Journal of Empirical Finance 17 (2010) 168–175

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

Trading activity, realized volatility and jumps

Pierre Giot [a,b,c], Sébastien Laurent [b,c], Mikael Petitjean [a,d,*]

[a] Louvain School of Management, Belgium
[b] CeReFiM, FUNDP, University of Namur, Belgium
[c] CORE, Université Catholique de Louvain, Belgium
[d] FUCaM, Catholic University of Mons, Belgium

ARTICLE INFO

Article history:
Received 24 April 2008
Received in revised form 1 July 2009
Accepted 6 July 2009
Available online 16 July 2009

JEL classification:
G10
G12
G13

Keywords:
Volume
Volatility
Transactions
Jumps
Bi-power variation

ABSTRACT

This paper takes a new look at the relation between volume and realized volatility. In contrast to prior studies, we decompose realized volatility into two major components: a continuously varying component and a discontinuous jump component. Our results confirm that the number of trades is the dominant factor shaping the volume–volatility relation, whatever the volatility component considered. However, we also show that the decomposition of realized volatility bears on the volume–volatility relation. Trade variables are positively related to the continuous component only. The well-documented positive volume–volatility relation does not hold for jumps.

1. Introduction

Volume and volatility convey extremely important implications for market participants. Not surprisingly, the relation between the two has been extensively studied in the past. In the early empirical literature, which is mostly based on monthly and weekly stock returns, volatility and trading volume are measured respectively by absolute returns and number of shares traded per equally time-spaced intervals. A positive contemporaneous relation between the two is generally documented, although it does not always appear to be sizeable (e.g. Karpoff, 1987).

In more recent theoretical studies, volume is decomposed into trade frequency (i.e. the number of trades) and trade size (i.e. the average number of shares per trade). A rough taxonomy of the theoretical models on the volume–volatility relation consists of three classes of models: competitive, strategic, and mixture of distribution models.

Competitive microstructure models are extensions of the Glosten and Milgrom (1985) sequential trading model in which market makers and uninformed investors experience adverse selection when trading with informed investors. When investors differ in their beliefs regarding the importance of information and act competitively, larger-sized trades tend to be executed by better-informed investors. Thus, competitive models suggest a positive relation between trade size and price volatility (Easley et al., 1997).

* Corresponding author.
E-mail address: mikael.petitjean@fucam.ac.be (M. Petitjean).

0927-5398/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2009.07.001

Strategic microstructure models also incorporate asymmetric information across agents but assume that informed investors engage in stealth trading by breaking up large trades into many smaller transactions. Therefore, the effect of the trade size on price volatility is attenuated and its impact may be transferred to the number of trades. Thus, strategic models predict a positive relation between volatility and number of trades and/or order imbalance (e.g. Kyle, 1985).

The third class of models relies on the mixture of distributions (MD) hypothesis (e.g. Harris, 1987). Although the MD models have been criticized on the grounds that they are primarily statistical models, recent empirical works have provided strong evidence coherent with the predictions of the MD models. For example, these models predict that average trade sizes should have no effect on price volatility. In contrast, it is the number of trades that should reflect the number of daily information arrivals.

Our own research builds on four empirical papers. First, Jones et al. (1994) decompose trading volume into its two components and find that stock price volatility is driven by the number of trades per equally time-spaced intervals. The average trade size offers no additional explanatory power. Second, Chan and Fong (2000) conclude that it is the order imbalance, rather than the number of trades, that drives the volume–volatility relation. Third, Huang and Masulis (2003) distinguish between small and large trades. For large trades, they confirm the Jones et al. (1994) findings, that only trade frequency affects price volatility. For small trades, the picture is roughly similar with the exception of small trades close to the maximum-guaranteed quoted depth, for which the trade frequency *and* the average trade size impact price volatility. Fourth, in contrast to the three above-mentioned papers, Chan and Fong (2006) use realized volatility in place of absolute returns (which give a very noisy estimator of the true latent volatility) as the volatility measure. Chan and Fong (2006) find that neither the trade size nor the order imbalance adds significantly more explanatory power to realized volatility beyond the number of trades.

This paper innovates in several ways. First, our sample covers a five-year period and includes the 100 largest stocks quoted on the New York Stock Exchange as of January 1, 1995. This is the most extensive data set used to study the relationship between the volume and the volatility. Second, we use midquotes to measure prices, in agreement with the recent literature on realized variance. Third, and most importantly, our study is the first to decompose realized volatility into its continuous and jump components based on the Barndorff-Nielsen Shephard (2004, 2006) bi-power variation method (see Section 2).

The fact that realized volatility can be decomposed prompts the following question: would the results of prior studies still hold if one decomposes realized volatility into its continuous, persistent part and its discontinuous, temporary, jump component? The motivation behind the decomposition of realized volatility into its two components relies on the following observation: the positive volatility–volume relation that is documented in the literature focuses exclusively on the *level* of volatility (i.e. low versus high volatility). The *nature* of volatility (i.e. continuous versus discontinuous volatility) is completely ignored, and one way to characterize the nature of volatility is precisely to introduce the concept of jumps. This is especially important since market participants usually care as much about the nature of volatility as about its level. For example, all traders make the distinction between 'good' and 'bad' volatilities. 'Good' volatility is directional, persistent, relatively easy to anticipate and accompanied by sufficiently high volume. 'Bad' volatility is *jumpy*, relatively difficult to foresee and associated with low volume. As such, 'good' and 'bad' volatilities can be respectively associated with the continuous, persistent part and the discontinuous, jump component of realized volatility.

In this work we show that the decomposition of realized volatility bears on the volume–volatility relation. Trade variables are positively related to the continuous, persistent component only. The positive relationship between volume and volatility does not hold for jumps, supporting the view of market participants according to which poor trading volume leads to more erratic price changes. This negative volume–jumps relation is revealed through the number of trades, which remains the dominant factor behind the volume–volatility relation. Beyond the number of trades, neither trade size nor order imbalance increases explanatory power significantly, whatever the volatility component considered.

The remainder of the paper is organized as follows. In Section 2, we explain the procedure used to identify the jump component of realized volatility. In Section 3, we describe the data and provide summary statistics. In Section 4, we report and interpret the empirical results. We conclude in Section 5.

2. The decomposition of realized volatility

To decompose realized volatility into its continuous and jump components, we consider the following continuous-time jump diffusion process $dp(t)=\mu(t)dt+\sigma(t)dW(t)+\kappa(t)dq(t)$, where $p(t)$ is a logarithmic asset price at time t, $\mu(t)$ is a continuous and locally bounded variation process, $\sigma(t)$ is a strictly positive stochastic volatility process with a sample path that is right continuous and has well defined limits, $W(t)$ is a standard Brownian motion, $q(t)$ is a pure jump process with intensity $\lambda(t)$ and $\kappa(t)$ is the jump size.

The so-called realized volatility, popularised by Andersen and Bollerslev (1998), is defined as the sum of intradaily squared returns, i.e., $\mathrm{RV}_t(\Delta)\equiv\sum_{j=1}^{1/\Delta} r_{t+j\Delta,\Delta}^2$, where $r_{t,\Delta}\equiv p(t)-p(t-\Delta)$ is the discretely sampled Δ-period return, where $1/\Delta$ is the number of intradaily periods.[1] It is now well established that $\mathrm{RV}_t(\Delta)\rightarrow\int_{t-1}^{t}\sigma^2(s)ds+\sum_{j=1}^{N_t}\kappa_{t,j}^2$, where N_t is the number of jumps on day t and $\kappa_{t,j}$ is the j-th jump size on that day.

In order to disentangle the continuous and the jump components of realized volatility, we need a robust to a jumps estimate of the integrated volatility, i.e., $\int_{t-1}^{t}\sigma^2(s)ds$. We use the realized bi-power variation of Barndorff-Nielsen and Shephard (2004)

[1] We use the same notation as in Andersen et al. (2006) and normalize the daily time interval to unity. We drop the Δ subscript for daily returns: $r_{t+1,1}\equiv r_{t+1}$.

which is defined as the sum of the product of adjacent absolute intradaily returns standardized by a constant, i.e., $\mathrm{BV}_t(\Delta)\equiv\mu_1^{-2}\sum_{j=2}^{1/\Delta}|r_{t+j\Delta,\Delta}||r_{t+(j-1)\Delta,\Delta}|$, where $\mu_1\equiv\sqrt{2/\pi}\simeq0.79788$. It can indeed be shown that even in the presence of jumps, $BV_t(\Delta)\rightarrow\int_{t-1}^{t}\sigma^2(s)ds$.

Thus, the difference between the realized volatility and the bi-power variation consistently estimates the jump contribution to the quadratic variation process. When $\Delta\rightarrow0$, $\mathrm{RV}_t(\Delta)-\mathrm{BV}_t(\Delta)\rightarrow\sum_{j=1}^{N_t}\kappa_{t,j}^2$. One might wish to select only statistically significant jumps, i.e., to consider very small jumps as being part of the continuous sample path rather than genuine discontinuities. For that we follow Huang and Tauchen (2005) and Andersen et al. (2006) and compute the Z statistic for jumps $Z_t(\Delta)\equiv\Delta^{-1/2}\frac{[RV_t(\Delta))-BV_t(\Delta)]RV_t(\Delta)^{-1}}{\left[(\mu_1^{-4}+2\mu_1^{-2}-5)\max\left\{1,TQ_t(\Delta)BV_t(\Delta)^{-2}\right\}\right]^{1/2}}$, where $\mathrm{TQ}_t(\Delta)$ is the tri-power quarticity,[2] a robust to jumps estimator of the integrated quarticity.

Under the null of no jumps, $Z_t(\Delta)$ is approximately the standard normal distributed and this test has been shown to have reasonable power against several plausible stochastic volatility jump diffusion models. Practically, we choose a significance level α and compute $J_{t,\alpha}(\Delta)=I_{t,\alpha}(\Delta)[\mathrm{RV}_t(\Delta)-\mathrm{BV}_t(\Delta)]$, where $I_{t,\alpha}(\Delta)\equiv I[Z_t(\Delta)>\Phi_\alpha]$. Of course, a smaller α means that we estimated fewer and larger jumps. Finally, as suggested by Andersen et al. (2006), we use 'staggered' versions of the bi-power variation and the tri-power quarticity measures to tackle microstructure noise that causes the high-frequency returns to be autocorrelated. To make sure that the jump component added to the continuous one equals the realized volatility, we define as an estimator of the integrated variance $C_{t,\alpha}(\Delta)=[1-I_{t,\alpha}(\Delta)]\mathrm{RV}_t(\Delta)+I_{t,\alpha}(\Delta)\mathrm{BV}_t(\Delta)$.

In this paper, we construct returns and realized volatility estimators based on 5-minute time intervals and use mid-quote prices, in agreement with the extant literature on high-frequency data. Finally, we use a conservative significance level ($\alpha=0.01\%$) to identify the number of economically meaningful jumps. For the sake of simplicity, we write: $\mathrm{RV}_t(\Delta)=\mathrm{RV}_t$, $C_{t,\alpha}(\Delta)=C_t$, and $J_{t,\alpha}(\Delta)=J_t$.

3. Data

The sample period covers a five-year period starting on January 1, 1995 and ending on September 30, 1999, which represents a total of 1199 trading days. The sample consists of the 100 largest stocks traded on the New York Stock Exchange as of January 1, 1995.

Data for this study is retrieved from the Trades and Quotes (TAQ) database. Following the literature, we only retain class A stocks and remove preferred stocks or shares, warrants, rights, derivatives, trusts, closed-end investment companies, American depositary receipts, units, shares of beneficial interest, holdings and realty trusts. We restrict our selection to stocks whose price is higher than $5 and lower than $999. The remaining stocks are then selected on the basis of market capitalization. We also apply the traditional filtering procedures (Chordia et al., 2001).

Following the early literature, we decompose the daily trading volume into number of trades and average trade size. As in Chan and Fong (2006), we also include order imbalance computed as the absolute value of the number of buyer-initiated trades minus the number of seller initiated trades for the day.[3]

Table 1 gives the cross-sectional means of the time series statistics for the 7 variables used in our study. Using the square root of time rule, the daily realized mean variance (RV) of 0.0242% translates into an annualized realized volatility of 24.6%. This is close to the annualized realized volatility of 27.5% reported by Chan and Fong (2006) for the Dow 30 stocks. The proportion of days with significant jumps at 0.01% (given by J in Table 1) is 26%, i.e. 313 day s over 1199 on average. RV and the daily realized continuous variance (C) display similar characteristics. All variables, except jumps, have similar autocorrelation patterns. Among the trade variables, average trade size (ATS) displays the lowest AR and CV coefficients. Daily number of trades (NT) exhibits the slowest decaying autocorrelation function while absolute order imbalance |OB| displays the highest CV.

Table 2 reports the correlations between volatility and transactions. We find very similar results to those reported by Chan and Fong (2006). More precisely, we confirm that: trading volume is highly correlated with the number of trades; the number of trades co-varies quite strongly with the order imbalance; realized volatility correlates more highly with the number of trades than with the average trade size or the order imbalance. Interestingly, the jump component (J) does not behave like the undecomposed measure of realized volatility (RV). First, the two measures are negatively correlated with one another. Second, volume (V) and the number of trades (NT) are negatively correlated with J, while they are positively correlated with RV. This suggests that the decomposition of realized volatility may bear on the volume–volatility relation. As pointed out in the previous studies, the overall level of realized volatility increases (decreases) when the trading volume (V) and, in particular, the number of trades (NT) increase (decrease). However, the decomposition of realized volatility seems to indicate that such a positive volume–volatility relation does not hold when the jump component of realized volatility is considered. It would thus only hold for the continuous part.

[2] $TQ_t(\Delta)\equiv\Delta^{-1}\mu_{4/3}^{-3}\sum_{j=3}^{1/\Delta}|r_{t+j\Delta,\Delta}|^{4/3}|r_{t+(j-1)\Delta,\Delta}|^{4/3}|r_{t+(j-2)\Delta,\Delta}|^{4/3}$, where $\mu_{4/3}\equiv2^{2/3}\Gamma(7/6)\Gamma(1/2)^{-1}$.

[3] To define the direction of a trade, we rely upon the widely-used Lee and Ready algorithm. We follow the recommendation contained in *SEC Rule 11Ac1-5*, assuming trades were recorded 5 s later than their actual execution time.

Table 1
Summary statistics.

	RV	J	C	V	ATS	NT	|OB|
Mean	2.42	1.04	2.38	1.40	1595	929	121
Standard deviation	2.56	1.16	2.48	1.00	643	578	153
Skewness	5.13	5.51	4.51	3.40	1.32	2.24	3.67
Kurtosis	56.80	54.69	43.75	27.63	49.67	13.34	31.64
Coefficient of variation	105.79	111.97	104.20	71.90	39.58	56.32	110.43
AR(1)	0.49	0.06	0.51	0.61	0.39	0.79	0.45
AR(12)	0.23	0.04	0.25	0.31	0.19	0.53	0.21

Realized variance (RV) and its two components, i.e. continuous variance (C) and significant jumps (J) at $\alpha = 0.01\%$, are multiplied by 10,000. V is trading volume (i.e. millions of shares traded per day), NT denotes daily number of trades, ATS is average trade size (numbers of shares traded each day divided by number of trades for the day), and |OB| is absolute order imbalance defined as the absolute value of the number of buyer-initiated trades minus number of seller initiated trades for the day.

Table 2
Correlation matrix.

	RV	J	C	V	ATS	NT	|OB|
RV	1.00						
J	−0.22	1.00					
C	0.92	−0.53	1.00				
V	0.54	−0.21	0.54	1.00			
ATS	0.05	0.02	0.03	0.49	1.00		
NT	0.62	−0.27	0.63	0.80	−0.04	1.00	
|OB|	0.22	0.07	0.21	0.34	−0.05	0.42	1.00

RV = realized variance; J = significant jumps at $\alpha = 0.01\%$; C = continuous variance with $\alpha = 0.01\%$; V = trading volume; NT = daily number of trades; ATS = average trade size; |OB| = absolute order imbalance.

4. Empirical analysis

In the following analysis, we estimate the volume–volatility relation for each of our 100 stocks. Following Chan and Fong (2006) (CF), we first estimate this relation by OLS, use Newey–West standard errors, and include a Monday dummy as well as 12 lags to account for some dynamics in the conditional expected return. Since Huang and Masulis (2003) (HM) advocate the use of the generalized method of moments (GMM), we also estimate the relationship by GMM. In addition, both the median and the robust regression techniques are employed to deal with non-normality and offer a more robust analysis. We follow the CF and HM approaches to regress realized volatility (RV) and its continuous component (C) on the number of shares traded (V), the number of trades (NT), the average number of shares per trade (ATS), and the absolute order imbalance (|OB|). To estimate the relationship between the jumps and the trading activity, we run TOBIT regressions since the population distribution of the jumps is spread over a large range of positive values, with a pileup at the value zero. Following Calzolari and Fiorentini (1998) and Lahaye et al. (2008), we also correct for heteroscedasticity by allowing for a GARCH(1,1) effect in the TOBIT regression. Tables 3 to 6 report the impact of the trade measures on realized volatility and its two components.[4]

Table 3 presents the results for the number of shares traded (i.e. trading volume). Results for Regression (1) with realized volatility as the dependent variable are qualitatively similar to those reported by Chan and Fong (2006) with a slightly higher R^2. The percentage of stocks for which volume is statistically significant rises from 95% to 100%. The mean ($\hat{\phi}$) is equal to 0.11, implying that an increase in the number of shares traded of 100,000 is accompanied by an increase in realized volatility of around 5%. The p-value is particularly low, equal to 0.05%. Moreover, the results are qualitatively similar when the GMM, robust, and median regressions are run (see Regressions (3)). Results are very similar when the continuous component of realized volatility is used as the dependent variable (in Regressions (2) and (4)). For the jumps, however, the picture is completely different. Regressions (5), i.e. the TOBIT and TOBIT–GARCH models, indicate that the jumps and the trading volume are significantly and negatively related for 76 and 78 stocks out of 100 respectively. In the TOBIT–GARCH regression, only 2 stocks display positive and significant coefficients for the trading volume at the 5% level.

Table 4 reports the results for the number of trades and shows that the number of trades explains a substantial portion of daily realized volatility. Overall, the use of the number of trades instead of the number of shares traded as explicative variable leads to a rise in the R^2. Again, the number of trades affects the jumps differently. Regressions (5) show that the jumps and the number of trades are significantly and negatively related for more than 80 stocks out of 100. Less than 10 stocks display positive and

[4] To test whether our results were induced by the upward trend in the trading activity, we added a time trend in each of the regressions. Since results were almost identical, we do not report them to save space. They are available upon request. We also applied Phillips–Perron unit root tests to test for stochastic trends in the de-trended series. There was no evidence of stochastic trends.

Table 3
Number of shares traded, realized volatility and its two components.

	RV				C				J	
	(1)	(3)			(2)	(4)			(5)	
	OLS	GMM	ROB	MED	OLS	GMM	ROB	MED	TOB	TOB–G
$\bar{\phi}$	0.110	0.176	0.114	0.139	0.106	0.179	0.118	0.141	−0.041	−0.053
$\overline{\sigma_{\phi}}$	0.022	0.033	0.005	0.011	0.021	0.032	0.005	0.012	0.013	0.014
% p-Value	0.050	0.220	0.000	0.000	0.052	0.132	0.000	0.014	8.894	7.268
% + Significant	100	99	100	100	100	100	100	100	4	2
% − Significant	0	0	0	0	0	0	0	0	76	78
% Adjusted R^2	40.22	22.48	33.85	–	41.46	23.04	33.65	–	0.52	1.87

$$\mathrm{RV}_{it} = \alpha_i + \alpha_{im}M_t + \sum_{j=1}^{12}\rho_{ij}\mathrm{RV}_{it-j} + \phi_i V_{it} + v_{it} \tag{1}$$

$$C_{it} = \alpha_i + \alpha_{im}M_t + \sum_{j=1}^{12}\rho_{ij}C_{it-j} + \phi_i V_{it} + v_{it} \tag{2}$$

$$\mathrm{RV}_{it} = \alpha_i + \alpha_{im}M_t + \phi_i V_{it} + v_{it} \tag{3}$$

$$C_{it} = \alpha_i + \alpha_{im}M_t + \phi_i V_{it} + v_{it} \tag{4}$$

$$J^*_{it} = \alpha_i + \alpha_{im}M_t + \phi_i V_{it} + v_{it}, \text{ where } J = \max(0, J^*) \tag{5}$$

The above regressions are run for the 100 largest stocks traded on the NYSE over the period January 1, 1995–September 30, 1999. $\mathrm{RV}_{it}/C_{it}/J_{it}$ is the realized variance/the continuous component of realized variance/the jump component of realized variance of stock i on day t, all multiplied by 10,000. M_t is a Monday dummy, V_{it} is the number of shares traded (divided by 100,000) for stock i on day t and ρ_{ij} measures the persistence of volatility shocks at lag j. Eqs. (1) and (2) are estimated by OLS while Eqs. (3) and (4) are estimated by the generalized method of moments (GMM), robust regression (ROB) using iteratively reweighted least squares, and median regression (MED). Eq. (5) is estimated by maximum likelihood using the TOBIT model with and without GARCH effect (TOB–G and TOB, respectively). For brevity, we only report the equally-weighted cross-sectional mean coefficient for the number of shares traded, with corresponding statistics. Newey–West standard errors and two-sided p-values across stocks are computed. We report the percentage of positive and negative $\hat{\phi}_i$ coefficients which are statistically different from zero at the 5% level. The last row reports the mean adjusted R^2s.

significant coefficients for the number of trades at the 5% level. The average p-values are smaller than 5% while they were between 5% and 10% in the specification with trading volume. Trade frequency thus appears to be more informative than the trading volume in explaining both the jumps and the continuous component of realized volatility.

Table 4
Number of trades, realized volatility and its two components.

	RV				C				J	
	(1)	(3)			(2)	(4)			(5)	
	OLS	GMM	ROB	MED	OLS	GMM	ROB	MED	TOB	TOB–G
$\bar{\beta}$	0.318	0.387	0.298	0.350	0.315	0.397	0.315	0.363	−0.129	−0.183
$\overline{\sigma_{\beta}}$	0.051	0.055	0.010	0.020	0.049	0.054	0.010	0.020	0.030	0.034
% p-Value	0.020	0.017	0.000	0.000	0.052	0.090	0.000	0.000	3.101	1.654
% + Significant	100	100	100	100	100	100	100	100	8	7
% − Significant	0	0	0	0	0	0	0	0	83	84
% Adjusted R^2	42.33	29.50	43.41	–	43.98	31.65	44.67	–	0.95	1.87

$$\mathrm{RV}_{it} = \alpha_i + \alpha_{im}M_t + \sum_{j=1}^{12}\rho_{ij}\mathrm{RV}_{it-j} + \beta_i \mathrm{NT}_{it} + v_{it} \tag{1}$$

$$C_{it} = \alpha_i + \alpha_{im}M_t + \sum_{j=1}^{12}\rho_{ij}C_{it-j} + \beta_i \mathrm{NT}_{it} + v_{it} \tag{2}$$

$$\mathrm{RV}_{it} = \alpha_i + \alpha_{im}M_t + \beta_i \mathrm{NT}_{it} + v_{it} \tag{3}$$

$$C_{it} = \alpha_i + \alpha_{im}M_t + \beta_i \mathrm{NT}_{it} + v_{it} \tag{4}$$

$$J^*_{it} = \alpha_i + \alpha_{im}M_t + \beta_i \mathrm{NT}_{it} + v_{it}, \text{ where } J = \max(0, J^*) \tag{5}$$

The above regressions are run for the 100 largest stocks traded on the NYSE over the period January 1, 1995–September 30, 1999. $\mathrm{RV}_{it}/C_{it}/J_{it}$ is the realized variance/the continuous component of realized variance/the jump component of realized variance of stock i on day t, all multiplied by 10,000. M_t is a Monday dummy, NT_{it} is the number of trades (divided by 100) for stock i on day t and ρ_{ij} measures the persistence of volatility shocks at lag j. Eqs. (1) and (2) are estimated by OLS while Eqs. (3) and (4) are estimated by the generalized method of moments (GMM), robust regression (ROB) using iteratively reweighted least squares, and median regression (MED). Eq. (5) is estimated by maximum likelihood using the TOBIT model with and without GARCH effect (TOB–G and TOB, respectively). For brevity, we only report the equally-weighted cross-sectional mean coefficient for number of trades, with corresponding statistics. Newey–West standard errors and two-sided p-values across stocks are computed. We also report the percentage of positive and negative $\hat{\beta}_i$ coefficients which are statistically different from zero at the 5% level. The last row indicates the mean adjusted R^2s.

Table 5
Absolute order imbalance, realized volatility and its two components.

	RV				C				J	
	(1)	(3)			(2)	(4)			(5)	
	OLS	GMM	ROB	MED	OLS	GMM	ROB	MED	TOB	TOB–G
$\bar{\lambda}$	0.435	1.688	0.419	0.5418	0.420	1.67	0.424	0.535	−0.147	−0.162
$\overline{\sigma_{\lambda}}$	0.162	0.611	0.047	0.093	0.160	0.59	0.049	0.095	0.110	0.103
% p-Value	5.667	9.362	3.040	6.156	6.854	9.08	4.366	6.749	16.950	18.976
% + Significant	83	79	88	86	82	73	85	83	18	12
% − Significant	0	1	3	1	0	1	4	3	46	48
% Adjusted R^2	34.13	2.26	10.13	–	35.52	2.77	10.17	–	0.66	2.24

$$\mathrm{RV}_{it} = \alpha_i + \alpha_{im}M_t + \sum_{j=1}^{12}\rho_{ij}\mathrm{RV}_{it-j} + \lambda_i|\mathrm{OB}|_{it} + v_{it} \tag{1}$$

$$C_{it} = \alpha_i + \alpha_{im}M_t + \sum_{j=1}^{12}\rho_{ij}C_{it-j} + \lambda_i|\mathrm{OB}|_{it} + v_{it} \tag{2}$$

$$\mathrm{RV}_{it} = \alpha_i + \alpha_{im}M_t + \lambda_i|\mathrm{OB}|_{it} + v_{it} \tag{3}$$

$$C_{it} = \alpha_i + \alpha_{im}M_t + \lambda_i|\mathrm{OB}|_{it} + v_{it} \tag{4}$$

$$J^*_{it} = \alpha_i + \alpha_{im}M_t + \lambda_i|\mathrm{OB}|_{it} + v_{it},\ \text{where} J = \max(0, J^*) \tag{5}$$

The above regressions are run for the 100 largest stocks traded on the NYSE over the period January 1, 1995–September 30, 1999. $\mathrm{RV}_{it}/C_{it}/J_{it}$ is the realized variance/the continuous component of realized variance/the jump component of realized variance of stock i on day t, all multiplied by 10,000. M_t is a Monday dummy, $|\mathrm{OB}|_{it}$ is the absolute order imbalance (divided by 100) for stock i on day t and ρ_{ij} measures the persistence of volatility shocks at lag j. Eqs. (1) and (2) are estimated by OLS while Eqs. (3) and (4) are estimated by the generalized method of moments (GMM), robust regression (ROB) using iteratively reweighted least squares, and median regression (MED). Eq. (5) is estimated by maximum likelihood using the TOBIT model with and without GARCH effect (TOB–G and TOB, respectively). For brevity, we only report the equally-weighted cross-sectional mean coefficient for absolute order imbalance, with corresponding statistics. Newey–West standard errors and two-sided p-values across stocks are computed. We also report the percentage of positive and negative $\hat{\lambda}_i$ coefficients which are statistically different from zero at the 5%level. The last row indicates the mean adjusted R^2s.

In line with the previous studies, the trade size (i.e. average number of shares per trade) has a hard time explaining realized volatility. The explanatory power of trade size is poor as no single p-value is lower than 25%. For the jumps, the average coefficient is positive but not significant, confirming the poor explanatory power of trade size. Adding average trade size to the number of trades has no significant effect on the adjusted R^2. Adding average trade size in the regression does not impact the explanatory power of the number of trades (which remains high as all coefficients for the number of trades remain significant). The results are even worst when the jumps are considered. Average trade size clearly plays no role in explaining the jumps and the number of trades remains the dominant factor.[5]

Table 5 reports the results for the absolute order imbalance as sole explicative variable. There are at least two reasons why the order imbalances can provide additional power beyond pure trading activity measures. First, a high absolute order imbalance can affect returns as market makers struggle to re-adjust their inventory. Second, order imbalances can signal excessive investor interest in a stock, and if this interest is autocorrelated, then the order imbalances could be related to future returns. Table 5 shows that the absolute order imbalance does have a role in explaining realized volatility. However, evidence is less convincing for order imbalance than for number of trades. First, the lowest average p-value is still above 5% (except for the robust regressions). Second, the number of stocks displaying positive and significant coefficients for the order imbalance is around 20% lower than the number of stocks displaying positive and significant coefficients for the number of trades (see Table 4). Third, and most importantly, the explanatory power of the order imbalance is much weaker once realized volatility is decomposed. This is particularly obvious for the jumps. While the number of trades was significant at the 5% level in the jump equation, this is not the case for the order imbalance. Finally, the percentage of stocks with negative and significant coefficients for the order imbalance in the TOBIT–GARCH model is 48 only: it was almost twice higher for the number of trades.

Table 6 reports the results when the absolute order imbalance and the number of trades are both included as regressors. Adding the absolute order imbalance to the number of trades does not increase the explanatory power of the regressions. First, the adjusted R^2s for the CF regressions are similar to those reported in Table 4. Second, every average p-value for the number of trades remains smaller than 5% while only one single average p-value for the order imbalance is smaller than 10%. Third, only one single stock exhibits a positive and significant coefficient for the order imbalance when the dependent variable is either realized volatility or its continuous component. Results are similar for the jumps as, once again, the number of trades dominates the order imbalance. The number of trades is, on average, significant at 5% while the order imbalance is not. The number of trades is also the most pervasive factor, as the jumps and the number of trades are significantly and negatively related for 93 stocks out of 100 in the TOBIT–GARCH model.

To sum things up, there is a strong evidence that trade frequency remains the dominant factor behind the volume–volatility relation. The decomposition of realized volatility even strengthens the dominant role played by the number of trades. However,

[5] Given the insignificant results, we do report the tables to save space. They are available upon request.

Table 6
Absolute order imbalance, number of trades, realized volatility and its two components.

	RV				C				J	
	(1)	(3)			(2)	(4)			(5)	
	OLS	GMM	ROB	MED	OLS	GMM	ROB	MED	TOB	TOB–G
$\bar{\lambda}$	−0.163	−0.753	−0.235	−0.230	−0.177	−0.762	−0.249	−0.264	0.186	0.172
$\overline{\sigma_\lambda}$	0.164	1.346	0.097	0.085	0.160	1.188	0.056	0.085	0.134	0.124
% p-Value	28.761	31.275	13.053	14.142	27.853	28.967	6.677	11.008	26.191	23.30
% + Significant	1	1	9	2	1	1	1	2	51	52
% − Significant	42	40	64	68	45	41	89	73	1	4
$\bar{\beta}$	0.347	0.483	0.487	0.379	0.347	0.500	0.344	0.397	−0.160	−0.215
$\overline{\sigma_\beta}$	0.052	0.150	0.024	0.022	0.049	0.135	0.012	0.023	0.035	0.039
% p-Value	0.049	3.339	0.000	0.000	0.056	2.423	0.000	0.000	4.143	1.367
% + Significant	100	94	100	100	100	95	100	100	2	2
% − Significant	0	0	0	0	0	0	0	0	89	93
% Adjusted R^2	42.90	17.32	31.30	–	44.60	20.90	46.12	–	1.23	1.91

$$\mathrm{RV}_{it} = \alpha_i + \alpha_{im}M_t + \sum_{j=1}^{12}\rho_{ij}\mathrm{RV}_{it-j} + \lambda_i|\mathrm{OB}|_{it} + \beta_i\mathrm{NT}_{it} + \nu_{it} \quad (1)$$

$$C_{it} = \alpha_i + \alpha_{im}M_t + \sum_{j=1}^{12}\rho_{ij}C_{it-j} + \lambda_i|\mathrm{OB}|_{it} + \beta_i\mathrm{NT}_{it} + \nu_{it} \quad (2)$$

$$\mathrm{RV}_{it} = \alpha_i + \alpha_{im}M_t + \lambda_i|\mathrm{OB}|_{it} + \beta_i\mathrm{NT}_{it} + \nu_{it} \quad (3)$$

$$C_{it} = \alpha_i + \alpha_{im}M_t + \lambda_i|\mathrm{OB}|_{it} + \beta_i\mathrm{NT}_{it} + \nu_{it} \quad (4)$$

$$J^*_{it} = \alpha_i + \alpha_{im}M_t + \lambda_i|\mathrm{OB}|_{it} + \beta_i\mathrm{NT}_{it} + \nu_{it},\ \text{where}\ J = \max(0, J^*) \quad (5)$$

The above regressions are run for the 100 largest stocks traded on the NYSE over the period January 1, 1995–September 30, 1999. $\mathrm{RV}_{it}/C_{it}/J_{it}$ is the realized variance/the continuous component of realized variance/the jump component of realized variance of stock i on day t, all multiplied by 10,000. M_t is a Monday dummy, $|\mathrm{OB}|_{it}$ is the absolute order imbalance (divided by 100), NT_{it} is the number of trades (divided by 100) for stock i on day t and ρ_{ij} measures the persistence of volatility shocks at lag j. Eqs. (1) and (2) are estimated by OLS while Eqs. (3) and (4) are estimated by the generalized method of moments (GMM), robust regression (ROB) using iteratively reweighted least squares, and median regression (MED). Eq. (5) is estimated by maximum likelihood using the TOBIT model with and without GARCH effect (TOB–G and TOB, respectively). For brevity, we only report the equally-weighted cross-sectional mean coefficient for absolute order imbalance, with corresponding statistics. Newey–West standard errors and two-sided p-values across stocks are computed. We also report the percentage of positive and negative $\hat{\lambda}_i$ coefficients which are statistically different from zero at the 5% level. The last row indicates the mean adjusted R^2s.

trade frequency affects the two components of realized volatility very differently: the relation between the trade frequency and the continuous component of realized volatility is positive while the relation between the trade frequency and the jump component is negative.

5. Conclusion

The ability to identify realized jumps has important implications in financial management, from portfolio and risk management to option and bond pricing and hedging. In particular, Eraker et al. (2003) show that the jump components command relatively larger risk premia than the continuous components, as their contribution to periods of market stress is greater. There is therefore a practical interest in identifying the jumps.

By decomposing realized volatility into its two components, we show that trading activity relates to volatility in a more subtle way than previously thought. Because volatility can be diffusive or discontinuous, stating that 'the greater the level of volume, the greater the volatility' hides a more complex reality. While trading activity relates positively to diffusive (or continuous) volatility, it relates negatively to the jumps (or discontinuous volatility). In other words, the positive relationship between the volume and the volatility that is documented in the literature only holds for diffusive volatility. This indicates that poor trading volume leads to more erratic volatility changes, as commonly argued in dealing rooms. This negative volume–jumps relation is revealed through the number of trades, which remains the dominant factor behind the volume–volatility relation. Neither trade size nor the order imbalance adds significantly more explanatory power beyond the number of trades, whatever the volatility component considered.

While Andersen et al. (2007) show that the jumps in exchange rates, stocks and bonds are linked to fundamentals, we still need to understand how the link really works. This will certainly help revive studies about the impact of news and economic events on financial markets.

Acknowledgements

We gratefully acknowledge the helpful comments from seminar participants at the 2009 EFM Symposium on Risk Management in Financial Institutions (Nantes, FR), the 2009 Humboldt–Copenhagen Conference on Recent Developments in Financial

Econometrics (Berlin, GE), the 11th Symposium on Finance, Banking, and Insurance (Karlsruhe, GE), and the 12th Belgian Financial Research Forum (Gent, BE). We also thank the editor, Richard Baillie, as well as an anonymous referee for their contributions.

References

Andersen, T.G., Bollerslev, T., 1998. DM–Dollar volatility: intraday activity patterns, macroeconomic announcements and longer run dependencies. The Journal of Finance 53, 219–265.
Andersen, T.G., Bollerslev, T., Diebold, F.X., 2006. Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility. Review of Economics and Statistics 89, 701–720.
Andersen, T., Bollerslev, T., Diebold, F., Vega, C., 2007. Real-time price discovery in stock, bond and foreign exchange markets. Journal of International Economics 73, 251–277.
Barndorff-Nielsen, O., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics 2, 1–48.
Barndorff-Nielsen, O., Shephard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 4, 1–30.
Calzolari, G., Fiorentini, G., 1998. A TOBIT model with GARCH errors. Econometric Reviews 17 (1), 85–104.
Chan, K., Fong, W., 2000. Trade size, order imbalance, and the volatility–volume relation. Journal of Financial Economics 57, 247–273.
Chan, K., Fong, W., 2006. Realized volatility and transactions. Journal of Banking and Finance 30, 2063–2085.
Chordia, T., Roll, R., Subrahmanyam, A., 2001. Market liquidity and trading activity. Journal of Finance 56, 501–530.
Easley, D., Kiefer, N., O'Hara, M., 1997. One day in the life of a very common stock. Review of Financial Studies 10, 805–835.
Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. The Journal of Finance 58, 1269–1300.
Glosten, L., Milgrom, P., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 14, 71–100.
Harris, L., 1987. Transaction data tests of the mixture of distributions hypothesis. Journal of Financial and Quantitative Analysis 22, 127–141.
Huang, R., Masulis, R., 2003. Trading activity and stock price volatility: evidence from the London Stock Exchange. Journal of Empirical Finance 10, 249–269.
Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variation. Journal of Financial Econometrics 3, 456–499.
Jones, C., Kaul, G., Lipson, M., 1994. Transactions, volume and volatility. Rewiew of Financial Studies 7, 631–651.
Karpoff, J., 1987. The relationship between price changes and trading volume: a survey. Journal of Financial and Quantitative Analysis 22, 109–126.
Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.
Lahaye, J., Laurent, S., Neely, C.J. (2008): Jumps, Cojumps and Macro Announcements. Mimeo.