The dynamics of co-jumps, volatility and correlation

**A.E. Clements and Y. Liao**
School of Economics and Finance, Queensland University of Technology, NCER.

**Abstract**
Understanding the dynamics of volatility and correlation is a crucially important issue. The literature has developed rapidly in recent years with more sophisticated estimates of volatility, and its associated jump and diffusion components. Previous work has found that jumps at an index level are not related to future volatility. Here we examine the links between co-jumps within a group of large stocks, the volatility of, and correlation between their returns. It is found that the occurrence of common, or co-jumps between the stocks are unrelated to the level of volatility or correlation. On the other hand, both volatility and correlation are lower subsequent to a co-jump. This indicates that co-jumps are a transient event but in contrast to earlier research have a greater impact that jumps at an index level.

**Keywords**
Realized volatility, correlation, jumps, co-jumps, point process

**JEL Classification Numbers**
C22, G00.

**Corresponding author**
Adam Clements
School of Economics and Finance
Queensland University of Technology
Brisbane, 4001
Qld, Australia

email a.clements@qut.edu.au
Understanding the volatility of, and correlation between asset returns is paramount. Many financial such as risk management, portfolio allocation and derivative pricing utilise estimates of volatility and correlation. As such there has been a vast literature relating to the estimation and forecasting of both volatility and correlation. Much of this literature has stemmed from the development of the univariate GARCH class of models, Engle (1982) and Bollerslev (1986) and subsequent multivariate models developed by Bollerslev (1990) and Engle (2002).

In recent years, this literature has benefited from the development of Realized Volatility (RV) by Andersen, Bollerslev, Diebold, and Labys (2001, 2003) which utilises intraday asset returns leading to improved measures of volatility at a daily frequency. Extending the basic notion of RV, more sophisticated and efficient sampling techniques have introduced, see for instance Bandi and Russell (2006, 2008), Hansen and Lunde (2008) and Zhang, Mykland, and Aït-Sahalia (2005), Zhang (2006). These estimates of RV converge to the total quadratic variation in the asset return process, which may be attributable to both a continuous diffusion and discrete jump process. Barndorff-Nielsen and Shephard (2006) and Andersen, Bollerslev, and Diebold (2007) develop methods for measuring the contribution to total quadratic variation from both sources. Andersen et al. (2007) propose a formal test for identifying when significant contributions from jump activity occur. By separating total volatility into jump and diffusion components, Andersen et al. (2007) find that almost all of predictability in volatility stems from the persistent diffusive component.

Given the need for pricing and managing the risks associated with holding a portfolio of assets, there is now a growing literature that attempts to document and study simultaneous discrete jumps across many assets (co-jumps). Progress on this front includes the development of tests for co-jumps in a pair of asset returns (Barndorff-Nielsen and Shephard (2003), Gobbi and Mancini (2007) and Jacod and Todorov (2009)), as well as a co-jump test developed by Bollerslev, Law, and Tauchen (2008) that is applicable to a large panel of high-frequency returns. These tests pave the way to identify the factors which drive the occurrence of co-jumps in different markets or assets and understand the dynamic properties of co-jumps themselves. Lahaye and Neely (2011) relate the co-jumps extracted from a panel of U.S. stocks to U.S. macroeconomic releases and formally modeled how news surprises explain co-jumps. Dungey and Hvozdyk (2011) studied the co-jumps in spot and futures prices in high frequency U.S Treasury data, and find that an anticipated macroeconomic news announcement is sufficient to change the probability of observing co-jumps. In contrast, there is little literature so far concerning about
formally modeling the dynamics of co-jumps themselves.

This paper seeks to extend the work of Andersen et al. (2007) and examine the relationship between co-jumps, the volatility of, and the correlation between the returns of a group of assets. Andersen et al. (2007) find that the predictability in volatility is mainly related to continuous diffusion and that jumps (at a univariate index level) are not important. In contrast to the univariate case, our results indicate the predictability in average volatility and correlation across a group of assets can be linked to the occurrence of co-jumps. In both cases, volatility and correlation rise subsequent to co-jumps occurring. On the other hand we find that the level of volatility does not influence the probability of a co-jump occurring.

The paper proceeds as follows. Section 2 describes the methods employed to estimate realized variances, correlations and identify co-jumps. Section 2.3 describes the stock return data used in the empirical analysis. Section 3 presents both the empirical methods employed and results. Section 4 provides concluding comments.

2 Theoretical Framework

In this section, we discuss the way in which we form realized variance and correlation estimates from high frequency financial data, and the way in which we detect the co-jumps in a panel of high frequency stock prices.

2.1 High Frequency Data, Realized Variance and Correlation

We assume that the logarithm of the asset price within the active part of the trading day evolves in continuous time as a standard jump-diffusion process given by

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

where $u(t)$ denotes the drift term that has continuous and locally bounded variation, $\sigma(t)$ is a strictly positive spot volatility process and $w(t)$ is a standard Brownian motion. The $\kappa(t)dq(t)$ term refers to a pure jump component, where $k(t)$ is the size of jump and $dq(t) = 1$ if there is a jump at time $t$ (and 0 otherwise). The corresponding discrete-time within-day geometric returns are

$$r_{t+j\Delta} = p(t + j/M) - p(t + (j - 1)/M), \quad j = 1, 2, ..., M, \tag{2}$$

where $M$ refers to the number of intraday equally spaced return observations over the trading day $t$, and $\Delta = 1/M$ denotes the sampling interval. As such, the daily return for the active part of the trading day equals $r_t = \sum_{j=1}^{M} r_{t+j\Delta}$. As noted in Andersen and Bollerslev (1998), Andersen et al. (2003), and Barndorff-Nielsen and Shephard (2002), the variance over the active
part of the trading day \( t \) can be measured by realized variance, which converges uniformly in probability to quadratic variation as the sampling frequency goes to infinity. Realized variance \((RV)\) is defined as the sum of the intraday squared returns, i.e.

\[
RV_{t+1}(\Delta) \equiv \sum_{j=1}^{M} r_{t+j\Delta}^2.
\]  

(3)

Moreover, the concept of realized variance can be extended to realized covariance in a multivariate setting. For example, with high frequency data available for the returns of two assets A and B, \( r_{t,j}^A \) and \( r_{t,j}^B \), the realized covariance can be defined as the summation of the cross products of intra-day returns of two assets A and B within a day \( t \)

\[
RCov_{t}^{AB} = \sum_{j=1}^{M} r_{t,j}^A r_{t,j}^B.
\]  

(4)

Then, the realized correlation of asset A and B can be defined as the ratio of realized covariance and product of realized standard deviations, which naturally follows as

\[
RCorr_{t}^{AB} = \frac{RCov_{t}^{AB}}{\sqrt{RV_t^A RV_t^B}}.
\]  

(5)

### 2.2 Identifying Co-jumps in Stock Prices

A number of statistical tests have been developed to detect the existence of jumps in asset prices. The jump test developed by (Barndorff-Nielsen and Shephard (2006), henceforth the BN-S test, plays a leading role in this literature, but is designed for a single asset. Many financial problems involve multiple assets and thus there is a need for tests that can detect simultaneous jumps in many assets (co-jump tests). The co-jump test developed recently by Bollerslev et al. (2008) (henceforth called the BLT test) that is applicable to a large panel of high-frequency returns fills this void. The intuition behind this test is that idiosyncratic noise in individual returns can hide the presence of a synchronous component. Therefore a test based on the cross products of returns in a panel can avoid this problem while still being sensitive to systematic movements across all stocks. In this study, we use the BLT test to detect the occurrence of co-jumps across a panel of stock prices.

Starting from a collection of \( n \) stock price processes \( \{p_{i,s}\}_{i=1}^{n} \) evolving in continuous time, Bollerslev et al. (2008) assume that each \( p_{i,s} \) evolves as

\[
dp_{i,s} = \mu_i(s)dt + \sigma_i(s)dW_i(s) + dL_i(s),
\]

where \( \mu_i(s) \) and \( \sigma_i(s) \) refer to the drift and local volatility, \( W_i(s) \) is a standard Brownian motion, and \( L_i(s) \) is a pure jump process. The price process is only observed at discrete time points,
so they consider a situation in which there are $M+1$ equidistant price observations each day. The $j$th within-day return of the $i$th log-price process on day $t$ is then

$$r_{i,t,j} = p_{i,(t-1)+\frac{j}{M}} - p_{i,(t-1)+\frac{j-1}{M}}, \quad i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, M.$$ 

and the $j$th within-day return on day $t$ of an equi-weighted portfolio of $n$ stocks is

$$r_{EQW,t,j} = \frac{1}{n} \sum_{i=1}^{n} r_{i,t,j},$$

The daily realized variance for this equi-weighted portfolio is given by

$$RV_{EQW,t} = \frac{1}{n} \sum_{j=1}^{n} \left( \frac{1}{n} \sum_{i=1}^{n} r_{i,t,j} \right)^2 = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{M} r_{i,t,j}^2 + \frac{1}{n^2} \sum_{i=1}^{n} \sum_{l=1, l \neq i}^{n} \sum_{j=1}^{M} r_{i,t,j} r_{l,t,j},$$

and when this is decomposed into its continuous and jump components, Bollerslev et al. (2008) show that most of the jump contribution to $RV_{EQW,t}$ originates from the covariation term (i.e. from within $\frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{M} r_{i,t,j}^2$ in (7)) when $n$ is large, while the effects of idiosyncratic jumps (i.e. that originate from the term) are essentially diversified away.

The notion that co-jumps can cause the price of a portfolio to jump when $n$ is large forms the basis for the co-jump test. Their derivation was based on an equi-weighted portfolio, but their conclusion that most of the information about co-jumps is contained in the covariation between stock returns is valid given any well-diversified portfolio.

The $z_{mcp}$ test statistic proposed by Bollerslev et al. (2008) is given by

$$z_{mcp,t,j} = \frac{mcp_{t,j} - \overline{mcp_t}}{s_{mcp,t}}, \quad j = 1, 2, \ldots, M,$$

where

$$mcp_{t,j} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{l=i+1}^{n} r_{i,t,j} r_{l,t,j}, \quad j = 1, 2, \ldots, M,$$

$$\overline{mcp_t} = \frac{1}{M} \sum_{j=1}^{M} mcp_{t,j} = \frac{1}{M} \left( \frac{n}{n-1} RV_{ew,t} - \frac{1}{n(n-1)} \sum_{i=1}^{n} RV_{i,t} \right),$$

$$s_{mcp,t} = \sqrt{\frac{1}{M-1} \sum_{j=1}^{M} (mcp_{t,j} - \overline{mcp_t})^2},$$

and it can be used as a test for common jumps because the jump (but not the continuous) component in the second term in $mcp_t = \frac{n}{n-1} RV_{ew,t} - \frac{1}{n(n-1)} \sum_{i=1}^{n} RV_{i,t}$ is diversified away as $n$ grows large.
It is easy to rearrange the expression for \( mcp_{t,j} \) to obtain

\[
mcp_{t,j} = \frac{n}{n-1} \left[ \frac{1}{n} \sum_{i=1}^{n} r_{i,t,j}^2 - \frac{1}{n^2} \sum_{i=1}^{n} r_{i,t,j}^2 \right] = \frac{n}{n-1} r_{ew,t,j}^2 - \frac{1}{n(n-1)} \sum_{i=1}^{n} r_{i,t,j}^2,
\]

and from this we can see that it is possible to calculate the test statistic directly from the squared returns of the equally weighted portfolio and the individual stocks. We work with this alternative expression for \( mcp_{t,j} \) in the subsequent empirical analysis.

The \( z_{mcp,t,j} \) statistic is not well approximated by any of the standard distribution\(^1\), but it is relatively straightforward to bootstrap its empirical distribution under the null hypothesis of no jumps, to find critical values that are relevant for a given application. Meanwhile, the \( zmcp \) co-jump test relies on three assumptions. Firstly, the studentization of the \( mcp_{t,j} \) test statistic each day relies on an assumption that the location and scale of this statistic remains approximately constant over the day. This assumption may be at odds with the well-known U-shaped pattern associated with intra-day stock volatility. Bollerslev et al. (2008) suggest a way to deal with this issue that we can scale the return for a stock over a certain time-interval by the reciprocal of the square root of the corresponding unconditional bi-power variation for that particular time-interval, in which it deflates returns near the beginning and the end of the day while inflating returns in the middle of the day. Then the intra-day pattern will be successfully removed from the \( zmcp \) co-jump test statistics calculated based on these scaled returns. We will implement this method in the later empirical analysis to construct co-jumps. Secondly, the \( mcp_{t,j} \) realizations are assumed to be serially uncorrelated, making it appropriate to simply standardized each of the within-day \( mcp \) statistic by using the corresponding daily sample standard deviation \( s_{mcp,t} \).

Lastly, it is important to note that the sample mean used in the \( zmcp \) test statistic incorporates the co-jump contribution relating to each day, and although the contribution of a few jumps on a day might be negligible, the contribution of several co-jumps is unlikely to be negligible and then relatively large intra-day \( mcp_{t,j} \) realizations might be masked by the correspondingly large sample mean \( \text{mean}_{mcp} \). Therefore, the test implies an assumption that co-jumps occur rarely, particularly no more than one co-jump over a day. This assumption finds empirical support in Bollerslev et al. (2008) and also in our study on 20 large stocks from the NYSE.

### 2.3 Data

The empirical data set comprises split-and-dividend-adjusted five minute intra-day prices for 20 large stocks trading on the NYSE for the period 22 July 2002 to 5 December 2011 (2357 trading days) collected from Thomson Reuters Tick History. Table 1 reports the company names and

---

\(^1\) Simulations conducted by Bollerslev et al. (2008) show that the distribution of the \( z_{mcp,t,j} \) statistic is centered to the left of zero and has a very strong right skew.
<table>
<thead>
<tr>
<th>Company</th>
<th>Ticker</th>
<th>$RV \times 10^{-4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>3M</td>
<td>MMM</td>
<td>1.9140</td>
</tr>
<tr>
<td>Alcoa</td>
<td>AA</td>
<td>5.8154</td>
</tr>
<tr>
<td>Bank of America</td>
<td>BAC</td>
<td>7.1787</td>
</tr>
<tr>
<td>Boeing</td>
<td>BA</td>
<td>2.8788</td>
</tr>
<tr>
<td>Caterpillar</td>
<td>CAT</td>
<td>3.6128</td>
</tr>
<tr>
<td>Chevron</td>
<td>CVX</td>
<td>2.4476</td>
</tr>
<tr>
<td>Cisco Systems</td>
<td>CSCO</td>
<td>3.5637</td>
</tr>
<tr>
<td>Coca-Cola</td>
<td>KO</td>
<td>1.3811</td>
</tr>
<tr>
<td>DuPont</td>
<td>DD</td>
<td>2.8471</td>
</tr>
<tr>
<td>ExxonMobil</td>
<td>XOM</td>
<td>2.2703</td>
</tr>
<tr>
<td>General Electric</td>
<td>GE</td>
<td>3.5086</td>
</tr>
<tr>
<td>Hewlett-Packard</td>
<td>HPQ</td>
<td>3.4980</td>
</tr>
<tr>
<td>The Home Depot</td>
<td>HD</td>
<td>3.2575</td>
</tr>
<tr>
<td>Intel</td>
<td>INTC</td>
<td>3.5229</td>
</tr>
<tr>
<td>IBM</td>
<td>IBM</td>
<td>1.8565</td>
</tr>
<tr>
<td>Johnson &amp; Johnson</td>
<td>JNJ</td>
<td>1.2867</td>
</tr>
<tr>
<td>JPMorgan Chase</td>
<td>JPM</td>
<td>5.5352</td>
</tr>
<tr>
<td>McDonald’s</td>
<td>MCD</td>
<td>2.1201</td>
</tr>
<tr>
<td>Microsoft</td>
<td>MSFT</td>
<td>2.4157</td>
</tr>
<tr>
<td>Pfizer</td>
<td>PFE</td>
<td>2.2550</td>
</tr>
</tbody>
</table>

| % of co-jumps          | 29.35% |
| Min duration           | 1 day  |
| Max duration           | 18 days|

Table 1: Descriptive statistics

associated tickers of the stocks considered, their respective average daily realised variances and basic statistics on co-jump activity. Table 1 shows that unsurprisingly, the banks BAC and JPM have two of the highest average volatilities (given the period spans the Global Financial Crisis) along with AA. Co-jumps are detected in 29.35% of trading days which is equivalent to 0.38% of five minute intervals as no day contains more than one co-jump. Co-jumps are detected on consecutive days, with a maximum duration of 18 days between co-jumps.

Figure 1 shows both the average realized volatility across the individual stocks (top panel) and the average realized correlation across pairs of stocks (bottom panel). It is clear that the average volatility of the portfolio of stocks follows a familiar pattern and is dominated by the Global Financial Crisis of 2008-09. The average pair-wise correlation is quite a persistent process and rises rapidly during the crisis period of 2008-09 and generally remains higher during the latter half of the sample. This relationship between volatility and correlation is often discussed in the
context of contagion across markets or countries and is also analysed in a formal manner within
a single market by Silvennoinen and Terasvirta (2013).

![Average realized volatility](image)

![Average realized correlation](image)

Figure 1: Average realized volatility estimates (top panel) and average pair-wise realized corre-
lations (bottom panel).

Figure 2 plots the autocorrelations in average realized variance, correlation and the co-jump
series (along with associated confidence intervals). It is clear the volatility exhibits the familiar
pattern of strong and slowly decaying persistence. Correlation is also strongly persistent and
behaves in a very similar manner to volatility. Co-jumps show little evidence of structure beyond
a small degree of persistence at the first lag.

3 Empirical Methods and Results

This section present both the empirical methods employed and presents the associated results.
Section 3.1 examines the behaviour of co-jumps and whether volatility and correlation across
the stocks influences the occurrence of co-jumps. On the other hand, Section 3.2 examines the
impact of co-jumps on volatility and correlation.
3.1 A model for co-jumps: The role of volatility and correlation

Here we model the occurrence of co-jumps as a point process. Such an approach is common when dealing with events in financial markets such as the arrival of trades or quotes, for an overview of the literature see Bauwens and Hautsch (2009). To begin we provide some definitions. Let \( \{t_i\}_{i=1,\ldots,n} \) be a random sequence of increasing event times \( 0 \geq t_1 > \cdots > t_n \) which describes a simple point process. \( N(t) := \sum_{i \geq 1} 1_{t_i \geq t} \) is a counting function. Given the counting process, we can view the conditional intensity, \( \lambda(t) \) as the expected change in \( N(t) \) (as a reflection of the probability of an event occurring) over a small time horizon,

\[
\lambda(t) = \lim_{s \downarrow t} \frac{1}{s-t} E[N_m(s) - N(t)]
\]  

Bauwens and Hautsch (2009) provides discussion of various specifications for \( \lambda(t) \) along with the relationships between such intensity models and hazard and duration models.

A common specification for \( \lambda(t) \) is the self-exciting Hawkes process, attributable to Hawkes (1971)

\[
\lambda(t) = \mu + \int_0^t w(t-u)dN(u) = \mu + \sum_{t_i < t} w(t - t_i)
\]  

where \( \mu \) is a constant and \( w() \) is a non-negative weight function. This process is self-exciting in the sense that \( \text{Cov}[N(a,b), N(b,c)] > 0 \) where \( 0 > a \geq b < c \). The weight function \( w() \) is a decreasing function of \( t - u \) meaning that subsequent to a spike the intensity decays.

Figure 2: Autocorrelations for Average realized volatility, average pair-wise realized correlations and co-jumps. The dashed lines represent the 95% confidence interval.
In the current context we estimate a discrete version of equation 7 where $\lambda_t$ will denote the discrete time conditional intensity for the $t$-th trading day. The intensity of the occurrence of co-jumps will be driven by past occurrences (to allow for self-excitation to capture short term persistence), and the level of volatility and correlation. This model allows us to examine whether co-jumps are more likely during periods of market stress (higher volatility and correlation). The model takes the following form

$$
\lambda_t = \mu + \gamma_1 \overline{RV}_t + \gamma_2 \overline{RV}_{t-1} + \gamma_3 \overline{RCorr}_t + \gamma_4 \overline{RCorr}_{t-1} + \left\{ \begin{array}{ll}
\beta \lambda_{t-1} & \text{if } dN_{t-1} = 0 \\
\beta \lambda_{t-1} + \alpha & \text{otherwise.} \end{array} \right. \quad (8)
$$

$\mu$ is a baseline intensity, $\lambda_{t-1}$ is the conditional intensity from the previous day, $\beta$ is the decay factor for the intensity from the day, $\lambda_{t-1}$ and $\alpha$ is the shock to the intensity on day $t$ if a co-jump occurred on day $t-1$ ($dN_{t-1,d} > 0$). A base self exciting model with no exogenous variables will also be estimated.

Table 2 reports the estimation results of co-jump intensity models, including the base self exciting model without covariates. We use both the level and the logarithm of the average realized variance and correlation in the model. Turning to the results reported in the first column of this table, the estimate for $\mu$, the baseline intensity of co-jumps, is very close to the proportion of trading days with co-jumps reported in Section 2.3. Meanwhile, the estimates $\beta$ and $\alpha$ confirm that co-jumps self-excite and exhibit a persistent intensity process. With an estimate of $\beta$ around 0.91 the half-life of the increase in intensity following a co-jump is around 7.3 days.

Interestingly, when we introduce the average realized volatility and realized correlation into the co-jump intensity model, the impact of the occurrence of past co-jumps continues to be important. Comparing the estimation results of the first column with the second and the fifth columns in the table, the estimates for $\mu$, $\beta$ and $\alpha$ are very close. Meanwhile, the coefficients of the average realized volatility, realized correlation (in either levels of logs) and their lags are insignificant at the 5% significance level, and the values of the log likelihood function remains the same. In other words, whereas the intensity of co-jumps appear to be self exciting with the level of co-jump intensity is not influenced by the market condition (the level of the average volatility and correlation). Average realized variance is highly correlated with the realized correlation, thus the insufficiency of the coefficients of on the two exogenous variables may be due the multi-collinearity. Therefore, we use the two variables separately in the model to investigate this issue, and report the relevant results in the third, the forth, the sixth and the seventh columns of Table 2. The insignificant coefficients of the average realized volatility and its lag,

---

2The role of jumps in individual stocks was also considered. Measures such as the proportion of the 20 stocks experiencing jumps were used as exogenous variables in the intensity model for co-jumps. None of these measures were found to be relevant.
or realized correlation and its lag are remain.

We also undertake a number of robustness checks. The analysis presented above employs co-jumps constructed from the original five-minute returns, based on the 5% significance level. To see whether the results are sensitive to the intraday pattern in volatility, we present the same results basing the zmcp co-jump test on the intraday pattern adjusted five-minute returns. These results are presented in Table 3 and are consistent with those in Table 2. Hence, the intraday pattern has no discernable impact on the detection of co-jumps. Meanwhile, we also use the co-jumps detected based on 1% and 10% significance level to re-estimate the set of models outlined above. We do not report the results here but they are very similar to those reported in Table 2 and Table 3.
A : \[ \lambda_t = \mu + \begin{cases} 
\beta \lambda_{t-1} & \text{if } dN_{t-1} = 0 \\
\beta \lambda_{t-1} + \alpha & \text{otherwise}
\end{cases} \]

B : \[ \lambda_t = \mu + \gamma_1 \overline{RV}_t + \gamma_2 \overline{RV}_{t-1} + \gamma_3 \overline{RCorr}_t + \gamma_4 \overline{RCorr}_{t-1} + \begin{cases} 
\beta \lambda_{t-1} & \text{if } dN_{t-1} = 0 \\
\beta \lambda_{t-1} + \alpha & \text{otherwise}
\end{cases} \]

C : \[ \lambda_t = \mu + \gamma_1 \log(\overline{RV}_t) + \gamma_2 \log(\overline{RV}_{t-1}) + \gamma_3 \log(\overline{RCorr}_t) + \gamma_4 \log(\overline{RCorr}_{t-1}) + \begin{cases} 
\beta \lambda_{t-1} & \text{if } dN_{t-1} = 0 \\
\beta \lambda_{t-1} + \alpha & \text{otherwise}
\end{cases} \]

\begin{tabular}{lccccccc}
\hline
\mu & 0.2211 & 0.2224 & 0.2224 & 0.2224 & 0.2224 & 0.2224 & 0.2224 \\
 & (0.0337) & (0.0514) & (0.0334) & (0.0540) & (0.1042) & (0.1076) & (0.0429) \\
\alpha & 0.0212 & 0.0212 & 0.0212 & 0.0212 & 0.0212 & 0.0212 & 0.0212 \\
 & (0.0111) & (0.0103) & (0.0110) & (0.0110) & (0.0144) & (0.0133) & (0.0093) \\
\beta & 0.9161 & 0.9146 & 0.9146 & 0.9146 & 0.9146 & 0.9146 & 0.9146 \\
 & (0.0587) & (0.0569) & (0.0597) & (0.0687) & (0.0876) & (0.0884) & (0.0480) \\
\gamma_1 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
 & (0.0007) & (0.0007) & (0.0007) & (0.0007) & (0.0007) & (0.0007) & (0.0007) \\
\gamma_2 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
 & (0.0178) & (0.0178) & (0.0178) & (0.0178) & (0.0178) & (0.0178) & (0.0178) \\
\gamma_3 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
 & (0.0036) & (0.0036) & (0.0036) & (0.0036) & (0.0036) & (0.0036) & (0.0036) \\
\gamma_4 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
 & (0.0074) & (0.0074) & (0.0074) & (0.0074) & (0.0074) & (0.0074) & (0.0074) \\
log likelihood & 1425.7 & 1425.7 & 1425.7 & 1425.7 & 1425.7 & 1425.7 & 1425.7 \\
\hline
\end{tabular}

**Note:** This table reports the MLE estimates for co-jump intensity model without covariates or with two covariates: average realized volatility and realized correlation. The co-jumps, average realized volatility and realized correlation are constructed from the original five-minute returns spanning from 22 July 2002 to 5 December 2011, for a total of 2357 trading days. The standard errors are reported in parentheses. The last row labeled log likelihood reports the negative value of the sum of log likelihood function across time evaluated at the estimates.

Table 2: Estimation results of co-jump intensity model: Part A
\[ A: \lambda_t = \mu + \begin{cases} 
\beta \lambda_{t-1} & \text{if } dN_{t-1} = 0 \\
\beta \lambda_{t-1} + \alpha & \text{otherwise} 
\end{cases} \]

\[ B: \lambda_t = \mu + \gamma_1 \bar{RV}_t + \gamma_2 \bar{RV}_{t-1} + \gamma_3 \bar{RCorr}_t + \gamma_4 \bar{RCorr}_{t-1} + \begin{cases} 
\beta \lambda_{t-1} & \text{if } dN_{t-1} = 0 \\
\beta \lambda_{t-1} + \alpha & \text{otherwise} 
\end{cases} \]

\[ C: \lambda_t = \mu + \gamma_1 \log(\bar{RV}_t) + \gamma_2 \log(\bar{RV}_{t-1}) + \gamma_3 \log(\bar{RCorr}_t) + \gamma_4 \log(\bar{RCorr}_{t-1}) + \begin{cases} 
\beta \lambda_{t-1} & \text{if } dN_{t-1} = 0 \\
\beta \lambda_{t-1} + \alpha & \text{otherwise} 
\end{cases} \]

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>B1</th>
<th>B2</th>
<th>C</th>
<th>C1</th>
<th>C2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>0.221</td>
<td>0.157</td>
<td>0.157</td>
<td>0.157</td>
<td>0.157</td>
<td>0.157</td>
<td>0.157</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>(0.0341)</td>
<td>(0.0702)</td>
<td>(0.0377)</td>
<td>(0.0717)</td>
<td>(0.0737)</td>
<td>(0.0769)</td>
<td>(0.0645)</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.0072</td>
<td>0.0055</td>
<td>0.0055</td>
<td>0.0055</td>
<td>0.0055</td>
<td>0.0055</td>
<td>0.0055</td>
</tr>
<tr>
<td>( \gamma_1 )</td>
<td>(0.0024)</td>
<td>(0.0020)</td>
<td>(0.0020)</td>
<td>(0.0021)</td>
<td>(0.0017)</td>
<td>(0.0013)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td>( \gamma_2 )</td>
<td>(0.0204)</td>
<td>(0.0259)</td>
<td>(0.0244)</td>
<td>(0.0189)</td>
<td>(0.0201)</td>
<td>(0.0328)</td>
<td>(0.0371)</td>
</tr>
<tr>
<td>( \gamma_3 )</td>
<td>(0.0189)</td>
<td>(0.0048)</td>
<td>(0.0048)</td>
<td>(0.00492)</td>
<td>(0.0068)</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>( \gamma_4 )</td>
<td>(0.0185)</td>
<td>(0.0339)</td>
<td>(0.1396)</td>
<td>(0.1407)</td>
<td>(0.1396)</td>
<td>(0.1407)</td>
<td>(0.1475)</td>
</tr>
<tr>
<td>log likelihood</td>
<td>1363.5</td>
<td>1363.5</td>
<td>1363.5</td>
<td>1363.5</td>
<td>1363.5</td>
<td>1363.5</td>
<td>1363.5</td>
</tr>
</tbody>
</table>

**Note:** This table reports the MLE estimates for co-jump intensity model without covariates or with two covariates: average realized volatility and realized correlation. The co-jumps are constructed from the intra-day pattern adjusted five-minute returns, and average realized volatility and realized correlation are constructed from the original five-minute returns, spanning from 22 July 2002 to 5 December 2011, for a total of 2357 trading days. The standard errors are reported in parentheses. The last row labeled log likelihood reports the negative value of the sum of log likelihood function across time evaluated at the estimates.

Table 3: Estimation results of co-jump intensity model: Part B
3.2 Models for volatility and correlation: The role of co-jumps

With the widespread availability of high-frequency data in financial markets, recent literature has focussed on realized volatility as a measure of time-varying financial volatility. A number of empirical studies have argued for the importance of long-memory dependence in realized volatility. Andersen et al. (2007) eschew such complicated fractionally integrated long-memory formulations and rely instead on the simple-to-estimate Heterogeneous Autoregressive Realized Volatility (HAR-RV) model proposed by Corsi (2009). The HAR-RV formulation is based on a straightforward extension of the so-called Heterogeneous ARCH, or HARCH, class of models analyzed by Muller, Dacorogna, Dave, Olsen, Pictet, and Weizsacker (1997), in which the conditional variance of the discretely sampled returns is parameterized as a linear function of the lagged squared returns over the identical return horizon together with the squared returns over longer and/or shorter return horizons. By separately including the continuous sample path and individual jump variability measures in this simple linear volatility forecasting model, Andersen et al. (2007) find that only the continuous part has predictive power, in turn resulting in significant gains relative to the simple realized volatility forecasting models advocated in some of the recent literature.

Here we extend the work of Andersen et al. (2007) and investigate whether the occurrence of co-jumps are related to future volatility and correlation. To do this, we rely on the HAR model as Andersen et al. (2007) did, but rather than individual jump component, we incorporate the co-jumps instead into the model to explore their role. Meanwhile, as co-jumps (more for common factor model later on).

The HAR-RV model assumes that market dynamics are completely determined by the behavior of the participants, and current realized volatility and the expectations of volatility over longer horizons can effectively form volatility expectations for the next period. Therefore, the HAR model specifies the realized volatility as a component that contains a daily, weekly and monthly realized volatility component, and can be expressed as

\[
RV^d_t = \alpha_0 + \alpha_1 RV^d_{t-1} + \alpha_2 RV^w_{t-1} + \alpha_3 RV^m_{t-1} + \varepsilon_t,
\]

where \(RV^d_t\) denotes the daily realized volatility, \(RV^w_{t-1} = \frac{1}{5} \sum_{i=1}^{5} RV_{t-i}\) is the weekly realized volatility and \(RV^m_{t-1} = \frac{1}{22} \sum_{i=1}^{22} RV_{t-i}\) represents the monthly realized volatility.

Andersen et al. (2007) turned to a simple extension of the HAR-RV, in which they incorporate the time series of significant jumps as additional explanatory variables in a straightforward linear fashion, resulting in a new HAR-RV-J model as

\[
RV^d_t = \alpha_0 + \alpha_1 RV^d_{t-1} + \alpha_2 RV^w_{t-1} + \alpha_3 RV^m_{t-1} + \alpha_4 J_{t-1} + \varepsilon_t,
\]
to explore the predictive power of jump component (in a univariate index) for the following day’s volatility. Meanwhile, Andersen et al. (2007) further extended this model by explicitly decomposing the realized volatilities into the continuous sample path variability and the jump variation utilizing the separate nonparametric measurements based on a statistical jump test to the HAR-RV-CJ model, which is expressed as

\[
RV_t^d = \alpha_0 + \alpha_1 C_{t-1}^d + \alpha_2 C_{t-1}^{uw} + \alpha_3 C_{t-1}^m + \alpha_4 J_{t-1}^d + \alpha_5 J_{t-1}^{uw} + \alpha_6 J_{t-1}^m + \varepsilon_t, \tag{11}
\]

where \(C_t\) and \(J_t\) are respectively the continuous sample path variation and the jump variation at time \(t\), and \(C_{t-1}^{uw} = \frac{1}{5} \sum_{i=1}^{5} C_{t-i}, C_{t-1}^m = \frac{1}{22} \sum_{i=1}^{22} C_{t-i}, J_{t-1}^{uw} = \frac{1}{5} \sum_{i=1}^{5} J_{t-i},\) and \(J_{t-1}^m = \frac{1}{22} \sum_{i=1}^{22} C_{t-i}\).

We extend the HAR-RV-CJ model to the HAR-COJUMP model as

\[
RV_t^d = \alpha_0 + \alpha_1 RV_{t-1}^d + \alpha_2 RV_{t-1}^{uw} + \alpha_3 RV_{t-1}^m + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d \varepsilon_t, \tag{12}
\]

or

\[
RCorr_t^d = \alpha_0 + \alpha_1 RCorr_{t-1}^d + \alpha_2 RCorr_{t-1}^{uw} + \alpha_3 RCorr_{t-1}^m + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d \varepsilon_t, \tag{13}
\]

to investigate the predictive power of co-jumps for the averaged realized variance and correlation, where \(CJ_t\) is the co-jump component at time \(t\). We use two time series of \(CJ\) in the model. The first is a binary time series which takes the value of 1 when there is an occurrence of co-jump and otherwise 0. The other is a time series of co-jump intensities derived from the model discussed in Section 3.1. The former helps to answer whether the future volatility tends to increase or decrease when there is an occurrence of co-jump today, and the latter is useful to answer whether the future volatility tends to increase or decrease when the probability of the occurrence of co-jump is high today.

Moreover, we also utilize nonlinear HAR-COJUMP models as a robustness check,

\[
\log(RV_t^d) = \alpha_0 + \alpha_1 \log(RV_{t-1}^d) + \alpha_2 \log(RV_{t-1}^{uw}) + \alpha_3 \log(RV_{t-1}^m) + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d \varepsilon_t, \tag{14}
\]

\[
(RV_t^d)^{1/2} = \alpha_0 + \alpha_1 (RV_{t-1}^d)^{1/2} + \alpha_2 (RV_{t-1}^{uw})^{1/2} + \alpha_3 (RV_{t-1}^m)^{1/2} + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d \varepsilon_t. \tag{15}
\]

or

\[
\log(RCorr_t^d) = \alpha_0 + \alpha_1 \log(RCorr_{t-1}^d) + \alpha_2 \log(RCorr_{t-1}^{uw}) + \alpha_3 \log(RCorr_{t-1}^m) + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d \varepsilon_t, \tag{16}
\]

\[
(RCorr_t^d)^{1/2} = \alpha_0 + \alpha_1 (RCorr_{t-1}^d)^{1/2} + \alpha_2 (RCorr_{t-1}^{uw})^{1/2} + \alpha_3 (RCorr_{t-1}^m)^{1/2} + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d \varepsilon_t. \tag{17}
\]
Table 4 reports the estimation results of the linear and nonlinear HAR-COJUMP models with co-jumps as a dummy variable. The estimates for $\alpha_1$, $\alpha_2$ and $\alpha_3$ confirm the existence of highly persistent volatility and correlation series. More importantly, the estimates of the co-jump component and the lag of co-jump component are systematically positive and negative across all models, and overwhelmingly significant. For example, an unit increase in the daily average realized volatility implies an average increase in the volatility on the following day of $0.2247 + 0.4741/5 + 0.2156/22 = 0.3293$ for days without co-jumps, whereas for days in which there is an occurrence of co-jumps (where $CJ^d = 1$), the increase in the volatility on the same day is $0.00004$ unit higher on average, and the increase in the volatility on the following day is $0.00002$ unit lower on average. In other words, the average realized volatility is relatively higher on the days with co-jumps. Meanwhile, if a co-jump occurs, volatility will on average be lower on the following day. The same impact of co-jumps is observed in the HAR-RC-COJUMP models.

It is also interesting to explore the impact of the likelihood of co-jumps on the level of volatility and correlation. We estimate the intensity of co-jumps based on the model described in Section 3.1, and replace the co-jump dummy variable with the estimated intensity in HAR-COJUMP model. Table 5 reports the estimation results. Interestingly, the coefficients of the intensity of co-jumps and its lag are significantly positive and negative across all models. It implies that whereas the more likely co-jumps occur, the lower the average volatility and correlation are on the same day, the average volatility and correlation tend to be higher on the following day. Given the intensity increases immediately after a co-jump occurs then one would expected a change in signs in the coefficients on the contemporaneous and lagged co-jumps in comparison to the previous specification.

4 Conclusion

Modelling and forecasting the volatility of, and correlation between asset returns is vitally important for many financial problems. This literature has benefited from advances in the measurement of volatility based on the availability of high frequency financial data. One recent advance permits the total volatility of returns to be decomposed into it continuous diffusion and discrete jump components. Methods have also been developed for identifying common jumps, or co-jumps within a group of asset returns. Previous research finds that the continuous diffusion component of index returns is important for forecasting volatility and that jumps in such series are not important. Here we have extended this work to examine the links between co-jumps within a group of stocks, the volatility of, and correlation between their returns. We find that the occurrence of co-jumps within a group of 20 large U.S. stocks is unrelated to
\[ y^d_t = \alpha_0 + \alpha_1 y^d_{t-1} + \alpha_2 y^w_{t-1} + \alpha_3 y^m_{t-1} + \alpha_4 C.J^d_t + \alpha_5 C.J^d_{t-1} \]

\[ \log(y^d_t) = \alpha_0 + \alpha_1 \log(y^d_{t-1}) + \alpha_2 \log(y^w_{t-1}) + \alpha_3 \log(y^m_{t-1}) + \alpha_4 C.J^d_t + \alpha_5 C.J^d_{t-1} \]

\[ (y^d_t)^{1/2} = \alpha_0 + \alpha_1 (y^d_{t-1})^{1/2} + \alpha_2 (y^w_{t-1})^{1/2} + \alpha_3 (y^m_{t-1})^{1/2} + \alpha_4 C.J^d_t + \alpha_5 C.J^d_{t-1} \]

<table>
<thead>
<tr>
<th>( y = \text{ARV} )</th>
<th>( y = \log(\text{ARV}) )</th>
<th>( y = (\text{ARV})^{1/2} )</th>
<th>( y = \text{RC} )</th>
<th>( y = \log(\text{RC}) )</th>
<th>( y = (\text{RC})^{1/2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_0 )</td>
<td>0.00002</td>
<td>-0.2989</td>
<td>0.0006</td>
<td>0.0293</td>
<td>-0.0668</td>
</tr>
<tr>
<td></td>
<td>(0.00001)</td>
<td>(0.0815)</td>
<td>(0.0002)</td>
<td>(0.0081)</td>
<td>(0.0160)</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.2247</td>
<td>0.5040</td>
<td>0.4481</td>
<td>0.2809</td>
<td>0.2660</td>
</tr>
<tr>
<td></td>
<td>(0.0250)</td>
<td>(0.0239)</td>
<td>(0.0241)</td>
<td>(0.0246)</td>
<td>(0.0246)</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.4741</td>
<td>0.3424</td>
<td>0.3659</td>
<td>0.4183</td>
<td>0.4285</td>
</tr>
<tr>
<td></td>
<td>(0.0421)</td>
<td>(0.0346)</td>
<td>(0.0361)</td>
<td>(0.0409)</td>
<td>(0.0411)</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>0.2156</td>
<td>0.1184</td>
<td>0.1438</td>
<td>0.2319</td>
<td>0.2284</td>
</tr>
<tr>
<td></td>
<td>(0.0370)</td>
<td>(0.0249)</td>
<td>(0.0274)</td>
<td>(0.0358)</td>
<td>(0.0364)</td>
</tr>
<tr>
<td>( \alpha_4 )</td>
<td>0.00004</td>
<td>0.0397</td>
<td>0.0005</td>
<td>0.0190</td>
<td>0.0416</td>
</tr>
<tr>
<td></td>
<td>(0.00002)</td>
<td>(0.0150)</td>
<td>(0.0002)</td>
<td>(0.0037)</td>
<td>(0.0086)</td>
</tr>
<tr>
<td>( \alpha_5 )</td>
<td>-0.00002</td>
<td>-0.0510</td>
<td>-0.0004</td>
<td>-0.0122</td>
<td>-0.0321</td>
</tr>
<tr>
<td></td>
<td>(0.00001)</td>
<td>(0.0150)</td>
<td>(0.0002)</td>
<td>(0.0037)</td>
<td>(0.0086)</td>
</tr>
</tbody>
</table>

| \( R^2 \) | 0.5433 | 0.8315 | 0.7822 | 0.5959 | 0.5670 | 0.5848 |
| adjusted \( R^2 \) | 0.5423 | 0.8311 | 0.7817 | 0.5950 | 0.5661 | 0.5839 |

**Note:** This table reports the OLS estimates for HAR-COJUMP model with co-jumps as a dummy variable. The co-jumps, average realized volatility and realized correlation are constructed from the original five-minute returns, spanning from 22 July 2002 to 5 December 2011, for a total of 2357 trading days. The standard errors are reported in parentheses. The last two rows labeled \( R^2 \) and adjusted \( R^2 \) are r-square and adjusted r-square of the regression.

Table 4: Estimation results of HAR-COJUMP model: Part A
\[
y_t^d = \alpha_0 + \alpha_1 y_{t-1}^d + \alpha_2 y_{t-1}^w + \alpha_3 y_{t-1}^m + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d
\]

\[
\log(y_t^d) = \alpha_0 + \alpha_1 \log(y_{t-1}^d) + \alpha_2 \log(y_{t-1}^w) + \alpha_3 \log(y_{t-1}^m) + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d
\]

\[
(y_t^d)^{1/2} = \alpha_0 + \alpha_1 (y_{t-1}^d)^{1/2} + \alpha_2 (y_{t-1}^w)^{1/2} + \alpha_3 (y_{t-1}^m)^{1/2} + \alpha_4 CJ_t^d + \alpha_5 CJ_{t-1}^d
\]

<table>
<thead>
<tr>
<th>(y = RV)</th>
<th>(y = \log(RV))</th>
<th>(y = (RV)^{1/2})</th>
<th>(y = RC)</th>
<th>(y = \log(RC))</th>
<th>(y = (RC)^{1/2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_0)</td>
<td>0.0001</td>
<td>-0.2074</td>
<td>0.0020</td>
<td>0.0582</td>
<td>-0.0101</td>
</tr>
<tr>
<td>(\alpha_0)</td>
<td>(0.00009)</td>
<td>(0.0999)</td>
<td>(0.0010)</td>
<td>(0.0211)</td>
<td>(0.0433)</td>
</tr>
<tr>
<td>(\alpha_1)</td>
<td>0.2222</td>
<td>0.5018</td>
<td>0.4452</td>
<td>0.2781</td>
<td>0.2637</td>
</tr>
<tr>
<td>(\alpha_1)</td>
<td>(0.0250)</td>
<td>(0.0239)</td>
<td>(0.0242)</td>
<td>(0.0248)</td>
<td>(0.0248)</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>0.4726</td>
<td>0.3402</td>
<td>0.3643</td>
<td>0.4136</td>
<td>0.4230</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>(0.0422)</td>
<td>(0.0347)</td>
<td>(0.0362)</td>
<td>(0.0411)</td>
<td>(0.0413)</td>
</tr>
<tr>
<td>(\alpha_3)</td>
<td>0.2141</td>
<td>0.1201</td>
<td>0.1449</td>
<td>0.2350</td>
<td>0.2325</td>
</tr>
<tr>
<td>(\alpha_3)</td>
<td>(0.0371)</td>
<td>(0.0250)</td>
<td>(0.0274)</td>
<td>(0.0360)</td>
<td>(0.0366)</td>
</tr>
<tr>
<td>(\alpha_4)</td>
<td>-0.0009</td>
<td>-2.2583</td>
<td>-0.0195</td>
<td>-0.5138</td>
<td>-1.3803</td>
</tr>
<tr>
<td>(\alpha_4)</td>
<td>(0.0005)</td>
<td>(0.7094)</td>
<td>(0.0086)</td>
<td>(0.1761)</td>
<td>(0.4079)</td>
</tr>
<tr>
<td>(\alpha_5)</td>
<td>0.0005</td>
<td>1.8607</td>
<td>0.0150</td>
<td>0.4296</td>
<td>1.1871</td>
</tr>
<tr>
<td>(\alpha_5)</td>
<td>(0.0002)</td>
<td>(0.7099)</td>
<td>(0.0050)</td>
<td>(0.1762)</td>
<td>(0.4082)</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.5422</td>
<td>0.8310</td>
<td>0.7814</td>
<td>0.5914</td>
<td>0.5627</td>
</tr>
<tr>
<td>(\text{adjusted } R^2)</td>
<td>0.5412</td>
<td>0.8306</td>
<td>0.7810</td>
<td>0.5905</td>
<td>0.5618</td>
</tr>
</tbody>
</table>

**Note:** This table reports the OLS estimates for HAR-COJUMP model with the intensity of co-jumps. The co-jumps intensity is estimated from the model described in Section 3.1. The average realized volatility is constructed from the original five-minute returns, spanning from 22 July 2002 to 5 December 2011, for a total of 2357 trading days. The standard errors are reported in parentheses. The last two rows labeled \(R^2\) and adjusted \(R^2\) are r-square and adjusted r-square of the regression.

Table 5: Estimation results of HAR-COJUMP model: Part B
prevailing market conditions captured by volatility and correlation. On the other hand we find that subsequent to a co-jump, both volatility and correlation are significantly lower. This is a much stronger effect than that identified in previous research. This appears to indicate that co-jumps across a range of assets are a more influential event than jumps in an index which may be driven a small number of assets.
References


