RELATIONSHIP BETWEEN VOLATILITY AND TRADING VOLUME: THE CASE OF HSI STOCK RETURNS DATA

Michaela Chocholatá
University of Economics Bratislava, Slovakia

Introduction (1)

- one of the characteristic features of stock returns is the time-varying volatility
- the pioneering work in the area of modelling volatility was presented by Engle [1982]
- nowadays a large number of modifications of the standard ARCH and GARCH models have been developed
- though the ARCH/GARCH – class models allow the volatility shocks to persist over time, they didn't provide the economic explanation for this phenomenon

Introduction (2)

- the paper of Lamoureux and Lastrapes [1990] offers the explanation for volatility persistence
- their approach has been applied in various studies to both individual stocks (stock-level analysis) and stock market indices (market-level analysis)
- they proved that the daily trading volume has a significant explanatory power regarding the variance of daily returns

The aim of the presentation:

- to analyse the relationship between the trading volume and the daily volatility of the Hong–Kong HSI stock returns data using the GJR-GARCH models and applying the approach of Lamoureux and Lastrapes [1990]

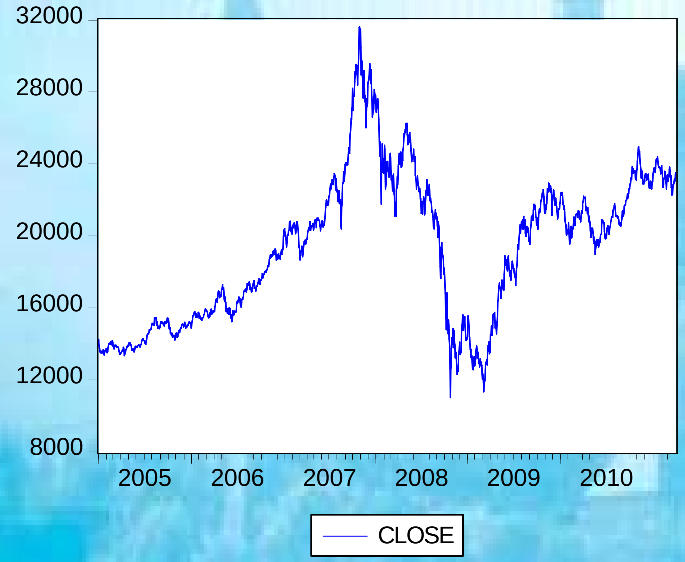
Data and Methodology

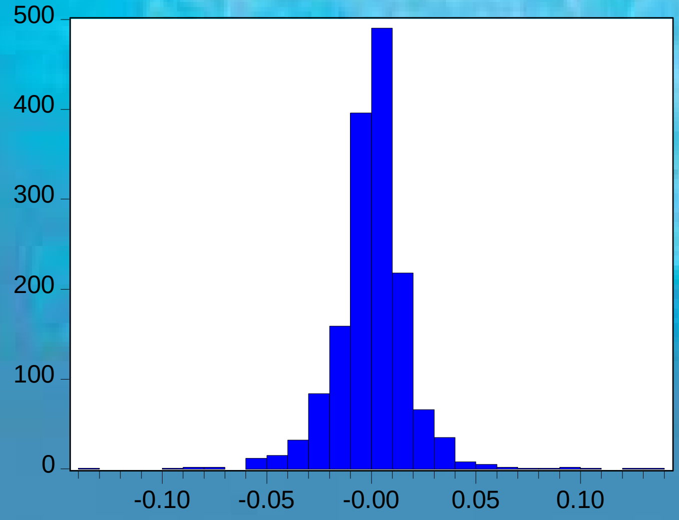
- the whole analysis was done on logarithmic transformation of daily index returns and daily trading volume
- the logarithmic stock returns are calculated as the logarithmic first difference of the daily closing values of the analyzed stock index, i.e.

$$r_t = d(\ln(P_t)) = \ln\left(\frac{P_t}{P_{t-1}}\right)$$

where P_t is the closing value of the stock index at time t and r_t denotes logarithm of the corresponding stock return

Closing values of the HSI stock index and descriptive statistics of the logarithmic return series





Series: DLCLOSE
Sample 1/03/2005 3/31/2011
Observations 1535

Mean	0.000328
Median	0.000947
Maximum	0.134068
Minimum	-0.135820
Std. Dev.	0.018115
Skewness	0.081094
Kurtosis	11.53212
Jarque-Bera	4657.657
Probability	0.000000

Methodology – conditional mean equation

- the logarithmic stock returns equation, i.e. the conditional mean equation, can be in general written as a Box-Jenkins ARMA(m,n) model of the form:

$$r_t = \omega_0 + \sum_{j=1}^{m} \phi_j r_{t-j} + \sum_{k=1}^{n} \theta_k \varepsilon_{t-k} + \varepsilon_t$$

where ω_0 is unknown constant, $\phi_j \ (j = 1, 2, \ldots m)$ and $\theta_k \ (k = 1, 2, \ldots n)$ are the parameters of the appropriate ARMA(m,n) model, ε_t is a disturbance term.

Methodology – conditional variance equation

- the conditional variance equation in case of a GJR-GARCH(p,q) model can be specified as:

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} \gamma_i \varepsilon_{t-i}^2 I_{t-i}^-$$

where from $I_{t-i}^- = \begin{cases} 1, if \ \varepsilon_{t-i} < 0 \\ 0, if \ \varepsilon_{t-i} > 0 \end{cases}$, it is clear the different impact of the positive shocks $\varepsilon_{t-i} > 0$ and negative shocks $\varepsilon_{t-i} < 0$ on the conditional variance

- to examine the effect of trading volume V_t on stock returns volatility, the following modification of the conditional variance equation is used

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i} + \sum_{i=1}^{q} \gamma_i \varepsilon_{t-i}^2 I_{t-i}^- + \delta V_t$$

Empirical results

- the analysis was done in two steps – without and with trading volume included into the conditional volatility equation
- the appropriate ARMA (m,n) model for logarithmic stock returns was estimated (table 1)
- the estimation results of conditional variance equations without and with the trading volume included using the GJR-GARCH(1,2) model are in table 2 and table 3, respectively

Table 1

Dependent Variable: D(LOG(CLOSE))
Method: Least Squares
Date: 11/20/11 Time: 18:59
Sample (adjusted): 1/04/2005 3/31/2011
Included observations: 1534 after adjustments
Convergence achieved after 5 iterations
Backcast: 12/21/2004 1/03/2005

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000328	0.000429	0.762914	0.4456
MA(10)	-0.070451	0.025503	-2.762422	0.0058

R-squared	0.005097	Mean dependent var	0.000327
Adjusted R-squared	0.004447	S.D. dependent var	0.018121
S.E. of regression	0.018080	Akaike info criterion	-5.186683
Sum squared resid	0.500808	Schwarz criterion	-5.179726
Log likelihood	3980.186	F-statistic	7.848098
Durbin-Watson stat	2.075733	Prob(F-statistic)	0.005151

Inverted MA Roots	.77	.62+.45i	.62-.45i	.24-.73i
	.24+.73i	-.24-.73i	-.24+.73i	-.62-.45i
	-.62+.45i	-.77		

Table 2

Dependent Variable: D(LOG(CLOSE))
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 11/20/11 Time: 19:03
Sample (adjusted): 1/04/2005 3/31/2011
Included observations: 1534 after adjustments
Convergence achieved after 14 iterations
MA backcast: 12/21/2004 1/03/2005, Variance backcast: ON
GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*RESID(-1)^2*(RESID(-1)<0)
+ C(6)*RESID(-2)^2 + C(7)*GARCH(-1)

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000376	0.000277	1.357983	0.1745
MA(10)	-0.014608	0.025503	-0.572803	0.5668
Variance Equation				
C	3.07E-06	6.99E-07	4.398775	0.0000
RESID(-1)^2	-0.064064	0.016873	-3.796743	0.0001
RESID(-1)^2*(RESID(-1)<0)	0.109838	0.016381	6.705210	0.0000
RESID(-2)^2	0.125669	0.025270	4.973077	0.0000
GARCH(-1)	0.871804	0.015290	57.01663	0.0000

R-squared	0.001888	Mean dependent var	0.000327
Adjusted R-squared	-0.002034	S.D. dependent var	0.018121
S.E. of regression	0.018139	Akaike info criterion	-5.772294
Sum squared resid	0.502423	Schwarz criterion	-5.747947
Log likelihood	4434.350	F-statistic	0.481319
Durbin-Watson stat	2.069974	Prob(F-statistic)	0.822649

Inverted MA Roots	.66	.53+.39i	.53-.39i	.20+.62i
	.20-.62i	-.20+.62i	-.20-.62i	-.53-.39i
	-.53+.39i	-.66		

Table 3

Dependent Variable: D(LOG(CLOSE))
Method: ML - ARCH (Marquardt) - Normal distribution
Date: 11/20/11 Time: 19:04
Sample (adjusted): 1/04/2005 3/31/2011
Included observations: 1534 after adjustments
Convergence achieved after 25 iterations
MA backcast: 12/21/2004 1/03/2005, Variance backcast: ON
GARCH = C(3) + C(4)*RESID(-1)^2 + C(5)*RESID(-1)^2*(RESID(-1)<0)
+ C(6)*RESID(-2)^2 + C(7)*GARCH(-1) + C(8)*LOG(VOLUME)

	Coefficient	Std. Error	z-Statistic	Prob.
C	0.000409	0.000268	1.522045	0.1280
MA(10)	-0.020115	0.023929	-0.840626	0.4006
		Variance Equation		
C	-0.000119	2.59E-05	-4.617316	0.0000
RESID(-1)^2	-0.068117	0.014880	-4.577800	0.0000
RESID(-1)^2*(RESID(-1)<0)	0.160458	0.022972	6.984846	0.0000
RESID(-2)^2	0.117919	0.023677	4.980276	0.0000
GARCH(-1)	0.816716	0.021827	37.41748	0.0000
LOG(VOLUME)	6.31E-06	1.34E-06	4.713631	0.0000

R-squared	0.002474	Mean dependent var	0.000327
Adjusted R-squared	-0.002101	S.D. dependent var	0.018121
S.E. of regression	0.018140	Akaike info criterion	-5.787474
Sum squared resid	0.502128	Schwarz criterion	-5.759648
Log likelihood	4446.992	F-statistic	0.540768
Durbin-Watson stat	2.070523	Prob(F-statistic)	0.803971

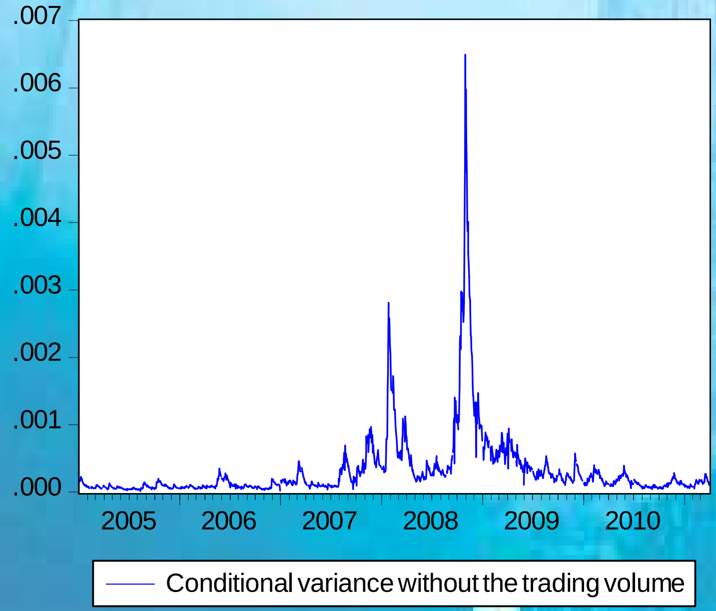
Inverted MA Roots	.68	.55-.40i	.55+.40i	.21+.64i
	.21-.64i	-.21+.64i	-.21-.64i	-.55-.40i
	-.55+.40i	-.68		

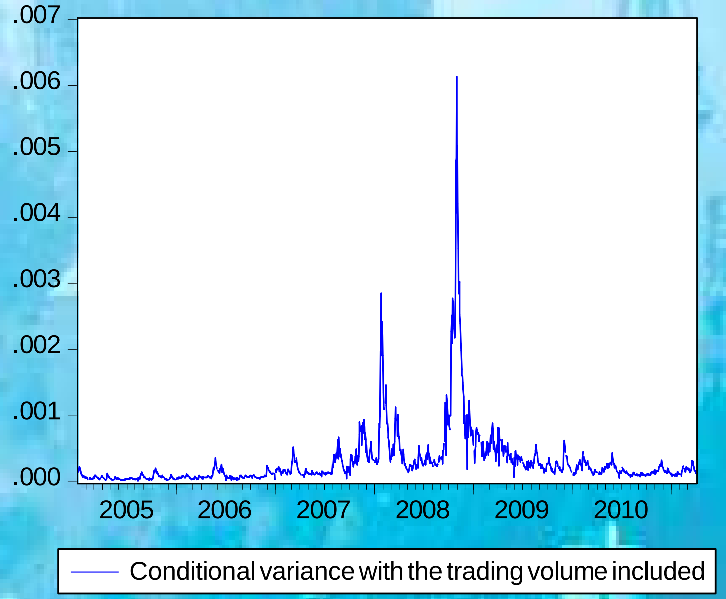
- the received results show quite high degree of the volatility persistence, since the sum $\sum_{i=1}^{q}\hat{\alpha}_i + \sum_{i=1}^{p}\hat{\beta}_i$ is high
- in model without trading volume variable it takes value of 0,933409 and besides this fact also the existence of the leverage effect was proved (since the corresponding parameter is statistically significant and positive)
- in model with trading volume variable the volatility persistence slowly declined to 0,866518

The diagnostic check statistics of the standardized residuals

- in order to have the information about adequacy of the presented estimates, we tested the standardized residuals
- the uncorrelatedness of the standardized residuals and squared standardized residuals was proved using the Ljung – Box Q – statistics and Q^2 – statistics, respectively
- the normality was not confirmed (Jarque – Bera test)

Conditional variance without and with the trading volume included





Concluding remarks

- the logarithm of the trading volume was included into the conditional volatility equation in order to investigate if it is a good proxy for information arrival
- taking into account some other papers (e.g. [Girard and Biswas 2007], [Gursoy et al. 2008], [Sharma et al. 1996]), the results of our analysis coincide with theirs, i.e. that the trading volume can be in general considered (in case of the market-level analysis) to be only a poor proxy for information flow

References (Extract)

- Engle R.F. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica, 50, No. 4.
- Girard E., Biswas R. (2007) Trading Volume and Market Volatility: Developed versus Emerging Stock Markets, The Financial Review, 42, pp. 429 – 459.
- Gursoy G., Yuksel A., Yuksel A. (2008) Trading volume and stock market volatility: evidence from emerging stock markets, Investment Management and Financial Innovations, Vol. 5, Issue 4, pp. 200 – 210.
- Lamoureux C., Lastrapes N. (1990) Heteroscedasticity in stock return data: volume versus GARCH Effects, The Journal of Finance, Vol. XLV, No. 1, pp. 221-229.
- Sharma J.L., Mougoue M., Kamath R. (1996) Heteroscedasticity in stock market indicator return data: volume versus GARCH effects, Applied Financial Economics, Vol. 6, pp. 337-342.