ANNALS OF ECONOMICS AND FINANCE **13-1**, 211–236 (2012)

The Relationship between Volatility and Trading Volume in the Chinese Stock Market: A Volatility Decomposition Perspective

Tianyi Wang

China Center for Economic Research, National School of Development
Peking University, 100871

and

Zhuo Huang*

China Center for Economic Research, National School of Development
Peking University, Beijing, 100871

We use heterogeneous autoregressive (HAR) model with high-frequency data of Hu-Shen 300 index to investigate the volatility-volume relationship via the volatility decomposition approach. Although we find that the continuous component of daily volatility is positively correlated with trading volume, the jump component reveals a significant and robust negative relation with volume. This result suggests that the jump component contains some "public information" while the continuous components are more likely driven by "private information". Discussion of the intertemporal relationship supports the information-driven trading hypothesis. Lagged realized skewness only significantly affects the continuous component.

Key Words: High frequency; Price jump; Trading volume.
JEL Classification Numbers: G10, G12, G14.

1. INTRODUCTION

Why do people trade and how do prices move? The trading volume and volatility are two key concepts in finance. By using high frequency data for Hu-Shen 300 index, we investigate the relationship between the trading volume and volatility in the Chinese stock market. The empirical results found in this paper shed light on the discussion of different finance theories on trading volume.

* Corresponding author. Email: zhuohuang@nsd.pku.edu.cn

1529-7373/2012

Since Black and Scholes (1973), the price process for a financial asset has often been modeled as a continuous diffusion process. However, the intraday data shows sharp price changes in some trading days. Those sharp changes are usually called price jumps since they cannot be described by a continuous process. Barndorff-Nielsen (2004) developed a rigorous theory based on semi-martingale processes and pointed out that daily volatility, known as quadratic variation, can be easily decomposed into a continuous component due to small price changes and a jump component contributed by large price movements. There have been plenty of researches discussing those two components themselves, but studies on the relationship between different volatility components and other financial variables are quite limited. Among those variables, we focus on trading volume which can provide great deal of information on price movement (Wang (2002)).

Studies since 1970's have indicated a strong positive contemporal correlation between volume and volatility[1], e.g. Karpoff (1987), Gallant et al. (1992), Zhao and Wang (2003), Yin (2010) etc. However, two very recent papers challenged this stylized fact using the volatility decomposition technique. Giot et al. (2010) finds that only the continuous component shows a positive contemporal volume-volatility relation, while the jump component shows negatively correlation. Amatyakul (2010) also presents the evidence showing similar negative correlation. To our best knowledge, there exists no such research on China's stock market. Using high-frequency data for Hu-Shen 300 index[2], we also find a significant and robust negative contemporal correlation between the jump component and trading volume, consistent with the finding for the US market.

Although researchers tend to agree that jumps have "information" implications, there is no clear conclusion on which type of information lies behind it. The negative correlation above provides a piece of empirical evidence which implies that the information behind price jump has "public" nature. This "public information" could be the release of macroeconomic information (Andersen et al. (2007)) or important market information exposure(Wang et al. (2011)). Even price jump itself can be a "public information", given market participants believe that there is important information to back it up. For the first type of information, such as a large deviation between expected and actual CPI, traders might have a consistent opinion on the valuation of influenced stocks resulting in a quick and

[1] Various theories have been proposed to explain this phenomenon such as mixed distribution hypothesis by Clark (1973), Tauchen and Pitts (1983), Andersen (1996); asymmetric information hypothesis by Kyle (1985) and Wang (1994); divergence of opinion hypothesis by Varian (1985).

[2] Hu-Shen 300 index is selected because:1)It is composed with large, high liquidity stocks from both Shanghai and Senzhen stock exchange which makes it a good representation of China's stock market; 2)It is the underline asset of index futures, therefore its data is much more reliable and available.

sharp price change without much trade. Hence there is a negative correlation. For the second type of information, such as a sharp drop or rise in stock price in the middle of the day, traders might temporarily stop trading to revaluate their portfolios. Such actions also result a negative correlation. On the contrary, if "private information" could provoke jumps, we are more likely to find a positive correlation, because "private information" needs large trading volume to reveal itself[3].

Besides the contemporal relations, we have also discussed the intertemporal volume-volatility relationship. Those results will provide evidence which can distinguish different driving forces behind trading [4]. Assume there are two kinds of traders. One of them (referred as "information trader") uses information (price, macroeconomic information, etc.) as a guide of their portfolio selection and the other (referred as "liquidity trader") trades only in need of liquidity. When liquidity traders sell their stocks due to exogenous reasons (such as an investing opportunity outside the stock market), people who are willing to take over those stocks require a compensation for loss of liquidity. This compensation is archived (in equilibrium) in a form of a lower current price and a higher future return. Since price movements in conjunction periods are opposite, the volatility is higher. Higher liquidity demand is followed by larger compensation and larger volatility in the following period. Therefore, as Campbell et al. (1993) showed, a positive correlation between current volatility and lagged trading volume is likely to be observed in liquidity trading. When informed traders trade their stocks due to private information, that information will spread over the market through price signal. For example, informed traders buy stock due to a piece of private good news and the stock price will rise. Other traders observe it and buy the same stock, resulting in a price rise afterwards. Since price movements in conjunction periods are in the same direction, the volatility is lower. The more the information exposed in the current period (means a higher volume), the lower the change in price in the following period and the lower the volatility accordingly. Therefore, when information-driven trading is dominated, we expect to observe a negative correlation between current volatility and lagged trading volume[5]. Wang

[3]Our result also shows a positive correlation between the continuous component and contemporary trading volume indicating that the continuous component is a result of "private information" spillover between traders.

[4]Most papers on the intertemporal relationship concentrate on forecasting, i.e. Tauchen et al. (1996), Rui et al. (2003),Chen et al. (2001). In this paper, we only focus on the logic behind this phenomena and leave the forecast implication for future research.

[5]Here is a numerical example. Assume there is no information. The stock price is 100 and volatility is 0. When there only exist liquidity trading, the current price drops to 98 and the following price returns to 100 (results a liquidity compensation equals 2). The following period's volatility is $2/98 > 0$. If the liquidity demand is stronger, we

(2004) refers the first scenario as the liquidity-driven trading hypothesis (LTD), while the second scenario is referred as the information-driven trading hypothesis (IDT). Our results indicate a significant negative correlation between current volatility and lagged trading volume after controlling dynamics of volatility itself and contemporal volume-volatility relation. This suggests that IDT is supported at least for our sample, Hu-Shen 300 index.

Furthermore, we decompose daily volatility into continuous and jump components using Hu-Shen 300 index high-frequency data and discuss the intertemporal relationship for different parts. Results show that only the continuous component and trading volume have a significant negative intertemporal correlation. As we have pointed out that the continuous component volatility is provoked by "private information", we conclude that "private information-driven trading" dominates the trading.

Our empirical analysis is based on the following building blocks. The volatility decomposition technique follows Barndorff-Nielsen and Shephard (2006). The volatility dynamics is modeled by heterogenous autoregressive model (HAR) proposed by Corsi (2009). Return's leverage effect is modeled by the same setup in Bollerslev et al. (2009). The treatment of trading volume is referred to Wang (2004).

We add a new term - "realized skewness" of return in our model. As earlier studies on return skewness usually use daily returns, they can only discuss skewness on the weekly, monthly or even quarterly frequency. Utilizing high-frequency data, we are able to discuss it on daily frequency. Realized skewness is included in our model since it reflects whether the price movement is "large rises and small drops" or "small rises and large drops", or more intuitively, the "market sentiment". Our results show skewness is negatively correlated with the total and continues component. It is irrelevant with the jump component. Therefore, the downside market sentiment will raise future volatility only through continues component.

As an attempt of using high frequency trading volume data, we utilized a simple notion of mean excess (intra-day) trading volume to characterize its dynamics around jumps. The results show that no significant illiquid or overliquid before jumps and a substantial and persistent volume shrinkage after jumps.

This paper differs from earlier studies on this issue in several ways. First, we use high-frequency data to calculate nonparametric volatility measures,

---

expect a sharper drop, say to 96, then the following period's volatility is 4/96, which is higher than 2/98. On the contrary, if there exist informational trading, the stock price will eventually rise to 104. However, since information is revealed by price change, the current price will not jump to 104 directly. It will jump to a smaller one, say to 102, then move to 104 in the following period. If the information is revealed more in the current period, the price will change more in the current period, say to 103, which leaves the return equals $1/103 < 2/102$ in the following period. In terms of volatility, it is smaller.

which are more accurate and leads to more reliable results. Second, we use volatility decomposition technique and discuss the volume-volatility relationship for both continuous and jump components. Although volatility decomposition technique has been applied to Chinese stock market by some researchers, no study has focused on the relationship between different components and other financial variables yet. Third, our econometric model is more flexible and capable to discuss the intertemporal correlation while controlling for volatility dynamics and price effect. Fourth, we have discussed realized skewness on daily frequency using high-frequency data. Last, a preliminary discussion using high frequency trading volume is provided.

The rest of the paper is organized as follows. Section 2 provides a brief introduction to volatility decomposition; section 3 discusses the data we used; section 4 focuses on model setups; section 5 summarizes main results and section 6 concludes.Selected nonlinear relationships and robustness check are also discussed.

2. VOLATILITY DECOMPOSITION

Volatility decomposition is based on quadratic variation theory. The instant price change dp_t follows a geometric Brownian motion with jumps.

$$p_t = \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s) + \sum_{i=1}^{N(t)} \kappa(s_j)$$

$\mu(t)$ is a continuous process with finite variance, $\sigma(t) > 0$ is a càdlàg instant volatility process, $W(t)$ is a Brownian motion and $N(t)$ is a poisson process with time varying intensity $\lambda(t)$, $\kappa(s_j)$ is the magnitude of j'th jump. The quadratic variation process $[p_t]$ is defined as:

$$[p_t] = \text{plim} \sum_{j=0}^{n-1} (p_{\tau_{j+1}} - p_{\tau_j})^2 \to \int_0^t \sigma^2(s)ds + \sum_{j=1}^{N(t)} \kappa^2(s_j) \quad n \to \infty$$

where τ_{j_n} is a partition of $[0, t]$. When $n \to \infty$, $\sup_j \{p_{\tau_{j+1}} - p_{\tau_j}\} \to 0$, the limit result provides us a new view of volatility. The first part is called integrated variance, which reflects the contribution of variation from the continuous price process. The second part is the sum of square of jumps' magnitude, which reflects the contribution of variation from the jump process. The second part will be zero in the absence of price jump.

In practice, price process can be sampled by different methods such as sampling with every k seconds (calendar time sampling, CTS), sampling with every k trades (tick time sampling, TTS) etc. We use CTS for this paper. Let $r_{t,j} = p_{t-1+j/M} - p_{t-1+(j-1)/M}$, where M is the size of daily sample. The sample counterpart of the above limit is

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2 \to \int_{t-1}^{t} \sigma^2(s)ds + \sum_{j=N(t-1)+1}^{N(t)} \kappa^2(s_j) \qquad M \to \infty$$

Literature documented this statistic as realized quadratic variation or realized variation(RV). RV converges to quadratic variation as sampling frequency increased or equivalently sampling interval goes to zero. RV is the sum of the continuous component and the jump component. We refer it as total variance that is equivalent to traditional volatility measures such as squared return, absolute return(Chow and Lawler (2003)) and GARCH type filtered volatility(Fabozzi et al. (2004)).

Barndorff-Nielsen (2004) proposed a consistent integrated variance estimator called bipower variation (BV):

$$BV = \mu_1^{-2} \sum_{j=2}^{M} |r_{t,j}||r_{t,j-1}| \to \int_{t-1}^{t} \sigma^2(s)ds \qquad M \to \infty$$

where $\mu_1 = \sqrt{\pi/2}$. When sampling frequency increases, price jump will not affect BV since at least one of $|r_{t,j}|$ and $|r_{t,j-1}|$ will shrink to zero as sampling interval shrinks to zero.

The natural implication from above is that the jump component can be measured non-parametrically by the difference between RV and BV. According to the simulation study by Huang and Tauchen (2005), the logarithm of RV and BV will deliver more stable results. Therefore, we use $J_t = \{\ln(RV_t) - \ln(BV_t)\}$ as the jump measure.

In theory, J_t cannot be negative. Since the sampling interval cannot approach to zero, it is possible to get a negative J_t. There are two possible methods to deal with it. One of then is to take $J_t < 0$ as measurement error (i.e. Bollerslev et al. (2009)). The other way is to determine jumps according to a critical threshold from the statistical distribution of J_t.

Barndorff-Nielsen and Shephard (2006) show that:

$$Z_t = \frac{\sqrt{M}}{\sqrt{(\mu_1^{-4} + 2\mu_1^{-2} - 5) IQ_t BV_t^{-2}}} \{\ln(RV_t) - \ln(BV_t)\} \to N(0,1)$$

where IQ_t is estimated via Quad-power variation:

$$IQ_t = M\mu_1^{-4} \sum_{i=4}^{M} |r_{t,i}||r_{t,i-1}||r_{t,i-2}||r_{t,i-3}| \to \int_{t-1}^{t} \sigma^2(s)ds$$

Therefore, we define $J_t^* = \{J_t | Z_t > \alpha\}$ as the continuous component, where α is the critical value. The limit results also indicate that J_t is a heteroskedastistic truncated normally distributed variable.

With non-truncated data, we can use the whole sample and it is free from selecting a critical value for jump test. Truncated data enables us to focus on statistically significant jumps but it will be subject to severe data loss. We use non-truncated data first then use truncated data to confirm results.

3. DATA

The raw data in this paper is Hu-Shen 300 index (SZ399300) 1 minute high-frequency data ranging from 01/04/2007 to 12/31/2010. The data source is Hexun database. After eliminating trading days with missing data, the final dataset contains 947 days.

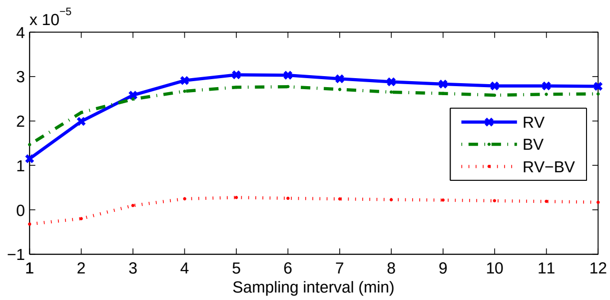

FIG. 1. Volatility signature plot for Hu-Shen 300 index (2007-2010)

In theory, sampling frequency should be as high as possible in order to accurately estimate RV. However, in the existence of microstructure noise, high sampling frequency will induce strong noise interference. Here, we use volatility signature plot to determine the optimal sampling frequency.

Specifically, we calculate RV, BV and J using sampling interval ranging from 1 to 12 minutes and present the result in Figure 1. All of the three variables are stable when the sampling intervals are larger than 5 minutes. Therefore, we use "5 minutes" for our sampling interval which is consistent with existing literatures. However, since the Hu-Shen 300 index is a non-tradeable asset, the signature plot is upward sloped, unlike the downward sloped curve for individual stocks. Similar results for S&P 500 cash index can be found in Huang et al. (2007). See Hansen and Lunde (2006) for further discussions. We also use the "rule of thumb" optimal (equal space) sample size proposed by Bandi et.al.(2008) to confirm the optimal sampling frequency. Specifically, the optimal sample size m^* is:

$$\hat{m_t^*} = \left\{ \frac{m_L}{3} \sum_{j=1}^{m_L} \hat{r}_{j,t}^4 \left(\frac{\sum_{t=1}^{T} \sum_{j=1}^{m} \hat{r}_{j,t}^2}{Tm} \right)^{-2} \right\}^{1/3}$$

where T is number of days in the sample, m is number of data points each day m_L correspond with a lower frequency. Bandi et al. (2008) suggests using "15 minutes" for m_L[6]. Substituting data into the formula, we have the optimal sampling frequency on average is 4.6782.

Before the opening of the morning market, there is a "call auction" period which is different from "continuous auction" used for the rest of the day. Therefore, we eliminate the first 1 minute from the first 5 minutes and leave the other 5 minutes' data intact, this procedure also eliminates the impact of overnight return on the calculation of realized measure [7]. Using volatility decomposition method mentioned above, we decompose Hu-Shen 300 index volatility into the continuous component (measured by BV) and the jump component (measured by J). The corresponding time series are plotted in Figure 2 and basic statistics are listed in Table 1.

TABLE 1.

Basic statistics of different volatility components

Series	mean	variance	median	skewness	kurtosis	Ljung-Box(10)	ADF(10) p
$\ln(BV_t)$	-8.3036	0.6956	-8.3216	0.1804	2.9039	2470.01	0.0003
$\ln(RV_t)$	-8.1832	0.6552	-8.2190	0.1813	2.8237	2484.75	0.0004
J_t	0.1204	0.0238	0.1050	0.6717	3.8956	52.56	0.0000
r_t	0.0004	0.0006	0.0025	-0.1805	5.6466	13.81	0.0000
$\ln(V_t)$	1.7799	0.2250	1.8253	-0.2302	2.4469	5412.71	0.0100

From basic statistics, we find that compared with RV and BV, the log version of them are not far from normal distribution in terms of skewness and kurtosis[8]. The mean and median of $\ln(RV)$ are larger than $\ln(BV)$ which is consistent with theory. $\ln(RV)$ and $\ln(BV)$ have strong persistent features while J is far less persistence, indicated by Ljung-Box statistics. Such feature was also documented in Bollerslev et al. (2009), Chen and Wang (2010). The fact that the jump component is less persistent implies that it is much more difficult to predict it.

[6]Bandi et al. (2008) show that the optimal sampling frequency is $(Q_t/(E(\epsilon_t^2))^2)^{1/3}$, where $Q_t = \frac{m}{3} \sum_{j=1}^{m} \hat{r_{j,t}^4}$ and suggest using 15 minutes for estimation.

[7]Lunch break is different from morning opening since it is much shorter and does not involve call auction.

[8]Other benefits of taking logarithm can be found in Barndorff-Nielsen and Shephard (2005).

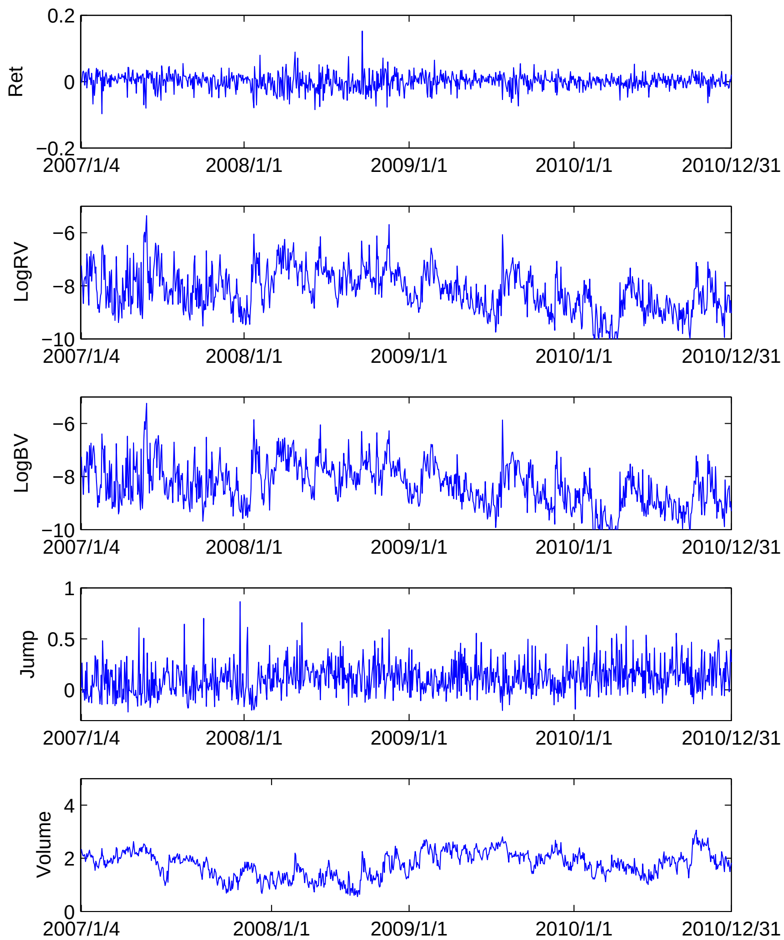

FIG. 2. Time series plot for data

Gallant et al. (1992) used dummies and time trends to preprocess trading volume series, since they found nonstationary components in the data. Our pre-test on stationary shows in our sample period, Hu-Shen 300 index is stationary at 2% level. Therefore, we leave trading volume intact.

4. MODEL SPECIFICATION

4.1. Volatility modeling

In this section, we use Corsi (2009) and Bollerslev et al. (2009) as building blocks with trading volume and realized skewness[9] to model volatility. Specifically, we model trading volume and its leverage effect as:

$$G(V_t, V_{t-1}, ...V_{t-m}) = \sum_{k=0}^{m} (\phi_{1,k} \ln(V_{t-k}) + \phi_{2,k} \ln(V_{t-k}) I\{r_{t-k} < 0\})$$

According to mixed distribution hypothesis (MDH), both price movements and trading volume are results of "information arrival" and literatures often treat trading volume as a proxy of information intensity. The second term is the leverage effect which enables our model to respond to trading volume associated with price rise or drop differently. It is common to consider a price rise as the result of "good news" and a price drop as the result of "bad news". When combined with trading volume, this leverage effect is a much more direct measure of "good news" and "bad news" than just return leverage.

Since the realized variance (total volatility) and the continuous component have fairly similar statistical properties, we use the same model to discuss both series. Here we take the continuous component volatility for example. Corsi (2009) proposes a simple heterogeneous autoregressive model (HAR) to model the long memory property of volatility dynamics. We add a GARCH structure in the residual of HAR model to make it more flexible.

$$\left.\begin{array}{rcl} \ln(BV)_t &=& \alpha_0 + \alpha_d \ln(BV)_{t-1} + \alpha_w \ln(BV)_{t-5:t-1} + \alpha_m \ln(BV)_{t-22:t-1} \\ && +\theta_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \theta_2 I\{r_{t-1} < 0\} + \theta_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I\{r_{t-1} < 0\} \\ && +\theta_4 RS_{t-1} + G(V_t, V_{t-1}, ...V_{t-m}) + \epsilon_t \\ \epsilon_t &=& \sqrt{h_t} u_t \\ h_t &=& \exp(\lambda_0 + \sum_{j=1}^{s} \lambda_j \ln(BV_{t-j})) + \sum_{j=1}^{p} \gamma_j h_{t-j} + \sum_{j=1}^{q} \beta_j \epsilon_{t-j}^2 \end{array}\right\} (M_{BV})^{10}$$

[9] Trading volume is a measures of trading intensity. Since we are using index data, we do not have other trading intensity measure (such as the number of trades, average number of shares traded per trade etc.(Jones et al. (1994), Chan and Fong (2000))) than trading volume.

[10] For the model of total volatility, we replace BV for RV and name it as M_{RV} correspondingly.

Where u_t follows generalized error distribution (GED). Compared with normal distribution, GED does better in fitting $\ln(BV)$'s unconditional distribution's tail thickness.

The mean equation contains three parts: the first part (terms associated with α_d, α_w, α_m) is the traditional HAR model terms of $\ln(BV)$. Specifically:

$$\ln(BV)_{t+1-k:t} = \frac{1}{k}\sum_{j=1}^{k}\ln(BV_{t-j})$$

By adding those terms, this model can capture the long memory property of $\ln(BV_t)$ indicated by Ljung-box statistic. We set k equals 5 and 22 for weekly and monthly average volatility respectively. The second part (terms associated with θ_1, θ_2, θ_3, θ_4) captures varieties of return effects. It includes: return level (θ_1) known as the scale effect measuring the effect of the lagged absolute return on current volatility, leverage effect (θ_2, θ_3) capturing the asymmetric effects of positive and negative returns on volatility, and realized skewness[11] effect (θ_4). The realized skewness is defined as:

$$RS_t = \sum_{t=1}^{M}\left(\frac{r_{t,j}}{\sqrt{RV_t}}\right)^3$$

It measures the asymmetry of intra-day returns. Considering "5 minutes" returns in a certain trading day, if the absolute value of ups are larger than the absolute value of downs, then there are a "fast rise, slow drop" scenario and a positive realized skewness. On the contrary, if the absolute value of downs are larger than the absolute value of ups, we have a "fast drop, slow rise" scenario and a negative realized skewness. In this sense, realized skewness is a measure of market sentiment. If market sentiment is optimistic, the skewness is positive and if the market sentiment is pessimistic, the skewness is negative. The last part ($G(\cdot)$) captures trading volume's effect on volatility.

Barndorff-Nielsen and Shephard (2005) points out that volatility-of-volatility is related with the magnitude of volatility. Therefore, we add $\ln(BV)$ into the variance equation. Unlike Bollerslev et al. (2009), we use multiplicative heteroskedasticity rather than additive heteroskedasticity for the stability of the solution since it guarantees a positive variance.

We model the jump component with HAR model rather than the simple autoregressive model used in Bollerslev et al. (2009) because the significant lags in AR(22) model all lie outside one week. HAR structure will make

[11] Here we use original moment rather than central moment to measure skewness since the mean of 5 minutes returns are close to zero. The t-statistics for mean zero in all trading days never exceeded 0.6. Amaya and Vasquez (2010) also uses this definition.

the model more concise. Specifically:

$$\left.\begin{array}{rcl}\ln\left(\frac{RV}{BV}\right)_t & = & \delta_0+\delta_d\ln\left(\frac{RV}{BV}\right)_{t-1}+\delta_w\ln\left(\frac{RV}{BV}\right)_{t-5:t-1}+\delta_m\ln\left(\frac{RV}{BV}\right)_{t-22:t-1}\\ & & +\theta_1\frac{|r_{t-1}|}{\sqrt{RV_{t-1}}}+\theta_2 I\{r_{t-1}<0\}+\theta_3\frac{|r_{t-1}|}{\sqrt{RV_{t-1}}}I\{r_{t-1}<0\}\\ & & +\theta_4 RS_{t-1}+G(V_t,V_{t-1},...V_{t-m})+\nu_t\end{array}\right\}(M_J)$$

where

$$\ln(\frac{RV}{BV})_{t+1-k:t}=\frac{1}{k}\sum_{j=1}^{k}\ln(\frac{RV}{BV})_{t-j}$$

Since $\ln\left(\frac{RV_t}{BV_t}\right)$'s distribution has heteroskedasticity, we use HSK robust standard error for inference. Like the continuous component's model, there are also three parts in mean equation: the first part captures serial correlations; the second part contains scale, leverage and skewness effects; the third part contains trading volume terms.

4.2. Calender effect on volatility

Amatyakul (2010) points out the importance of considering calender effect in the research in volume-volatility relationship. Extensive literatures have cast interest on "calender effect" in China's stock market, Zhou and Chen (2004) shows that the volatility on Monday is higher than the average level on Shanghai stock market. Gao and Kling (2005) found Fridays are profitable in term of returns in Shanghai stock market. Zhang et al. (2005) show that the calender effect changes violently through time depending on whether traders formulate their portfolios based on calender effect or not. Since former researches show that calender effects are highly time and asset sensitive, we briefly discuss it in this section.

By adding dummy in simple HAR model, we have:

$$\begin{array}{rcl}\ln(BV_t) & = & \alpha_0+\alpha_d\ln(BV)_{t-1}+\alpha_w\ln(BV)_{t-5:t-1}\\ & & +\alpha_m\ln(BV)_{t-22:t-1}+\delta D_{J,t}+\mu_t\\ \ln(\frac{RV}{BV})_t & = & \delta_0+\delta_d\ln\left(\frac{RV}{BV}\right)_{t-1}+\delta_w\ln\left(\frac{RV}{BV}\right)_{t-5:t-1}\\ & & +\delta_m\ln\left(\frac{RV}{BV}\right)_{t-22:t-1}+\delta D_{J,t}+\nu_t\end{array}$$

where $J=$\{Mon,...,Fri\}. Based on rolling window estimation with a window set to 250 days, we present t-statistics for corresponding daily effect in Figure 3. The dot line is ± 1.64 which corresponds to significant level equals 10%.

We only report weekdays that have t-statistics cross significant boundaries. For the continuous component, only Monday and Tuesday's dummies

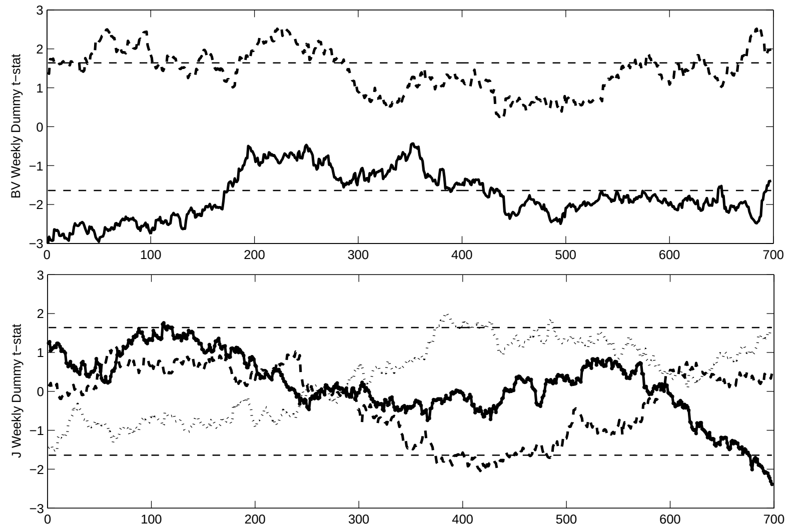

FIG. 3. Rolling windows estimation (U:Continuous component,L:Jump component)

cross the significant boundaries for a considerable period of time and never change sign during the sample period. For total volatility RV, we use the same setup as BV and replace BV with RV. Since the results are almost the same as BV's, we do not include them for saving space. On the contrast, the jump component does not have recognizable calendar effect although some dummies cross significant boundaries occasionally (but all of which cross boundaries change signs during sampling period). Therefore, we include Monday and Tuesday dummies in the continuous component and total volatility model, and do not include any weekday dummy in the jump component model.

5. ESTIMATION RESULTS

In this section, we focus on results of the continuous component's model M_{BV} and the jump component's model M_J. Since results of the total volatility model M_{RV} is consistent with former literatures (e.g. Wang (2004)) and similar to the continuous component's model, we only report results related to dynamics volume-volatility relation.

5.1. M_{BV} (The continuous component) model

There are four order parameters in this model: the lag order of trading volume m, the lag order of GARCH model (p, q) and the lag order of $\ln(BV)$ in variance equation s. Here, we set m equals to 1 and 2 to discuss contemporal and intertemporal volume-volatility relation. For given m, BIC criteria shows $(p, q, s) = (1, 1, 1)$ as a reasonable and concise choice. Results are presented in Table 2 with t-statistics in parentheses. (1), (2) and (3) are models without calender effect dummies while $(1^{'})$, $(2^{'})$ and $(3^{'})$ are models with calender effect dummies.

TABLE 2.

Volume-volatility relation: the continuous component

	(1)		$(1^{'})$		(2)		$(2^{'})$		(3)		$(3^{'})$	
α_0	-0.942	(3.95)	-0.899	(3.79)	-1.260	(5.63)	-1.253	(5.63)	-1.268	(5.66)	-1.243	(5.56)
α_d	0.427	(10.78)	0.428	(10.51)	0.374	(10.26)	0.374	(10.19)	0.501	(14.54)	0.500	(14.45)
α_w	0.317	(5.64)	0.315	(5.60)	0.328	(6.52)	0.328	(6.48)	0.273	(5.95)	0.274	(5.95)
α_m	0.159	(3.32)	0.162	(3.38)	0.233	(5.08)	0.233	(5.10)	0.160	(3.69)	0.161	(3.70)
θ_1	0.086	(3.27)	0.077	(2.91)	0.061	(2.51)	0.052	(2.13)	0.079	(3.42)	0.075	(3.19)
θ_2	-0.097	(1.84)	-0.099	(1.90)	-0.065	(1.33)	-0.061	(1.23)	0.246	(2.20)	0.237	(2.11)
θ_3	0.152	(3.45)	0.155	(3.51)	0.233	(5.66)	0.236	(5.72)	0.233	(6.07)	0.234	(6.05)
θ_4	-0.616	(3.53)	-0.592	(3.40)	-0.306	(1.88)	-0.255	(1.58)	-0.515	(3.27)	-0.496	(3.15)
$\phi_{1,0}$					0.235	(6.39)	0.238	(6.46)	0.963	(14.87)	0.964	(14.77)
$\phi_{2,0}$					0.194	(11.49)	0.191	(11.29)	0.220	(13.20)	0.217	(13.03)
$\phi_{1,1}$									-0.734	(10.50)	-0.736	(10.45)
$\phi_{2,1}$									-0.187	(3.12)	-0.181	(3.00)
λ_0	-8.882	(4.02)	-8.865	(3.96)	-7.665	(3.69)	-7.651	(3.65)	-9.845	(3.89)	-9.693	(3.83)
λ_1	-0.545	(2.40)	-0.540	(2.35)	-0.404	(1.93)	-0.401	(1.91)	-0.521	(2.00)	-0.503	(1.95)
β_1	0.068	(2.60)	0.064	(2.50)	0.065	(2.35)	0.063	(2.27)	0.036	(2.79)	0.035	(2.75)
γ_1	0.882	(21.26)	0.887	(21.06)	0.875	(16.52)	0.878	(16.27)	0.942	(45.99)	0.943	(45.48)
$\ln(\nu)$	0.419	(6.87)	0.418	(6.83)	0.406	(6.46)	0.400	(6.33)	0.566	(8.62)	0.571	(8.76)
BIC	1492.1		1498.8		1349.3		1258.6		1235.5		1246.0	
LogL.	-701.7		-698.2		-623.4		-621.3		-559.7		-558.1	

Absolute t statistics in parentheses

Firstly, α_d, α_w and α_m are significantly positive indicating a highly persistent feature which is consistent with Ljung-Box statistics. Secondly, the autocorrelation is declining along with time, i.e. $\alpha_d > \alpha_w > \alpha_m$ which is different from US market's results (α_w is larger than α_w and α_m) reported in Bollerslev et al. (2009). Furthermore, the coefficient of lagged monthly average volatility is much larger than it is in US market indicating the "long memory" in Hu-Shen 300 index is stronger.

Secondly, we focus on return effects on volatility through θ_1, θ_2, θ_3. θ_1 is positively significant at 5% level suggesting that a higher current absolute return will result in a higher future volatility even after controlling the dynamics of volatility by HAR structure[12]. θ_2 is not significant in most models and changes sign occasionally. Therefore, the "sign" effect is not stable. θ_3 measures the traditional leverage effect and it is positively significant, which coincides with extensively documented leverage effect in literatures.

θ_4 indicates how lagged skewness affects current volatility. It is significant at 12% level when we only control contemporal trading volume and it is significant at 1% under all other settings. The negative coefficient means that an increase in lagged skewness will reduce current volatility while a decrease in lagged skewness will increase current volatility. As we discussed above, skewness is an indicator of relative speed between ups and downs, that is related to market sentiment. If traders are in high sprite, the market price will move up shapely and drop slowly, that results in an increase in skewness and a decrease in volatility.

$\phi_{1,0}$ and $\phi_{2,0}$ capture trading volume and its leverage term's effect on the continuous component volatility. Results suggest both of them are positively significant at 1% level. The positive relation highlighted by $\phi_{1,0}$ confirms long standing literature on contemporal volume-volatility relation. $\phi_{2,0}$ indicates a significant leverage effect even when we have controlled conventional leverage effect defined on returns.

The most interesting part is the dynamic volume-volatility relation summarized in $\phi_{1,1}$ (volume level) and $\phi_{2,1}$ (volume leverage). Results indicate that at least on 1% both of them are negatively significant, which means higher lagged trading volume will reduce current volatility. And volume related to price drop will reduce volatility more than it will when associated with price rise. This also indicates volatility is more sensitive to negative price movement.

5.2. M_{RV} (total volatility) model results

The orders selection result is the same for M_{RV} which supports $(p, q, s) = (1, 1, 1)$. Results can be found in Table 3. We only report mean equation to save space. Model (1) does not contain weekday dummies while model (2) does.

Like BV models, we still have α_d, α_w and α_m positively significant with decreasing order $\alpha_d > \alpha_w > \alpha_m$ indicating a decrease in autocorrelation. We also find a positive contemporal volume-volatility relation as well as a negative relation between skewness and volatility. Importantly, the current volatility is negatively correlated with lagged trading volume after control-

[12]Generally specking, absolute return is a noisy proxy of daily volatility. The fact that θ_1 is significant, even after we control lagged volatility indicates that "noise" also contains information in volatility. We also add lagged jump components in our model and find θ_1 remains significant, which means θ_1 is significant not just because it is correlated with jumps.

TABLE 3.

Total variance and trading volume

	(1)		(2)	
α_0	-1.193	(5.49)	-1.169	(5.41)
α_1	0.517	(14.71)	0.515	(14.64)
α_w	0.262	(5.53)	0.263	(5.53)
α_m	0.158	(3.64)	0.159	(3.46)
θ_1	0.090	(3.69)	0.086	(2.07)
θ_2	0.230	(2.15)	0.223	(5.39)
θ_3	0.212	(5.43)	0.213	(4.26)
θ_4	-0.665	(4.33)	-0.655	(4.26)
$\phi_{1,0}$	0.888	(13.70)	0.886	(13.59)
$\phi_{2,0}$	0.218	(13.42)	0.216	(13.26)
$\phi_{1,1}$	-0.681	(9.73)	-0.682	(9.67)
$\phi_{2,1}$	-0.177	(3.05)	-0.172	(2.95)
BIC	1158.4		1168.2	
LogL	-521.16		-519.23	

Absolute t statistics in parentheses

ling volatility dynamics and contemporal trading volume. According to Wang (2004), this result suggests that the driving force behind Hu-Shen 300 stocks is mainly information. This result, combined with similar results in the continuous component volatility and results being presented in section 5.3, indicates no significant intertemporal volume-volatility relation in the jump component, we conclude that "private information" is the information which drives Hu-Shen 300 trading.

5.3. Nonlinear dynamic volume-volatility relationship

Barclay and Warner (1993) proposed "stealth trading hypothesis" concluding that informed trader will use median trading volume to maximize their private information's value. The logic lies that heavy trading volume will expose their favored information too quickly while light trading volume will induce high transaction cost. Therefore, if the information-driven trading hypothesis holds, we expect to find a more negative relation in median trading volume. We use M_{RV} and M_{BV} model to test this implication[13]. Specifically, we re-define $G(\cdot)$ as:

$$\phi_{1,0}\ln(V_t) + \phi_{1,1}\ln(V_{t-1}) + \phi_{2,1}\left\{\ln(V_{t-1})\right\}^2$$

[13]Since no dynamic volume-jump volatility is found, we do not discuss nonlinear relation between them.

In Table 4, (1) is for total volatility and (2) is for the continuous component volatility.

TABLE 4.

Nonlinear volume-volatility relation

	(1)		(2)	
α_0	-0.317	(12.89)	-0.321	(1.17)
α_1	0.499	(5.07)	0.493	(12.97)
α_w	0.271	(3.41)	0.273	(5.32)
α_m	0.163	(2.90)	0.167	(3.51)
θ_1	0.084	(1.57)	0.075	(2.70)
θ_2	-0.073	(4.42)	-0.077	(1.58)
θ_3	0.189	(3.93)	0.206	(4.86)
θ_4	-0.648	(10.90)	-0.541	(3.12)
$\phi_{1,0}$	0.761	(6.91)	0.827	(11.92)
$\phi_{1,1}$	-1.397	(4.09)	-1.482	(7.14)
$\phi_{2,1}$	0.230	(2.29)	0.239	(4.09)
BIC	1318.7		1385.5	
LogL	-597.9		-631.3	

Absolute t statistics in parentheses

Results show that both volatilities have significantly positive coefficients on lagged quadratic trading volume, which means there exists a minimum variance trading volume. This reinforces stealth trading hypothesis and information-driven trading.

5.4. M_J (The jump component) model

There is only one order parameter needed to be chosen in M_J model and as before, we set m equals 1 and 2. We do not include any calender dummy in M_J model since there is no recognizable calender effect. Furthermore, we use HSK robust standard error for inference in model (1)-(3) because of the heteroskedasticity in $\ln(RV/BV)$. In model (4), we use the jackknife standard error as a robustness check to guard against large jumps (outliers). See Table 5 for results.

δ_d, δ_w and δ_m capture the autocorrelation in the jump component volatility. Among them, δ_m is the only significant coefficient since it is significantly larger than the others. Similar results can be found in Bollerslev et al. (2009), Chen and Wang (2010). This phenomena indicates that the volatility clustering of the jump component is much weaker than the continuous component and it is unlikely to observe two large jumps happening in conjunction days.

Return effects illustrated by $\{\psi_1, \ldots, \psi_4\}$ are all insignificant suggesting that (at least in our sample) return information is not informative for

TABLE 5.

Jump component and trading volume

	(1)		(2)		(3)		(4)	
ψ_0	0.048	(3.36)	0.113	(4.50)	0.125	(3.89)	0.125	(3.78)
δ_d	0.004	(0.12)	0.005	(0.13)	0.009	(0.23)	0.009	(0.22)
δ_w	-0.026	(0.26)	-0.050	(0.50)	-0.052	(0.51)	-0.052	(0.50)
δ_m	0.557	(4.40)	0.554	(4.46)	0.564	(4.49)	0.564	(4.46)
θ_1	0.003	(0.36)	0.006	(0.63)	0.006	(0.64)	0.006	(0.64)
θ_2	0.020	(1.26)	0.016	(1.01)	-0.021	(0.52)	-0.021	(0.52)
θ_3	-0.008	(0.51)	-0.015	(0.99)	-0.017	(1.08)	-0.017	(1.04)
θ_4	0.022	(0.40)	0.001	(0.01)	0.005	(0.09)	0.005	(0.09)
$\phi_{1,0}$			-0.031	(2.84)	-0.064	(2.60)	-0.064	(2.57)
$\phi_{2,0}$			-0.005	(0.89)	-0.006	(1.04)	-0.006	(1.05)
$\phi_{1,1}$					0.027	(0.95)	0.027	(0.93)
$\phi_{2,1}$					0.021	(0.98)	0.021	(0.98)
adj. R^2	0.027		0.034		0.035		0.035	

Absolute t statistics in parentheses

predicting the next period's jump component. Former researches show contradictory results on return terms: Bollerslev et al. (2009) found a negative ψ_1 for S&P500 future index while Chen and Wang (2010) found a positive ψ_1 in Hu-Shen 300 index during 2006-2008. Therefore, we conclude that return effects on the next period's jump component are unstable.

$\phi_{1,0}$ and $\phi_{2,0}$ capture current trading volume and its leverage effect. The former is negatively significant at least 5% level showing that higher trading volume will reduce the current jump component volatility. The latter is insignificant indicating there is no leverage effect in volume on the jump component. $\phi_{1,1}$ and $\phi_{2,1}$ are insignificant suggesting that lagged trading volume is non-informative. Negative contemporal volume-volatility in the jump component is contradictory to traditional results. Such results are not unique in China. Giot et al. (2010) reported such negative relation in the largest 100 S&P 500 stocks. Although results depend on model setups, the worth case still shows 76% negatively significant relations and only 4% is positively significant.

5.5. Dynamics of trading volume around jumps

The negative correlation between jump component and trading volume seems contradictory to common sense. How can price move without substantially increase in trading volume? The answer, we believe, lies in the high frequency dynamics of trading volume around jumps. Since the jump test proposed by Barndorff-Nielsen and Shephard (2006) can not identify exactly which 5min return in a day contains jump (intra-day jumps), here we use truncation based jump test with Time-of-Day (TOD) volatility pattern as benchmark to identify them:

$$TOD_j = \frac{M \sum_{t=1}^{T} |r_{t,j}|^2 1\left(|r_{t,j}| \leq \tau\sqrt{BV_t \wedge RV_t} M^{-\varpi}\right)}{\sum_{t,j} |r_{t,j}|^2 1\left(|r_{t,j}| \leq \tau\sqrt{BV_t \wedge RV_t} M^{-\varpi}\right)}$$

Where $j = 1, \ldots, M$, $\tau > 0$ and $\varpi \in (0, 0.5)$, and truncation threshold is

$$\alpha_{t,j} = \tau\sqrt{(BV_t \wedge RV_t) \times TOD_j M^{-\varpi}}$$

A 5min return $r_{t,j}$ is "jump" when $|r_{t,j}| > \alpha_{t,j}$. In practice, we use $\tau = 2.5$ and $\varpi = 0.49$ suggested by Bollerslev et al. (2011).

After identified intra-day jumps, we calculate the mean 5min trading volume before and after jumps across trading days. When we enter the intra-day level, empirical evidence show that trading volume exhibits a U-shape patten during opening hours. Therefore, meaningful comparison relay on a reasonable estimator of local level of volume. We define a simple volume pattern estimator as:

$$V_{patten} = \frac{1}{T(j)} \sum_{t=1}^{T} \ln V_{t,j} \times 1(|r_{t,j}| \leq \alpha_{t,j})$$

Where $T(j) = \#\{\ln V_{t,j} \neq 0 | t = 1, \ldots, T\}$.

Also we define the trading volume $5 \times k$ minutes apart from jumps as:

$$\begin{aligned} \widetilde{\ln V}_{t,j}(k) \;=\; & \ln V_{t,j} \times 1(|r_{t,j+k}| > \alpha_{t,j+k}, \\ & \text{No jumps between } j \text{ and } j+k,\ j+k \in [1, 48]) \end{aligned}$$

This simply means that a 5min trading volume satisfies: 1) $5 \times k$ minutes from a jump in the corresponding day, 2) there is no other jumps in this $5 \times k$ minutes interval. The second criteria rules out possible compound effect when there is more than one jump in a single day.

We measure the dynamics using mean excess trading volume (MEV) defined as:

$$MEV(k) = \frac{1}{N(k)} \sum_{t=1}^{T} \sum_{j=1}^{M} \left(\frac{\widetilde{\ln V}_{t,j}(k)}{V_{patten}(j)} - 1 \right) \times 100\%$$

Where $N(k) = \#\{\widetilde{\ln V}_{t,j} \neq 0 | t = 1, \ldots, T; j = 1, \ldots, M\}$.

Since stock market opens 4 hours a day, we plot this measure in Figure 4 for $k \in [-24, 24]$. The patten is quite clear: 1) Before jump, MEV is close to

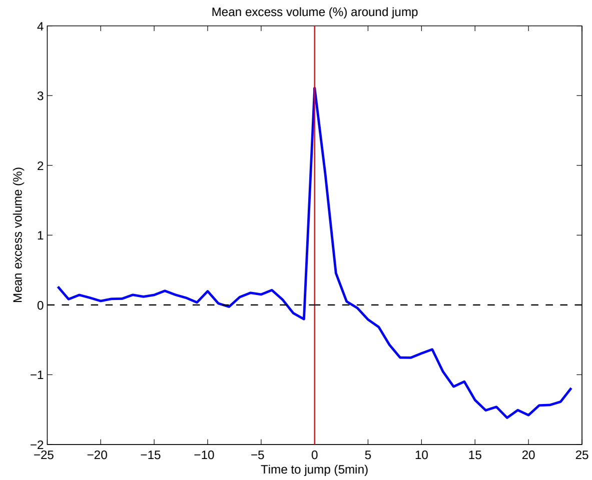

FIG. 4. Volume dynamics around jumps (T-120min to T+120min)

zero and there is no deterministic sign of illiquid or liquidity dry out. 2) At jump, MEV indicates a sharp increase in trading volume which is coincide with common sense that volume goes hand in hand with volatility. 3) After jumps, trading volume shrinks substantially for at least half trading day and such shrinkage is responsible for the negative correlation between jump volatility and trading volume.

Amatyakul (2010) proposed three possible mechanisms for this negative correlation: 1)Lack of liquidity: stocks are considered illiquid when trading volume is light. In that situation where there are extremely few trades, almost every price movement will be detected as jumps. 2)Trader's behavior: people might halt their trading activities to reevaluate their investing options after seeing large price movements, therefore a negative correlation. 3)Public information-driven trading: if there is a piece of public information in the market and traders hold a similar opinion of the effect of this information (therefore, the similar opinion of the price of stock), the price change will be done in short time without a lot of trade. Hence, we will observe a negative correlation between the two.

In the light of intra-day volume dynamics around jumps, the first mechanism is not supported. Also, Hu-Shen 300 index is formulated based on large and liquid stocks, it is hard to believe they are illiquid as a whole. The other two are not exclusive. As "public information" is not clearly defined, macroeconomic information release and important market information release (explanation 3), even price jump itself (explanation 2) will serve as "public information". Both of them suggests a volume shrinkage after jumps which coincides with the patten presented in Figure 4.

Evidence on the first channel relies on research on traders and cannot be fully investigated with aggregate index data. More appropriate data is high frequency data for individual stocks or even for traders' actions. Direct evidence on the second channel is also inadequate and we only have indirect evidence. Andersen et al. (2007) found that the jump component volatility rises around the release of macroeconomic information which implies that at least some jumps are related to public information release. Similar research in China's market can be found in Wang et al. (2011) which pointed out a high correlation between some large jumps and important market information exposure on Shanghai stock exchange index during 2007-2008.

6. ROBUSTNESS CHECK OF RESULTS

6.1. Subsample estimation

In this section, we divide full sample into two subsamples: sample A (Jan/2007 - Dec/2008) and Sample B (Jan/2009 - Dec/2010).

TABLE 6.

Subsample estimations

	(1a)		(1b)			(2a)		(2b)	
α_0	-2.510	(4.94)	-2.961	(7.00)	δ_0	0.143	(3.30)	0.260	(4.01)
α_d	0.476	(10.04)	0.441	(7.62)	δ_d	0.078	(1.42)	-0.070	(1.41)
α_w	0.294	(4.73)	0.289	(4.22)	δ_w	-0.142	(0.97)	-0.014	(0.10)
α_m	-0.005	(0.06)	0.103	(1.85)	δ_m	0.550	(3.26)	0.233	(1.20)
θ_1	0.042	(1.31)	0.049	(1.38)	θ_1	0.023	(1.66)	-0.005	(0.43)
θ_2	-0.057	(0.36)	0.474	(2.31)	θ_2	-0.017	(0.30)	0.002	(0.02)
θ_3	0.338	(7.40)	0.104	(1.73)	θ_3	-0.043	(1.91)	0.009	(0.35)
θ_4	-0.585	(2.65)	-0.346	(1.18)	θ_4	-0.022	(0.31)	0.025	(0.29)
$\phi_{1,0}$	1.063	(11.76)	0.104	(9.34)	$\phi_{1,0}$	-0.075	(2.10)	-0.080	(2.17)
$\phi_{2,0}$	0.260	(9.17)	0.048	(9.76)	$\phi_{2,0}$	-0.006	(0.60)	-0.008	(1.21)
$\phi_{1,1}$	-0.829	(8.82)	-0.059	(3.61)	$\phi_{1,1}$	0.011	(0.25)	0.013	(0.34)
$\phi_{2,1}$	0.051	(0.51)	-0.007	(2.06)	$\phi_{2,1}$	0.033	(0.89)	-0.005	(0.16)
$D_{M,T}$	Y		Y			N		N	
LogL	-269.9		-245.5		adj.R^2	0.046		0.027	

Absolute t statistics in parentheses

Table 6 presents the results on the continuous as well as the jump component volatility in which (1) represents the continuous component model, (2) represents the jump component model, (a) represents the sub-sample A and (b) represents the sub-sample B. To save space, we only report mean equations. The following results are sub-sample robust.

For the continuous component volatility: 1)$\alpha_d > \alpha_w > |\alpha_m|$ still holds. 2)contemporal volume-volatility relation is still significantly positive. 3)intertemporal volume-volatility relation is still significantly negative. For the jump component volatility: 4)contemporal volume-volatility relation is still significantly negative.

The following results are sub-sample unstable. For the continuous component volatility: 1)Lagged monthly volatility is insignificant in the first sub-sample and even the sign changes between two sub-samples indicating volatility is more persistent in the second sub-sample. 2)The leverage effect in the second sub-sample is insignificant while the skewness effect is significant only in the first sub-sample. No significant serial correlation is found in the second sub-sample.

Figure 7 shows results on the nonlinearity between volatility and trading volume in sub-samples. Again, we only report mean equation. (1), (2), (a), (b) are the same as above.

TABLE 7.

Nonlinearity dynamic volume-volatility relation: sub-sample estimation

	(1a)		(1b)		(2a)		(2b)	
α_0	-1.334	(2.41)	-0.900	(1.51)	-1.699	(2.89)	-0.914	(1.48)
α_d	0.524	(9.76)	0.413	(6.71)	0.521	(10.60)	0.387	(6.12)
α_w	0.245	(3.41)	0.339	(4.33)	0.240	(3.71)	0.347	(4.26)
α_m	0.068	(0.85)	0.081	(1.33)	0.028	(0.35)	0.096	(1.50)
θ_1	0.079	(1.80)	0.034	(0.86)	0.063	(1.64)	0.025	(0.59)
θ_2	-0.095	(1.47)	0.037	(0.57)	-0.100	(1.56)	0.039	(0.55)
θ_3	0.261	(4.47)	0.064	(0.84)	0.283	(5.54)	0.077	(0.94)
θ_4	-0.749	(3.36)	-0.560	(2.14)	-0.622	(2.62)	-0.463	(1.65)
$\phi_{1,0}$	0.840	(8.53)	0.687	(6.67)	0.941	(10.11)	0.793	(7.03)
$\phi_{1,1}$	-1.183	(3.58)	-1.697	(4.15)	-1.346	(4.03)	-1.855	(4.09)
$\phi_{2,1}$	-0.177	(1.67)	-0.348	(3.46)	0.206	(1.92)	0.371	(3.29)
LogL	-302.8		-264.8		-305.7		-291.6	

Absolute t statistics in parentheses

In both samples, the quadratic form of lagged trading volume is significantly positive at 10% level. This indicates the existence of minimum volatility trading volume in both sub-sample. The results are more significant in the second sub-sample. Also, the results are more significant in the continuous component volatility than in total volatility since there is no significant relation between the jump component and trading volume.

6.2. Truncated jump component

Former discussion is based on the non-truncated jump component $\ln\left(\frac{RV_t}{BV_t}\right)$ which treats negative value as measurement error. In this section we use jump test proposed by Barndorff-Nielsen and Shephard (2006) (Z statistic) to identify statistically significant jumps (critical value set as $\alpha = 0.05$) and re-defined the jump component volatility as:

$$J_t = \{\ln(RV_t) - \ln(BV_t)\} I\{Z_t > c_\alpha\}$$

The benefit of this is to ensure the nonnegativity of the jump component volatility, but the results may depend on the selection of the critical value of the test. Since the newly defined the jump component is truncated, we use Tobit model following Giot et al. (2010)[14].

TABLE 8.

Tobit model results

	(1)		(2)		(3)	
δ_0	-0.113	(1.23)	-0.176	(1.31)	0.263	(1.72)
δ_d	0.031	(0.30)	0.194	(1.14)	-0.096	(0.74)
δ_w	-0.416	(1.55)	-0.542	(1.24)	-0.400	(1.18)
δ_m	1.286	(3.05)	1.063	(1.47)	0.683	(1.27)
ψ_1	0.014	(0.57)	0.053	(1.36)	-0.002	(0.06)
ψ_2	0.026	(0.22)	0.026	(0.15)	0.007	(0.04)
ψ_3	-0.029	(0.70)	-0.061	(1.02)	-0.018	(0.29)
ψ_4	-0.008	(0.05)	0.044	(0.20)	-0.109	(0.49)
$\phi_{1,0}$	-0.133	(1.97)	-0.152	(1.47)	-0.198	(2.10)
$\phi_{2,0}$	-0.011	(0.67)	0.001	(0.05)	-0.021	(1.13)
$\phi_{1,1}$	0.075	(0.99)	0.073	(0.62)	0.032	(0.31)
$\phi_{2,1}$	0.036	(0.59)	0.026	(0.24)	-0.000	(0.00)
σ	0.355	(29.89)	0.383	(19.21)	0.326	(23.27)
Log likelihood.	-457.5		-222.5		-225.3	

Absolute t statistics in parentheses

The regression equation of latent variable is the same as M_J with truncated J instead of non-truncated J. Table 8 includes three models depending on different samples: model (1) is for full sample, model (2) is for sub-sample A and model (3) is for sub-sample B. In the full sample, although it is not as significant as the linear model, Tobit model shows a significantly positive lagged monthly jump component volatility. Contemporal

[14] It is worth pointing out Tobit model is sensitive to HSK and non-normality of the latent variable. Therefore, we suggest that the results gained by linear regression model is at least equally important.

negative volume-volatility relation on the jump component volatility is also pronounced at 5% level in full and the second sub-sample. It is significant at 15% level in the first sub-sample. No significant trading volume effect is found on the jump component volatility.

7. CONCLUSION

We use the heterogenous autocorrelation (HAR) model to investigate the volume-volatility relationship in China's stock market via the volatility decomposition technique. Results suggest that the total volatility and the continuous component volatility are positively correlated with current trading volume, which coincides with the findings in the long lasting literatures. However, the jump component volatility reveals a significant and robust negative correlation with current trading volume. Considering the data we use, this phenomena implies that the jump component contains some kind of "public information" such as the macroeconomic news, market information release or (given the common knowledge "jump contains information") price jump itself. Evidence from intra-day trading volume suggests that jumps is unlikely provoked by liquidity dry out. The dynamics patten of mean excess intra-day trading volume around jumps coincides with "public information" implications.

Results about the intertemporal relationship shows that the total volatility is negatively correlated with lagged trading volume even after controlling the volatility dynamics and contemporal volume. Such findings support the information driven trading (IDT) hypothesis discussed in Wang (2004). More detailed discussion shows that negative intertemporal correlation can only be found for the continuous component volatility. It indicates "private information" is the information that drives Hu-Shen 300 trading. We also find that the lagged price skewness is only negatively correlated with the total and continuous volatility. This shows that market sentiment has influential power on volatility only through the continuous volatility. Jumps are not affected by market sentiment.

A potential research topic is to discuss the volume-volatility relationship for individual stocks as well as portfolios. By doing this, we can discuss how characteristics of stocks affect the current findings. Relationship between trading volume and other trading intensity measures is also an interesting topic that is worth investigating.

REFERENCES

Amatyakul, P., 2010. The Relationship Between Trading Volume and Jump. Processes in Financial Markets. *Duke Journal of Economics* **XXII**.

Amaya, D. and A. Vasquez, 2010. Skewness from High-Frequency Data Predicts the Cross-Section of Stock Returns. *Working Paper*.

Andersen, T. G., 1996. Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility. *Journal of Finance* **51**, 169-204.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega, 2007. Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics* **73**, 251-277.

Bandi, F. M., J. R. Russell, and C. Yang, 2008. Realized volatility forecasting and option pricing. *Journal of Econometrics* **147**, 34-46.

Barclay, M. J. and J. B. Warner, 1993. Stealth trading and volatility : Which trades move prices? *Journal of Financial Economics* **34**, 281–305.

Barndorff-Nielsen, O. E., 2004. Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics* **2**, 1-37.

Barndorff-Nielsen, O. E. and N. Shephard, 2005. How accurate is the asymptotic approximation to the distribution of realised volatility? *Andrews, D., Powell, J., Ruud, P., Stock, J. (Eds.), Identification and Inference for Econometric Models. A Festschrift for Tom Rothenberg, Cambridge University Press, Cambridge,*, 306-331.

———, 2006. Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation. *Journal of Financial Econometrics* **4(1)**, 1-30.

Black, F. and M. S. Scholes, 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* **81**, 637-54.

Bollerslev, T., U. Kretschmer, C. Pigorsch, and G. Tauchen, 2009. A discrete-time model for daily S&P500 returns and realized variations: Jumps and leverage effects. *Journal of Econometrics* **150**, 151-166.

Bollerslev, T., V. Todorov, and S. Z. Li, 2011. Jump Tails, Extreme Dependencies, and the Distribution of Stock Returns. *forthcoming in Journal of Econometics*.

Campbell, J. Y., S. J. Grossman, and J. Wang, 1993. Trading Volume and Serial Correlation in Stock Returns. *The Quarterly Journal of Economics* **108**, 905-39.

Chan, K. and W.-M. Fong, 2000. Trade size, order imbalance, and the volatility-volume relation. *Journal of Financial Economics* **57**, 247-273.

Chen, G., M. Firth, and O. M. Rui, 2001. The Dynamic Relation between Stock Returns, Trading Volume, and Volatility. *The Financial Review* **36**, 153-73.

Chen, G. and Z. Wang, 2010. Continuous volatility and jump volatility in China's stock market. *Systems Engineering - Theory & Practice* **30(9)**, 1554-1562.

Chow, G. C. and C. C. Lawler, 2003. A Time Series Analysis of the Shanghai and New York Stock Price Indices. *Annals of Economics and Finance* **4**, 17-35.

Clark, P. K., 1973. A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica* **41**, 135-55.

Corsi, F., 2009. A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics* **7**, 174-196.

Fabozzi, F. J., R. Tunaru, and T. Wu, 2004. Modeling Volatility for the Chinese Equity Markets. *Annals of Economics and Finance* **5**, 79-92.

Gallant, A. R., P. E. Rossi, and G. Tauchen, 1992. Stock Prices and Volume. *Review of Financial Studies* **5**, 199-242.

Gao, L. and G. Kling, 2005. Calendar Effects in Chinese Stock Market. *Annals of Economics and Finance* **6**, 75-88.

Giot, P., S. Laurent, and M. Petitjean, 2010. Trading activity, realized volatility and jumps. *Journal of Empirical Finance* **17(1)**, 168-175.

Hansen, P. R. and A. Lunde, 2006. Realized Variance and Market Microstructure Noise. *Journal of Business & Economic Statistics* **24**, 127-161.

Huang, S. J., Q. Liu, and J. Yu, 2007. Realized Daily Variance of S&P 500 Cash Index: A Revaluation of Stylized Facts. *Annals of Economics and Finance* **8**, 33-56.

Huang, X. and G. Tauchen, 2005. The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics* **3**, 456-499.

Jones, C., G. Kaul, and M. Lipson, 1994. Transactions, volume, and volatility. *Review of Financial Studies* **7(4)**, 631-651.

Karpoff, J. M., 1987. The Relation between Price Changes and Trading Volume: A Survey. *Journal of Financial and Quantitative Analysis* **22**, 109-126.

Kyle, A. S., 1985. Continuous Auctions and Insider Trading. *Econometrica* **53**, 1315-35.

Rui, M., Y. Song, and Q. Wang, 2003. The Research on Whether Trade Volume of Shares in Chinese Stock Market Contains the Information to Project Shares' Earnings. *Statistical Research* **issue 3**, 54-59.

Tauchen, G., H. Zhang, and M. Liu, 1996. Volume, volatility, and leverage: A dynamic analysis. *Journal of Econometrics* **74**, 177-208.

Tauchen, G. E. and M. Pitts, 1983. The Price Variability-Volume Relationship on Speculative Markets. *Econometrica* **51**, 485-505.

Varian, H. R., 1985. Divergence of Opinion in Complete Markets: A Note. *Journal of Finance* **40**, 309-17.

Wang, C., P. Hao, and Z. Fang, 2011. Investigation of the Efficiency of Information into the Security Market Based on the Jumps' Features. *Journal of Beijing Institute of Technology(Social Sciences Edition)* **13(1)**, 1-5.

Wang, H., 2004. Dynamic Volume-Volatility Relation. *Working Paper*.

Wang, J., 1994. A Model of Competitive Stock Trading Volume. *Journal of Political Economy* **102**, 127-68.

———, 2002. Trading Volume and Asset Prices. *Annals of Economics and Finance* **3**, 299-359.

Yin, W., 2010. An empirical research on China's stock market's volume-volatility relationship. *World Economic Ourlook* **3**, 66-79.

Zhang, B., M. Zhou, and L. Li, 2005. Asymmetric Effect in Chinese Stock Market Based on Rolling Sample Tests. *Economic Management* **4**, 85-90.

Zhao, L. and Y. Wang, 2003. Hushen stock markets' volume, return and volatility correlations: evidence from empirical data analysis. *Economic Science* **2**, 57-67.

Zhou, S. and Q. Chen, 2004. Testing for the weekday effect of return volatility in the Shanghai stock market. *Application of Statistics and Management* **23(3)**, 57-60.