The Relationship Between Trading Volume and Jump Processes in Financial Markets

Pongpitch Amatyakul

Professor George Tauchen, Faculty Advisor

- The Duke Community Standard was upheld in the completion of this paper.
- Honors Thesis to be submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College, Duke University.
- Duke University, Durham, North Carolina, 2010
Acknowledgements

I would like to thank my thesis advisor, Professor George Tauchen, for his guidance and advice throughout this research process. I am grateful for feedback I received from Professor Tim Bollerslev and Bruno Feunou throughout Econ 201. I would like to thank my classmates Sam Lim, Derek Song, Matt Rognlie, and Ab Sawant for their comments and suggestions.
Abstract

This paper explored the relationship between volume and jumps in stock prices using high frequency, minute by minute, data from the S&P 500 stocks. It has been known that stock prices compose of a continuous part that follows geometric Brownian motion and a jump part. There have been jump-detection tests developed by various financial economists. These jump tests were performed, analyzed, and regressions had been run to find possible relationships between the trading volume and jump detection. The relationship tends to be slightly negative. Economically, this gives evidence towards jumps being a result of common knowledge shocks. Common knowledge shocks are information interpreted the same way resulting in price movements without heavy trading.
1. Introduction

Jumps are rare financial events in which asset prices move drastically from one moment in time to the next. This paper aims to find what type of information flow causes these jumps to occur in the financial market. Jumps can play an important role in risk and portfolio management, and jump detection has been a hot area of research in the past decade due to more accessible high frequency data. Although continuous models remain prevalent decades later, potentials for jumps were first noticed as early as 1976 by Merton; he figured out that pricing of assets should not just incorporate the continuous process, but should also take into account the jumps that occur. If asset prices follow purely continuous processes, then it should be possible to draw a smooth line between prices at different times without lifting a pen or pencil, but that is not the case from observations of real data. Since then, we have fully established that jumps do exist, and researchers have been trying to study and gather more information on them in order to find better ways to manage risk.

Trading volume is defined to be the number of shares traded over a given day. Trading volume is easily observable in financial data, and it can provide insight to the structure of financial markets. It can tell us the amount of market activity at any given time interval. However, the volume alone doesn’t give away much information, so it is necessary to investigate the relationship between volume and other observable data in the market, such as price. The relationship of price and volume can give us the rate of information flow in a market. According to the model developed by Tauchen and Pitts (1983), volume and price changes are related. Price changes form the basis of estimators used in jump tests, so there could be some relationship between jumps and volume. This relationship can be explored by implementing different jump tests from literature.

This paper is the first to explore the economic interpretation of jumps by observing the relationship between trading volume and jumps. It is known that volume is indicative of information flow. Epps and Epps (1976) built a model in which a larger amount of disagreement among traders results in a higher volume. An example of this would be a rumor of a merger between company A and company B. If a group of investors believe that this is beneficial to company A, then they will price company A’s stock at a higher price, and they would be willing to buy stocks at the original price or
lower. They would buy company A’s stocks from those investors who believe that the merger is going to harm company A. This results in high volume of trading. If jumps are indeed a result of disagreements, then it should have a positive relationship with volume. On the other hand, jumps can result from common knowledge shocks. A common knowledge shock is a result of information being interpreted exactly the same way by most of the traders. Following the example above, if all investors believe that the merger between company A and company B will help company A operate more efficiently, then everybody suddenly values company A’s stocks at a higher price, the price will jump up instantaneously without much change in volume. If we assume that regular price changes without jumps have some positive relationship with volume, this common knowledge shock could lower that trend between jumps and volume. If jumps occur on a particular day, the volume on that day could be lower than other days with similar price changes. So there are two ways that a jump can occur; the relationship between the jump component and volume can give us key insights as to which of the two theories dominate the jump process.

On a similar note, this paper will also discuss and test the relationship of volume and realized variance. Realized variance is the best way to estimate the variance of stock price movements. Realized variance can be accurately approximated using high frequency data. There has been literature such as Tauchen and Pitts (1983) and Andersen (1996) that include models and empirical data that indicated a positive relationship between the variance and volume at least for longer time-spans. This paper will test whether those relationships exist in high frequency data covering the last 12 years.

The next logical step is to place the all the variables into the same regression. Since we already have the individual relationships, it would be interesting to see how they interact. This can separate out their individual effects and could yield more insights to the dynamics of volume, jumps, and variance in general.

This paper is organized in the following way: Section 2 lays out the background for the model of equity prices which the jump tests are based upon. Section 3 describes each of the three jump tests and how to calculate the daily jump test statistics, which will eventually be used in the regressions. Section 4 discusses the market microstructure noise, the effects it has on high frequency data, and how to minimize those effects.
Section 5 introduces the data that was used, and section 6 describes the regressions that were performed in order to determine the relationship between jumps and volume. The results of the regressions are explained in section 7, and interpreted in section 8. Finally, section 9 summarizes the findings and highlights important conclusions. The tables and figures that accompany the paper can be found in section 10 and the appendix.

2. Movements of Equity Prices

Before discussing jump tests, it is important to fully comprehend the theory of equity prices and returns. This section will first introduce the necessary variables (2.1) and will then go on to discuss the theoretical processes that underlie stock prices and returns (2.2). This will provide a basis for what is used to determine jump test statistics in section 3.

2.1 Basic Variables

The basic variables that are necessary to understand the movement of stock prices will be defined here. The first variable will be the logarithmic price.

\[ p(t_i) = \log(s(t_i)) \]  

(1)

Where \( s(t_i) \) is the real stock price at time \( t_i \).

Following that equation, the returns at each time period will be defined by:

\[ r_i = p(t_i) - p(t_{i-1}) \]  

(2)

2.2 Model of Equity Prices

The theory behind security price movements will be developed here. Prior to Merton (1976), stock price returns were thought to have followed Brownian motion with mean \( \mu \) and standard deviation \( \sigma \). The stochastic equation for the stock price return \( r \) is the following:

\[ dr(t) = \mu(t)dt + \sigma(t)dW(t) \]  

(3)

where \( W \) is the normal distribution.

The model above is deemed insufficient by Merton (1976). He claimed that the stock price returns are a mixture of continuous stochastic processes and Poisson-driven ‘jump processes.’ Thus, the original model has been replaced by the continuous time
jump diffusion model. This model introduces one more term to the equation in the form of a jump process. A jump may or may not occur in any given time interval. The model is as follows:

\[ dr(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t) dq(t) \]  \hspace{1cm} (4)

where \( \kappa \) is the size of the jump, and \( q(t) \) is the number of jumps up until time \( t \).

3. Jump Tests and Jump Test statistics

In recent years, with readily-accessible high frequency data and greater processing power, several tests have been developed to assess whether stock price returns contain discontinuous processes. The most notable ones include Mancini (2006), Lee and Mykland (2006), Barndorff-Nielsen and Shephard (2006), Jiang and Oomen (2008), and Aït-Sahalia and Jacod (2008). The last three were selected for this paper to explore the relationship between volume and jumps.

3.1 Barndorff-Nielsen and Shephard Test

The Barndorff-Nielsen and Shephard test (2004, 2006) is one of the most commonly used jump tests in literature. It uses the relationship between realized variance and bipower variation in order to detect rare jumps in stock prices. If there are no jumps, bipower variation and realized variance asymptotically approach the integrated variance as the sampling frequency reaches infinity.

If each sampled data is represented by the letter \( i \) and there are \( M \) number of sampled data per day, the daily realized variance is defined as the following.

\[ RV_i = \sum_{i=1}^{M} r_{i,i}^2 \]  \hspace{1cm} (5)

In the limit, the realized variance goes to the integrated variance plus the jump component, \( \kappa \).

\[ \lim(M \to \infty)RV_i = \int_{t-1}^{t} \sigma^2(s)ds + \sum_{i=1}^{M} \kappa^2(t_i) \]  \hspace{1cm} (6)

Bipower variation is defined to be the following.

\[ BV_i = \frac{\pi}{2} \left( \frac{M}{M-1} \right) \sum_{i=2}^{M} |r_{i-1}||r_i| \]  \hspace{1cm} (7)
In the limit, this bipower variation goes just to the integrated variance.

\[
\lim(M \to \infty)BV_t = \int_{t-1}^{t} \sigma^2(s) ds \tag{8}
\]

These asymptotic properties only hold at high sampling frequencies, which is the case for the data used in this paper. Our interest lies in the jump component and so we will try to isolate this component. This could be achieved by obtaining the difference between the realized variance and the bipower variation. Huang and Tauchen (2005) used this variable called relative jump.

\[
RJ_t = \frac{RV_t - BV_t}{RV_t} \tag{9}
\]

It is necessary to studentize this value so that we can compare the statistic to the normal distribution. After studentizing, the numbers will be scale-free and can be used in hypothesis testing. As a result, it is necessary to find the standard deviation. Barndorff-Nielsen and Shephard (2006) recommended using the quadpower quarticity term, defined here.

\[
QP_t = M\mu_1^{-1}\left(\frac{M}{M - 3}\right)\sum_{j=4}^{M} |r_{i,j-3}| |r_{i,j-2}| |r_{i,j-1}| |r_{i,j}| \tag{10}
\]

where \( \mu_1 = \sqrt{\frac{2}{\pi}} \)

In the limit, this becomes the integrated quarticity.

\[
\lim(M \to \infty)QP_t = \int_{t-1}^{t} \sigma^4_s ds \tag{11}
\]

With this information, Barndorff-Nielsen and Shephard (2006) suggested the ratio max-adjusted test statistic which is a one sided normal with variance one.

\[
\frac{RJ_t}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right) \frac{1}{M} \max(1, \frac{QP_t}{BV_t^2})}} \tag{12}
\]

This is the test statistic that will be used later on in order to find the relationship between volume and jump days; the higher the statistic, the higher probability that the day contains a jump.
3.2 Jiang and Oomen Test

Based partially on the work of Barndorff-Nielsen and Shephard, Jiang and Oomen (2008) devised their own method to test for jump days. It is based on a variable in which they called ‘swap-variance.’ It is motivated by an observation in finance that, in the absence of jumps, the difference between simple return and log returns captures one half of the integrated variance. This observation is well known in financial literature and is the basis of the variance swap replication strategy, a strategy that will only work in the absence of jumps. Specifically, this is the equivalent of a delta hedged log contract when there is no discontinuity.

Let us define some variables. Capitalized R is the geometric return of the real price.

\[
R_{t,i} = \frac{s_{t,i} - s_{t,i-1}}{s_{t,i-1}}
\]  

Swap variance is defined to be

\[
SwV_i = 2 \sum_{i=1}^{M} R_{t,i} + 2 \ln \frac{S_T}{S_0}
\]  

Where \(S_T\) is the price at the end of the day and \(S_0\) is the opening price.

In the probability limit, swap variance and realized variance (defined in equation 5) should be equal if there is no discontinuity. In the absence of jumps, the only difference between \(RV\) and \(SwV\) should come as a result of discretization, but this value is likely to be small and negligible. With jumps, a noticeable difference should be observed.

The sign of the difference between swap variance and realized variance is also significant. A negative result means that a negative jump occurred and a positive result means that a positive jump occurred.

The task is again to studentize the test statistic so it is distributed normally and thus the test statistic is easy to interpret. Jiang and Oomen came up with three simple test statistics that are distributed as standard normals. This paper picked only one of the three: the ratio test. The formula of the ratio test is as follows:

\[
Z_{JO} = \frac{BV_i M}{\sqrt{\Omega_{SwV}}} \left( 1 - \frac{RV_i}{SwV_i} \right)
\]  

where M is the number of sampled data per day.
The formula uses the fact that the ratio of realized variance over the swap variance is supposed to be 1. Studentization was completed by multiplying the bipower variation and dividing through by standard deviation, which corresponds to the square root of the omega term, defined below

\[ \Omega_{SwV_t} = \frac{\mu_0}{9} \int (\sigma_u^2)^3 du = \frac{\mu_0}{9} \frac{M^2 \mu_{3/2}}{M - 3} \sum_{i=4}^{M} \left| r_{i,t} \right|^{3/2} \left| r_{i,t-1} \right|^{3/2} \left| r_{i,t-2} \right|^{3/2} \left| r_{i,t-3} \right|^{3/2} \]

(16)

where \( \mu = E(|x|^p) \) and \( x \) is normally distributed with mean 0 and variance 1.

The test statistic that will be used to find the relationship between volume and jump will be the absolute value of equation 15. Since a jump will occur if the absolute value of \( Z_{JO} \) is greater than some threshold. It is also noted that this paper does not explore the information regarding negative and positive jumps, and thus it is not necessary to retain the sign of equation 15.

### 3.3 Aït-Sahalia and Jacod Test

Aït-Sahalia and Jacod (2008) introduced a novel method to detect jumps. They take the sum of the absolute returns and put it to a high power (\( p \)). The test itself uses two different sampling intervals: one at \( \Delta \) and the other one at \( k\Delta \). This paper uses \( \Delta \) equals to five, \( k \) equals to two, and \( p \) equals four.

The estimator \( B \) is defined to be:

\[ B(p, \Delta)_t = \sum_{i=1}^{M} |r_{i,t}|^{p} \]

(17)

The test statistic is:

\[ S_t = \frac{B(p, \Delta)_t}{B(p, k\Delta)_t} \]

(18)

The test statistic \( S \) can be interpreted with regards to jumps as follows: it should converge to \( k^{p/2-1} \) or 2 for the specifications done in this paper, if there are no jumps. On the other hand, it will converge to 1 if jumps exist. Under the null hypothesis of no jump, the test statistic is less than another statistic \( c \). In order to build the \( c \) statistic, it is easiest to define some more variables.

The variable \( O \) is defined as:
\[ O(p, k) = \frac{1}{\mu_p^2} \mu_{z_p} k^{p-2}(1 + k) + k^{p-2}(k - 1) \mu_p^2 - 2k^{p/2-1} \mu_{k-p} \]

\( \mu_a \) is defined the same way as in Equation 16, \( \mu_{a,b} = E([X]'[X + \sqrt{k+1}Y]) \), where both \( X \) and \( Y \) are standard normals independent of each other.

The following variable \( A \), is the equivalent of Barndorff-Nielsen and Shephard's multipower variation
\[
A(v, q, \Delta) = \frac{\Delta^{1-qv/2} \prod_{i=0}^{q+1} \sum_{j=0}^{q-f} |r_{i+j-1}|^v}{\mu_r^q}
\]
where \( v=p/(p+1) \) and \( q=p+1 \), where \( p \) is defined earlier (at the beginning of section 3.3).
\( \mu \) was defined following equation 16.

Another variable \( V \) is defined to be:
\[
V_t = \frac{\Delta O(p, k) A(v, 2q, \Delta)}{A(v, q, \Delta)^2}
\]

With those variables defined, we can now define \( c \)
\[
c_t = k^{p/2-1} - z_\alpha \sqrt{V_t}
\]
where \( z_\alpha \) is the standard normal point corresponding to confidence interval alpha, which is defined to be 95% in this paper.

If the statistic \( S \) in equation 18 is less than \( c \) in equation 22, then a jump is detected. So the statistic chosen to find the relationship between volume and jump day is \( c-S \). The higher the \( c-S \) statistic, the higher chance that the stock will jump on that day.\(^1\)

### 4 Market Microstructure Noise

This section is about market microstructure noise. It is divided into 2 subsections.
The first subsection describes mathematically what market microstructure noise is and the process to minimize its effects on the findings. The second subsection investigates the effect of changes in tick size on market microstructure noise and determines whether or not sampling intervals should be adjusted to correspond to tick size changes.

\(^1\) Please note that this statistic is analogous to the studentizing process that was applied to the Barndorff-Nielsen and Shephard and Jiang and Oomen tests. If we had subtracted \( s \) by \( k^{p/2-1} \) and divide by the square root of \( V \), the format would have been very similar.
4.1 Minimizing the effects of market microstructure noise

Market microstructure noise is a phenomenon in which the market price does not exactly reflect the real price of a security. In Hansen and Lunde (2006) the market microstructure noise is defined to be

\[ u(t) = p(t) - p^*(t) \] (23)

where \( p^*(t) \) is the latent real log price and \( p(t) \) is the observable log price in the market at time \( t \). Only short term deviations in prices such as the bid-ask spread, instantaneous information asymmetry, and tick sizes are considered market microstructure noise. Long-term deviations such as those arising from irregular investor behaviors do not contribute to market microstructure noise. For many of the estimators used in high frequency analysis, such as realized variance, it would be ideal to use the latent real log price. However, it is not possible to know what that is, and all the information observable is the change in prices:

\[ p(t + \delta) - p(t) = p^*(t + \delta) - p^*(t) + u(t + \delta) - u(t) \] (24)

Now let us analyze equation 24 by decreasing the size of the interval, \( \delta \). As \( \delta \) is decreased, the magnitude of the change in latent real log price decreases. This is because, when the interval shortens, there is less new information for the investors to process, and that results in a smaller price movement. However, the change in the noise is independent of \( \delta \) as the noise term is, in theory, independent and identically drawn. Noise is inherently build into the price and is supposed to be constant in magnitude over time. This means that as we make \( \delta \) closer and closer to zero, the change in latent real log price will be minimal, and all we observe is the change in noise.

From the discussion above, it could be seen why market microstructure noise poses a very large problem to high frequency data. High frequency corresponds to the case where \( \delta \) approaches zero. With noise, some of the vital estimators such as realized variance and bipower variation could be biased. According to Merton (1981), the best way to estimate realized variance is to observe continuous price movements without noise. In theory, if there is no noise, then it is best to sample at the highest possible frequency. However, in reality, this market microstructure noise play such a big role at the highest frequencies that this bias is intolerable.
A simple graphical tool called “signature volatility plot” was created by Andersen, Bollerslev, Diebold, and Labys (1999) and used to get a sense of how large of a role the noise factor plays at different sampling frequencies. The idea here is that variance is independent of the sampling frequency at which prices are observed. In signature volatility plots, average realized variance (or the equivalent annualized standard deviation) is plotted against different sampling frequencies. These should be the same if noise does not play a role. Since the data in the real financial world is not ideal and contains noise, there will be some trends to observe. Andersen et al (2000) found that if stocks are very liquid, as is the case in all of the stocks used in this paper, then the shape of the volatility signature plots will be downward sloping. Realized variance is the sum of the squares of the returns. The returns are the change in price as represented by equation 24. As the sampling frequency decreases, the \( p^*(t+\delta) - p^*(t) \) term dominates the noise term \( u(t+\delta) - u(t) \), and RV is no longer biased.

Figure 1 is an example of a volatility signature plot from one of the ten stocks chosen for this paper: Coca Cola. The average of each day’s realized variance is converted into annualized standard deviation and plotted against sampling frequency, which ranges from the finest possible given the data (1-minute) to 20 minutes. Based on this and several other plots using other stocks, it was decided that the sampling frequency should be ten minutes in order to lower the noise and still be able to obtain all the basic price movements.

4.2 Tick Size and Volatility Signature Plots

This sub-section aims to find out whether the sampling interval should change adaptively, as there are two changes in the tick size during the time period, or whether they should be constant. Price discreteness, or rounding, is one source of market microstructure noise and is directly affected by tick size. Tick size is the minimal increment in which the stock price can move. In theory, if the price increment is \( \lambda \), the maximum possible error due to rounding is one-half of \( \lambda \). If the real price distribution is uniform and can take any value in the increment, then the average error due to rounding will be one-fourth of \( \lambda \). Thus, if lambda is lowered, there should be a corresponding reduction in market microstructure noise. This is confirmed by literature, as Hansen and
Lunde (2006) found that noise plays a much smaller role after the tick size dropped from 1/16 of a dollar to one cent in 2001.

Signature volatility plots of a sample of stocks at different time intervals were plotted and analyzed. There were two tick size changes in the data. The first one was only two months into the data, in June 1997, when the tick size was changed from 1/8 of a dollar to 1/16 of a dollar. The second time was in January 2001, when the tick size was decreased to 1 cent. This naturally divides the data into three parts. Coca-Cola was the stock chosen, along with Procter and Gamble, to show the volatility signature plots for the different time periods. It is difficult to compare the three plots even on the same graph because there existed many natural changes in volatility during the three time periods, and thus the decision was made to tabulate the data. The change in volatility, in terms of the percent change from the volatility at one minute sampling interval, seemed like the most obvious calculation. It can tell relatively what the trend and how flat the signature volatility plots are. This result is displayed in table 1.

Let us first focus on the result of Coca Cola. Table 1 showed that drop-off of volatility is most prevalent where the tick size is highest and that the drop-off is much lower in the last 8 years where the tick size is in cents. Just changing from sampling at 1 minute intervals to 5 minute intervals decreased volatility by 37 percent in the first time interval in 1997, compared to just about 10 percent for the whole set of data. However, this could be misleading because the data set is smallest in the first interval, only 36 days, as compared to over 1900 days from the interval from 2001 to 2009, but the conclusion can be made because the same trend exists when the increment decreased from 1/16 of a dollar to 1 cent.

Similar trends were observed from the results of Proctor and Gamble as well. For PG, the drop-off difference between when the tick size was 1/8 of a dollar from 1 minute to 5 minutes is relatively low (6.9%), and the drop-off is greater for 1/16 dollar (9.1%), again, possibly due to the small sample size in the first interval. Despite that, the conclusion is clear for both of the stocks chosen; the minimum changes in volatility for all the sampling frequencies occur after the tick size had been reduce to 1 cent. That is, the volatility signature plots remained relatively flat when the tick size is small and steepened at the beginning when the tick size is large. In term of market microstructure
noise, this result confirms the theoretical assumption that reducing the tick size directly reduces the noise.

A decision had to be made regarding the sampling interval that will be used. Although it might be better to approximate the different estimators such as RV using a higher sampling frequency after each tick size change, market microstructure noise consists of more than just rounding error and tick size changes. There are other parts that are not fully captured such as the bid-ask spread and information asymmetry. Therefore, in this paper, the sampling interval of ten minutes was used, unless stated otherwise.

5. Data

5.1 Raw Data

The data used for this paper are based on the minute-by-minute price quotes from the commercial vendor, price-data.com. The stocks chosen were the 10 largest companies (based on market capitalization) in the S&P 100 index as of 31 December 2008 that have the full set of data for the last 12 years. The ten stocks included Procter and Gamble (symbol: PG), General Electric (symbol: GE), AT&T (symbol: T), Johnson and Johnson (symbol: JNJ), Microsoft (symbol: MSFT), Pfizer (symbol: PFE), JP Morgan Chase (symbol: JPM), International Business Machines (symbol: IBM), Cisco Systems (symbol: CSCO), and Coca Cola (symbol: KO). The time period was between 9:35am to 4:00pm of every trading day from April 1997 to January 2009. The first five minutes of the day from 9:30 to 9:35 were discarded because the prices have yet to stabilize from the previous closing. There were 385 observations in a day. For the analysis in this paper, the data was sampled either every five minutes or every ten minutes, resulting in either 77 or 38 data points per day. Sub-sampling was not used and thus the rest of the information was discarded. The number of days ranges from 2921 days to 2925 days, depending on the individual stocks.

The volume data are obtained from Google finance. It is the number of shares that were traded per day, everyday, in alignment with the price quotes.

5.2 Notes on volume
There are volume trends that should be observed as it could explain several phenomena that occurred in the regressions. After all of the volumes versus time plots were generated for all the ten stocks, it could be seen that normally volume is in the range of millions to tens of millions per day. This means that all the stocks have very high liquidities, as expected since they are the top stocks in the S&P 100. One of the trends observed is that the volumes spike up at the end of 2008 and early 2009. This is true for all the stocks except IBM. The spikes correspond directly to the economic crisis that occurred following the housing bubble burst. Overall, more stocks have an upward volume trend through the sample.

There are a few things to note for individual stocks. Pfizer and AT&T seemed to have a more prominent upward trend in the data set. IBM’s volumes spiked up the most around 2000 and never reach that height again. This, no doubt, is the result of the tech bubble that occurred earlier in the decade. Microsoft and Cisco’s volume trends are very similar to each other; they have very high volatilities and look the same throughout the sample except for the last few months, in which its volume and volatilities spiked up just like the other normal stocks.

It was of interest to see what the volume trends were for each day of the week. This was motivated by findings by Jain (1988) that volume averages are different for each day of the week. They reported that volume is lowest on Monday, increases monotonically until Wednesday, and then decreases monotonically until Friday. This paper aims to see what sort of trends exist for the stocks chosen. This was plotted in figure 2 for four randomly selected stocks: Coca Cola, Johnson & Johnson, Procter and Gamble and AT&T. The only main trend seen here is that volume is lowest on Monday. It seemed like the volume is randomly distributed on the other days. Some simple regression of volume on each of the days of the week was done for several of the stocks and the results seemed to confirm that. The only coefficient that is almost always statistically significant is Monday, and that coefficient is always negative. This will play a role in the regression.

6. Regression Methodology
This section lists the different regressions performed to find the relationships between the jump statistics and trading volume. This will be an exploratory data analysis, and thus there will be many regressions with different regressors. The aim is to gather as much information as possible for analysis.

The first model will be the most simple, just a one variable regression of the log of volume on the jump statistics. The volume (v) was logged in order to find the effect of one unit change in the statistic to percent volume. The test statistics include the daily Z statistic for the Barndorff-Nielsen-Shephard and the Jiang-Oomen test, as well as the daily \(c - S\) statistic for the Aït-Sahalia and Jacod test. The details of these statistics can be found in section 3. All these statistics represent the likelihood of jumps; with a higher number meaning a higher chance. So the regressions, in terms of the variables, are as follows:

\[
\log(v_t) = \beta_1 Z_{BNS} + \beta_2 
\]

(25)

\[
\log(v_t) = \beta_1 |Z_{JO}| + \beta_2 
\]

(26)

\[
\log(v_t) = \beta_1 (c_t - S_t) + \beta_2 
\]

(27)

The realized variance had long been known to be correlated with volume and it would be interesting to include it into a regression with volume and jump statistic. There have been multiple findings in literature including Tauchen and Pitts (1983) and Andersen (1996) that suggest a relationships between volume and realized variance, at least for observations over a longer time span. Before adding it to the regression with jump statistic, it is wise to test these findings in literature to see whether it fits into the data set explored in this paper. Thus, the following regressions’ results can give us a general idea of what the relationships exist between volume and realized variance. The coefficient \(\beta_1\) would give us the corresponding percent change in volume after a one-percent increase in the realized variance.

\[
\log(v_t) = \beta_1 \log(RV_t) + \beta_2 
\]

(28)

A more complicated model will be used for the following regressions. It was observed that there are more factors that affect volume, and volumes are not independent
of each other. From the plot of volume versus time, it was observed that volume tend to cluster similar to volatility clustering; high volume days tend to follow high volume days. It was also argued in the data section that Monday probably has an effect on volume and should be included as one of the regressors. As a result, new models incorporating these variables are created. Following Corsi (2009) and his HAR model, these models include the volume of the day before, the average volume of the week before, and the average volume of the month before. The dummy (0 or 1) variable for Monday is also included following the discussion in section 5. The week before is defined to be 5 working days and the month before is defined to be 22 working days, just like Corsi’s HAR model:

\[
\log(v_t) = \beta_1 Z_{BNS} + \beta_2 \log(v_{t-1}) + \beta_3 \log(v_{t,5}) + \beta_4 \log(v_{t,22}) + \beta_5 Monday + \beta_6 \\
\log(v_t) = \beta_1 |Z_{JO}| + \beta_2 \log(v_{t-1}) + \beta_3 \log(v_{t,5}) + \beta_4 \log(v_{t,22}) + \beta_5 Monday + \beta_6 \\
\log(v_t) = \beta_1 (c_t - S_t) + \beta_2 \log(v_{t-1}) + \beta_3 \log(v_{t,5}) + \beta_4 \log(v_{t,22}) + \beta_5 Monday + \beta_6
\]

The last regression’s aim is to explore which part causes the volume to change, whether it is the realized variance or the jump statistic itself. After performing the regressions corresponding to equations 25-28, we will get to observe the impacts of the realized volatility and jumps on volume. In order to explore how these factors interact, it is necessary to have regressions where both terms appear on the right hand side of the regressions. These regressions are almost identical to the one above, with the realized variance added to the regressor:

\[
\log(v_t) = \beta_1 Z_{BNS} + \beta_2 \log(v_{t-1}) + \beta_3 \log(v_{t,5}) + \beta_4 \log(v_{t,22}) + \beta_5 Monday + \beta_6 \log(RV) + \beta_7 \\
\log(v_t) = \beta_1 |Z_{JO}| + \beta_2 \log(v_{t-1}) + \beta_3 \log(v_{t,5}) + \beta_4 \log(v_{t,22}) + \beta_5 Monday + \beta_6 \log(RV) + \beta_7 \\
\log(v_t) = \beta_1 (c_t - S_t) + \beta_2 \log(v_{t-1}) + \beta_3 \log(v_{t,5}) + \beta_4 \log(v_{t,22}) + \beta_5 Monday + \beta_6 \log(RV) + \beta_7
\]

7. Results
The results of the simple regressions corresponding to equations 25 through 27 are tabulated in table 2. For the Barndorff-Nielsen and Shephard test, five of the ten stocks have a negative relationship between volume and jump statistic that is significant at the 5 percent level. The other five have coefficients that are not significant. As for the Jiang-Oomen swap variance test, all the coefficients are significant except for Microsoft. Seven showed a negative relationship between the probability of jumping and volume on that day, and two showed a positive relationship. For the last test, the Aït-Sahalia-Jacod test, nine of the relationships were significant and all pointed towards a negative correlation between jump statistic and volume. Overall, a simple regression suggests a negative relationship between jump statistic and volume.

In table 3, we see the results of the regression described by equation 26. For the 5% significance level, six of the ten stocks showed a positive relationship between volume and realized variance, and two showed a negative relationship. This matches results in existing literature. To summarize, the simple basic relationship is negative for volume and jump statistic but positive between volume and realized variance.

Equations 29-34 were performed, and the results are included in the appendix. In almost all the cases, the R-squared values significant improved from that in the simple regressions as a result of the additional variables. Let’s first analyze the newly introduced variables that weren’t included in the simple regression. The Monday variable behaved as expected, at the 95% confidence level, it was negatively correlated with volume in all of the stocks except Cisco and Microsoft. Since the volume data of these two stocks had high volatilities and lacked observable trends (as noted in section 5.2), any regression results will diverge from that of the other stocks. Generally in the other stocks, if the trading day is Monday, the volume can be expected to be roughly 8 to 12 percent lower.

The average volume the day before, the week before, and the month before are all positively correlated with the current day’s volume, as expected. The stocks for which these coefficients are sometimes not significant are those in the hi-tech industry. Other than Cisco and Microsoft, the other tech company in the data is IBM. The result suggests that tech stock volumes are less dependent on prior volumes.
As for the last three regressions corresponding to equations 32-34, looking only at the coefficients of realized variance, most of them remained positively correlated with volume. The exceptions, again, comes from Cisco and Microsoft.

Generally, equations 32-34 change very little of the coefficient that relates jump statistics and volume as compared to equations 29-31. A summary table is provided in table 4. For the Barndorff-Nielsen and Shephard Test, its statistic does not seem to be correlated with volume. Without the realized variance variable, one stock showed a positive relationship and the other showed a positive relationship. The positive relationship disappeared when the realized variance variable is included. As for the Ait-Sahalia-Jacod Test, its statistic is like the Barndorff-Nielsen and Shephard in that there is inconclusive evidence of a relationship between jump statistic and volume. There are three stocks that showed a negative relationship but there was 1 that showed the opposite. Lastly, for the Jiang-Oomen test, the result is clearer. The relationship between jump statistic and volume is negative.

8. Interpretation

For the simple regressions, the results showed a negative relationship between volume and jump statistic for all of the jump tests. This is consistent with the theory that jumps are common knowledge shocks. The gist of this theory is that common knowledge drive prices up or down with little change in volume. Compared to a normal day, volume on days with common-knowledge driven jumps should have lower volume. If for example, we all learn that gas prices have gone up in the past hour in Dubai, the price of Chevron stock would go up with little change in volume. Everybody would immediately value it higher, without the need to trade; so the price immediately “jumps” up without any change in volume. In this theory, trading happens when there is information asymmetry; that is, when one trader values a stock higher than another trader, then he starts to buy because he feels that it is relatively cheaper.

One could argue that liquidity could be the cause behind a mildly negative relationship between volume and jump statistic. Stocks are considered illiquid when trading volume is light. In that case where there are extremely few trades, almost every price movement will be detected as jumps. However, this explanation does not hold for
stocks analyzed in this paper since they are the most liquid stocks in the S&P 100 and are traded every minute, if not every second.

Another possible explanation for the slightly negative relationship is related to how people act in response to seeing a jump. It is reasonable to assume that people might halt their trading activities after seeing large price movements to reevaluate their investing options. Therefore, even though this paper aims to find the economic reasoning behind the cause of jumps, it may have uncovered the effects of jumps instead.

The results from the regressions with more variables are not as pronounced as those from the simple one. Only the Jiang-Oomen test still showed the same relationship that lower volume leads to a higher chance of a jump occurring. The extra variables reduced the significance of the relationships. If it is true that all days with lower volumes are more likely to be days with jumps, then when the Monday variable is included, we could be eliminating a lot of the jump days and label them as a different regressor.

9. Conclusions

This paper is the first to explore the relationship between the jump components from different jump tests with daily trading volume. With that information, we can make economic interpretations as to what type of information shocks are detected as jumps by the jump tests. The main result is that volume and jump statistic are negatively related when a simple regression was performed. This means that the common knowledge shocks contribute more to the jump components than disagreements in price. Hence, jumps consist of mainly information shocks that are similarly interpreted by most investors.

This result could be turned the other way. Not only that we can infer that the main component of jumps comes from common knowledge shocks, we can also predict the likelihood of a jump occurring. When a day has low trading volume, there is a higher chance that jumps have occurred.

When the lagged volume terms were introduced, the relationship became less clear. The exception comes from Jiang-Oomen test, which still showed a negative relationship, and so common knowledge shocks dominate. A further investigation as to how the variables such as lagged volume terms, Monday variables, and realized variance
interact could possibly yield a satisfactory explanation as to why the relationship between jump statistic and volume disappeared when more regressors were added.

This paper also confirms that the relationship between realized variance and volume is indeed positive using high frequency data from the last 12 years. When both realized variance were added in addition to the jump statistic and lagged volume terms, and regressed against volume, the results were mainly unchanged. There are negative relationships between jumps and volume only for the Jiang-Oomen test, while all the other tests are inconclusive.

To summarize, from the regressions performed, the relationship between jumps and volume is either negative or non-existent. This adds to the current literature of detecting and understanding jump components. This paper provides support for the theory that jumps are results of common knowledge shocks.
10. Figures and Tables

Figure 1

Sample volatility signature plot (KC)

Figure 1a: Average volume for each day of the week for JNJ

Figure 1b: Average volume for each day of the week for T

Figure 1c: Average volume for each day of the week for PG

Figure 1d: Average volume for each day of the week for KD
Table 1: Effect of tick size on the signature volatility plots

<table>
<thead>
<tr>
<th></th>
<th>data interval</th>
<th>5 min</th>
<th>10 min</th>
<th>20 min</th>
<th>30 min</th>
</tr>
</thead>
<tbody>
<tr>
<td>KO</td>
<td>full length</td>
<td>-9.84%</td>
<td>-13.40%</td>
<td>-16.79%</td>
<td>-21.13%</td>
</tr>
<tr>
<td></td>
<td>Apr-May '97</td>
<td>-36.98%</td>
<td>-45.14%</td>
<td>-49.29%</td>
<td>-54.01%</td>
</tr>
<tr>
<td></td>
<td>Jun '97-Jan '01</td>
<td>-12.43%</td>
<td>-16.08%</td>
<td>-18.91%</td>
<td>-22.45%</td>
</tr>
<tr>
<td></td>
<td>Feb '01-Jan '09</td>
<td>-5.57%</td>
<td>-8.86%</td>
<td>-12.81%</td>
<td>-18.00%</td>
</tr>
<tr>
<td>PG</td>
<td>Full length</td>
<td>-4.29%</td>
<td>-7.10%</td>
<td>-9.62%</td>
<td>-12.04%</td>
</tr>
<tr>
<td></td>
<td>Apr-May '97</td>
<td>-6.91%</td>
<td>-11.96%</td>
<td>-14.24%</td>
<td>-21.37%</td>
</tr>
<tr>
<td></td>
<td>Jun '97-Jan '01</td>
<td>-9.10%</td>
<td>-13.72%</td>
<td>-18.31%</td>
<td>-21.31%</td>
</tr>
<tr>
<td></td>
<td>Feb '01-Jan '09</td>
<td>-1.86%</td>
<td>-3.79%</td>
<td>-5.38%</td>
<td>-7.50%</td>
</tr>
</tbody>
</table>

Table 2: The simple regressions: volume vs. jump statistics

<table>
<thead>
<tr>
<th></th>
<th>BNS-Stat</th>
<th>R²</th>
<th>JiangStat</th>
<th>R²</th>
<th>JacodStat</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSCO</td>
<td>-0.0241**</td>
<td>0.002</td>
<td>0.0030**</td>
<td>0.002</td>
<td>0.0055</td>
<td>0.000</td>
</tr>
<tr>
<td>GE</td>
<td>-0.0286**</td>
<td>0.003</td>
<td>-0.0111**</td>
<td>0.001</td>
<td>-0.0522**</td>
<td>0.046</td>
</tr>
<tr>
<td>IBM</td>
<td>0.0006</td>
<td>0.000</td>
<td>1.566**</td>
<td>0.000</td>
<td>-0.518**</td>
<td>0.071</td>
</tr>
<tr>
<td>JNJ</td>
<td>-0.0199**</td>
<td>0.002</td>
<td>-0.0077**</td>
<td>0.003</td>
<td>-0.048**</td>
<td>0.050</td>
</tr>
<tr>
<td>JPM</td>
<td>-0.0217**</td>
<td>0.001</td>
<td>-0.0092**</td>
<td>0.000</td>
<td>-0.075**</td>
<td>0.078</td>
</tr>
<tr>
<td>KO</td>
<td>-0.0187**</td>
<td>0.002</td>
<td>-0.0093**</td>
<td>0.039</td>
<td>-0.047**</td>
<td>0.004</td>
</tr>
<tr>
<td>MSFT</td>
<td>-0.0001</td>
<td>0.000</td>
<td>3.937</td>
<td>0.000</td>
<td>0.010**</td>
<td>0.002</td>
</tr>
<tr>
<td>PFE</td>
<td>-0.0176</td>
<td>0.001</td>
<td>-0.0020**</td>
<td>0.002</td>
<td>-0.029**</td>
<td>0.012</td>
</tr>
<tr>
<td>PG</td>
<td>-0.0053</td>
<td>0.000</td>
<td>-0.0082**</td>
<td>0.000</td>
<td>-0.030**</td>
<td>0.015</td>
</tr>
<tr>
<td>T</td>
<td>-0.0733**</td>
<td>0.011</td>
<td>-0.016**</td>
<td>0.007</td>
<td>-0.062**</td>
<td>0.043</td>
</tr>
</tbody>
</table>

-coefficient *10^-3 for Jiang Oomen
* = significant at 10%
** = significant at 5%

Table 3: Volume and Realized Variance

<table>
<thead>
<tr>
<th></th>
<th>RV</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSCO</td>
<td>-0.0690</td>
<td>0.001</td>
</tr>
<tr>
<td>GE</td>
<td>0.1058**</td>
<td>0.043</td>
</tr>
<tr>
<td>IBM</td>
<td>0.2558**</td>
<td>0.396</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.0353**</td>
<td>0.006</td>
</tr>
<tr>
<td>JPM</td>
<td>0.2314**</td>
<td>0.161</td>
</tr>
<tr>
<td>KO</td>
<td>0.0178**</td>
<td>0.001</td>
</tr>
<tr>
<td>MSFT</td>
<td>-0.0173</td>
<td>0.001</td>
</tr>
<tr>
<td>PFE</td>
<td>-0.0876**</td>
<td>0.016</td>
</tr>
<tr>
<td>PG</td>
<td>0.0487**</td>
<td>0.010</td>
</tr>
<tr>
<td>T</td>
<td>-0.0510**</td>
<td>0.004</td>
</tr>
</tbody>
</table>

*= significant at 10%
**= significant at 5%

Table 4: Summary Table is on the next page
<table>
<thead>
<tr>
<th></th>
<th>CSCO</th>
<th>GE</th>
<th>IBM</th>
<th>JNJ</th>
<th>JPM</th>
<th>KO</th>
<th>MSFT</th>
<th>PFE</th>
<th>PG</th>
<th>T</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNS-Stat</td>
<td>negative</td>
<td>negative</td>
<td>n/a</td>
<td>negative</td>
<td>n/a</td>
<td>negative</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>negative</td>
<td>5 negatives</td>
</tr>
<tr>
<td>Jacob-Stat</td>
<td>n/a</td>
<td>negative</td>
<td>negative</td>
<td>negative</td>
<td>n/a</td>
<td>negative</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>negative</td>
<td>8 negatives</td>
</tr>
<tr>
<td>JO-Stat</td>
<td>positive</td>
<td>negative</td>
<td>positive</td>
<td>negative</td>
<td>negative</td>
<td>negative</td>
<td>positive</td>
<td>negative</td>
<td>negative</td>
<td>negative</td>
<td>7 negatives, 3 positive</td>
</tr>
<tr>
<td>BNS+others</td>
<td>negative</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>positive</td>
<td>n/a</td>
<td>n/a</td>
<td>1 negative, 1 positive</td>
</tr>
<tr>
<td>Jacob+others</td>
<td>positive</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>2 negatives, 1 positive</td>
</tr>
<tr>
<td>JO+others</td>
<td>n/a</td>
<td>negative</td>
<td>negative</td>
<td>negative</td>
<td>negative</td>
<td>positive</td>
<td>negative</td>
<td>negative</td>
<td>negative</td>
<td>negative</td>
<td>8 negatives, 1 positive</td>
</tr>
<tr>
<td>BNS+RV+others</td>
<td>negative</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>1 negative</td>
</tr>
<tr>
<td>Jacob+RV+others</td>
<td>positive</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>negative</td>
<td>negative</td>
<td>3 negatives, 1 positive</td>
</tr>
<tr>
<td>JO+RV+others</td>
<td>n/a</td>
<td>negative</td>
<td>n/a</td>
<td>negative</td>
<td>n/a</td>
<td>negative</td>
<td>negative</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>6 negatives</td>
</tr>
</tbody>
</table>

|                | 3    | 5    | 2    | 5    | 6    | 5    | 0    | 5    | 6    | 5     |         |

This table is a summary of the regressions corresponding to equations 23-25 and 27-32. It tells the relationships between volume and jump statistics for each of the regressions. The first 3 rows are the simple regressions and the next three is the regression with lagged volume and the Monday dummy variable. The last three are the ones with all the variables, including the log of the realized variance. The threshold for the relationship is using 95 percent confidence.
### Appendix

#### Table 5: Regressions corresponding to equations 29-31

<table>
<thead>
<tr>
<th></th>
<th>BNS-Stat</th>
<th>vol(t-1)</th>
<th>vol(t-5)</th>
<th>vol(t-22)</th>
<th>Monday</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSCO</td>
<td>-0.0194**</td>
<td>0.0328</td>
<td>0.3595**</td>
<td>0.0739</td>
<td>-0.0233</td>
<td>0.030</td>
</tr>
<tr>
<td>GE</td>
<td>-0.0038</td>
<td>0.3474**</td>
<td>0.3455**</td>
<td>0.2415**</td>
<td>-0.0961**</td>
<td>0.659</td>
</tr>
<tr>
<td>IBM</td>
<td>0.0033</td>
<td>0.4984**</td>
<td>0.1714**</td>
<td>0.1926**</td>
<td>-0.1077</td>
<td>0.507</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.0006</td>
<td>0.4307**</td>
<td>0.2159**</td>
<td>0.2824**</td>
<td>-0.1015**</td>
<td>0.633</td>
</tr>
<tr>
<td>JPM</td>
<td>-0.001</td>
<td>0.3965**</td>
<td>0.2682**</td>
<td>0.3061**</td>
<td>-0.1099**</td>
<td>0.784</td>
</tr>
<tr>
<td>KO</td>
<td>0.0087</td>
<td>0.4639**</td>
<td>0.2184**</td>
<td>0.2184**</td>
<td>0.2557**</td>
<td>0.656</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.0023</td>
<td>0.0930**</td>
<td>0.1278</td>
<td>0.2343*</td>
<td>-0.0474*</td>
<td>0.022</td>
</tr>
<tr>
<td>PFE</td>
<td>0.0121**</td>
<td>0.3789**</td>
<td>0.2764**</td>
<td>0.3134**</td>
<td>-0.0994**</td>
<td>0.775</td>
</tr>
<tr>
<td>PG</td>
<td>0.0082</td>
<td>0.4401**</td>
<td>0.2097**</td>
<td>0.2872**</td>
<td>-0.1129**</td>
<td>0.666</td>
</tr>
<tr>
<td>T</td>
<td>0.0062</td>
<td>0.4771**</td>
<td>0.1491**</td>
<td>0.3604**</td>
<td>-0.0848**</td>
<td>0.882</td>
</tr>
<tr>
<td></td>
<td>Jaod-Stat</td>
<td>vol(t-1)</td>
<td>vol(t-5)</td>
<td>vol(t-22)</td>
<td>Monday</td>
<td>$R^2$</td>
</tr>
<tr>
<td>CSCO</td>
<td>0.0102*</td>
<td>0.0304</td>
<td>0.3711**</td>
<td>0.0883</td>
<td>-0.0258</td>
<td>0.030</td>
</tr>
<tr>
<td>GE</td>
<td>-0.0178**</td>
<td>0.3410**</td>
<td>0.3451**</td>
<td>0.2423**</td>
<td>-0.0899**</td>
<td>0.664</td>
</tr>
<tr>
<td>IBM</td>
<td>-0.0207**</td>
<td>0.4827**</td>
<td>0.1692**</td>
<td>0.1656**</td>
<td>-0.1048**</td>
<td>0.518</td>
</tr>
<tr>
<td>JNJ</td>
<td>-0.0225**</td>
<td>0.4217**</td>
<td>0.2070**</td>
<td>0.2836**</td>
<td>-0.0976**</td>
<td>0.643</td>
</tr>
<tr>
<td>JPM</td>
<td>-0.0163**</td>
<td>0.3924**</td>
<td>0.2582**</td>
<td>0.3034**</td>
<td>-0.1059**</td>
<td>0.787</td>
</tr>
<tr>
<td>KO</td>
<td>-0.0155**</td>
<td>0.4525**</td>
<td>0.2163**</td>
<td>0.2567**</td>
<td>-0.1069**</td>
<td>0.660</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.0123**</td>
<td>0.0916**</td>
<td>0.1381</td>
<td>0.2290*</td>
<td>-0.0486*</td>
<td>0.023</td>
</tr>
<tr>
<td>PFE</td>
<td>-0.0181**</td>
<td>0.3719**</td>
<td>0.2665**</td>
<td>0.3271**</td>
<td>-0.0983**</td>
<td>0.779</td>
</tr>
<tr>
<td>PG</td>
<td>-0.0164**</td>
<td>0.4293**</td>
<td>0.2102**</td>
<td>0.2033**</td>
<td>-0.1114**</td>
<td>0.661</td>
</tr>
<tr>
<td>T</td>
<td>-0.0075**</td>
<td>0.4727**</td>
<td>0.1479**</td>
<td>0.3605**</td>
<td>-0.0848**</td>
<td>0.882</td>
</tr>
</tbody>
</table>

*= significant at 10%

**=significant at 5%

-coefficient *10^-6 for Jiang Oomen
Table 6: Regressions corresponding to equations 32-34

<table>
<thead>
<tr>
<th></th>
<th>BNS-Stat</th>
<th>vol(t-1)</th>
<th>vol(t-5)</th>
<th>vol(t-22)</th>
<th>Monday</th>
<th>RV</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSCO</td>
<td>-0.0198**</td>
<td>0.0290</td>
<td>0.3719**</td>
<td>0.0995</td>
<td>-0.0268</td>
<td>-0.0239</td>
<td>0.032</td>
</tr>
<tr>
<td>GE</td>
<td>-0.0045</td>
<td>0.3422**</td>
<td>0.3233**</td>
<td>0.2561**</td>
<td>-0.0900**</td>
<td>0.0461**</td>
<td>0.667</td>
</tr>
<tr>
<td>IBM</td>
<td>-0.0005</td>
<td>0.4189**</td>
<td>0.1215**</td>
<td>0.0141</td>
<td>-0.0878**</td>
<td>0.1436**</td>
<td>0.589</td>
</tr>
<tr>
<td>JNJ</td>
<td>0.0013</td>
<td>0.4171**</td>
<td>0.2011**</td>
<td>0.3239**</td>
<td>-0.0983**</td>
<td>0.0528**</td>
<td>0.646</td>
</tr>
<tr>
<td>JPM</td>
<td>-0.0019</td>
<td>0.3796**</td>
<td>0.2462**</td>
<td>0.3064**</td>
<td>-0.0993**</td>
<td>0.0589**</td>
<td>0.792</td>
</tr>
<tr>
<td>KO</td>
<td>0.0066</td>
<td>0.4514**</td>
<td>0.2126**</td>
<td>0.2898**</td>
<td>-0.1068**</td>
<td>0.0495**</td>
<td>0.666</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.0020</td>
<td>0.0904**</td>
<td>0.1363</td>
<td>0.2446**</td>
<td>-0.0499**</td>
<td>-0.0209</td>
<td>0.023</td>
</tr>
<tr>
<td>PFE</td>
<td>0.0094*</td>
<td>0.3604**</td>
<td>0.2621**</td>
<td>0.3838**</td>
<td>-0.0920**</td>
<td>0.0759**</td>
<td>0.786</td>
</tr>
<tr>
<td>PG</td>
<td>0.0044</td>
<td>0.4280**</td>
<td>0.2053**</td>
<td>0.3098**</td>
<td>-0.1070**</td>
<td>0.0537**</td>
<td>0.668</td>
</tr>
<tr>
<td>T</td>
<td>0.0063</td>
<td>0.4652**</td>
<td>0.1406**</td>
<td>0.3891**</td>
<td>-0.0822**</td>
<td>0.0448**</td>
<td>0.884</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>vol(t-1)</th>
<th>vol(t-5)</th>
<th>vol(t-22)</th>
<th>Monday</th>
<th>RV</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSCO</td>
<td>2.35**</td>
<td>0.0289</td>
<td>0.3742**</td>
<td>0.1034</td>
<td>-0.0267</td>
<td>-0.0233</td>
</tr>
<tr>
<td>GE</td>
<td>-4.68</td>
<td>0.3418**</td>
<td>0.3229**</td>
<td>0.2571**</td>
<td>-0.0902**</td>
<td>0.0461**</td>
</tr>
<tr>
<td>IBM</td>
<td>98.8</td>
<td>0.4186**</td>
<td>0.1220**</td>
<td>0.0141</td>
<td>-0.0878**</td>
<td>0.1458**</td>
</tr>
<tr>
<td>JNJ</td>
<td>-1.12</td>
<td>0.4175**</td>
<td>0.2008**</td>
<td>0.3231**</td>
<td>-0.0985**</td>
<td>0.0528**</td>
</tr>
<tr>
<td>JPM</td>
<td>-4.69**</td>
<td>0.3798**</td>
<td>0.2469**</td>
<td>0.3058**</td>
<td>-0.0987**</td>
<td>0.0587**</td>
</tr>
<tr>
<td>KO</td>
<td>0.300</td>
<td>0.4504**</td>
<td>0.2142**</td>
<td>0.2888**</td>
<td>-0.1071**</td>
<td>0.0499**</td>
</tr>
<tr>
<td>MSFT</td>
<td>-304</td>
<td>0.0909**</td>
<td>0.1361</td>
<td>0.2424**</td>
<td>-0.0501</td>
<td>-0.0990</td>
</tr>
<tr>
<td>PFE</td>
<td>-874**</td>
<td>0.3587**</td>
<td>0.2583**</td>
<td>0.3839**</td>
<td>-0.0911**</td>
<td>0.0763**</td>
</tr>
<tr>
<td>PG</td>
<td>-5.87**</td>
<td>0.4286**</td>
<td>0.2039**</td>
<td>0.3100**</td>
<td>-0.1060**</td>
<td>0.0541**</td>
</tr>
<tr>
<td>T</td>
<td>-0.079</td>
<td>0.4647**</td>
<td>0.1403**</td>
<td>0.3889**</td>
<td>-0.0821**</td>
<td>0.0447**</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>vol(t-1)</th>
<th>vol(t-5)</th>
<th>vol(t-22)</th>
<th>Monday</th>
<th>RV</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSCO</td>
<td>0.0077</td>
<td>0.0289</td>
<td>0.3757**</td>
<td>0.1000</td>
<td>-0.0271</td>
<td>-0.0132</td>
</tr>
<tr>
<td>GE</td>
<td>-0.0097**</td>
<td>0.3398**</td>
<td>0.3241**</td>
<td>0.2513**</td>
<td>-0.0878**</td>
<td>0.0368**</td>
</tr>
<tr>
<td>IBM</td>
<td>0.0023</td>
<td>0.4186**</td>
<td>0.1220**</td>
<td>0.0141</td>
<td>-0.0878**</td>
<td>0.1458**</td>
</tr>
<tr>
<td>JNJ</td>
<td>-0.0146**</td>
<td>0.4147**</td>
<td>0.1989**</td>
<td>0.3145**</td>
<td>-0.0960**</td>
<td>0.0401**</td>
</tr>
<tr>
<td>JPM</td>
<td>-0.0054*</td>
<td>0.3798**</td>
<td>0.2452**</td>
<td>0.3054**</td>
<td>-0.0995**</td>
<td>0.0533**</td>
</tr>
<tr>
<td>KO</td>
<td>-0.0087**</td>
<td>0.4639**</td>
<td>0.2125**</td>
<td>0.2856**</td>
<td>-0.1058**</td>
<td>0.0432**</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.0103*</td>
<td>0.0908**</td>
<td>0.1400</td>
<td>0.2341**</td>
<td>-0.0494**</td>
<td>-0.0090</td>
</tr>
<tr>
<td>PFE</td>
<td>-0.0087**</td>
<td>0.3587**</td>
<td>0.2583**</td>
<td>0.3830**</td>
<td>-0.0923**</td>
<td>0.0673**</td>
</tr>
<tr>
<td>PG</td>
<td>-0.0080**</td>
<td>0.4240**</td>
<td>0.2059**</td>
<td>0.2059**</td>
<td>-0.1068**</td>
<td>0.0483**</td>
</tr>
<tr>
<td>T</td>
<td>-0.0012</td>
<td>0.4644**</td>
<td>0.1404**</td>
<td>0.3881**</td>
<td>-0.0822**</td>
<td>0.0434**</td>
</tr>
</tbody>
</table>

* = significant at 10%
** = significant at 5%
-coefficient *10^-6 for Jiang Oomen
References