Relative Contribution of Common Jumps in Realized Correlation

Kyu Won Choi

April 25, 2012

Professor Tim Bollerslev
Professor George Tauchen
Professor Arlie O. Petters

Honors thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Mathematics in Trinity College of Duke University.

Acknowledgements

I am grateful to Professors Tim Bollerslev, Arlie O. Petters, and George Tauchen for their comments, guidance and encouragement throughout the process. I am also thankful for Professor Jake Bouvrie for his great comments. I would like to thank Professor David Kraines for organizing PRUV fellowship program. Finally, I want to give credits to my peers in the Honors Senior Workshop: Vivek Bhattacharya, Hao Sun, and Grace Shuting Wei. Their constructive feedbacks were helpful in completing the final draft of my thesis.

Abstract

This paper studies common intraday jumps and relative contribution of these common jumps in realized correlation between individual stocks and market index, using high-frequency price data. In introducing stochastic models for stock price returns, we show that discrete time model (binomial tree) converges to geometric brownian motion in continuous time. We find that the common jumps significantly contribute in realized correlation at different threshold cut-offs and both common jumps and realized correlation are relatively consistent across time period including financial crisis. However, we observe a statistically significant difference in realized correlation and suggestive difference in contribution of common jumps between financial and food industry. In addition, we find a weak, positive relationship between relative contribution of common jumps and realized correlation, when we further sample high-frequency data into a year. We also observe that the volatility index and market index reveal the strongest relationship.

1 Introduction

Volatility is a central field of study in financial economics; it plays a vital role in a number of applications such as asset pricing, asset allocation and diversification, risk management, and risk forecasting. For instance, estimates of volatility are considered as inputs to the Black-Scholes option pricing model (Black and Scholes 1973), an important model in finance. Since Merton (1976) observed discontinuities in asset prices, various studies have supported the claim and discussed the importance of discontinuities or jumps that contribute toward the volatility of assets. Drost, Nijman, and Werker (1998) found that a continuous time diffusion model cannot explain the time series of dollar exchange rates clearly. Similarly, Andersen, Benzoni, and Lund (2002) noted that it is critical to include discrete jump components in any equity return continuous time model.

As several studies proposed, jumps play an essential role in financial economics. Maheu and McCurdy (2004) showed that a model incorporating jumps improves volatility forecast. Tauchen and Zhou (2011) and Lee and Mykland (2008) explained that characterizing the distribution and causes of jumps can improve asset pricing models. Then, where do these jumps come from? The most common cause of the presence of jumps is a sudden availability of new information (Merton 1976). Jumps occur in an efficient market where very significant, unanticipated, new information is instantaneously incorporated into the price.

Given the theoretical significance of jumps, it is important to be able to detect them in data. In recent year, the availability of high-frequency data has resulted in a significant improvement in the accuracy of jumps detection. Since high-frequency data enables statistical inference on discontinuous components and jump components separately, it has facilitated the research on jumps in securities. In fact, several authors have proposed nonparametric statistical tests to determine whether an individual price movement over a given time interval is likely to reflect a jump. Barndorff-Nielsen and Shephard (2004, 2006) introduced two measures of volatility that are jump robust and non-jump robust and employed these two estimates to determine whether a data series contains jumps. Ait-Sahalia and Jacod (2007) observed the difference between higher-order moments computed at two different frequencies. Lee and Mykland (2008) compared the magnitude of each price change with a sliding-window measure of local volatility to identify whether individual price changes are jumps. However, recent work suggests that the jump tests produce inconsistent results due to an intraday pattern in volatility. Thus, for this study, we employ the Bollerslev, Todorov, Li (2011) jump test that takes into account of the effect of U-shaped intraday pattern in volatility.

Further, Huang and Tauchen (2005) performed theoretical and Monte Carlo analysis of high frequency returns on the S&P cash and futures index and noted the empirical importance of jumps as a source of stock market price variance. They defined the relative jump (RJ) measure as an indicator of the distribution of jumps to within day total variance of the process. The relative jump (RJ) is the ratio of difference between jump-robust and non-jumps robust measures of variances to non-jump robust measure. As a result, $100 * RJ$ is a direct measure of the percentage contribution of jumps to total price variance (Huang and Tauchen, 2005). This study finds evidence of jumps which account for about 4.5 percent to 7.0 percent of the total daily variance of the S&P cash or futures index. This relative average contribution of 4.5 percent to 7.0 percent of total daily price variance motivates the study in a bivariate setting, where we can study common jumps between two securities, instead of jumps in one security.

Bivariate volatility modeling, where we study the interaction between two securities, is important in the area of risk management, portfolio allocation, and others, since many issues in financial market depend on covariance risk. In fact, it is possible to extend several concepts in an univariate setting to the similar concepts in a bivariate setting. First, non-jump robust estimator of daily volatility in a univariate setting is extended to the covariance estimator that includes both diffusive components and common jump components. This non-jump robust estimator for daily covariance is obtained by summing up the intraday cross products of high-frequency vectors of returns within a day. As the number of cross products between high-frequency returns in a day goes to infinity, the estimator is consistent to the daily covariance of the underlying assets (Barndorff-Nielsen, Shephard, 2004). Similarly, it is possible to expand the concept of jumps in individual security into common jumps between two securities. In bivariate setting, common jumps between two assets are classified, when jumps for two assets are identified to occur in the same direction during the same time interval. However, it is not so simple to extend jump robust estimator of daily volatility in a univariate setting to the corresponding concept in a bivariate setting. Instead, we employ a different approach to obtain common jump robust estimator of daily covariance.

Obtaining the covariation estimator and common jumps in a bivariate setting motivates to question the relative average contribution of the common jumps in total covariance between two securities. In fact, disentangling diffusive covariation and common jumps has important applications in finance such as model selection, forecasting, option pricing, and risk management. The paper therefore studies the relative contribution of common jumps by separating the two components: diffusive covariation and the common jumps. Here, the common intraday jumps are classified using the Bollerslev, Todorov and Li test (2011) and they are removed from the daily covariance estimator that contains both diffusive and common jump components. Excluding common jumps results in the covariance estimator that is consisted of diffusive components and idiosyncratic jumps only. The relative difference between them enables to study relative contribution of common jumps.

In particular, the paper studies common jumps between the proxy for the market [1] and individual securities. Employing a proxy for the market portfolio enables us to observe the common jumps that are concerned with a measurement of the overall market. It is also possible to explore the central issues such as how markets process information, since jumps are an important mechanism for incorporating systematic news into prices. The contribution of this paper is to empirically investigate the importance of common jumps in an equity and the market using an intraday nonparametric jump test. The study also briefly discusses the general differences observed between food and finance industry. In addition, the paper looks at the the relative contribution of common jumps across different sub-periods to determine whether the contributions of jumps are significantly higher during financial crisis and to observe how closely relative contribution of common jumps are related to realized correlation.

The remaining of the paper is structured as follows. In section 2, we introduce the mathematical background; in particular, we study both continuous and discrete stochastic processes. In section 3, we discuss the underlying asset pricing model, jump-robust and non-jump-robust estimators of variance, and the covariance in bivariate setting. In section 4, we examine the statistical methods employed throughout the paper. The methods include intraday jump test proposed by Bollerslev, Todorov and Li (2011), and the relative contribution of common intraday jumps detected from the test. Then, section 5 discusses the data source and the market friction called "microstructure noise" that contaminates the estimators of variance at the very high frequency data. Here, we select the appropriate sampling frequency interval that maximizes the usage of the data and minimizes the market noise. In section 5, we introduce the empirical results and carry out analysis by industry and year. Finally, section 6 ends the paper with the conclusion drawn from the study.

[1] The Standard and Poor Futures Market Index (SPFU) is used as a proxy for the market. The index will be discussed in detail later in the data section.

2 Mathematical Background

It is necessary to study mathematical theories and equations before we understand the economic models. In particular, stochastic process of probability theory is closely related to the stock price model. Variables that change over time in an unpredictable way are assumed to follow a stochastic process. The process may be discrete or continuous in time. In a discrete-time, variables change only at certain fixed point in time, while they can change at any time in continuous-time stochastic processes. Further, the process may take discrete or continuous variables. Only discrete values are possible in a discrete variable process, while variables can take any value in a certain range in continuous-variable process. We model stock price using continuous-variable, continuous-time stochastic process. We further study discrete-variable, discrete-time process with binomial tree.

2.1 Wiener Process (Brownian Motion)

Markov process is a stochastic process in which only the present value of a variable is relevant for predicting the future. The past history of the variables is irrelevant. Wiener process is a particular type of Markov stochastic process with a mean change of zero and a variance rate of 1 per year. It is sometimes referred to as a Brownian motion. Formally stated, a variable W is said to follow Wiener process (Brownian motion), if the variable satisfies two properties below.

Property 1. The change, ΔW during a small period of time Δt is

$$\Delta W = \epsilon\sqrt{\Delta t},$$

where ϵ has a standardized normal distribution $\phi(0,1)$. $\phi(m,v)$ indicates a normal distribution with mean m and variance v^2.

Property 2. For $t_0 < t_1 < \cdots < t_k$, change ΔW of any two different short intervals of time,

$$W_{t_1} - W_{t_0}, \cdots, W_{t_k} - W_{t_{k-1}}$$

are independent.

The first property suggests that ΔW has a normal distribution with mean 0, and variance Δt. The second property confirms that W follows a Markov process.

2.1.1 Generalized Wiener Process

Mean change per unit time for a stochastic process is known as the drift rate and the variance per unit time is known as the variance rate. The basic Wiener process, previously introduced ΔW has a drift rate of zero and a variance rate of 1. From here, we use dx corresponding to the process during infinitely small time interval, dt, instead of Δx and Δt. A generalized Wiener process for a variable x is defined as

$$dx = adt + bdW = adt + b\epsilon dt,$$

where a and b are constants. In form of stochastic integral, it is

$$x = x_0 + a\int_0^t ds + b\int_0^t dW.$$

The term, adt implies that x has an expected drift rate of a per unit of time. The second term, bdW can be interpereted as adding noise or variability to the path followed by x. The amount of this noise or variability is b times a Wiener process. Thus, it appears that dx has a normal distribution is with mean adt and variance b^2dt. In general, Itô Process is a Wiener process in which the parameters a and b are functions of the values of the underlying variable x and time t. Algebraically,

$$dx = a(x,t)dt + b(x,t)dW$$

in which both expected drift rate and variance rate of an Itô process depend on change over time.

2.2 Continuous-time Model: Geometric Brownian Motion

Stock price is assumed to follow a Geometric Brownian motion

$$S(t) = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W(t)},$$

which is a solution of stochastic differential equation below,

$$\frac{dS}{S} = (r - \frac{\sigma^2}{2})dt + \sigma dW.$$

$S_0 > 0$ is the initial value of stock. $(r - \frac{\sigma^2}{2})$ is a drift (with interest rate r), and σ is volatility of stock or standard deviation of stock. Finally, $W(t)$ is a Brownian motion. The variability of the stock price occurs from the randomness of the underlying Brownian motion. The stock price from this Geometric Brownian Motion is always non-negative. The model is also reasonable, using economic principles for stock prices in an ideal non-arbitrage world, in which it is not possible to make a profit with certainty by observing the past values $S(u) : 0 \leq u \leq t$ of the stock. Geometric Brownian Motion is a Markov process: the future given the present state is independent of the past. Given $S(t)$, present state at time t, $S(t+h)$, h time units after time t, is independent of $S(u) : 0 \leq u < t$, the past before time t.

$$\begin{aligned} S(t+h) &= S_0 e^{X(t+h)} \\ &= S_0 e^{X(t)+X(t+h)-X(t)} \\ &= S_0 e^{X(t)} e^{X(t+h)-X(t)} \\ &= S(t) e^{X(t+h)-X(t)} \end{aligned}$$

Thus, given $S(t)$, the future $S(t+h)$ only depends on the future increment of the Brownian Motion, $X(t+h) - X(t)$. Since increments of Brownian Motion are independent as stated in property 2, the future value is independent of the past.

2.3 Discrete-time Model: Binomial tree

This subsection introduces binomial model of stock price, which assumes discrete states and discrete time. The number of possible stock prices and time steps are both finite. We start with the set-up of binomial tree model and finally show that this binomial tree converges to the distribution of the geometric Brownian motion as the time step becomes smaller.

2.3.1 Binomial Tree Set-up

We suppose that an initial stock price is $S_0 > 0$. The fixed time horizon [0 t] is divided into large positive integer, n equal intervals with small duration Δt. At each step Δt, stock price either moves up or down by a factor of $u > 1$ or $0 \leq d < 1$ respectively. Underlying volatility is assumed to be $\sigma > 0$ and interest rate $r \geq 0$. At time Δt, the stock price, $S_{\Delta t}$ is determined as follows.

$$S_{\Delta t} = \begin{cases} S_0 u & \text{with probability p,} \\ S_0 d & \text{with probability 1-p.} \end{cases}$$

We now observe expectation and variance of the variable, $\frac{S_{\Delta t}}{S_0}$.

$$E\left(\frac{S_{\Delta t}}{S_0}\right) \equiv pu + (1-p)d = e^{r\Delta t}$$

$$E\left(\left[\frac{S_{\Delta t}}{S_0}\right]^2\right) \equiv pu^2 + (1-p)d^2$$

$$\begin{aligned} Var\left(\frac{S_{\Delta t}}{S_0}\right) &\equiv E\left(\left[\frac{S_{\Delta t}}{S_0}\right]^2\right) - \left[E\left(\frac{S_{\Delta t}}{S_0}\right)\right]^2 \\ &= pu^2 + (1-p)d^2 - e^{2r\Delta t} = \sigma^2\Delta t \end{aligned}$$

Rearranging the terms in expectation of $\frac{S_{\Delta t}}{S_0}$, we find

$$p = \frac{e^{r\Delta t} - d}{u-d}, \qquad\qquad 1-p = \frac{u - e^{r\Delta t}}{u-d}.$$

Replacing p and $(1-p)$ with above values into the variance of $\frac{S_{\Delta t}}{S_0}$ results in the equation below.

$$\begin{aligned} \sigma^2\Delta t &= \left(\frac{e^{r\Delta t} - d}{u-d}\right)u^2 + \left(\frac{u - e^{r\Delta t}}{u-d}\right)d^2 - e^{2r\Delta t} \\ &= \frac{e^{r\Delta t}(u^2-d^2) - du^2 + ud^2}{u-d} - e^{2r\Delta t} \\ &= \frac{(u-d)[e^{r\Delta t}(u+d) - ud]}{u-d} - e^{2r\Delta t} \\ &= e^{r\Delta t}(u+d) - ud - e^{2r\Delta t}. \end{aligned}$$

In order to find u and d, one more assumption is necessary. We employ the general binomial tree, Cox, Ross, Rubinstein (1979) model, which assumes that $ud = 1$. Replacing ud with 1 and d with $\frac{1}{u}$ yields

$$\sigma^2\Delta t = e^{r\Delta t}\left(u + \frac{1}{u}\right) - 1 - e^{2r\Delta t}.$$

For further derivation, we employ perturbation analysis of $1+\sigma^2\Delta t \approx e^{\sigma^2\Delta t}$, since $\Delta t \approx 0$.[2]

$$ue^{\sigma^2\Delta t} = e^{r\Delta t}(u^2+1) - ue^{2r\Delta t}$$

$$u^2 - (e^{r\Delta t} + e^{\sigma^2\Delta t - r\Delta t})u + 1 = 0$$

$$u^2 - (1 + r\Delta t + 1 + \sigma^2\Delta - r\Delta t)u + 1 = 0$$

$$u^2 - (1 + e^{\sigma^2\Delta t})u + 1 = 0$$

We solve the quadratic equation of u.

$$\begin{aligned} u &= \frac{1}{2}\left((1+e^{\sigma^2\Delta t}) \pm \sqrt{(1+e^{\sigma^2\Delta t})^2 - 4}\right) \\ &\approx \frac{1}{2}\left((1+e^{\sigma^2\Delta t}) \pm \sqrt{\left(1+1+\sigma^2\Delta t\right)^2 - 4}\right) \\ &\approx \frac{1}{2}\left((1+e^{\sigma^2\Delta t}) \pm \sqrt{4\sigma^2\Delta t + (\sigma^2\Delta t)^2}\right) \\ &\approx \frac{1}{2}(1+1+\sigma^2\Delta t \pm 2\sigma\sqrt{\Delta t}) \\ &= 1 \pm \sigma\sqrt{\Delta t} + \frac{\sigma^2\Delta t}{2} \approx e^{\pm\sigma\sqrt{\Delta t}} \end{aligned}$$

Since $u > 1$, we have $u = e^{\sigma\sqrt{\Delta t}}$ and $d = e^{-\sigma\sqrt{\Delta t}}$.

During n steps, let U be the total number of times that stock price goes up, and D be the total number of times that stock price goes down. Then, $n = U + D$. Further, suppose that $X_1, X_2, \dots X_n$ are independent identically distributed random variables whose probability of $X_i = 1$ is p and $X_i = -1$ is $1-p$. By defining $M_n \equiv \sum_{i=1}^{i=n} X_i$, we have $M_n = U - D$. Thus,

$$U = \frac{1}{2}(n + M_n), \qquad \text{and} \qquad D = \frac{1}{2}(n - M_n).$$

[2]Note that the taylor series expansion of exponential function is

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots && \text{for all } x \\ &= 1 + x + O(x^2) \quad \approx 1 + x && \text{for } x \to 0. \end{aligned}$$

2.3.2 Convergence to Geometric Brownian Motion

With the expressions derived in previous subsection, we finally have stock price at the time t as

$$\begin{aligned} S_t &= S_0 u^U d^D \\ &= S_0 u^{\frac{1}{2}(n+M_n)} d^{\frac{1}{2}(n-M_n)} \\ &= S_0 e^{\frac{\sigma\sqrt{\Delta t}}{2}(n+M_n)} e^{\frac{-\sigma\sqrt{\Delta t}}{2}(n-M_n)} \\ &= S_0 e^{\sigma\sqrt{\Delta t} M_n} \end{aligned}$$

The goal is to show that the stock price at time t, $S_0 e^{\sigma\sqrt{\Delta t} M_n}$ under the n-step binomial model converges to the Geometric Brownian motion, $S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W(t)}$ at time t, as $n \to \infty$, that is $\Delta t \to 0$.

$$\begin{aligned} E\left(e^{u\sigma\sqrt{\Delta t} M_n}\right) &= E\left(\prod_{i=1}^{n} e^{u\sigma\sqrt{\Delta t} X_i}\right) = \prod_{i=1}^{n} E\left(e^{u\sigma\sqrt{\Delta t} X_i}\right) \\ &= \prod_{i=1}^{n} (e^{u\sigma\sqrt{\Delta t}} p + e^{-u\sigma\sqrt{\Delta t}}(1-p)) = (e^{u\sigma\sqrt{\Delta t}} p + e^{-u\sigma\sqrt{\Delta t}}(1-p))^n \\ &= \left[e^{u\sigma\sqrt{\Delta t}}\left(\frac{e^{r\Delta t}-d}{u-d}\right) + e^{-u\sigma\sqrt{\Delta t}}\left(\frac{u-e^{r\Delta t}}{u-d}\right)\right]^n \\ &= \left[e^{u\sigma\sqrt{\Delta t}}\left(\frac{e^{r\Delta t}-e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}}-e^{-\sigma\sqrt{\Delta t}}}\right) + e^{-u\sigma\sqrt{\Delta t}}\left(\frac{e^{\sigma\sqrt{\Delta t}}-e^{r\Delta t}}{e^{\sigma\sqrt{\Delta t}}-e^{-\sigma\sqrt{\Delta t}}}\right)\right]^n \end{aligned}$$

Taking logarithm on both sides of equation lead to,

$$\begin{aligned} \log\left[E\left(e^{u\sigma\sqrt{\Delta t} M_n}\right)\right] &= n\log\left[e^{u\sigma\sqrt{\Delta t}}\left(\frac{e^{r\Delta t}-e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}}-e^{-\sigma\sqrt{\Delta t}}}\right) + e^{-u\sigma\sqrt{\Delta t}}\left(\frac{e^{\sigma\sqrt{\Delta t}}-e^{r\Delta t}}{e^{\sigma\sqrt{\Delta t}}-e^{-\sigma\sqrt{\Delta t}}}\right)\right] \\ &= \frac{t}{\Delta t}\log\left[\frac{e^{u\sigma\sqrt{\Delta t}}(e^{r\Delta t}-e^{-\sigma\sqrt{\Delta t}}) + e^{-u\sigma\sqrt{\Delta t}}(e^{\sigma\sqrt{\Delta t}}-e^{r\Delta t})}{e^{\sigma\sqrt{\Delta t}}-e^{-\sigma\sqrt{\Delta t}}}\right] \\ &= \frac{t}{\Delta t}\log\left[\frac{e^{u\sigma\sqrt{\Delta t}}(e^{r\Delta t}-e^{-\sigma\sqrt{\Delta t}}) + e^{-u\sigma\sqrt{\Delta t}}(e^{\sigma\sqrt{\Delta t}}-e^{r\Delta t})}{2\sinh(\sigma\sqrt{\Delta t})}\right] \\ &= \frac{t}{\Delta t}\log\left[\frac{e^{r\Delta t}(e^{u\sigma\sqrt{\Delta t}}-e^{-u\sigma\sqrt{\Delta t}}) - (e^{u\sigma\sqrt{\Delta t}-\sigma\sqrt{\Delta t}}-e^{-u\sigma\sqrt{\Delta t}+\sigma\sqrt{\Delta t}})}{2\sinh(\sigma\sqrt{\Delta t})}\right] \\ &= \frac{t}{\Delta t}\log\left[\frac{e^{r\Delta t}\sinh(u\sigma\sqrt{\Delta t}) - \sinh(u\sigma\sqrt{\Delta t}-\sigma\sqrt{\Delta t})}{\sinh(\sigma\sqrt{\Delta t})}\right]. \end{aligned}$$

Note that we have used the fact that $\sinh(x) = \frac{e^x - e^{-x}}{2}$. We can further proceed it using $\sinh(x-y) = \sinh(x)\cosh(y) - \cosh(x)\sinh(y)$.

$$\begin{aligned}
\log\left[E\left(e^{u\sigma\sqrt{\Delta t}M_n}\right)\right] &= \frac{t}{\Delta t}\log\left[\frac{e^{r\Delta t}\sinh(u\sigma\sqrt{\Delta t}) - \sinh(u\sigma\sqrt{\Delta t})\cosh(\sigma\sqrt{\Delta t}) + \cosh(u\sigma\sqrt{\Delta t})\sinh(\sigma\sqrt{\Delta t})}{\sinh(\sigma\sqrt{\Delta t})}\right] \\
&= \frac{t}{\Delta t}\log\left[\cosh(u\sigma\sqrt{\Delta t}) + \frac{\sinh(u\sigma\sqrt{\Delta t})(e^{r\Delta t} - \cosh(\sigma\sqrt{\Delta t}))}{\sinh(\sigma\sqrt{\Delta t})}\right]
\end{aligned}$$

Finally, we employ the perturbation technique that we have previously used. [3] Therefore, as $\Delta t \to 0$,

$$\begin{aligned}
\log\left[E\left(e^{u\sigma\sqrt{\Delta t}M_n}\right)\right] &= \frac{t}{\Delta t}\log\left[1 + \frac{1}{2}u^2\sigma^2\Delta t + \frac{u\sigma\sqrt{\Delta t}(1 + r\Delta t - 1 - \frac{1}{2}\sigma^2\Delta t)}{\sigma\sqrt{\Delta t}} + O((\Delta t)^2)\right] \\
&= \frac{t}{\Delta t}\log\left[1 + \frac{1}{2}u^2\sigma^2\Delta t + ru\Delta t - \frac{1}{2}\sigma^2 u\Delta t + O((\Delta t)^2)\right] \\
&\approx \frac{t}{\Delta t}\left[\frac{1}{2}u^2\sigma^2\Delta t + ru\Delta t - \frac{1}{2}\sigma^2 u\Delta t + O((\Delta t)^2)\right] \\
&\approx \frac{1}{2}u^2\sigma^2 t + rut - \frac{1}{2}\sigma^2 ut + O(\Delta t) \\
&\approx \frac{t\sigma^2u^2}{2} + \left(r - \frac{\sigma^2}{2}\right)tu
\end{aligned}$$

The result shows the function that generates the moment for a normal random variable with mean $\left(r - \frac{\sigma^2 u}{2}\right)t$ and variance $\sigma^2 ut$. Therefore, we have shown that distribution of $e^{\sigma\sqrt{\Delta t}M_n}$ converges to a distribution of geometric Brownian motion $e^{(r-\frac{\sigma^2}{2})t+\sigma W(t)}$ at time t.

[3] Note that the taylor series expansions of $\sinh(x)$ and $\cosh(x)$ are

$$\begin{aligned}
\sinh(x) &= \sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \qquad \text{for all } x \\
&= x + O(x^3) \quad \approx x \qquad \text{for } x \to 0.
\end{aligned}$$

$$\begin{aligned}
\cosh(x) &= \sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \qquad \text{for all } x \\
&= 1 + \frac{1}{2}x^2 + O(x^3) \quad \approx 1 + \frac{1}{2}x^2 \qquad \text{for } x \to 0.
\end{aligned}$$

3 Theoretical Background

3.1 Stochastic Models for Returns

The section begins with the theoretical process that describes an asset price. The following differential equation for logarithmic price $p(t)$ is a standard stochastic model for asset price movements

$$dp(t) = \mu(t)dt + \sigma(t)dw(t), \tag{1}$$

where $\mu(t)$ corresponds to the time-varying drift component of prices and $\sigma(t)dw(t)$ corresponds to the time-varying volatility components. $\sigma(t)$ is the volatility level and $w(t)$ indicates a standard Brownian motion. The model assumes that all sample path of logarithmic stock prices are continuous in nature and follows a diffusion motion with mean $\mu(t)$ and standard deviation $\sigma(t)$. This standard Brownian motion model enables us to study the Black-Scholes equations and derive the value of call and put options easily.

Although the continuous-time model has some advantages, Merton (1976) pointed out that this model is inconsistent with discontinuities that have been observed in various asset classes. The modern literature on stock price movement further suggests that discontinuities in prices are an essential process in asset pricing, portfolio management, risk management and others. Drost, Nijman, and Werker (1998) discussed that a continuous time diffusion model alone was not enough to clearly explain the time series of dollar exchange rates. Liu et al (2003) explained that including jump components into the price movement results in different optimal portfolio strategies. These discontinuities or "jumps" are observed, when the new, unanticipated information becomes available and is instantaneously reflected into the price. Incorporating the discontinuities components of the price into the model therefore results in the continuous-time jump diffusion model that includes both continuous and jump processes

$$dp(t) = \mu(t)dt + \sigma(t)dw(t) + k(t)dq(t). \tag{2}$$

Here, $k(t)$ indicates the magnitude of jump at time t and $q(t)$ is a Possion counting process. $dq(t)$ is a binary indicator of jumps in the infinitesimal interval dt. In order to separate these jump components from continuous variations, researchers have developed various statistical estimators.

3.2 Estimators of Variance

The price model in equation (2) indicates that the volatility of price results from two processes: diffusive volatility process, $\sigma(t)$ and jumps process, $k(t)$. The quadratic variation (QV) is a measure that aggregates both components

$$QV_t \equiv \int_{t-1}^{t} \sigma^2(\tau)d\tau + \sum_{j=1}^{M} k^2(t_j), \tag{3}$$

while the integrated variance (IV) measures diffusive volatility component only

$$IV_t \equiv \int_{t-1}^{t} \sigma^2(\tau)d\tau. \tag{4}$$

Since only the prices are observed discretely, it is not possible to obtain the QV and IV in equation (3) and (4). Instead, researchers have tried to approximate the measures. Barndorff-Nielsen and Shephard (2004, 2006) proposed two estimators of variance. The first estimator is the realized variance (RV)

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2, \tag{5}$$

where $r_{t,j}$ is the intraday geometric return between the $(j-1)^{th}$ and j^{th} price observations in a day, defined as $r_{t,j} \equiv p_{t,j} - p_{t,j-1}$ and M is the total number of price returns in one day. Summing up these intraday squared returns results in RV_t. As the number of observations or sampling frequency approaches infinity, the realized variance converges to the quadratic variance in equation (3). Intuitively, each squared intraday geometric return is an estimator of the true variance of the particular interval. Summing up all the squared geometric returns for infinitely small intervals therefore results in an asymptotically consistent estimator of the daily quadratic variance.

The second measure is a "jump-robust" estimator called bipower variation (BV). For a day t, BV is defined as

$$BV_t = \mu_1^{-2} \frac{M}{(M-1)} \sum_{j=2}^{M} |r_{t,j-1}||r_{t,j}|, \tag{6}$$

where $\mu_1^{-2} = \frac{\pi}{2}$. As the time interval between observations approaches zero, the bipower variance converges to the integrated variance in equation (4). Intuitively, when the magnitude of the price return $r_{t,j}$ is considerably larger than other returns in the same day, it is likely to be a jump. The bipower estimator reduces the effect of large price change by multiplying the price return with a neighboring return value that tends to be smaller. Thus, the bipower variation is an asymptotically consistent estimator of the integrated variance.

Based on the realized variance and bipower variance estimators, it is possible to estimate the volatility due to jump process. The relative contribution of jumps for a day t (RJ_t) is therefore defined as

$$RJ_t \equiv \frac{RV_t - BV_t}{RV_t}. \tag{7}$$

In words, the relative contribution of jump is the ratio of difference between realized variation and bipower variance to realized variation. It is an asymptotically consistent estimator for $\frac{QV_t - IV_t}{QV_t}$, which indicates a ratio of jumps to total within day variance of the process. As Huang and Tauchen (2005) pointed out, $100 * RJ$ is a direct measure of the percentage contribution of jumps to total price variance.

3.3 Realized Correlation

Assets in the financial market are not independent of each other; their returns may move together in the same direction or in the opposite direction. The multivariate volatility modeling is particularly important in pricing financial instruments, risk management, portfolio allocations, managerial decision making and many others. In particular, the covariance risk is important in studying interactions between assets. The concept of realized variance is extended to as realized covariance ($RCov$) in bivariate setting. The realized covariance for a day t of assets A and B is obtained in a similar manner to the realized variance which is resulted from summing over the squared intraday returns in equation (5). Instead of squared intraday returns on one security, the cross products of intraday returns for two assets A and B are summed over within a day t

$$RCov_t^{AB} = \sum_{j=1}^{M} r_{t,j}^A r_{t,j}^B. \tag{8}$$

As previously introduced, $r_{t,j}^k$ is an intraday geometric return for security k defined as $r_{t,j}^k \equiv p_{t,j}^k - p_{t,j-1}^k$ and M is the total number of price returns in one day. Intuitively, each cross product of intraday geometric returns is an estimator of the true covariance between two securities of the particular interval. The realized covariance converges to quadratic covariation that aggregate both integrated covariation and co-jumps as the sampling frequency goes to infinity. It is a consistent estimator of the daily quadratic covariation containing the co-jumps.

Then, the realized correlation ($RCorr_t$) for stock A and stock B naturally follows as:

$$RCorr_t^{AB} = \frac{RCov_t^{AB}}{\sqrt{RV_t^A RV_t^B}}. \tag{9}$$

Realized correlation is the ratio of realized covariances and product of realized standard deviations. The realized correlation is a Pearson correlation, which is only sensitive to linear relations between the returns of two assets. As Barndorff-Nielsen and Shephard (2004) pointed out, the correlation in equation (9) is susceptible to jumps in underlying prices.

4 Statistical Methods

4.1 Bollerslev, Todorov, Li Jump Test (2011)

Researchers have proposed various nonparametric statistical tests that determine whether an individual price change in a particular interval is a jump or a diffusive movement. For example, test developed by Lee and Mykland (2008) compares the magnitude of each price change with a measure of local volatility based on sliding window of returns. Ait-Sahalia and Jacod (2007) observed the difference between higher-order moments computed at two different frequencies to identify jumps. However, recent work suggests that the jump tests produce inconsistent results. Schwert (2009) found that different tests identify different days with jumps. He also found that tests are even inconsistent with themselves, detecting different jumps depending on the sampling frequency. In particular, Van Tassel (2008) showed that the test proposed by Lee and Mykland (2008) exaggerates the number of statistically significant jumps in the early morning and understates jumps in the middle of the day, yielding incoherent results. Rognlie (2010) discussed that the inconsistencies can be explained by dramatic intraday changes in volatility. In fact, he proved that a significant number of detected jumps occurs due to this behavior.

The intraday pattern in volatility have widely been studied in the literature. Wood, McInish, and Ord (1985) introduced U-shaped intraday volatility pattern and Andersen and Bollerslev (1997) confirmed this pattern. The intraday volatility is generally the greatest at the beginning of the trading day and declines until a minimum is reached in the early afternoon. The volatility rises again during the rest of the day. In fact, it is found that the average volatility at the peak is twice or more than the minimum average volatility. In addition to the work by Van Tassel (2008), Rognlie (2010) also suggested that these dramatic changes in volatility pattern throughout the day are directly responsible for detecting biased number of jumps. The U-shaped volatility pattern distorts the classification of jumps, since the same price movement may or may not be identified as a jump, depending on the local volatility. The price movement is more likely to be detected as a jump in the morning and in the end of the trading day, when the volatility is high. Thus, the intraday volatility exaggerates the number of statistically significant jumps during the beginning and the end of the day. However, the same price movement is less likely to be identified as a jump in the middle of the day, when the volatility is low. Thus, it underrepresents jumps in the middle of the day. The typical U-shaped volatility pattern is illustrated in Figure 1 computed based on the data which will be introduced later in the section.

Therefore, it is essential to consider the effect of intraday volatility in evaluating jump tests. To obtain correct results, Rognlie (2010) proposed that the morning should be adjusted by a higher volatility measure than returns in the afternoon. Therefore, the jump test by Bollerslev, Todorov, Li (2011) is a reliable test that takes into account of this U-shaped intraday pattern in volatility. The test estimates Time-of-Day (TOD) volatility pattern for each asset, based on the daily variation measures of RV and BV introduced in equation (5) and (6)

$$TOD_i = \frac{n \sum_{t=1}^{T} r_{i_t,n}^2 \mathbb{I}(|r_{i_t,n}| \le \tau \sqrt{BV_t \wedge RV_t} n^{-0.49})}{\sum_{s=1}^{nT} r_{s,n}^2 \mathbb{I}(|r_{s,n}| \le \tau \sqrt{BV_{s/n} \wedge RV_{s/n}} n^{-0.49})}. \qquad i_t = (t-1)n + i, \qquad i = 1, ..., n \tag{10}$$

$\mathbb{I}$ is the binary indicator function that equals 1, when the absolute geometric intraday return is less than τ standard deviation of a local estimator and 0, otherwise. TOD_i measures the ratio of the diffusive variation over different parts of the day relative to its average value for the day. The truncation level introduced in below equation takes the TOD volatility pattern into account

$$\alpha_i = \tau \sqrt{(BV_{i/n} \wedge RV_{i/n}) * TOD_{i-[i/n]n}}, \qquad i = 1, .., nT \tag{11}$$

and it is used to separate the realized jumps from diffusive price movement. Finally, the high-frequency intraday geometric return for each stock is identified as a realized jump, if

$$|r_{i,n}| \ge \alpha_i n^{-0.49}. \tag{12}$$

In words, the particular interval are said to contain a realized jump, if the magnitude of the intraday geometric return previously defined as $r_{t,j} \equiv p_{t,j} - p_{t,j-1}$ is greater than or equal to the adjusted TOD_i threshold.

4.2 Intervals with Co-Jumps

Prior to study common jumps and realized correlation simultaneously, this subsection elaborates on intraday time intervals that are detected to contain jumps. In order to determine whether each intraday time interval contains a jump, we first implement the Bollerslev, Todorov and Li jump test (2011) in previous subsection. The jump test identifies the time intervals that are determined to include a realized jump, if the magnitude of the intraday geometric return for the particular time interval is greater than or equal to the adjusted TOD_i threshold as discussed in equation (10). After classifying the intraday time intervals with realized jumps for individual security, these detected time intervals of individual asset are compared with the corresponding time intervals of proxy for market index. If the particular time interval of individual stock is identified to contain a realized jump and the corresponding time interval of market index is also detected to contain a realized jump occurring in the same direction, that particular time interval is considered to be the "interval with co-jump". However, one exception remains for market volatility index; [4] the volatility index is the market's expectation of the volatility that moves in the opposite direction with the market index. Since greater volatility index indicates a higher expectation of volatility and the fear in the market, it leads to the decline in stock prices and market index. Thus, "intervals with co-jumps" between market index and volatility index are defined, when two realized jumps for the same intraday time interval are classified to occur in the opposite directions.

4.3 Relative Contribution of Co-Jumps in Realized Correlation

This subsection finally combines realized correlation and intervals with co-jumps that have previously been studied. In order to observe the effect of co-jumps in the realized correlation, we would like to observe the realized correlation that excludes the intervals with co-jumps. The first step goes back to the Bollerslev, Todorov, and Li jump test which identifies intervals that include intraday realized jumps of asset A and B. For these identified intervals with co-jumps, the cross products between intraday returns, $r^A_{t,j}$ and $r^B_{t,j}$ in equation(8) are set to zero. However, the cross products of intraday returns for any other intraday time intervals remain the same as before. The below summarizes the first step

$$\widehat{r^A_{t,j} r^B_{t,j}} = \begin{cases} 0 & \text{if interval } [\frac{j-1}{M}, \frac{j}{M}] \text{ detected to contain a co-jump,} \\ r^A_{t,j} r^B_{t,j} & \text{otherwise.} \end{cases} \tag{13}$$

[4]The Standard and Poor Futures Volatility Index (VIX) is used as a market volatility index.

Using the cross products of returns in equation (13), it is then possible to obtain a new realized covariance measure that excludes intervals with co-jumps. This new realized covariance estimate would be different from the previous realized covariance in equation (8), if the day contains at least one interval with co-jump. The realized covariance that excludes intervals with co-jumps is denoted as $\widehat{RCov_t^{AB}}$ and is obtained in following way

$$\widehat{RCov_t^{AB}} = \sum_{j=1}^{M} \widehat{r_{t,j}^A r_{t,j}^B}. \tag{14}$$

Using the new realized covariance results in a modified realized correlation denoted as $\widehat{RCorr_t^{AB}}$ that also excludes intervals with co-jumps

$$\widehat{RCorr_t^{AB}} = \frac{\widehat{RCov_t^{AB}}}{\sqrt{RV_t^A RV_t^B}}. \tag{15}$$

The asymptotic difference between realized correlation measure in equation (9) and the modified realized correlation measure without co-jumps in equation (15) indicates the co-jump components. The ratio of the asymptotic difference and $RCorr_t^{AB}$ therefore represents the relative contribution of co-jumps in realized correlation ($RCCJ$) of two stocks

$$RCCJ_t^{AB} = \frac{RCorr_t^{AB} - \widehat{RCorr_t^{AB}}}{RCorr_t^{AB}}. \tag{16}$$

In fact, denominators of two realized correlations remain the same: the denominators are simply the product of realized standard deviation of each asset. Thus, relative contribution of co-jumps in realized correlation ($RCCJ$) of two stocks is the same as

$$RCCJ_t^{AB} = \frac{RCov_t^{AB} - \widehat{RCov_t^{AB}}}{RCov_t^{AB}}. \tag{17}$$

The equation (17) helps better understand the arithmetic behind the idea. If the intraday time interval is classified to contain a common jump that occurs in the same direction, the product of two returns is always greater than zero. However, equation (13) sets the product of two returns as zero for $\widehat{r_{t,j}^A r_{t,j}^B}$, when the time interval is detected to contain a common jump. Therefore, it follows that

$$r_{t,j}^A r_{t,j}^B \geq \widehat{r_{t,j}^A r_{t,j}^B}, \tag{18}$$

$$RCov_t^{AB} \geq \widehat{RCov_t^{AB}}.$$

The numerator of equation (17) is always non-negative and depending on the realized covariance between two securities - whether covariance between them is positive or negative - the sign of $RCCJ_t^{AB}$ is determined. However, it is important to note one exception we defined in section 3.2: the common jump between volatility index and market index is defined to occur in the opposite direction. That is, the cross product of two returns is always less than zero for the interval with co-jumps. Thus, the inequality signs in equation (18) should be reversed.

5 Data

5.1 Background

We used a set of minute-by-minute price data for 8 commonly traded stocks (BAC, C, CPB, HNZ, JPM, PEP, SLE, WPC) along with the S&P500 Futures market index (SPFU) and the S&P500 Futures volatility index (VIX) for this study. Although no particular quantitative criteria was used to divide stocks into different industry, we regard CPB, HNZ, PEP, SLE stocks to be part of food industry and BAC, C, JPM, WFC to be part of finance industry. The price data is obtained from the commercial data vendor called 'price-data.com'. The data for individuals stocks and the S&P500 Futures market index begins from January, 1998, ends on December, 2010 and includes 3219 trading days. However, the S&P500 Futures volatility index is only available from January, 2004 to December, 2008 that corresponds to 1169 trading days. Each trading day includes every minute price data from 9:30AM to 4:00PM. However, the first five minutes price data and overnight returns are excluded, since the price data during these periods behave differently from other intraday price data. Prior to the study, the data is adjusted for stock splits based on Yahoo! Finance and other abnormalities such as missing price values are corrected. Further, the data is aligned so that all the series include the same set of data. Table 1 shows the list of all the stock indexes, tickers, date ranges and number of trading days.

For the purpose of this study, the S&P500 Futures market index (SPFU) is used as a proxy for the market. The Standard and Poor's 500 index (S&P 500) contains a number of large companies, so the movement in the direction of the S&P Futures can serve as an indicator of overall short-term market direction. If the S&P Futures are up, it is an indication that there is upward pressure on the market and the stock market will rise. On the other hand, if the S&P Futures are down, it is a sign that there is downward pressure on the market and it will likely fall. Therefore, the S&P Futures index is a good proxy to market price movements, as it effectively trades the basket of 500 stocks in the S&P500 index in a single bundle. Similarly, the S&P500 Futures volatility index (VIX) is used as an indication of the market volatility. The index reflects market's expectation of the volatility.

5.2 Microstructure Noise

It is important to choose an appropriate sampling interval of the data, since the sampling frequency affects the estimators of variance introduced in section 2.2. Although asymptotic consistency of the estimators suggest that the prices should be sampled as frequently as possible, the estimators are contaminated with "microstructure noise" or market frictions at very high frequency. The microstructure noise is a short-term deviation between the market efficient price and observed price, which usually occurs due to bid-ask spread and discretization error. To determine the appropriate sampling interval that maintains a balance between maximizing the use of available information and limiting the impact of microstructure noise, Andersen, Bollerslev, Diebold and Labys (2000) proposed a "volatility signature plot". The graph plots average realized variance against sampling intervals. Figure 2 represents volatility signature plot for S&P500 Futures market index. As shown, microstructure noise contributes to the high average realized variance at higher frequencies (1-, 2-minutes). The average realized variance starts to flatten out, approaching a constant value for longer intervals (10-minutes). The constant value at lower frequencies (longer intervals) indicates that disregarding any more information would no longer compensate for microstructure noise. For the purpose of this study, a 5-minute time interval is selected as an appropriate trade off between the use of additional information and the microstructure noise.

6 Results and Analysis

6.1 Empirical Results

The daily realized correlation between the first eight equities, the S&P500 Futures volatility index (VIX) in Table 1 and the S&P500 Futures market index (SPFU) is calculated, and Table 2 reports the mean daily realized correlation over the entire sample period. In order to better understand what the correlation numbers indicate and how they vary, we plot the daily realized correlation over the sample period in Figure 3. As expected, the eight individual equities reveal the positive correlation with the S&P500 Futures market index. The prices of individual stocks move in the same direction as the overall market. However, the volatility index (VIX) at the right bottom indicates the negative correlation, since VIX is the market's expectation of the volatility that moves in the opposite direction with the market index. The range of mean realized correlation of eight equities is between 0.30 to 0.50. It is noteworthy to recognize that correlation between stocks from food industry (CPB, HNZ, PEP, SLE) and the market index is around 0.33, while the stocks from finance industry (BAC, C, JPM, WFC) have correlation value around 0.47. The finance industry has a stronger correlation with market. The volatility index reveals the strong negative correlation, -0.70 with the S&P500 Futures market index, because the volatility index implies the fear in the market, therefore being closely related to S&P500 Futures market index than other individual stocks.

Next, we observe the results obtained from intraday intervals that are classified to contain common jumps. Since there are 76 number of 5-minute intervals per day and the data spans for 3219 days, the total $244,644$ number of 5-minute intervals are present for the study between individual stocks and the market. The Bollerslev, Todorov and Li jump test (2011) has been implemented at $\tau = 2.0$ and $\tau = 2.5$. That is, the time interval is considered to contain a realized jump, if the magnitude of the intraday geometric return is greater than or equal to 2.0 or 2.5 standard deviations of adjusted threshold respectively. Table 4 shows the number of intervals that are classified to contain realized co-jumps with S&P500 Futures market index. For each stock, there are about 2100 to 3000 intervals with co-jumps at $\tau = 2.0$, and about 700 to 1050 at $\tau = 2.5$. The larger the value of τ is, the less the number of intervals is identified to include realized jumps. Accordingly, the number of intervals with co-jumps decreases for greater τ. Once again, it is interesting to note that stocks from food industry (CPB, HNZ, PEP, SLE) have around 2300 intervals with co-jumps when $\tau = 2.0$, while equities in finance industry (BAC, C, JPM, WFC) have about 2900 number of intervals with co-jumps for the same τ value. Similar results hold for $\tau = 2.5$. Further, it is critical to note that the total number of intervals for the volatility index is only $88,844$ due to shorter data length that only spans 1169 trading days. However, the number of intervals detected to include co-jumps for the volatility index is still similar to the number intervals with co-jumps for other individual stocks. Therefore, the volatility index has the highest percentage of intervals with co-jumps with the market index compared to the percentage of intervals for other individual stocks.

Finally, we incorporate the concept of intervals with common jumps into the realized correlation. This practice results in the relative contribution of common jumps in realized correlations as represented in Table 5. Figure 4 plots the daily relative contribution common jumps at adjusted threshold $\tau = 2.0$. Table 5 suggests that the intervals with common jumps contribute about 17 percent to 19 percent of realized correlation when $\tau = 2.0$, while intervals with common jumps account for about 7 percent to 8 percent of realized correlation when $\tau = 2.5$. One half of the increase in τ, a standard deviation of adjusted threshold in identifying the intraday realized jumps, decreases the percentage in the relative contribution of co-jumps into more than half. Relative contribution of common jumps seem to be fairly consistent across eight individual equities. However, intervals that are classified to contain common jumps between the volatility index and S&P500 Futures market index account for large percentage in realized correlation. The co-jumps contribute to about 25 percent and 13 percent of realized correlation for $\tau = 2.0$ and $\tau = 2.5$ respectively.

6.2 Analysis by Industry

The previous section suggests that stocks from finance industry (BAC, C, JPM, WFC) have a higher association with the overall market than stocks from food industry (CPB, HNZ, PEP, SLE). Also, finance related stocks appear to have a greater number of intervals that include common jumps with the market than food related stocks. The first suggestion is based on mean realized correlation of stocks over thirteen years period in Table 2; four mean values of finance related stocks are compared with four mean values of food related stocks. Similarly, the second suggestion comes from the total number of common jumps in Table 4; four contribution values of each industry are compared with each other. We now sub-sample the price data into thirteen time periods, each corresponding to one year and calculate mean realized correlation and mean relative contribution of common jumps of individual stock for each year. This results in thirteen mean values for each stock and fifty-two mean values for each industry. With the numbers reported in Table 7 and Table 8, we implement the statistical analysis that compares two industries.

As a simple two-group comparison statistical test, we use the 't-test' based on the t-distribution. Figure 5 summarizes the test results. Assuming equal variances, the mean of fifty-two realized correlation values of finance industry is 0.4725, while the mean of correlation values of food industry is 0.3389. Under the null hypothesis that there is no difference in the realized correlations between two industries, the p-value ($p < 0.0001$) based on t-distribution suggests an evidence against the null hypothesis. Therefore, we conclude that the difference between two means is statistically significant; financial stocks are more closely related to overall market than the stocks from food industry. We further carry out the same analysis with the values of relative contribution of common jumps at $\tau = 2.5$. Figure 6 shows that the mean of fifty-two common jumps contributions of finance industry is 0.0807 and the mean of common jumps contributions of food industry is 0.0714 with the same standard error of 0.00398. It is hard to conclude whether we should reject the null hypothesis that assumes no difference in contributions of common jumps or not, because the p-values from the test are on the verge of statistical significance. While the one-sided p-value from the test is 0.0508, two-sided p-value is 0.1016. Instead of clear conclusion, in this case, we get an idea that the difference in relative contributions of common jumps between two industries is present, yet it is not as statistically significant as the difference in realized correlations between two industries.

6.3 Analysis by Year

In addition to the comparison between two industries, we study the realized correlation and contribution of common jumps values among thirteen years. We exclude S&P500 Futures volatility index due to its relatively short data length compared to other individual stocks. Table 3 shows the mean values of realized correlation across eight individual stocks for each year. On average, the eight individual stocks have the lowest mean correlation of 0.2532 with the market on year 2000, while they are much more strongly related to the market, with mean correlation 0.4968 during year 2003. Realized correlation values stay in these ranges and show small variations across different years. Similarly, Table 6 reveals that the average contribution of common jumps between eight individual stock and the market at τ=2.5, two and a half standard deviations of adjusted threshold for each year. The average contribution of common jumps is the smallest, around 5.35 percent in year 2000. The common jumps between the stocks and market contribute the most about 9.25 percent in realized correlation on year 2004. In an attempt to look for a particular pattern or difference in the correlations and contributions during the thirteen-year period, we carry out 'F-test', analysis of variance test (ANOVA) among thirteen years using the realized correlation and contribution of common jumps values in Table 7 and Table 8. Figure 7 and 8 present F-statistics; they indicate that their means are not the same across years. Despite of their statistically significant differences, we do not see the particular pattern or significant outlier (peak or trough) in the correlation and contribution of common jumps values. In fact, while we expected to observe significant fluctuations in these values, especially during the 2008 financial crisis, they are relatively stable with numbers from other years. The mean correlation of realized correlation with the market in 2008 is 0.4709, a little bit higher among other years, and the common jumps contribute 6.27 percent in realized correlation.

6.4 Relationship between Realized Correlation and RCCJ

Having analyzed realized correlation and relative contribution of common jumps by industry and year, we finally question the presence of a relationship between correlation and contribution of common jumps. Figure 9 plots relative contribution of common jumps at $\tau = 2.5$ against realized correlation of fifty-two points that are based on individual stock in every year. The graph seems to suggest a possibility of a weak positive linear relationship: the larger the realized correlation value, the greater contribution the common jumps is. The linear regression in Figure 9 proposes an estimated correlation coefficient of 0.0989 with standard error 0.0266. P-value of 0.0003 implies that there is an evidence against the null hypothesis of no association between realized correlation and relative contribution of common jumps.

7 Conclusion

Prior to introducing economic models of the study, we briefly study stochastic processes that are both continuous and discrete. In particular, geometric brownian motion and discrete time binomial tree are introduced as stock price models. We also show the convergence of binomial tree distribution to geometric brownian motion. The study then extends the concept of relative contribution of jumps in realized variance of a single security to a bivariate setting. We first study realized covariance between two securities, as analogous to realized variance in a univariate setting. We also observe common jumps that occur during the same intraday time interval, by comparing the presence of jumps in two corresponding intraday intervals. By considering realized covariances and common jumps together, we introduce a new covariance measure that is designed to exclude the intervals with common jumps. The ratio of the difference between the new covariance estimate and the previous covariance measure to the previous realized covariance is defined as relative contribution of common jumps in realized covariance or realized correlation. In particular, the study focuses on the relation between the S&P Futures market index, a proxy for market and individual equities in addition to S&P Futures volatility index. We implement the Bollerslev, Todorov, Li jump test (2011) that effectively takes into account of intraday volatility pattern, which otherwise distorts the number of jumps. Using this jump test, 5-minute intervals that are considered to contain market proxy realized jumps are successfully identified; similarly, the test also detects 5-minute intervals that are identified to contain realized jumps of individual stock and volatility index. We compare these intervals with realized jumps and exclude intraday intervals that are classified to contain common jumps in calculating new realized covariance, so that the effect of these realized common jumps is better examined.

As previously stated, Huang and Tauchen (2005) found that relative jump in the S&P Futures market index is around 7 percent. This study finds that the common jumps between the S&P Futures market index and individual equity account for about 17 to 19 percent, when the magnitude of price change is greater than or equal to two standard deviations of adjusted threshold. Even when the jump threshold increases to two and a half standard deviations of adjusted threshold, the relative contribution of co-jumps still accounts for 7 to 9 percent. These findings lead us to conclude that the common jumps play a significant role in studying correlation between the market and the stock.

By comparing realized correlation and the relative contribution of common jumps between stocks from food and financial industry, we find that financial stocks have higher correlation with the overall market than stocks from food industry. It also seems that common jumps in correlation between financial stocks and market are greater than those between food stocks and market, although the statistical significance remains inconclusive from the t-test. We further study the realized correlation and the relative contribution of common jumps on yearly basis by dividing samples into the sub-periods. Although they are different from each year, we do not observe significant outliers that particularly stand out from the data, even in times of financial crisis. The realized correlation and the relative contribution of co-jumps remain quite stable throughout the sample period. Through this further analysis, we come to a conclusion that the relative contribution of common jumps tend to be higher for the equities that also have higher realized correlation with the market. Even though the pattern between realized correlation and relative contribution of common jumps is weak, the relationship between them still reveals positive slope.

8 Figures and Tables

Figure 1: **Intraday Volatility Pattern for S&P Futures Market Index.** The figure shows minute-by-minute pattern in average absolute returns. The pattern is consistent with the widespread intraday volatility pattern in literature. Average volatility starts high at the beginning of the day, falls to a midday trough and gradually rises until the close.

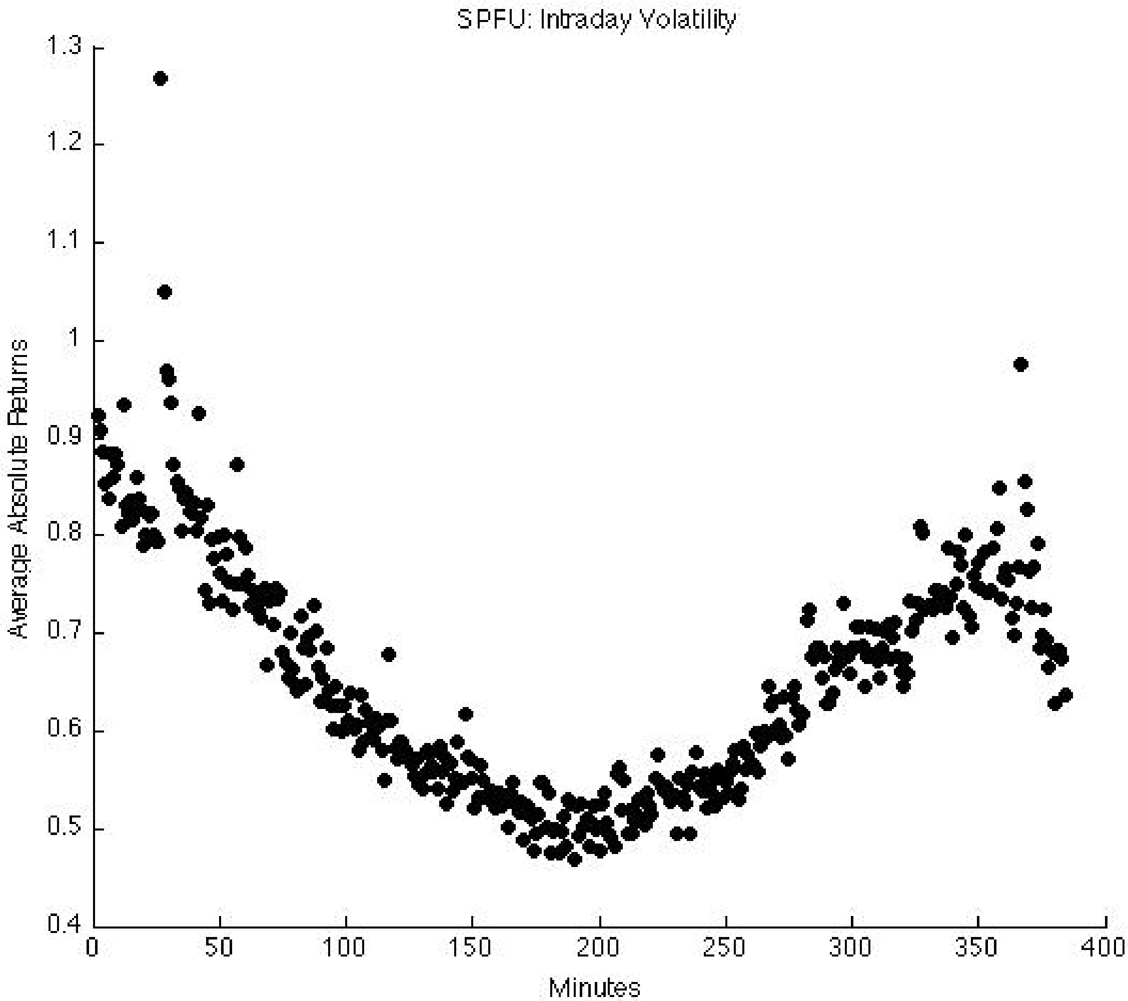


Figure 2: **Volatility Signature Plot for S&P Futures Market Index.** The graph indicates the average realized variance for every minute sampling frequency. As shown in the graph, the smaller the sampling interval (the higher the sampling frequency), the higher the average RV, since more microstructure noise contributes to RV measurement. The plot flattens out as the sampling interval becomes longer. Since 5-minute interval seems to capture enough variation and limits microstructure noise, the 5-minute interval is selected for this study.

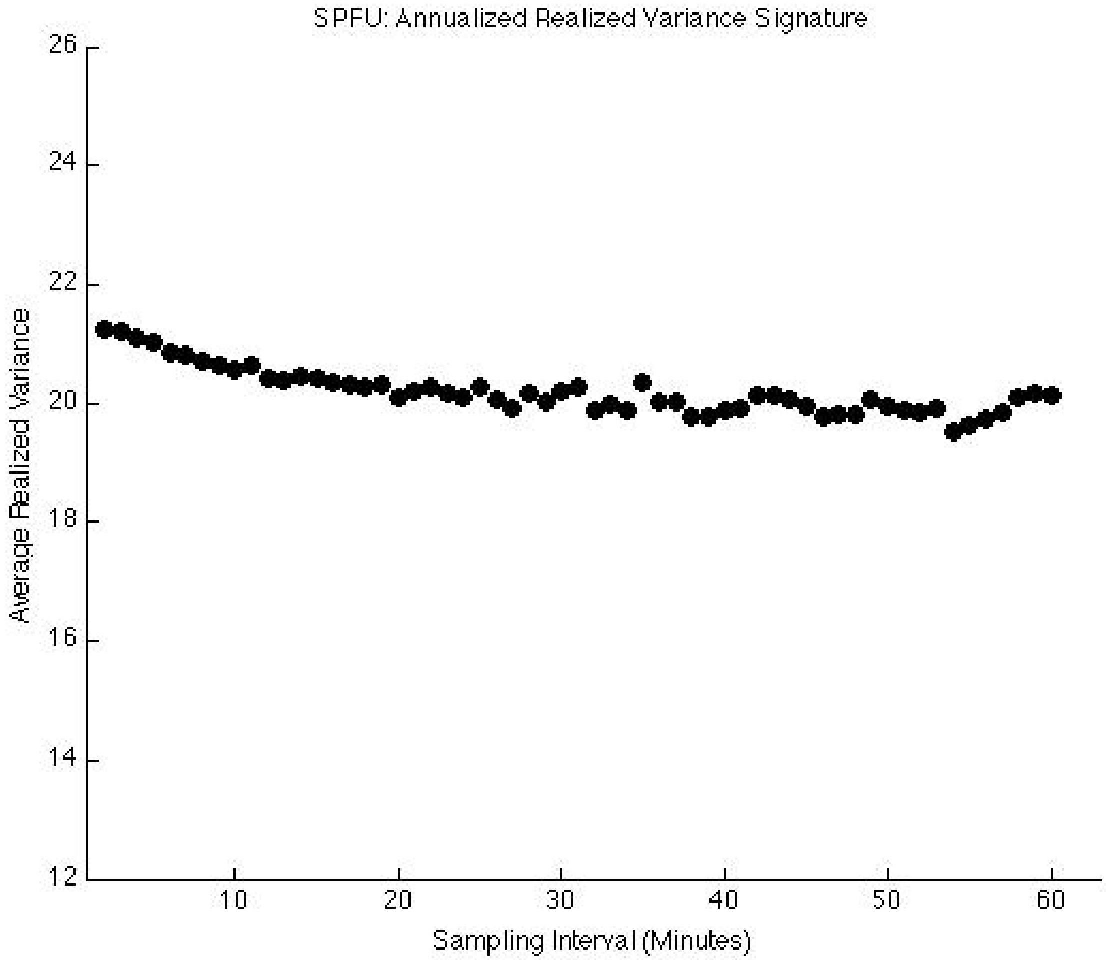


Figure 3: **Daily Realized Correlation** The realized correlation between S&P Futures market index and individuals stocks ranges from January, 1998 to December, 2010. It compares the individual stock performance to overall market performance over 13 years period. The last graph indicates the daily realized correlation between the market index and S&P Futures volatility index from January, 2004 to December, 2008 .

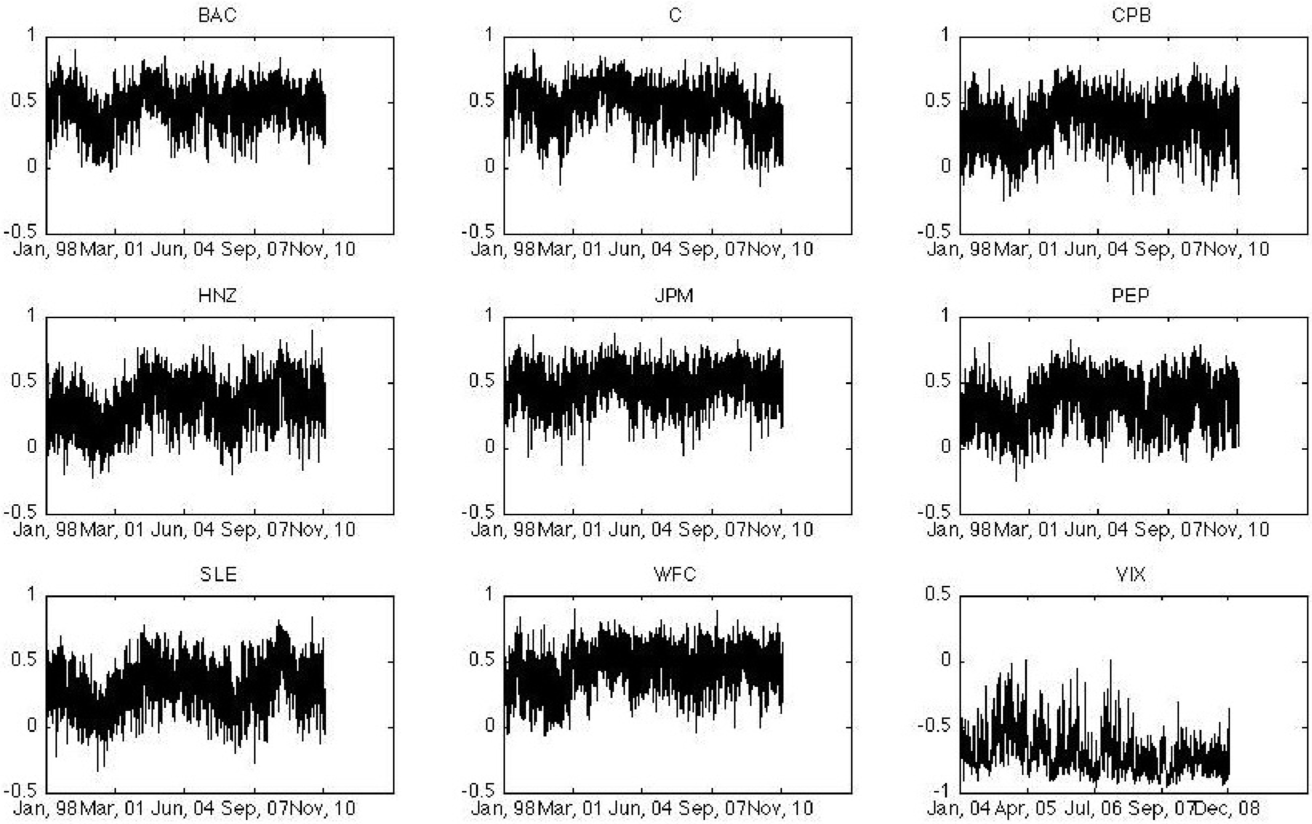


Figure 4: **Relative Contribution of Co-Jumps at Adjusted Threshold $\tau = 2$.** We select 5-minute interval for the Bollerslev, Todorov and Li (2011) jump test, and classify whether an interval contains a jump. Common jumps between the market and individual stocks as well as volatility index are further detected and relative contribution of co-jumps are obtained

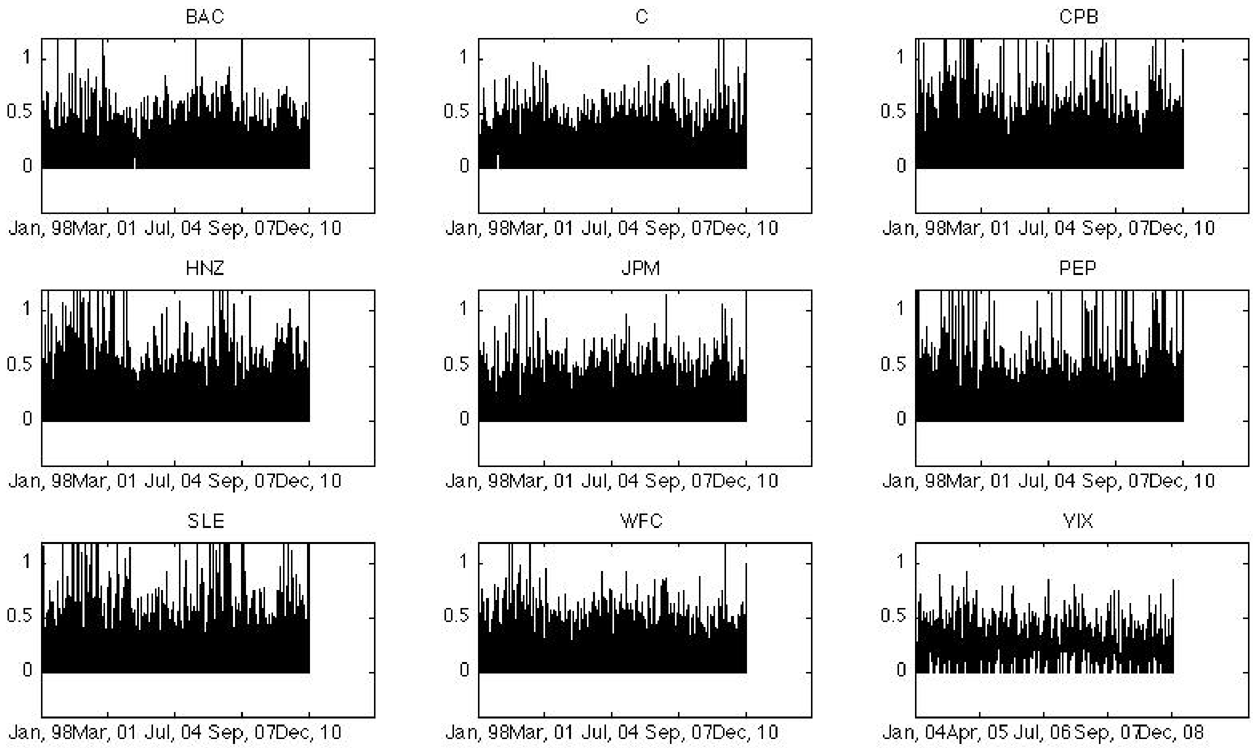
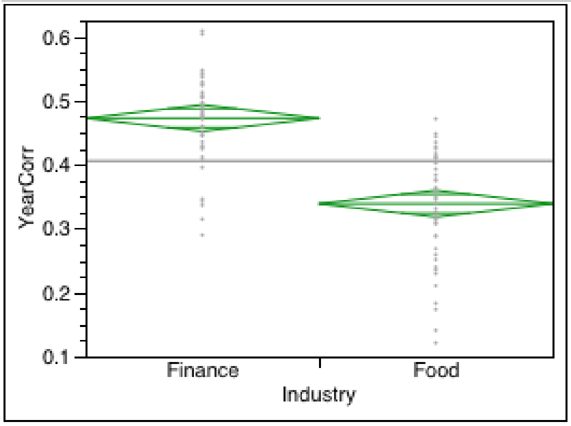


Figure 5: **Analysis of Mean Realized Correlation by Industry.** The figure plots yearly average realized correlation values of eight individual stocks by industry. The green line in the middle is the mean realized correlation of each industry. Upper and lower intervals together form a diamond shape.

Oneway Analysis of YearCorr By Industry



Missing Rows 3115

Oneway Anova

Summary of Fit

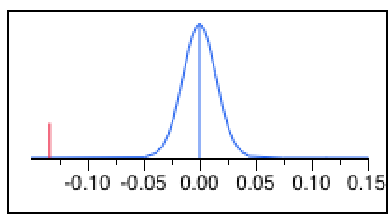
Rsquare	0.442575
Adj Rsquare	0.43711
Root Mean Square Error	0.075719
Mean of Response	0.40568
Observations (or Sum Wgts)	104

t Test

Food-Finance

Assuming equal variances

Difference	-0.13363	t Ratio	-8.99913
Std Err Dif	0.01485	DF	102
Upper CL Dif	-0.10418	Prob > |t|	<.0001*
Lower CL Dif	-0.16309	Prob > t	1.0000
Confidence	0.95	Prob < t	<.0001*



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Industry	1	0.4643142	0.464314	80.9843	<.0001*
Error	102	0.5848051	0.005733		
C. Total	103	1.0491193			

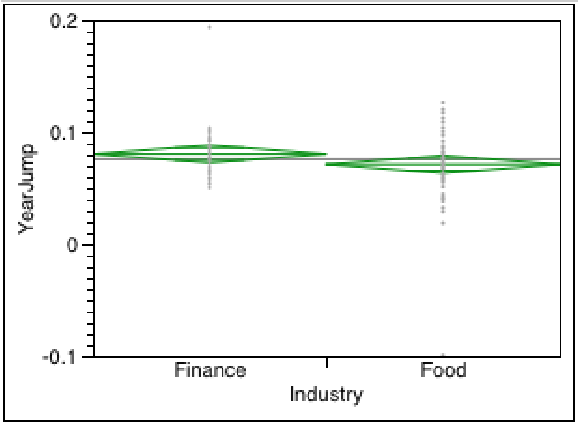
Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
Finance	52	0.472497	0.01050	0.45167	0.49332
Food	52	0.338862	0.01050	0.31803	0.35969

Std Error uses a pooled estimate of error variance

Figure 6: **Analysis of Mean Relative Contribution of Common Jumps by Industry.** The figure plots yearly average relative contribution of common jump values of eight individual stocks by industry. The green line in the middle is the mean relative contribution of common jumps of each industry. Upper and lower intervals together form a diamond shape.

Oneway Analysis of YearJump By Industry



Missing Rows 3115

Oneway Anova

Summary of Fit

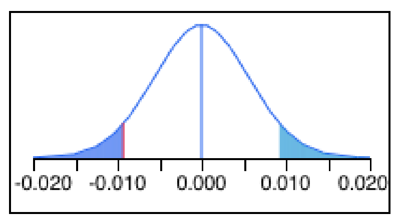
Rsquare	0.026064
Adj Rsquare	0.016516
Root Mean Square Error	0.028676
Mean of Response	0.076009
Observations (or Sum Wgts)	104

t Test

Food-Finance

Assuming equal variances

Difference	-0.00929	t Ratio	-1.65218
Std Err Dif	0.00562	DF	102
Upper CL Dif	0.00186	Prob > |t|	0.1016
Lower CL Dif	-0.02045	Prob > t	0.9492
Confidence	0.95	Prob < t	0.0508



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Industry	1	0.00224460	0.002245	2.7297	0.1016
Error	102	0.08387396	0.000822		
C. Total	103	0.08611857			

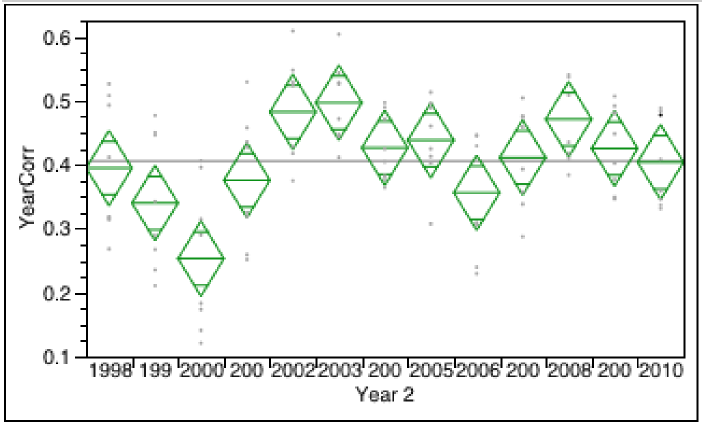
Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
Finance	52	0.080655	0.00398	0.07277	0.08854
Food	52	0.071363	0.00398	0.06348	0.07925

Std Error uses a pooled estimate of error variance

Figure 7: **Analysis of Mean Realized Correlation by Year.** The figure plots realized correlation values of eight individual stocks by year. The green line in the middle is the mean realized correlation of each year. Upper and lower intervals together form a diamond shape.

Oneway Analysis of YearCorr By Year 2



Missing Rows	3109
Excluded Rows	6

Oneway Anova

Summary of Fit

Rsquare	0.391637
Adj Rsquare	0.311413
Root Mean Square Error	0.083748
Mean of Response	0.40568
Observations (or Sum Wgts)	104

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Year 2	12	0.4108740	0.034240	4.8818	<.0001*
Error	91	0.6382453	0.007014		
C. Total	103	1.0491193			

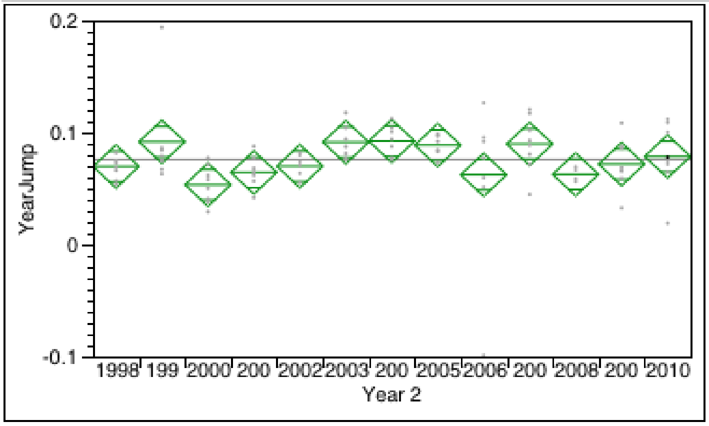
Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1998	8	0.394547	0.02961	0.33573	0.45336
1999	8	0.339934	0.02961	0.28112	0.39875
2000	8	0.253199	0.02961	0.19438	0.31201
2001	8	0.375588	0.02961	0.31677	0.43440
2002	8	0.482201	0.02961	0.42339	0.54102
2003	8	0.496754	0.02961	0.43794	0.55557
2004	8	0.426248	0.02961	0.36743	0.48506
2005	8	0.438148	0.02961	0.37933	0.49696
2006	8	0.356087	0.02961	0.29727	0.41490
2007	8	0.411125	0.02961	0.35231	0.46994
2008	8	0.470906	0.02961	0.41209	0.52972
2009	8	0.425244	0.02961	0.36643	0.48406
2010	8	0.403854	0.02961	0.34504	0.46267

Std Error uses a pooled estimate of error variance

Figure 8: **Analysis of Mean Relative Contribution of Common Jumps by Year.** The figure plots relative contribution of common jumps values of eight individual stocks by year. The green line in the middle is the mean relative contribution of common jumps of each year. Upper and lower intervals together form a diamond shape.

Oneway Analysis of YearJump By Year 2



Missing Rows	3109
Excluded Rows	6

Oneway Anova

Summary of Fit

Rsquare	0.206604
Adj Rsquare	0.101981
Root Mean Square Error	0.027401
Mean of Response	0.076009
Observations (or Sum Wgts)	104

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Year 2	12	0.01779247	0.001483	1.9747	0.0355*
Error	91	0.06832609	0.000751		
C. Total	103	0.08611857			

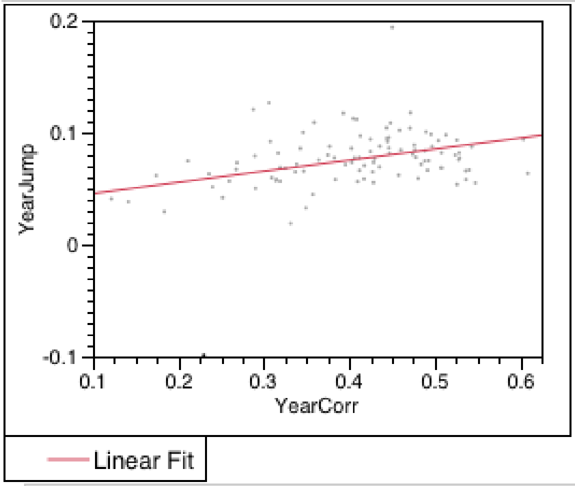
Means for Oneway Anova

Level	Number	Mean	Std Error	Lower 95%	Upper 95%
1998	8	0.069779	0.00969	0.05054	0.08902
1999	8	0.091866	0.00969	0.07262	0.11111
2000	8	0.053484	0.00969	0.03424	0.07273
2001	8	0.064467	0.00969	0.04522	0.08371
2002	8	0.070049	0.00969	0.05081	0.08929
2003	8	0.091525	0.00969	0.07228	0.11077
2004	8	0.092494	0.00969	0.07325	0.11174
2005	8	0.088846	0.00969	0.06960	0.10809
2006	8	0.062362	0.00969	0.04312	0.08161
2007	8	0.089913	0.00969	0.07067	0.10916
2008	8	0.062737	0.00969	0.04349	0.08198
2009	8	0.071857	0.00969	0.05261	0.09110
2010	8	0.078740	0.00969	0.05950	0.09798

Std Error uses a pooled estimate of error variance

Figure 9: **Realized Correlation and its Relative Contribution of Common Jumps.** The figure plots all corresponding realized correlation and relative contribution of common jumps for eight each stock in every year (in table 5 and 6). Best linear fit is introduced and the line appears to have a weak positive slope.

Bivariate Fit of YearJump By YearCorr



Linear Fit

YearJump = 0.035886 + 0.0989035*YearCorr

Summary of Fit

RSquare	0.119166
RSquare Adj	0.11053
Root Mean Square Error	0.027271
Mean of Response	0.076009
Observations (or Sum Wgts)	104

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	0.01026239	0.010262	13.7993
Error	102	0.07585618	0.000744	**Prob > F**
C. Total	103	0.08611857		0.0003*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.035886	0.011127	3.23	0.0017*
YearCorr	0.0989035	0.026625	3.71	0.0003*

Table 1: **Ticker, Data Range and Number of Trading Days.**

Ticker	Full Name	Industry	Data Range	Number of Trading Days
BAC	Bank of America Corp	Finance	1/2/98 to 12/30/10	3219
C	Citigroup Inc	Finance	1/2/98 to 12/30/10	3219
CPB	Campbell Soup Co	Food	1/2/98 to 12/30/10	3219
HNZ	Heinz, H.J. Co	Food	1/2/98 to 12/30/10	3219
JPM	JP Morgan Chase Co	Finance	1/2/98 to 12/30/10	3219
PEP	PepsiCo Inc	Food	1/2/98 to 12/30/10	3219
SLE	Sara Lee Corp	Food	1/2/98 to 12/30/10	3219
WFC	Wells Fargo Co	Finance	1/2/98 to 12/30/10	3219
VIX	S&P500 Futures Volatility Index	Market	1/5/04 to 12/31/08	1169
SPFU	S&P500 Futures Market Index	Market	1/2/98 to 12/30/10	3219

Table 2: **Mean Daily Realized Correlation.** Table shows the average daily realized correlation between individual stocks and the market index. The correlation is positive, while volatility index (VIX) is negatively correlated to market.

Stock	Mean
BAC	0.4716
C	0.4819
CPB	0.3512
HNZ	0.3437
JPM	0.4850
PEP	0.3497
SLE	0.3109
WFC	0.4514
VIX	-0.7001

Table 3: **Mean Realized Correlation.** Table shows the average realized correlation between eight stocks and the overall market for each year. Due to shorter data length, VIX is excluded in calculating the mean value.

Year	Mean
1998	0.3945
1999	0.3399
2000	0.2532
2001	0.3756
2002	0.4822
2003	0.4968
2004	0.4262
2005	0.4381
2006	0.3561
2007	0.4111
2008	0.4709
2009	0.4252
2010	0.4039

Table 4: **Number of Co-Jump Intervals for τ =2 and τ=2.5.** Table indicates the number of common jumps between individual stocks and the market index at two different threshold cut-offs, detected from Bollerslev, Todorov, and Li jump test.

Stock	$\tau = 2$	$\tau = 2.5$
BAC	2901	1003
C	2903	970
CPB	2511	833
HNZ	2385	823
JPM	2980	1040
PEP	2249	737
SLE	2125	715
WFC	2853	1014
VIX	2180	857

Table 5: **Mean Relative Contribution of Co-Jumps.** Table indicates the relative contribution of common jumps between individual stocks and the market. The larger the threshold cut-off ($\tau = 2.5$), the less the contribution of common jumps is.

Stock	$\tau = 2.0$	$\tau = 2.5$
BAC	0.1828	0.0853
C	0.1723	0.0755
CPB	0.1853	0.0751
HNZ	0.1779	0.0791
JPM	0.1764	0.0808
PEP	0.1923	0.0745
SLE	0.1669	0.0567
WFC	0.1769	0.0809
VIX	0.2500	0.1298

Table 6: **Mean Daily Relative Contribution at $\tau = 2.5$.** Table shows the average relative contribution of common jumps between eight stocks and the overall market for each year. Due to shorter data length, VIX is excluded in calculating the mean value.

Year	Mean
1998	0.0698
1999	0.0919
2000	0.0535
2001	0.0645
2002	0.0700
2003	0.0915
2004	0.0925
2005	0.0888
2006	0.0624
2007	0.0899
2008	0.0627
2009	0.0719
2010	0.0787

Table 7: **Mean Realized Correlation**

	BAC	C	JPM	WPC	CPB	HNZ	PEP	SLE
1998	0.5079	0.5260	0.4927	0.4119	0.3189	0.2684	0.3168	0.3138
1999	0.4503	0.4767	0.4464	0.3426	0.2894	0.2356	0.2676	0.2109
2000	0.3147	0.3955	0.4058	0.2900	0.1738	0.1411	0.1833	0.1214
2001	0.4272	0.5292	0.4577	0.4356	0.3251	0.2594	0.3189	0.2517
2002	0.5475	0.6088	0.5287	0.5238	0.4282	0.4176	0.4281	0.3749
2003	0.5280	0.6037	0.5432	0.5258	0.4715	0.4435	0.4469	0.4115
2004	0.4247	0.4969	0.4897	0.4709	0.4044	0.3824	0.3768	0.3641
2005	0.5130	0.4885	0.4602	0.4957	0.4029	0.4130	0.4247	0.3073
2006	0.4439	0.4463	0.4447	0.4293	0.3091	0.2397	0.3059	0.2297
2007	0.4760	0.4739	0.5042	0.4589	0.3382	0.3929	0.3574	0.2876
2008	0.5404	0.5365	0.5363	0.5083	0.4097	0.4172	0.3837	0.4352
2009	0.4805	0.3465	0.5069	0.4916	0.3490	0.4482	0.3768	0.4024
2010	0.4776	0.3367	0.4883	0.4840	0.3457	0.4085	0.3589	0.3312

Table 8: **Mean Relative Contribution of co-jumps at τ=2.5**

	BAC	C	JPM	WPC	CPB	HNZ	PEP	SLE
1998	0.0821	0.0538	0.0751	0.0663	0.0689	0.0731	0.0818	0.0571
1999	0.1939	0.0845	0.0863	0.0859	0.0792	0.0632	0.0672	0.0748
2000	0.0582	0.0716	0.0776	0.0503	0.0616	0.0383	0.0294	0.0409
2001	0.0649	0.0775	0.0620	0.0878	0.0682	0.0568	0.0567	0.0420
2002	0.0551	0.0637	0.0821	0.0800	0.0735	0.0707	0.0554	0.0800
2003	0.0759	0.0941	0.0875	0.0935	0.1177	0.1043	0.0816	0.0777
2004	0.0942	0.0876	0.1007	0.1040	0.1125	0.0776	0.0877	0.0756
2005	0.0980	0.0838	0.0845	0.0985	0.0735	0.0972	0.0831	0.0921
2006	0.0949	0.0957	0.0921	0.0770	0.0598	0.0515	0.1266	-0.0987
2007	0.0889	0.0807	0.0930	0.1021	0.0718	0.1173	0.0448	0.1206
2008	0.0667	0.0581	0.0659	0.0686	0.0564	0.0587	0.0581	0.0694
2009	0.0592	0.0655	0.0688	0.0665	0.0328	0.1086	0.0876	0.0860
2010	0.0781	0.0649	0.0748	0.0718	0.1003	0.1118	0.1091	0.0190

.

References

[1] Ait-Sahalia, Y., and J.Jacod (2007). "Testing for Jumps in Discretely Observed Process," *Annals of Statistics*, 35(1), 355-392.

[2] Andersen, T., L. Benzoni, and J. Lund (2002). "An Empirical Investigation of Continuous-Time Equity Return Models," *Journal of Finance*, 57(3), 1239-1284.

[3] Andersen, T., and T. Bollerslev (1997). "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4(2-3), 115-158.

[4] Andersen, T., T. Bollerslev, and F. Diebold and P. Labys (2000). "Exchange Rate Returns Standardized by Realized Volatility are (Nearly) Gaussian," *Multinational Finance Journal*, 4(3), 159-179.

[5] Barndorff-Nielsen, O., and N. Shephard (2004). "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2(1), 1-37.

[6] Barndorff-Nielsen, O., and N. Shephard (2006). "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4(1), 1-30.

[7] Black, F., and M. Scholes (1973). "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81(3), 637-654.

[8] Bollerslev, T., V. Todorov, and S.Z. Li (2011). "Jump Tails, Extreme Dependencies, and the Distribution of Stock Returns," Working Paper.

[9] Dong, X., and A. Petters (2012). "Mathematical Finance with Applications", in preparation.

[10] Drost, F., T. Nijman, and B. Werker (1998). "Estimation and Testing in Models Containing Both Jumps and Conditional Heteroskedasticity," *Journal of Business Economics and Statistics*, 16(2), 237-243.

[11] Huang, X., and G. Tauchen (2005). "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3(4), 256-499.

[12] Hull, J. *Options, Futures, and other Derivatives,* Pearson: Prentice Hall, 2009.

[13] Lee, S., and P. Mykland (2008). "Jumps in Financial Markets: A New Nonparmetric Test and Jump Dynamics," *Review of Financial Studies*, 21(6), 2535-2563.

[14] Liu, J., A. Longstaff, and J. Pan (2003). "Dynamic Asset Allocation with Event Risk," *Journal of Finance*, 58(1), 231-259.

[15] Maheu, J. M., and T. H. McCurdy (2004). "News Arrival, jump Dynamics, and Volatility Components for Individual Stock Returns," *Journal of Finance*, 59(2), 755-793.

[16] Merton, R. (1976). "Option Pricing when Underlying Stock Returns are Discontinuous," *Journal of Financial Econometrics*, 3(1), 125-144.

[17] Rognlie (2010). "Spurious Jump Detection and Intraday Changes in Volatility," Duke University Economics Honors Thesis.

[18] Schwert, M. (2009). "Hop, Skip and Jump - What Are Modern "Jumps" Tests Finding in Stock Returns?," Duke University Economics Honors Thesis.

[19] Tauchen, G., and H. Zhou (2011). "Realized Jumps on Financial Markets and Predicting Credit Spreads," *Journal of Econometrics*, 160(1), 102-118.

[20] Van Tassel, P. (2008). "Patterns Within the Trading Day: Volatility and Jump Discontinuities in High Frequency Equity Price Series," Duke University Economics Honors Thesis.

[21] Wood, R., McInish, and J.Ord (1985). "An Investigation of Transaction Data for NYSE Stocks," *Journal of Finance*, 40(3), 723-739.