New Evidence on the Relation between Trading Volume and Volatility

Sibel ÇELİK (Corresponding author)
Dept. of Insurance and Risk Management, Dumlupinar University, Turkey
Tel: 90-274-265-2031-4664  E-mail: sibelcelik1@gmail.com

Received: February 5, 2013    Accepted: February 19, 2013
doi:10.5296/ber.v3i1.3222    URL: http://dx.doi.org/10.5296/ber.v3i1.3222

Abstract
The paper tests the relationship between trading volume and return volatility within the scope of Mixture of Distribution Hypothesis and Sequential Information Arrival Hypothesis in Istanbul Stock Exchange (ISE) by using intraday data for the period between 04.02.2005 to 30.04.2010. Two sub-samples are used to consider the effect of Global crisis. The results show that findings differ across two sub-samples. The findings support the Mixture of Distribution Hypothesis in pre-crisis period. However, the evidence is mixed in crisis period. Although we reject the Mixture of Distribution Hypothesis in crisis period, we can not strongly reject the Sequential Information Hypothesis.

Keywords: Trading volume, Volatility, Intraday data

1. Introduction
The volume-volatility relationship has attract a great deal of attention from researchers during the past three decades. There are many reasons to examine volume-volatility relation. First, this relation provides information relating the structure of financial markets. Second, the price-volume relation is important for event studies that use price and volume relation data to draw inferences. Third, price-volume relation is critical for discussions on the distribution of speculative prices. Fourth price-volume relation has significant implications for researches in future markets (Karpoff, 1987).
is contemporaneous positive relation between volume and volatility. Because, all traders simultaneously receive the new information. Thus, new equilibrium is established and there will be no intermediate equilibrium. Since, the variables contemporaneously change in response to new information, it should be impossible to use past volatility data to forecast volume (vice versa).

Contrary to MDH, SIAH assumes that all traders receive new information in a sequential. In other words, traders change their positions when the new information arrives the market. However, all traders do not receive the new information at the same time. Therefore, the response of each traders to the new information establish an incomplete equilibrium. The final equilibrium is established when all traders receive and response to the new information. Thus, SIAH suggests that there should be a lead-lag relation between volume and volatility. Stated in other words, lagged values of volume may be used to forecast current volatility and vice versa.

In the previous literature, numerous papers investigate the price-volume relation within the frame of MDH and SIAH. While some evidence show a contemporaneous and positive relation between volume and volatility supporting MDH (Clark, 1973; Epps and Epps, 1976; Tauchen and Pitts, 1983; Harris 1986; Karpoff, 1987; Andersen, 1996; Alsubaie and Najand, 2009; Choi et al. 2012), there are also findings that are inconsistent with MDH (Najand and Yung 1991; Bessembinder and Seguin 1992, 1993; Darrat et al., 2003; Aggarwal and Mougoue, 2011). Similar with MDH, evidence for the validity of SIAH is mixed in the literature. Some of the studies support the SIAH (Brooks, 1998; Girma and Mougoue; 2002; Darrat et al., 2003; Darrat et. al, 2007; Chiang et al., 2010) while there is also evidence inconsistent with SIAH (Alsubaie and Najand, 2009; Boubaker and Makram, 2011; Choi et al., 2012) in the previous literature.

The availability of high frequency data makes it possible to examine volume-volatility relation by using high frequency based volatility or volume measures. However, the literature on the volume-volatility relation by using high frequency based measure is still sparse. For example, Gwilym et al. (1999) examine the intraday behavior of five-minute FTSE-100, Short Sterling and Long Gilt LIFFE futures returns volatility and volume. They find strong evidence of bi-directional causality on the basis of Granger causality test. Darrat et al. (2007) examine the dynamic relation between intraday trading volume and return volatility of large and small NYSE stocks and conclude that there is bi-directional causality between volume and volatility in accordance with the SIAH. Giot et al. (2010) investigate the relation between volume and realized volatility and find that number of trades is the dominant factor shaping the volume-volatility relation. Hatrick et al. (2011) investigate the causal relationship among realized volatility, the number of transactions and volume with the intraday time intervals of 10, 20 and 30 minute by using vector autoregressive models. They find that realized volatility reacts positively to the lagged average trade size. However, the realized volatility forms a negative relationship with the first few lagged number of trades. Halova (2012) tests the volume-volatility relation by using intraday oil and gas futures prices and volume traded in the NYMEX. As a result of Granger-causality tests, Halova (2012) concludes that past values of volume help explain volatility which agrees with the SIAH.

The volume-volatility relation is examined by various papers in Turkey as well. Okan et al.
(2009) examine the volume-volatility relationship (dynamic and casual) for the ISE-30 index futures using daily data for the period 2006-2008 by employing GARCH, EGARCH and VAR models. Their findings are consistent with SIAH, however they reject the MDH in Turkish stock index futures. Kırán (2010) tests the relation between volume and volatility of ISE-100 index by using GARCH, EGARCH and TGARCH models. Kırán (2010) concludes that MDH and SIAH are not valid in Istanbul Stock Exchange (ISE). Boyacığlu et al. (2010) examine volume-volatility relationship of ISE National 100 index by applying unstructured VAR model for the period 1997-2009. They find that there is long-term negative relation from trading volume to volatility. In addition, there is bidirectional relationship between return volatility and trading volume. In conclusion, they reject the MDH and SIAH in ISE.

This paper aims to test the MDH and SIAH hypotheses in Istanbul Stock Exchange (ISE). The paper contributes to existing literature in two ways. First, in the literature, researchers mostly use daily volatility measure as a proxy for volatility. However, previous literature suggest to use intraday data in order to obtain accurate volatility estimates (Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2001, 2002). High frequency based volatility measure is still used sparsely in volume-volatility relation studies. This paper uses high frequency based volatility measure- realized volatility- as a proxy for volatility. Second, previous literature mostly used GARCH models to examine the relation between volume and volatility. This paper differs from the previous literature in terms of using a high frequency based volatility forecasting model- HARX-RV. In similar with international literature, in Turkish literature also use daily volatility measure and GARCH models to investigate volume-volatility relation. As well as we know this paper is the first which use high frequency data based volatility measure and high frequency based volatility forecasting model in the studies of volume-volatility relationship. Therefore the preliminary findings are worth to consider.

The remainder of the paper proceeds as follows. We describe the data and research design in section 2. Section 3 presents the empirical findings. Section 4 concludes.

2. Data and Research Design

The dataset comprises of intradaily ISE-30 index data and daily trade volume data as a proxy for information arrival running from 04.02.2005 to 30.04.2010. The optimal sampling choice is very important in high frequency data analysis. We use 5-minute frequency to calculate the realized volatility following various papers in the literature (Andersen et al., 2001). The number of intraday observation is 80 and we have 104880 observation in total. In the literature, it is found that structural breaks (financial crisis) affects both volatility dynamics (Dungey et al., 2011) and the volatility-volume relationship (Karanasos and Kyrtsou, 2011). Therefore, we analyse two sub-sample considering the last global financial crisis. There are many studies used 17 July 2007 as a starting point of global crisis (Dungey,2009). Thus, we take 17 July 2007 as a starting date of global crisis and analyse volume –volatility relationship for two sub-sample (pre-crisis period runs from 04.02.2005 to 16.07.2007 and crisis period runs from 17.07.2007 to 30.04.2010).

If we denote the $\delta$ - period returns by, $r_{t,\delta} = p(t) - p(t - \delta)$ the daily realized volatility will be
computed as the summing corresponding $1/\delta$ high frequency intradaily squared returns as in Equation [1]:

$$RV_{t+1}(\delta) = \sum_{j=1}^{1/\delta} r_{t+j, \delta, \delta}^2$$  \[1\]

Firstly, we investigate the contemporaneous effect of information arrival on volatility by applying the HAR-RV model of Corsi (2004). HAR-RV model of Corsi (2004) based on the Heterogeneous Market Hypothesis (Müller et al., 1993) which proposes the presence of heterogeneity in the traders. The heterogeneity may result from differences in degree of information, geographical location or time horizons and so on. If we take different time horizons, traders with different time horizons will perceive, react and cause different types of volatility components. Therefore, short term, medium term and long term volatility components will exist.

HAR-RV model can be described as in Equation [2].

$$RV_{t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5, t} + \beta_M RV_{t-22, t} + \xi_{t+1}$$  \[2\]

$RV_t, RV_{t-5}$ and $RV_{t-22}$ denotes the daily, weekly and monthly realized volatility respectively.

Multi period realized volatility components such as weekly and monthly realized volatility is calculated as,

$$RV_{t,t+h} = h^{-1}[RV_{t+1} + RV_{t+2} + .... + RV_{t+h}]$$  \[3\]

We take $h=5$ and $h=22$ as the weekly and monthly volatility, respectively.

We extend the HAR-RV model of Corsi (2004) as HARX-RV model following Aguilar and Ringgenberg (2011) by adding the trade volume as a proxy for the information arrival in HAR-RV model. HARX-RV model is denoted as in Equation [5].

$$RV_{t+1} = \beta_0 + \beta_D RV_t + \beta_W RV_{t-5, t} + \beta_M RV_{t-22, t} + \alpha Vol_{t+1} + \xi_{t+1}$$  \[5\]

In Equation [5], $Vol_{t+1}$ is the trade volume as a proxy for information arrival for day, t+1. If information arrival affects the volatility, we would expect the coefficient of $Vol_{t+1}, \alpha$ to be positive and statistically significant.

After testing the contemporaneous effect of information arrival on volatility, we test the causality relationship between volume and volatility by applying Vector Auto Regression
(VAR) Granger Causality model. Before applying VAR model, we test the long run relationship between volatility and volume by using Johansen Procedure (1990). In the case of cointegration relationship between variables, it is necessary to apply Vector Error Correction Model (VECM) to test whether the long run equilibrium relationship between variables are present. Last, we apply VECM Granger Causality test to test the causality relation between volume and volatility. VECM Granger Causality test can be defined as in Equation [6].

\[
\Delta RV_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \Delta RV_{t-i} + \sum_{i=1}^{p} \theta_i \Delta Vol_{t-i} + \nu_i \delta_{t-i} + \varphi_i \\
\Delta Vol_t = \omega_0 + \sum_{i=1}^{p} \eta_i \Delta Vol_{t-i} + \sum_{i=1}^{p} \theta_i \Delta RV_{t-i} + \delta_i \epsilon_{t-i} + \zeta_i \tag{6}
\]

The null hypotheses are “RV does not Granger cause Vol” and “Vol does not Granger cause RV.

3. Empirical Findings

Table 1. shows the summary statistics of realized volatility and trade volume variables

Table 1. The Summary Statistics of Variables

<table>
<thead>
<tr>
<th></th>
<th>Pre-Crisis Period</th>
<th>Crisis Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>13.9675</td>
<td>14.314</td>
</tr>
<tr>
<td>Median</td>
<td>13.9978</td>
<td>14.3461</td>
</tr>
<tr>
<td>Maximum</td>
<td>14.9141</td>
<td>15.1764</td>
</tr>
<tr>
<td>Minimum</td>
<td>13.0165</td>
<td>12.2667</td>
</tr>
<tr>
<td>Std.Dev.</td>
<td>0.3294</td>
<td>0.4112</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.0596</td>
<td>-0.4425</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>2.6996</td>
<td>52.4549</td>
</tr>
<tr>
<td>JB-Stat</td>
<td>2.6846</td>
<td>75189.74***</td>
</tr>
</tbody>
</table>

Note: Volume is the logarithm of trade volume variable. RV represents realized volatility. *** indicates the %1 significance level.

It is clear in Table. 1 that both the mean values of volume and RV variables are greater in crisis period. Also, the maximum values and and the standard deviations of volume and volatility are greater in crisis period. The volume has left skewed distribution while the volatility has right skew distribution. Except volume in pre-crisis period, the variables do not distribute normally.

First, we estimate model [2] that does not include volume variable. The results are given in Table 2.

Table 2. The Result of HAR-RV Model

<table>
<thead>
<tr>
<th></th>
<th>Pre-Crisis Period</th>
<th>Crisis Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 )</td>
<td>0.0001***</td>
<td>0.0001**</td>
</tr>
</tbody>
</table>
Note: We use the Newey-West standard errors to overcome autocorrelation, or correlation, and heteroskedasticity in the error terms. ***, ** and * indicates the %1, %5 and %10 significance level, respectively.

In Table 2 the sum of parameter estimates ($\beta_1 + \beta_2 + \beta_3$) indicates the persistence. Then we estimate HARX-RV model that includes volume variable and give the results in Table 3.

Table 3. The Result of HARX-RV Model

<table>
<thead>
<tr>
<th></th>
<th>Pre-Crisis Period</th>
<th>Crisis Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$</td>
<td>-0.0039***</td>
<td>-0.0040***</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.0211</td>
<td>-0.0347</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.2468***</td>
<td>0.6117**</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.2875**</td>
<td>0.2518*</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.0002***</td>
<td>0.0002***</td>
</tr>
<tr>
<td>Adjusted-$R^2$</td>
<td>0.1238</td>
<td>0.2860</td>
</tr>
<tr>
<td>$F$</td>
<td>21.9996***</td>
<td>68.2181***</td>
</tr>
<tr>
<td>$\beta_1 + \beta_2 + \beta_3$</td>
<td>0.5554</td>
<td>0.8288</td>
</tr>
</tbody>
</table>

Note: We use the Newey-West standard errors to overcome autocorrelation, or correlation, and heteroskedasticity in the error terms. ***, ** and * indicates the %1, %5 and %10 significance level, respectively.

If trading volume is considered as a proxy of the arrival of new information, we expect that $\alpha > 0$. Similarly, if information arrival affects volatility, we expect that the coefficient of volume, $\alpha$ is to be positive and statistically significant in HARX-RV model estimation. In addition, the sum of parameter estimates ($\beta_1 + \beta_2 + \beta_3$) should be smaller when trading volume is included in model. In this case, we can say that volatility persistence tends to disappear. Both in pre-crisis period and crisis period, $\alpha$ is positive and statistically significant supporting that trading volume is considered as a proxy of the arrival of new information and there is positive relation between information arrival and volatility. Including trading volume into model reduce the persistence of volatility terms only in pre-crisis period. However, the persistence is greater in HARX-RV model than HAR-RV model in crisis period.

At a later stage, we test the causality relationship between volatility and volume. For this purpose, firstly we investigate the long run relationship between variables. According to Johansen procedure (1990) we find two cointegrating vector both in pre-crisis and crisis period. Therefore, we estimate VECM model instead of VAR model (With 3 lag according to Schwarz Criteria). For brevity, we only present the results of VECM Granger Causality test in Table 4.

Table 4. The Results of VECM Granger Causality Test
Panel A presents the causality between volume and volatility in pre-crisis period. According to the findings, we can not reject the null hypotheses of both “Vol does not Granger Cause RV” and “RV does not Granger Cause Vol” for pre-crisis period. This result means that there is no causality relation between volatility and volume in pre-crisis period. Panel B shows the causality between volatility and volume in crisis period. We reject the null hypotheses of “Vol does not Granger Cause RV” at 5% significance level and “RV does not Granger Cause Vol” at %10 significance level. Therefore, in crisis period, there is bidirectional relationship between volatility and volume.

In sum, the results differ across two sub-periods. In pre-crisis period, trading volume coefficient in HARX-RV model is positive and significant. In addition to this, including trading volume into model reduce the persistence of volatility terms. The another important finding is that there is not causality relation between volume and volatility. Since MDH assumes contemporaneous relation volume and volatility, there should be no causality relation between these two variables. Thus, the findings in pre-crisis period supports MDH suggesting that there is contemporaneous relation between volume and volatility. The findings of crisis period is rather mixed. The trading volume coefficient in HARX-RV model is positive and significant in crisis period and there is bidirectional relation between volume and volatility supporting SIAH. However, including trading volume into model does not reduce the persistence of volatility terms in crisis period. Thus, the findings do not fully support SIAH in crisis period.

4. Concluding Remarks

The paper aims to test the MDH and SIAH hypothesis by using high frequency based volatility measure and model in ISE. For this purpose we use intradaily ISE-30 index data and daily trade volume data as a proxy for information arrival running from 04.02.2005 to 30.04.2010. We apply two different methods. First we test whether there is contemporaneous relation between volume and volatility by applying HARX-RV model. As a result of HARX-RV model, we find positive and statistically significant relation between information arrival and volatility both in pre-crisis and crisis period. When trading volume is included, the persistence of volatility decrease only in pre-crisis period. At a later stage, we test the causality relationship between volatility and volume by using cointegration and VECM Granger causality model. Although we find no causality relation between volatility and volume in pre-crisis period, bidirectional relationship between volatility and volume is found in crisis period. Therefore, our results differ across two sub-periods. We conclude that MDH is valid only in pre-crisis period. However we can not strongly reject the SIAH in crisis period. These findings are not consistent
with those of Kıran (2010) who conclude that MDH and SIAH are not valid in Istanbul Stock Exchange (ISE) and Boyacıoğlu et al. (2010) who reject the MDH and SIAH in ISE. Overall the results are supportive of MDH in pre-crisis period. This may reflect that the arrival of new information is simultaneously available to all traders and so volume and volatility moves synchronously.

The paper can be extended in several ways. First, different variables can be used for as a proxy for information arrival such as the number of trades. Second, the paper can be extended using other high frequency based measures of volatility, such as two time scale estimator (Zhang et al., 2005) or realized kernel (Barndorff Neilsen et al., 2008).

References


Appendix 1. The Graphs of Variables for pre-crisis and crisis period

![Figure 1. The Graph of Variables in Pre-Crisis Period](image1)

![Figure 2. The Graph of Variables in Crisis Period](image2)

Copyright Disclaimer

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).