	-11.097160	0.000000
D(1545)	-0.005514	0.000487	-11.332310	0.000000
D(1600)	-0.005285	0.000486	-10.881290	0.000000
D(1615)	-0.005252	0.000486	-10.805250	0.000000
D(1630)	-0.005416	0.000485	-11.162450	0.000000
D(1645)	-0.004281	0.000485	-8.826298	0.000000
R-squared	0.099206	Mean dependent var		0.005032
Adjusted R-squared	0.092654	S.D. dependent var		0.004616
S.E. of regression	0.004397	Akaike info criterion		-10.846170
Sum squared resid	0.071769	Schwartz criterion		-10.799550
Log likelihood	15003.51	F-statistic		15.141040
Durbin-Watson stat	0.649974	Prob(F-statistic)		0.000000

Note:

1. D(Mon)=1 for Monday and 0 otherwise; D(Tue)=1 for Tuesday and 0 otherwise; other weekdays are defined similarly.
2. D(0945)=1 for 0930-0945 interval and 0 otherwise; D(1000)=1 for 0945-1000 interval and 0 otherwise; other 15-minute intervals are defined similarly.

Table 4.3  
Regression Results on Specification 3

LS // Dependent variable is V  
Included observations: 3740

Variable	Coefficient	Std. Error	T-Statistic	Prob.
a	0.007636	0.000592	12.904120	0.000000
R	0.033365	0.008022	4.159108	0.000000
D  R	0.055428	0.017099	3.241541	0.001200
D(MON)	0.000668	0.000320	2.084424	0.037200
D(TUE)	0.000213	0.000306	0.696129	0.486400
D(WED)	0.000489	0.000305	1.606147	0.108300
D(THU)	0.000163	0.000314	0.516953	0.605200
D(0945)	-0.003076	0.000718	-4.285939	0.000000
D(1000)	-0.003219	0.000750	-4.290300	0.000000
D(1015)	-0.002476	0.000735	-3.367152	0.000800
D(1030)	-0.002971	0.000733	-4.053379	0.000100
D(1045)	-0.003925	0.000736	-5.330838	0.000000
D(1100)	-0.003768	0.000752	-5.012964	0.000000
D(1115)	-0.003740	0.000743	-5.033342	0.000000
D(1130)	-0.003977	0.000740	-5.375108	0.000000
D(1145)	-0.003626	0.000733	-4.948319	0.000000
D(1200)	-0.003749	0.000717	-5.227439	0.000000
D(1215)	-0.004372	0.000748	-5.846230	0.000000
D(1230)	-0.004101	0.000722	-5.678491	0.000000
D(1445)	-0.002175	0.000722	-3.014358	0.002600
D(1500)	-0.002532	0.000726	-3.488126	0.000500
D(1515)	-0.003270	0.000719	-4.549136	0.000000
D(1530)	-0.003870	0.000743	-5.206477	0.000000
D(1545)	-0.003973	0.000733	-5.418759	0.000000
D(1600)	-0.004027	0.000716	-5.622652	0.000000
D(1615)	-0.004258	0.000727	-5.854262	0.000000
D(1630)	-0.004035	0.000730	-5.526450	0.000000
D(1645)	-0.002925	0.000749	-3.907401	0.000100
D(MON)  R	0.001047	0.007648	0.136851	0.891200
D(TUE)  R	0.003329	0.008169	0.407508	0.683700
D(WED)  R	-0.005239	0.006657	-0.786862	0.431400
D(THU)  R	0.007514	0.007548	0.995572	0.319500
D(0945)  R	-0.011362	0.008499	-1.336914	0.181300
D(1000)  R	-0.012026	0.012111	-0.992990	0.320800
D(1015)  R	-0.037409	0.014286	-2.618619	0.008900
D(1030)  R	-0.040020	0.013846	-2.890290	0.003900
D(1045)  R	-0.005982	0.015018	-0.398324	0.690400
D(1100)  R	0.016549	0.022394	0.738993	0.460000
D(1115)  R	-0.010263	0.017921	-0.572699	0.566900
D(1130)  R	-0.036109	0.017309	-2.086153	0.037000
D(1145)  R	-0.023003	0.017802	-1.292148	0.196400
D(1200)  R	-0.023876	0.017848	-1.337762	0.181100
D(1215)  R	0.007709	0.018076	0.426464	0.669800

Table 4.3 (Continuation)  
 Regression Results on Specification 3

LS // Dependent variable is V				
Included observations: 3740				
D(1230)  R	-0.012891	0.018992	-0.678773	0.497300
D(1445)  R	-0.005551	0.015719	-0.353152	0.724000
D(1500)  R	-0.037764	0.014106	-2.677136	0.007500
D(1515)  R	-0.023366	0.013799	-1.693327	0.090500
D(1530)  R	-0.028873	0.013482	-2.141665	0.032300
D(1545)  R	-0.031950	0.017703	-1.804734	0.071200
D(1600)  R	-0.014221	0.014350	-0.991064	0.321700
D(1615)  R	-0.013637	0.015525	-0.878391	0.379800
D(1630)  R	-0.027851	0.014563	-1.912424	0.055900
D(1645)  R	-0.016743	0.015863	-1.055433	0.291300
D(MON)  R  D	-0.005779	0.009653	-0.598654	0.549400
D(TUE)  R  D	-0.006013	0.009948	-0.604416	0.545600
D(WED)  R  D	-0.003530	0.009507	-0.371362	0.710400
D(THU)  R  D	0.000173	0.009813	0.017642	0.985900
D(0945)  R  D	-0.064209	0.019004	-3.378782	0.000700
D(1000)  R  D	-0.067102	0.019617	-3.420538	0.000600
D(1015)  R  D	-0.069110	0.021077	-3.279004	0.001100
D(1030)  R  D	-0.066732	0.021634	-3.084542	0.002100
D(1045)  R  D	-0.070856	0.021700	-3.265332	0.001100
D(1100)  R  D	-0.103514	0.026059	-3.972357	0.000100
D(1115)  R  D	-0.082482	0.024864	-3.317295	0.000900
D(1130)  R  D	-0.046045	0.025384	-1.813906	0.069800
D(1145)  R  D	-0.085968	0.025405	-3.383907	0.000700
D(1200)  R  D	-0.080279	0.024282	-3.306164	0.001000
D(1215)  R  D	-0.086759	0.024630	-3.522548	0.000400
D(1230)  R  D	-0.074888	0.025585	-2.926994	0.003400
D(1445)  R  D	-0.083305	0.023167	-3.595824	0.000300
D(1500)  R  D	-0.053107	0.020682	-2.567768	0.010300
D(1515)  R  D	-0.054952	0.020986	-2.618464	0.008900
D(1530)  R  D	-0.037615	0.020908	-1.799057	0.072100
D(1545)  R  D	-0.031438	0.023563	-1.334200	0.182200
D(1600)  R  D	-0.045496	0.021604	-2.105950	0.035300
D(1615)  R  D	-0.031763	0.022086	-1.438149	0.150500
D(1630)  R  D	-0.018872	0.023535	-0.801848	0.422700
D(1645)  R  D	-0.047305	0.021979	-2.152296	0.031400
Resquared	0.120339	Mean dependent var	0.005032	
Adjusted R-squared	0.101842	S.D. dependent var	0.004616	
S.E. of regression	0.004375	Akaike info criterion	-10.843170	
Sum squared resid	0.070085	Schwartz criterion	-10.713310	
Log likelihood	15047.91	F-statistic	6.506053	
Durbin-Watson stat	0.671887	Prob(F-statistic)	0.000000	

### 4.3 Causality Test

Apart from the contemporaneous relationship between return and volume, it would also be interesting to examine the causal relation or the lead-lag structure of these two variables. Ho et al (1993) gave two reasons why the causal relation should be examined : firstly, the specification given in the contemporaneous equation implicitly assumes return causes volume; if this assumption is not true, specification error would be incurred in the estimation of the relationship; secondly, technical analysts usually take an opposite view as they look at the volume to predict or to confirm future price movements. Thus, in this section, Grange-causality test is used to investigate whether the return and volume are causally related. The causality test here involves 3740 intraday data set (made up of 170 trading days of 22 intraday 15-minute sessions) of return and volume which covers from 22 July 1992 till 15 April 1993.

Consider the following unrestricted model for return (R) and volume (V):

$$R_t = \sum_{i=1}^m \alpha_i V_{t-i} + \sum_{j=1}^n \beta_j R_{t-j} + \mu_{1t} \text{ -----(1)}$$

$$V_t = \sum_{i=1}^p \lambda_i V_{t-i} + \sum_{j=1}^q \delta_j R_{t-j} + \mu_{2t} \text{ -----(2)}$$

where m, n, p and q are the lag lengths for the respective series.  $\mu_{1t}$  and  $\mu_{2t}$  are the error terms in equations 1 and 2, and it is assumed that they are not correlated.

The Granger-causality test examines the dynamic relationship between the market return and the trading volume. Before the causality test is performed, the unit root test is used to determine the stationarity of both return and volume series. The return and volume series here refer to the average return for the KLSE CI and the volume ratio, respectively. As presented in Tables 4.4 and 4.5, the Augmented Dicky-Fuller (ADF) statistics for both return and volume series are, in absolute term, greater than the critical value provided by Mckinnon. This suggests that unit roots do not exist in both series and that they are stationary at level.

Both series are also tested for white noise by using Ljung-Box Q-statistics. Up to 30 lags are used. The results are presented in

Table 4.4  
 Augmented Dickey-Fuller Unit Root Test on R

ADF Test Statistic	-25.58806	1% Critical Value	-2.5663	
		5% Critical Value	-1.9394	
		10% Critical Value	-1.6156	
*MacKinnon critical values for rejection of hypothesis of a unit root.				
Augmented Dickey-Fuller Test Equation				
LS // Dependent variable is D(R)				
Included observations: 3740				
Variable	Coefficient	Std. Error	T-Statistic	Prob
R(- 1)	-0.789854	0.030868	-25.58806	0.000000
D(R(- 1))	-0.086099	0.028284	-3.044074	0.002400
D(R(-2))	-0.004671	0.025373	-0.184076	0.854000
D(R(-3))	0.017547	0.021785	0.805465	0.420600
D(R(-4))	0.004626	0.016388	0.282262	0.777800
R-squared	0.436043	Mean dependent var		4.49E-05
Adjusted R-squared	0.435439	S.D. dependent var		0.065072
S.E. of regression	0.048893	Akaike info criterion		-6.034891
Sum squared resid	8.928743	Schwartz criterion		-6.026567
Log likelihood	5983.417	F-statistic		721.961300
Durbin-Watson stat	1.996964	Prob(F-statistic)		0.000000

Table 4.5  
 Augmented Dickey-Fuller Unit Root Test on V

ADF Test Statistic	-6.677887	1% Critical Value	-2.5663	
		5% Critical Value	-1.9394	
		10% Critical Value	-1.6156	
*MacKinnon critical values for rejection of hypothesis of a unit root.				
Augmented Dickey-Fuller Test Equation				
LS // Dependent variable is D(V)				
Included observations: 3740				
Variable	Coefficient	Std. Error	T-Statistic	Prob
V(- 1)	-0.056402	0.008446	-6.677887	0.000000
D(V(- 1))	-0.561935	0.017352	-32.384920	0.000000
D(V(-2))	-0.349114	0.019064	-18.312920	0.000000
D(V(-3))	-0.243955	0.018759	-13.004580	0.000000
D(V(-4))	-0.127440	0.016272	-7.832055	0.000000
R-squared	0.282239	Mean dependent var		-2.07E-06
Adjusted R-squared	0.281471	S.D. dependent var		0.003850
S.E. of regression	0.003264	Akaike info criterion		-11.448560
Sum squared resid	0.039780	Schwartz criterion		-11.440230
Log likelihood	16106.98	F-statistic		367.17130
Durbin-Watson stat	2.030844	Prob(F-statistic)		0.000000



Tables 4.6 and 4.7. It is found that the null hypotheses of all the autocorrelations are zero have to be rejected for both return and volume series. This suggests that the series are not white noise. The first order autocorrelation of 0.652 implies that about 43 per cent of the intraday volume variation is predictable by using the preceding intraday volume. For the return series, the first order autocorrelation of 0.137 shows that only 2 per cent in return figures can be predicted by the previous intraday return. Both figures of return and volume are much smaller than those observed in Finnish stock market where Martikainen et al (1994) obtained figures of 10 per cent and 87 per cent in the return and volume series, respectively.

The determination of the number of lags for the variables which are to be included in the model is based on the Akaike Information Criterion (AIC). Up to 30 lags have been tested on both return and volume series. In equation 1, The lowest AIC value is recorded at lag 25 for both return and volume series. In equation 2, the minimum AIC value is achieved at lag 25 for return series and at lag 2 for volume series.

Table 4.8 shows the Granger-causality results. The highly significant F-statistic suggests that both the null hypotheses of

Table 4.6  
Correlogram of R

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.137	0.137	70.713	0.000
		2	0.100	0.082	107.87	0.000
		3	0.044	0.021	115.20	0.000
		4	0.002	-0.014	115.22	0.000
		5	-0.000	-0.005	115.22	0.000
		6	0.011	0.012	115.63	0.000
		7	0.014	0.012	116.34	0.000
		8	0.036	0.032	121.25	0.000
		9	0.059	0.049	134.16	0.000
		10	0.037	0.018	139.43	0.000
		11	0.072	0.056	159.09	0.000
		12	0.048	0.025	167.59	0.000
		13	0.039	0.019	173.25	0.000
		14	-0.009	-0.026	173.53	0.000
		15	0.003	0.000	173.56	0.000
		16	0.015	0.015	174.38	0.000
		17	0.029	0.024	177.57	0.000
		18	0.018	0.004	178.75	0.000
		19	0.010	-0.005	179.11	0.000
		20	0.048	0.037	187.79	0.000
		21	0.068	0.053	205.29	0.000
		22	0.099	0.074	241.86	0.000
		23	0.061	0.026	255.92	0.000
		24	0.005	-0.028	256.02	0.000
		25	-0.031	-0.043	259.61	0.000
		26	0.005	0.012	259.69	0.000
		27	-0.028	-0.025	262.61	0.000
		28	-0.024	-0.026	264.83	0.000
		29	0.008	0.005	265.07	0.000
		30	0.015	0.007	265.89	0.000

Table 4.7  
Correlogram of V

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.652	0.652	1590.4	0.000
		2 0.606	0.314	2963.7	0.000
		3 0.561	0.165	4143.7	0.000
		4 0.544	0.137	5252.7	0.000
		5 0.527	0.104	6293.6	0.000
		6 0.523	0.103	7319.6	0.000
		7 0.506	0.065	8280.6	0.000
		8 0.501	0.068	9222.7	0.000
		9 0.494	0.058	10137.	0.000
		10 0.490	0.057	11040.	0.000
		11 0.477	0.033	11895.	0.000
		12 0.466	0.025	12712.	0.000
		13 0.494	0.100	13626.	0.000
		14 0.470	0.016	14456.	0.000
		15 0.438	-0.035	15176.	0.000
		16 0.449	0.041	15932.	0.000
		17 0.433	0.006	16637.	0.000
		18 0.434	0.024	17345.	0.000
		19 0.433	0.026	18051.	0.000
		20 0.451	0.068	18818.	0.000
		21 0.481	0.109	19688.	0.000
		22 0.523	0.148	20717.	0.000
		23 0.453	-0.073	21490.	0.000
		24 0.411	-0.096	22126.	0.000
		25 0.384	-0.066	22681.	0.000
		26 0.388	-0.013	23249.	0.000
		27 0.379	-0.016	23791.	0.000
		28 0.370	-0.013	24307.	0.000
		29 0.359	-0.018	24793.	0.000
		30 0.371	0.032	25313.	0.000

Table 4.8  
Granger-Causality Test

	Hypothesis	Wald F-statistic
Equation 1	$\sum_{i=1}^m \alpha_i=0$	29.25*
	$\sum_{j=1}^n \beta_j=0$	35.74*
	$\sum_{i=1}^m \alpha_i=0, \sum_{j=1}^n \beta_j=0$	37.38*
Equation 2	$\sum_{i=1}^p \lambda_i=0$	299.39*
	$\sum_{j=1}^q \delta_j=0$	6.86*
	$\sum_{i=1}^p \lambda_i=0, \sum_{j=1}^q \delta_j=0$	162.22*

Note:

m=25, n=25, p=25, q=2

\* Significant at 1 per cent level

volume does not Granger cause return and return does not Granger cause volume have to be rejected. This may indicate that feedback or bilateral causality exists in the KLSE over the study period. Such phenomenon is consistent with Martikainen et al (1994) and Smirlock and Starks (1988). The finding is, however, different from what was found by Ho et al (1993) in the Hong Kong Stock Exchange and Jain and Joh (1988) in the New York Stock Exchange where the relation was unidirectional in that only volume causes return in the respective markets.

The two different sets of results would probably suggest that feedback tends to exist in smaller markets like Finland and Malaysia while the more developed markets such as New York and Hong Kong tend to exhibit unidirectional causality.

The lag lengths for return and volume series which are generally greater than the 24-daily-intraday-sessions show that the memory or historical information of the KLSE cannot be absorbed within a day. However, in the KLSE, volume seems to contain useful information for forecasting return and vice versa.