Testing for cojumps in high-frequency asset prices

Markus Kösler*[a] and Hans Manner†[a]

[a]University of Cologne, Institute of Econometrics and Statistics

February 12, 2016

Abstract

We consider the problem of testing the synchronicity of jumps, i.e., the presence of cojumps, in financial high-frequency data. Our approach requires a univariate jump detection procedure that tests for and locates jumps and we propose a refinement of the approach by Lee and Mykland [*Review of Financial Studies*, 21(6), 2535–2563, 2008] via an iterative updating algorithm for the critical values in finite samples. We propose several test statistics for the null hypothesis of independent jump processes against the alternative of dependent jump processes. The test statistics compare the number of coexceedances with their expected number under the null of independent jump processes. Furthermore, we extend the tests to the context of non-homogeneous intensities using Hawkes processes. We derive the asymptotic distributions of these test statistics under double asymptotics, increasing both the observation frequency and the length of the sample. It is shown that the asymptotic distributions do not depend on the jump detection and the estimation of the unknown jump intensity. For finite samples we suggest a block bootstrap procedure that can mitigate the size distortions of the tests. A comprehensive Monte Carlo study examines the finite sample properties of the tests in a realistic market scenario and an application to financial high-frequency data illustrates its practical use.

JEL Classification: C12, C14, C32, G10, G11

Keywords: Cojumps, jump detection, high-frequency data, Hawkes processes

1 Introduction

Today, jumps are a widely accepted characteristic of financial markets. Though rare, they tend to recur and can be devastating to investors. In the recent decade or so, a large field of literature has emerged that studies their presence, dynamics and dependence. The consequences of the presence of jumps are manifold. The most fundamental implication is that markets are no longer complete, as self-financing replication strategies for derivatives break down. Thus, jumps may, in part, explain the mere presence of derivates markets, which would be redundant under price diffusions with stochastic volatility only. Cont and Tankov (2004) offer a detailed analysis of the conditions and present alternative hedging strategies under market incompleteness. Eraker et al. (2003) study implications on option prices and find that jumps in prices and volatility steepen

*Corresponding author. Tel.: +49 2214704267; E-mail: markus.koesler@statistik.uni-koeln.de
†Tel.: +49 2214704130; E-mail: manner@statistik.uni-koeln.de

the slope of the implied volatility curve and drive up its level, respectively. Pan (2002) introduces jumps as an additional factor in the variance risk of an asset and isolate a seperate jump-risk premium of about 3.5% excess return. Bollerslev and Todorov (2011) further strengthen these results and find that fear of extreme tail events attributes for more than half of the variance risk premium in aggregate market returns. The time dynamics of jump-risk premia are studied by Todorov (2010). He finds that large jumps trigger a persistent increase in the jump-risk premium that slowly mean-reverts and suggests investors adjusting their risk attitudes as an explanation.

Several approaches have been developed to test for the presence of jumps. Barndorff-Nielsen and Shepard (2004, 2006) were the first to suggest a formal jump detection by identifying the jump contribution to integrated volatility (IV) under fill-in asymptotics through the difference of realized variation measures and Bipower Variation (BPV), which is a jump-robust IV estimator. Several subsequent contribution fall into the same class of IV disentanglement tests, but improve upon BPV in efficiency and small sample properties, including Mancini (2009), Corsi et al. (2010), Andersen et al. (2012) and Bos and Janus (2013). Bibinger and Reiß (2014) propose a spectral estimator for integrated (co)volatility, which readily extends to the bivariate case.

Andersen et al. (2007) and Lee and Mykland (2008) rely on standardizing returns by a local jump-robust volatility estimate. Their approaches differ in the subsequent determination of critical values using Bonferroni correction and Extreme Value Theory, respectively. Aït-Sahalia and Jacod (2009) consider the convergence of a ratio of BPVs on different time scales , while Jiang and Oomen (2008) exploit the difference of simple and logarithmic returns under jumps. A detailed review of univariate jump tests can be found in Dumitru and Urga (2012).

In the multivariate setting, jumps do not lose their relevance. On the contrary, one cannot hope for jumps to dilute in a large portfolio except in the special case when jumps' sizes, signs and timing are independent. If extreme events arrive synchronously as cojumps, they constitute a form of systematic risk that cannot be diversified or hedged away. Cojumps have gained attention in related fields, which approach the topic from different angles. Bormann and Schaumburg (2016) propose a test for higher-order tail risk. Grothe et al. (2014) develop a peak-over-threshold model with self-exciting intensity to model dependencies and clustering of extreme events. Bivariate tail dependencies between aggregate market jumps and simultaneous idiosyncratic jumps are estimated by Bollerslev et al. (2013).

Albeit the relevance, the multivariate cojump literature is scarcer. In an early approach, Bollerslev et al. (2008) show that the BPV test isolates cojump contributions in an equiweight portfolio for the number of assets going to infinity. They also propose a simulative procedure based on the distribution of the mean cross-product of returns. A drawback of this approach is that it relies on the estimation of local covariance matrices. Therein, one faces a trade-off between the three features jump-robustness, robustness to microstructure and asychronicity, and the required positive semi-definiteness of the estimated covariance matrix. Jacod and Todorov (2009) propose bivariate tests for both the null hypothesis of disjoint jumps and for the null hypothesis of joint jumps. Mancini and Gobbi (2012) generalize the original BPV test by identifying bivariate jumps as the difference of realized covariation and a threshold estimator of integrated covariation. A spectral approach for the bivariate case is proposed by Bibinger and Winkelmann (2015). More recently, Caporin et al. (2015) propose a test for the presence of d-dimensional cojumps versus the alternative of cojumps of dimensions zero up to d by

comparing two appropriate measures of power variations that, however, is sensible to the choice of a bandwidth parameter. Gilder et al. (2014) propose a simple coexceedance rule, wherever two or more univariate jump detections in the same time interval are considered cojumps. Our approach builds on this latter approach, but formalizes it by testing the statistical significance of the coexceedances.

This article provides the following contributions to the literature. We propose two refinements for the univariate jump detection procedure by Lee and Mykland (2008) (henceforth LM). First, instead of BPV we employ the more robust Median-RV estimator for local volatility by Andersen et al. (2012) and show that the asymptotic properties of the procedure are preserved. Second, we develop a novel iterative algorithm which determines the jump classification threshold that jointly minimizes spurious detections and detection failures in a data driven manner. Third, we propose a testing scheme that jointly tests for cojumps in any subset of a (possibly large) d-variate portfolio, while still allowing assets to have (partly) idiosyncratic jump processes. The tests can easily be adapted to only test against, e.g., the alternative of cojumps of a certain degree. Our approach relies on univariately identifying jump times and comparing the distribution of joint detections over discrete observation intervals (coexceedances) to their expected distribution under the null of independent Poisson jump processes. By that, we provide a distributional theory for the coexceedance approach by Gilder et al. (2014). Recent works by Aït-Sahalia et al. (2015) and Bormetti et al. (2015) cast doubt on the assumption of homogeneity of jump processes. For this reason we generalize our procedure to allow for self-exciting (multivariate) Hawkes processes for the jump intensities. We propose several test statistics and derive their asymptotic distribution under double asymptotics, letting both observation frequency and sample size go to infinity. As a finite sample refinement we consider a (block) bootstrap algorithm, which delivers promising results in Monte Carlo simulations. Finally, an empirical application to a large data set for stocks in the S & P 500 index illustrates our approach and shows how it can be used to construct portfolios with reduced cojump risk. At the same time we show that it appears impossible to completely remove the risk of cojumps by diversification.

The paper is organized in four sections. Section 2 introduces the theoretical framework and our test statistics. In Section 3 Monte Carlo simulations illustrate the size and power in finite samples. Section 4 highlights the empirical usefulness of our test in an empirical application, whereas Section 5 concludes. Proofs are collected in the Appendix.

2 Theoretical framework

In this section we develop the methodology for testing for co-jumps. Section 2.2 reviews the univariate jump detection procedure suggested by Lee and Mykland (2008) and proposes an adjustment for finite samples. Section 2.3 presents a set of test statistics for the null hypothesis of independent jump times. Possible generalizations for the case of non-homogeneous jump processes are considered in section 2.5.

2.1 The underlying model

For the univariate case, we denote by $Y_t^{(i)}$ the log price of asset i, with $i = 1, \dots, d$, in the observation period $t \in [0, T]$ at continuous time t. We assume that the log price follows a

Brownian semimartingale with finite activity jumps[1] (as, e.g., in Barndorff-Nielsen et al., 2006)

$$Y_t^{(i)} = Y_0^{(i)} + \int_0^t a_s^{(i)} ds + \int_0^t \sigma_s^{(i)\prime} dW_s + \sum_{k=1}^{N_t^{(i)}} c_k^{(i)}, \tag{1}$$

where $Y_0^{(i)}$ is the log price at the beginning of the observation period[2], a_t is a càdlàg adapted drift term, σ_t is a possibly vector-valued càdlàg, adapted stochastic volatility (SV) process bounded away from 0 and ∞ and W_t is a vector of Brownian Motion of suitable dimension with covariance matrix Σ_t. The c_k are real-valued random variables and, for now, we assume $N_t^{(i)}$ to be a homogeneous Poisson process of finite intensity λ counting the number of jump arrivals up to time t, such that $\mathbb{E}[N(T)] = \lambda T < \infty$ for all $T < \infty$. Note that the last term in (1) constitutes a *compound* Poisson process if the c_k are i.i.d. Also, if $c_k = 0, \forall k$ or $N(t) = 0$ for all t, then we are in the case of no jumps and (1) constitutes a regular drift-diffusion process. The stochastic volatility (SV) process is given by

$$\sigma_t^{(i)} = \sigma_0^{(i)} + \int_0^t a_s^{(i)} ds + \int_0^t \sigma_s^{(i)\prime} dW_s + \int_0^t \vartheta_s^{(i)\prime} dZ_s, \tag{2}$$

where σ_0 is a constant, a is drift term, $\sigma(s)$ and $\vartheta(s)$ are adapted and càdlàg volatility-of-volatility processes, Z_t is a Brownian Motion vector with dimensions suiting $\vartheta(s)$ with arbitrary covariance structure independent of W_t. This allows the scale factor σ_t of the driving Brownian Motion in (1) to depend in some fashion on the driver itself. Therefore the model potentially captures leverage between price increments and their volatility. Observe that by assuming (2) we do not allow for jumps in the SV process.

For the multivariate case, the above notation extends to

$$\mathbf{Y}_t = \mathbf{Y}_0 + \int_0^t \mathbf{a}_s ds + \int_0^t \boldsymbol{\sigma}_s dW_s + \mathcal{J}_t \tag{3}$$

with $\mathbf{Y}_t$ the d-dimensional vector of asset prices, $\mathcal{J}_t = \left(\sum_{k=1}^{N_t^{(1)}} c_k^{(1)}, ..., \sum_{k=1}^{N_t^{(d)}} c_k^{(d)} \right)'$ is the d-dimensional jump process. Note that we allow for the possibility that $\Delta N_{t-}^{(i_1)} = \Delta N_{t-}^{(i_2)} = 1$ for some $i_1 \neq i_2$ and t, i.e. assets may cojump. The multivariate SV process now reads as,

$$\boldsymbol{\sigma}_t = \boldsymbol{\sigma}_0 + \int_0^t \boldsymbol{a}_s' ds + \int_0^t \boldsymbol{\sigma}_s' dW_s + \int_0^t \boldsymbol{\vartheta}_s dZ_s, \tag{4}$$

where all terms are as in (2), but $\boldsymbol{\sigma}_0, \boldsymbol{a}', \boldsymbol{\sigma}'$ are now arrays of appropriate dimensions.

In the remainder of this paper we assume to have synchronous observations, that is $M = \lfloor \frac{T}{\delta} \rfloor$ observations on an equidistant partition of the observation period $[0, T]$ with step size δ for all d assets and $\lfloor \cdot \rfloor$ is the floor function which returns the largest integer smaller or equal to its argument. The symbol $\circ$ denotes the Hadamard product (elementwise vector/matrix product). In this setting we consider three types of asymptotic arguments: Firstly, the δ-*asymptotics*, where for fixed T, $\delta \to 0$, i.e., the observation frequency going to infinity, which implies $M = \lfloor \frac{T}{\delta} \rfloor \to \infty$. Secondly, the T-*asymptotics*, where $T \to \infty$, but δ remains constant. Note that T-asymptotics

[1] Also called jump-diffusion

[2] Henceforth this constant will be set to zero without loss of generality as an identical analysis could be conducted on the transformed log price $\tilde{y}_t := y_t - y_0$

imply $M = \frac{T}{\delta} \to \infty$ as a function of T. Lastly, the *joint asymptotics*, where both $\delta \to 0$ and $T \to \infty$, but such that $\delta T \to 0$, that is δ approaches zero faster than T goes to infinity. We denote "$\xrightarrow{p}$" for convergence in probability, "$\xrightarrow{d}$" for convergence in distribution, "$\overset{asy}{=}$" for asymptotic equivalence and "$\overset{asy.law}{=}$" for asymptotic equivalence in law. Throughout the text we impose the following assumption, which ensures sufficient "smoothness" of the drift and volatility process:

Assumption 1.

$$\sup_i \sup_{t_i \le u \le t_{i+1}} |a_u - a_{t_i}| = O_p(\delta^{1/2-\epsilon})$$
$$\sup_i \sup_{t_i \le u \le t_{i+1}} |\sigma_u - \sigma_{t_i}| = O_p(\delta^{1/2-\epsilon}).$$

2.2 Univariate jump detection

Lee and Mykland (2008) propose detecting jumps by identifying "unusually" large returns after standardization by a jump-robust local volatility estimate. The test statistic for the i-th asset on which the jump detection is based reads as

$$\mathcal{L}_j^{(i)} = \frac{\left| y_j^{(i)} \right|}{\hat{\sigma}_j^{(i)}}, \tag{5}$$

where $y_j^{(i)} = Y_{j\delta}^{(i)} - Y_{(j-1)\delta}^{(i)}$ is the j-th log return of the i-th asset and $\hat{\sigma}_j^{(i)}$ is a jump robust local volatility estimator calculated from the K observations preceding the respective return, with $K = O_p(\delta^a),\ a \in (-1, -\frac{1}{2})$.[3] For the normalization Lee and Mykland (2008) originally rely on realized bipower variation (BPV) by Barndorff-Nielsen and Shepard (2006), but due to the upward bias of BPV in finite samples, we propose a robustification by employing the Median Realized Variance (MED-RV) estimator of Andersen et al. (2012). Then the local volatility estimator is given by

$$\hat{\sigma}_j^{(i)} = \sqrt{\vartheta \frac{K}{(K-1)(K-2)} \sum_{l=0}^{K-2} \text{med}(|y_{j-l}^{(i)}|, |y_{j-l-1}^{(i)}|, |y_{j-l-2}^{(i)}|)^2},$$

where $\vartheta = \frac{\pi}{6-4\sqrt{3}+\pi}$. Note that, as the volatility of the j-th return is calculated over the K preceeding returns, no local volatility estimate can be obtained for the first K observations, such that the effective sample for which jumps can be detected is truncated to $[(K+1)\delta, T]$. Throughout the remainder of this paper we denote with $\widetilde{M} := M - K$ the number of available standardized returns.

We summarize the asymptotic behavior of $\mathcal{L}_j^{(i)}$ when using the MED-RV estimator in Lemma 1.

Lemma 1. *If $Y^{(i)}$ follows (1) we have under the δ-asymptotic for $j = K+1, ..., M$*

$$\textit{under } H_0^{(1)} \textit{ of no jump:} \qquad \mathcal{L}_j^{(i)} \to_d \left| U_j^{(i)} \right|$$
$$\textit{under } H_1^{(1)} \textit{ of jumps in } \big((j-1)\delta, j\delta\big]: \qquad \mathcal{L}_j^{(i)} \to \left| U_j^{(i)} \right| + \frac{\left| c_{(j-\epsilon)\delta}^{(i)} \right|}{\int_{(j-1)\delta}^{j\delta} \sigma_s^{(i)} ds \sqrt{\delta}} \to \infty.$$

[3] In accordance with this recommendation we chose $K = 312$ when working with 1-minute returns.

where $U_j^{(i)} \overset{iid}{\sim} \mathcal{N}(0,1)$, $\epsilon \in [0,1)$ *such that* $c_{j-\epsilon\delta}^{(i)}$ *denotes the size of the jump in the i-th asset that occurred in the (infinitesimally small) sample interval* $\big((j-1)\delta, j\delta\big]$ *and* $\sigma_{j\delta}^{(i)}\sqrt{\delta}$ *is the approximate standard deviation of the diffusive increment* $\int_{(j-1)\delta}^{j\delta} \sigma_s^{(i)} dW_s^{(i)}$.

Since we test for jumps at every observation in the sample, the overall critical value needs to be suitably adjusted. Hence, the rejection region is based on the distribution of the maximum of the test statistics $\mathcal{L}_i$, which under a suitable standardization converges to a standard Gumbel variable (Galambos, 1987). Then, in the absence of jumps we have under the δ-asymptotics

$$\frac{\max_j |\mathcal{L}_j^{(i)}| - C_{\widetilde{M}}}{S_{\widetilde{M}}} \xrightarrow{d} \xi, \qquad j \in \{K+1, ..., M\}, \tag{6}$$

where

$$\begin{aligned} F_\xi(x) &= P(\xi \leq x) = \exp(-e^{-x}), \\ C_{\widetilde{M}} &= \frac{(2\ln \widetilde{M})^{1/2}}{\mu_1} - \frac{\ln 4\pi + \ln(\ln \widetilde{M})}{2\mu_1 (2\ln \widetilde{M})^{1/2}}, \\ \text{and} \qquad S_{\widetilde{M}} &= \frac{1}{(2\ln \widetilde{M})^{1/2}}, \end{aligned}$$

and $\mu_1 = \mathbb{E}(|U_j^{(i)}|) = \sqrt{\frac{2}{\pi}}$.

Therefore, the univariate jump detection procedure considers the normalized absolute returns

$$LM_j^{(i)} = \frac{|\mathcal{L}_j^{(i)}| - C_M}{S_M}. \tag{7}$$

Whenever $LM_j^{(i)}$ exceeds the appropriate quantile of F_ξ, denoted by β, the corresponding observations are classified as jumps. Formally, the jump indicator is defined as

$$\widehat{\Delta N_j^{(i)}}(\beta) = \begin{cases} 1 & \text{if } LM_j^{(i)} > \beta \\ 0 & \text{if } LM_j^{(i)} \leq \beta. \end{cases} \tag{8}$$

Intuitively speaking, all returns that are larger than some quantile of the maximum distribution of test statistics under the null are considered jumps.

A challenge to our approach of testing for cojumps is the stacking of different testing procedures. In particular, our approach relies on the univariately detected jumps which underlie Type-I and Type II-errors. Therefore, a correct overall size of the univariate jump detection is less relevant than an optimal balance of both error types. Asymptotically in δ, the probability of spurious detection (type-I error) is of order $e^{-\beta}$, while for $\beta \to \infty$ the probability of failing to detect a jumpe (type-II error) vanishes slower than $\sqrt{\frac{T}{\delta}\ln\frac{T}{\delta}}$. This means that for $\delta \to \infty$ we can let the critical value for jump detection β go to infinity with an appropriate rate and achieve perfect jump identification δ-asymptotically. In finite samples the minimizer of the sum of the two error probabilities is given by (Lee and Mykland, 2008)

$$\beta^* = \underset{\beta}{\text{argmin}} \Big(P(\text{Type-I error}) + P(\text{Type-II error}) \Big)$$

$$= -\log\left(\frac{\sigma\mu_1^2\sqrt{T}N}{\sqrt{2\widetilde{M}\ln\widetilde{M}}}\right), \tag{9}$$

where $\sigma = \sqrt{\frac{1}{T}\int_0^T \sigma_s^2 ds}$ is the unconditional volatility. This implies that in finite samples the error minimizing β^* depends on the unknown number of realized jumps N and is therefore infeasible to choose from the outset. To improve upon an arbitrary choice of β we propose an iterative updating procedure for β to minimize the joint misclassification probability in finite samples.

The algorithm works as follows:

1. Obtain a jump robust estimate $\hat{\sigma}$ of the volatility σ (e.g. by MED-RV) and compute the statistic LM_j for $j = K+1, \ldots, M$.

2. Choose arbitrary first critical value $\beta_0 \in [-\infty, \max_{j\in\{K+1,\ldots,M\}} LM_j^{(i)})$.

3. Run univariate jump detection and compute the detected number of jumps for given β_0 as

$$\widehat{N}_0 = \widetilde{M}(1 - \mathbb{F}_m(\beta_0)), \tag{10}$$

   where $\mathbb{F}_m$ is the empirical distribution function of $LM_j^{(i)}$ under the jump alternative[4].

4. Update β as

$$\beta_{n+1} = -\log\left(\frac{\hat{\sigma}\sqrt{T}\widehat{N}_n}{\sqrt{2\widetilde{M}\ln\widetilde{M}}}\right) = -\log\left(\frac{\hat{\sigma}\sqrt{T}\widetilde{M}(1-\mathbb{F}_m(\beta_n))}{\sqrt{2\widetilde{M}\ln\widetilde{M}}}\right), \tag{11}$$

5. Stop if $\beta_n = \beta_{n-1}$ for some n and call this final value $\hat{\beta}^*$.

Proposition 1. *If prices follow* (1) *and jump sizes follow a bimodal normal mixture with symmetric means and variance, then*

$$\beta_n \to \beta^*,$$

as $\delta \to 0, n \to \infty$.

Remark The assumption of a symmetric bimodal normal mixture for the jump size distributions is justified by our empirical finding; see Figure 6 in Section 4. Also note that even though Proposition 1 only provides an asymptotic result with the observation frequency approaching infinity, in finite samples the critical value $\hat{\beta}^*$ that is chosen by our algorithm approximates the unknown β^*. Simulations show that this works surprisingly well in finite samples and does not only improve univariate jump detection substantially, but also leads to much better results for our cojump testing procedure that we introduce in the next section. Table 1 illustrates the performance of our algorithm results compared to fixed quantiles from the Gumbel distribution.

[4] In particular, under univariate jumps, $LM_j^{(i)}$ follows a mixture distribution composed of the null distribution of normalized diffusive increments and the distribution of the returns that are affected by jumps with weights $\frac{\widetilde{M}-N_T^{(i)}}{\widetilde{M}}$ and $\frac{N_T^{(i)}}{\widetilde{M}}$, respectively.

Table 1: Bias, MSE and Concordance for iteratively updated jump classification threshold versus fixed quantiles.

	λ	1				3				10			
κ		$\hat{\beta}^*$	$\beta_{0.9}$	$\beta_{0.95}$	$\beta_{0.99}$	$\hat{\beta}^*$	$\beta_{0.9}$	$\beta_{0.95}$	$\beta_{0.99}$	$\hat{\beta}^*$	$\beta_{0.9}$	$\beta_{0.95}$	$\beta_{0.99}$
	Bias	1.79	0.77	1.49	3.12	3.19	1.90	2.62	4.25	4.45	3.13	3.85	5.48
2	MSE	4.14	0.75	2.37	9.88	11.59	3.76	7.01	18.20	21.21	9.94	14.96	30.17
	Con	0.00	0.01	0.01	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.00
	Bias	1.83	0.79	1.51	3.14	1.61	1.87	2.59	4.22	1.46	3.10	3.82	5.45
4	MSE	3.52	0.79	2.44	10.03	2.67	3.65	6.87	17.98	2.18	9.78	14.76	29.88
	Con	0.17	0.20	0.18	0.13	0.20	0.19	0.17	0.13	0.22	0.17	0.15	0.11
	Bias	0.27	0.78	1.50	3.13	0.22	1.89	2.61	4.24	0.15	3.08	3.80	5.43
8	MSE	0.08	0.76	2.40	9.95	0.05	3.73	6.98	18.15	0.03	9.66	14.62	29.67
	Con	0.76	0.74	0.72	0.67	0.78	0.73	0.71	0.66	0.78	0.68	0.65	0.60
	Bias	0.02	0.79	1.51	3.14	-0.01	1.91	2.63	4.26	-0.18	3.08	3.80	5.43
16	MSE	0.00	0.79	2.45	10.04	0.00	3.79	7.06	18.28	0.03	9.63	14.57	29.61
	Con	0.96	0.96	0.96	0.95	0.96	0.96	0.95	0.94	0.95	0.93	0.92	0.90
	Bias	-0.01	0.78	1.50	3.13	-0.09	1.88	2.60	4.23	-0.52	3.10	3.82	5.45
32	MSE	0.00	0.76	2.39	9.92	0.01	3.68	6.90	18.03	0.28	9.77	14.76	29.87
	Con	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.97	0.97	0.96

Results simulated under model described in section 3. Concordance is defined as the intersection of true (simulated) jumptimes and detected jumptimes based on $\hat{\beta}^*, \beta_{0.9}, \beta_{0.95}$ and $\beta_{0.99}$, respectively, divided by their union. The (1-α)-quantiles β_α of the standard Gumbel distribution are given by $-\log(-\log(1-\alpha))$ such that $\beta_{0.9} = 2.2504, \beta_{0.95} = 2.9702$ and $\beta_{0.99} = 4.6001$.

2.3 Testing for cojumps

For now we assume the univariate jump time processes $N_t = (N_t^{(1)}, ..., N_t^{(d)})$ to be a homogeneous Poisson processes with constant intensity $\lambda = (\lambda^{(1)}, ..., \lambda^{(d)})$. We relax this assumption in Section 2.5. Denoting the iteratively updated critical values defined above by $\hat{\beta}_i^*$,

$$\widehat{\Delta \boldsymbol{N}_j}(\hat{\boldsymbol{\beta}}^*) := \left[\widehat{\Delta N_j^{(1)}}(\hat{\beta}_1^*), .., \widehat{\Delta N_j^{(d)}}(\hat{\beta}_d^*)\right] = \left[\mathbb{1}_{\{LM_j^{(1)} > \hat{\beta}_1^*\}}, ..., \mathbb{1}_{\{LM_j^{(d)} > \hat{\beta}_d^*\}}\right] \tag{12}$$

is a vector indicating univariate LM-test results in the j-th sample interval, $((j-1)\delta, j\delta]$,i.e., a d-variate jump indicator, while

$$\widehat{\boldsymbol{J}} = \left[\widehat{\Delta \boldsymbol{N}_1}(\hat{\boldsymbol{\beta}}^*)', ..., \widehat{\Delta \boldsymbol{N}_{\widetilde{M}}}(\hat{\boldsymbol{\beta}}^*)'\right]' \tag{13}$$

is an $\widetilde{M} \times d$ matrix of estimated jump *times*. We are interested in testing the null hypothesis of independent jump processes:

$$\begin{aligned}
&H_0^{(2)} : \text{Jump times are independent} \\
&\quad \forall t : \ \Delta N_{t-}^{(1)} \neq \Delta N_{t-}^{(2)} \neq ... \neq \Delta N_{t-}^{(d)}, \\
&H_1^{(2)} : \text{Jump times are dependent} \\
&\quad \exists t : \ \Delta N_{t-}^{(i_1)} = \Delta N_{t-}^{(i_2)} = ... = \Delta N_{t-}^{(i_k)} = 1 \\
&\qquad \text{for } \{i_1, ..., i_k\} \text{ some index set } I \text{ with } I \subseteq \{1, ..., d\} \text{ and } |I| = k.
\end{aligned}$$

Table 2: Univariate Test Results

asset\period	K+1,	K+2,	...,	j,	...,	$M = \lfloor \frac{T}{\delta} \rfloor$	rate
asset 1	0			0	$\cdots$	1	λ_1
asset 2	0			1			
				$\vdots$			
$\vdots$	$\vdots$			0			$\vdots$
				$\vdots$			
asset d	0			1	$\cdots$	0	λ_d
column sum	g_1			g_j		$g_{\widetilde{M}}$	

Table 2 visualizes the information available after the univariate jump detection (collected in $\hat{\boldsymbol{J}}$). The cells indicate the intervals for which jumps have been detected for the corresponding assets. Let g_j be the column sums, counting the number of jump detections in the same interval,

$$g_j := \sum_{i=1}^{d} \widehat{\Delta N_j^{(i)}} = \begin{cases} d & \text{systematic cojump} \\ d-1 & \\ \vdots & \text{partial cojumps} \\ 2 & \\ 1 & \text{idiosyncratic jump} \\ 0 & \text{no jump.} \end{cases} \tag{14}$$

Then $g_j = 0$ means no jump occurred in the j-th interval, while $g_j = 1$ indicates an idiosyncratic jump, $2 \leq g_j < d$ a partial cojump and $g_j = d$ a systematic (or market-wide) cojump. In general, if $g_j = k,\ k \in \{0, 1, ..., d\}$, we say we detected a *cojump of extent k (or k-cojump)* at interval $((j-1)\delta, j\delta]$. Gilder et al. (2014) study cojumps under the assumption that a joint exceedance of LM statistics above their respective jump classification threshold is equivalent to the presence of cojumps in the respective assets, i.e., they assume $P\left(\Delta N_{t-}^{(i_1)} > 0, ..., \Delta N_{t-}^{(i_m)} > 0 \middle| g_j = k\right) = 1,\ t \in ((j-1)\delta, j\delta]$ for some $\delta > 0$. This may be justified asymptotically, as Lee and Mykland (2008) show the probability of global misclassification (spurious detection or failure to detect) vanishes in large samples as long as the critical values β are allowed to increase as $\delta \to 0$. In finite samples with data observed over small discrete intervals, however, there is always a positive probability of observing jumps in different assets in the same observation interval by chance alone. Therefore, we propose to test for the presence of cojumps by comparing the relative frequency of coexceedances with their probability under the null hypothesis of independent jumps times. Our basic (infeasible) test statistic is given by

$$Z_0 = \sqrt{n} \sum_{k=2}^{d} (\mathbb{P}_k - P_k), \tag{15}$$

where P_k is the (unknown) probability of a k-cojump and $\mathbb{P}_k$ is the empirical frequency of

k-cojumps detected in the sample,

$$\mathbb{P}_k = \frac{1}{\widetilde{M}} \sum_{j=1}^{\widetilde{M}} \mathbb{1}_{\{g_j = k\}}. \tag{16}$$

Intuitively, Z_0 measures the combined probability mass exceeding the expected probabilities under the null in the right tail of the empirical distribution of g. Since the $P_i, i \geq 2$ are typically small for rare jumps, cases of empirical "shortcomings", i.e., large-extent cojumps happening less often than expected can hardly affect the test statistic.

To compute P_k, note that g_j is the sum of d independent Bernoulli experiments, each with probability p_i of success (asset i jumps in period j), such that for a single Bernoulli experiment with $i = 1, ..., d$

$$\Delta N_j^{(i)} = \begin{cases} 1 & \text{with prob } p_i \\ 0 & \text{with prob } 1 - p_i, \end{cases} \tag{17}$$

where $p_i = P\left(N_{j\delta}^{(i)} - N_{(j-1)\delta}^{(i)} > 0\right) = 1 - P\left(N_{j\delta}^{(i)} - N_{(j-1)\delta}^{(i)} = 0\right) = 1 - (1 - e^{-\lambda_i \delta}) \approx \lambda_i \delta$ if $\lambda_i \delta$ is small. By the memorylessness property of Poisson processes, the $\Delta N_j^{(i)}$ have the same probability generating function (PGF) for all j:

$$G_{\Delta N_j^{(i)}}(z) = \mathbb{E}[z^{\Delta N_j^{(i)}}] = (1 - p_i) + p_i z = (1 - \delta\lambda_i) + \delta\lambda_i z. \tag{18}$$

As g_j is a sum of independent random variables $\Delta N_j^{(i)}$ under the null, the PGF of g_j is the product of the PGFs of its summands

$$G_{g_j}(z) = \prod_{i=1}^{d} G_{\Delta N_j^{(i)}}(z) = \prod_{i=1}^{d} \Big((1 - \delta\lambda_i) + \delta\lambda_i z \Big) \tag{19}$$

and we obtain the probability P_k of a k-cojump, $k \in \{0, ..., d\}$, under the null as the k-th derivative of the PGF,

$$P_k = P(g_j = k) = \frac{1}{k!} \frac{\partial^k G_{g_j}(z)}{\partial z^k}\bigg|_{z=0}. \tag{20}$$

This probability depends on the unknown intensities λ_i, for which we employ th simple MLE estimator $\hat{\lambda}_i = \frac{\widehat{N}_i}{T}$, resulting in the estimated probabilities $\hat{P}_k$.

An attractive feature of the first-step jump detection is that their global probability of misclassification converges to zero quickly as $\delta \to 0$; see Theorems 3 and 4 in Lee and Mykland (2008). That means that for high observation frequencies we can treat the jumps as perfectly observed. The following proposition states sufficient conditions for identifying P_k.

Proposition 2. *If prices follow* (1)*, we have and as long as* $T, M \to \infty$ *such that* $T/M \to 0$*, Proposition 1 applies and thus*

$$\frac{\widehat{N}_i}{T} =: \hat{\lambda}_i \xrightarrow{p} \lambda_i,$$

and therefore

$$\hat{P}_k \xrightarrow{p} P_k,\ k = 0, ..., d.$$

This allows us to define the feasible test statistic

$$Z = \sqrt{\widetilde{M}} \sum_{k=2}^{d} (\mathbb{P}_k - \hat{P}_k). \tag{21}$$

There are several alternative test statistics that we consider. Whereas Z compares the relative frequency of cojumps to their theoretical counterparts, Z_1 relies on the idiosyncratic jumps instead:

$$Z_1 = \sqrt{\widetilde{M}} \big(\mathbb{P}_1 - \hat{P}_1\big) \tag{22}$$

Intuitively, when there are too many cojumps there should be fewer idiosyncratic jumps than expected under the null and Z_1 is designed to detect this. Next, the statistic Z^2 is a conventional chi-square statistic that considers all types of jumps,

$$Z^2 = \widetilde{M} \sum_{k=0}^{d} \left(\frac{\mathbb{P}_k - \hat{P}_k}{\sqrt{\hat{P}_k}} \right)^2. \tag{23}$$

The last two statistics are defined similarly, but weigh the deviations of the empirical frequencies from their expected values either by the standard deviation or the variance. In this way more weight is attached to empirical deviations of high extent.

$$Z_{sd}^2 = \widetilde{M} \sum_{k=0}^{d} \left(\frac{\mathbb{P}_k - \hat{P}_k}{\sqrt{\hat{P}_k(1 - \hat{P}_k)}} \right)^2 \tag{24}$$

$$Z_{var}^2 = \widetilde{M} \sum_{k=0}^{d} \left(\frac{\mathbb{P}_k - \hat{P}_k}{\hat{P}_k(1 - \hat{P}_k)} \right)^2. \tag{25}$$

For convenience of notation let $P = (P_0, ..., P_d)$ and $\mathbb{P} = (\mathbb{P}_0, ..., \mathbb{P}_d)$. Furthermore, let the distribution function of g_j under $H_0^{(2)}$ be

$$F(k) = P(g_j \leq k) = \sum_{i=0}^{k} P_k, \tag{26}$$

The asymptotic distributions of our test statistics are summarized in the following proposition.

Proposition 3. *Let* $\boldsymbol{Y}$ *follow* (3) *and as long as* $T, M \to \infty$ *such that* $\delta T \to 0$*, asymptotically under* $H_0^{(2)}$

(i) $Z \sim \mathcal{N}\Big(0, F(1)(1 - F(1))\Big)$

(ii) $Z_1 \sim \mathcal{N}\Big(0, F(0)(1 - F(0)) + F(1)(1 - F(1)) - 2(F(0) - F(0)F(1))\Big)$

(iii) $Z^2 \sim \chi_d^2$

(iv) $Z_{sd}^2 \sim \sum_{i=0}^{d} \gamma_i^{sd} U_i$, *where* γ_i^{sd} *are the eigenvalues of* Ω_{sd}

(v) $Z^2_{var} \sim \sum_{i=0}^{d} \gamma_i^{var} U_i$ *, where* γ_i^{var} *are the eigenvalues of* Ω_{var}

where $U_i \sim N(0,1)$ *and*

$$\Omega_{sd} = Cov\left(\frac{\mathbb{P}-P}{\sqrt{P\circ(1-P)}}\right) = \frac{1}{\sqrt{P\circ(1-P)}}\frac{1}{\sqrt{P\circ(1-P)}}' \circ \Sigma$$
$$\Omega_{var} = Cov\left(\frac{\mathbb{P}-P}{P\circ(1-P)}\right) = \frac{1}{P\circ(1-P)}\frac{1}{P\circ(1-P)}' \circ \Sigma$$
$$\Sigma = \begin{pmatrix} P_1(1-P_1) & -P_1P_2 & \cdots & -P_1P_d \\ -P_2P_1 & & & \\ \vdots & & \ddots & \vdots \\ -P_dP_1 & \cdots & & P_d(1-P_d) \end{pmatrix}.$$

Remark The covariance matrices in Proposition 3 (*iv,v*) have eigenvalues different from zero and one, such that their distribution is not χ^2, but depends on d, λ and δ in a highly non-linear fashion. However, as λ and δ are fixed within any given sample, as soon as an estimate of λ has been obtained, it is straightforward to simulate critical values for Z^2_{sd} and Z^2_{var}.

Our test statistics (except Z_1) consider a standardized difference between cojump frequencies and their estimated probabilities for all extents from $k = 2, ...d$ ($k = 0, ..., d$ for squared statistics). That is, they are affected by cojumps of all extents. If one is interested however, say, only in large-extent cojumps, one can modify the test statistics to disregard small-extent cojumps and be sensitive only to cojumps of extents greater or equal to a specified minimum extent. Their definitions and asymptotic distribution is collected in the following corollary.

Corollary 1. *Define the* minimum extent *test statistics by for* $m \geq 2$*:*

$$Z_m = \sqrt{\widetilde{M}} \sum_{k=m}^{d} (\mathbb{P}_k - \hat{P}_k), \qquad Z^2_m = \widetilde{M} \sum_{k=m}^{d} \left(\frac{\mathbb{P}_k - \hat{P}_k}{\sqrt{\hat{P}_k}}\right)^2,$$
$$Z^2_{m,sd} = \widetilde{M} \sum_{k=m}^{d} \left(\frac{\mathbb{P}_k - \hat{P}_k}{\sqrt{\hat{P}_k(1-\hat{P}_k)}}\right)^2, \qquad Z^2_{m,var} = \widetilde{M} \sum_{k=m}^{d} \left(\frac{\mathbb{P}_k - \hat{P}_k}{\hat{P}_k(1-\hat{P}_k)}\right)^2.$$

Let Y *follow* (3)*. Under* $H_0^{(2)}$*, we have under the joint asymptotics* ($T, M \to \infty$, $T/M \to 0$)*,*

(i) $Z_m \sim \mathcal{N}\big(0, F(m-1)(1-F(m-1))\big)$

(ii) $Z^2_m \sim \chi^2_{d-m}$

(iii) $Z^2_{m,sd} \sim \sum_{i=0}^{d-m} \gamma_i^{sd} U_i^2$*, where* γ_i^{sd} *are the eigenvalues of* Ω^m_{sd}

(iv) $Z^2_{m,var} \sim \sum_{i=0}^{d-m} \gamma_i^{var} U_i^2$ *, where* γ_i^{var} *are the eigenvalues of* Ω^m_{var}

where $U_i \sim N(0,1)$ *and with* $\Omega_{sd}, \Omega_{var}$ *and* Σ *as in Proposition 3 and* $\Omega^m_{(\cdot)} = \big(\Omega_{(\cdot),ij}\big)_{i,j=m,...,d}$ *the covariance matrix excluding the first* m *rows and columns.*

2.4 Bootstrap

Our test procedure contains two sources of errors in finite samples that do not enter the asymptotic distributions in Proposition 3. First, the error when estimating the number of jumps in

the sample N_T vanishes as $\delta \to 0$ and, second, the error when estimating λ that vanishes under joint asymptotics ($T, M \to \infty$, $\delta T \to 0$). Thus the jump detection and the estimation of the intensity λ introduce nuisance parameters that only disappear asymptotically. However, in finite sample we expect these errors to significantly affect the proposed test statistics, rendering the asymptotic results to be of limited applicability in finite samples. We expect these distortions to be particularly relevant when λ is small, i.e., when jumps are relatively rare events. We suggest to base inference on the following block bootstrap procedure as proposed in Politis and Romano (1994) that takes into account the additional variation introduced by these two sources of error. Our Monte Carlo simulations below have shown that this bootstrap is able to correct the size distortions that are observed in finite samples.

1. For each asset, construct univariate bootstrap sample by blockwise resample returns from original observations with deterministic or random block length with rate $O(K^{1+\epsilon})$.
2. From the multivariate bootstrap sample, calculate $\hat{\sigma}$, $\widehat{\Delta \boldsymbol{N}}(\hat{\boldsymbol{\beta}}^*_{emp})$, $\hat{\boldsymbol{\lambda}}$, $\mathbb{P}$, P and finally Z, Z_1, $Z^2_{(\cdot)}$.
3. Repeat 1)-2) to obtain bootstrap distribution.

Observe that the blockwise resampling of returns (nearly) preserves the autoregressive structure of the volatility process with only a few breaks in the SV process at the block borders. This however does not violate Assumption 1 such that the bootstrap sample nearly resembles the same DGP. Furthermore, bootstrapping each series individually destroys all dependence across assets, in particular removes possible cojumps, and therefore ensures that the bootstrap samples obey $H_0^{(2)}$.

2.5 Non-iid case

Assuming constant intensity for the jump processes may be rather stringent depending on the considered observation frequency and the application of interest, especially in financial markets where extreme events are typically observed to cluster in time; see, e.g., Chavez-Demoulin and McGill (2012). Therefore we generalize our approach by modeling the intensity using self-exciting Hawkes processes introduced by Hawkes (1971) and applied in finance by, e.g., in Chavez-Demoulin et al. (2005), McNeil et al. (2005), Embrechts et al. (2011) or Bormetti et al. (2015). In the univariate case, the conditional intensity $\lambda_t^{(i)}$ of the i-th jump process $N_t^{(i)}$ is given by

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \int_{-\infty}^{t} d_i(\Delta Y_s^{(i)}) g_i(t-s) dN_s^{(i)}, \tag{27}$$

where $\lambda_0^{(i)}$ is the i-th assets baseline intensity, $d_i(\cdot)$ is the impact function, specifying the influence of jumps $\Delta Y_s^{(i)}$ on the conditional intensity and $g_i(\cdot)$ is the decay function, which determines the decay of the intensity back to its baseline level in the absence of further events and therefore determines the degree of memory of the process. Following Grothe et al. (2014), we specify $g_i(t-s) = e^{-\beta_i(t-s)}$. A popular choice for the impact function is $d_i(Y_s^{(i)}) = e^{\gamma Y_s^{(i)}}$, $\gamma > 0$,as, e.g., in Chavez-Demoulin and McGill (2012), but as shown in Grothe et al. (2014) this choice does not guarantee stationarity of the process. In this paper we choose the simple impact function $d_i(\Delta Y_{i,s^-}) = \alpha_i$, i.e., a constant impact irrespective of the size of the jump. Then in

discrete time a univariate Hawkes model for the time-varying intensity in asset i is given by

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \sum_{t_k<t} \alpha_i e^{-\beta_i(t-t_k)}. \tag{28}$$

The model can also be generalized to the multivariate case by allowing the intensity in asset i not only to be driven by its own jumps, but also by jumps in the other assets and therefore allowing for (dynamic) spillover effect. In the most general case the intensity is given by

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \sum_{n=1}^{d*} \sum_{t_{kn}<t} \alpha_{in} e^{-\beta_{in}(t-t_{kn})}, \tag{29}$$

where $d^* \in \{1, ..., d\}$. For $d^* = 1$ we recover the univariate Hawkes process for which the intensity of asset i depends only on its own jumps. For the other extreme, $d^* = d$, jumps in all assets may affect asset i's intensity. For d large one may typically have to place restrictions on the spillover dynamics by allowing spillover only between subsets of the assets as otherwise the identification of the parameters is likely to be very difficult.

Estimation of the model can be achieved by maximum likelihood estimation. The associated log-likelihood function is

$$\mathcal{LL}(T) = \sum_{i=1}^{d} \mathcal{LL}^{(i)}(T) \text{ with} \tag{30}$$

$$\mathcal{LL}^{(i)}(T) = (1-\lambda_0)T + \sum_{t_j^i<T} \log\left(\lambda_0^i + \sum_{m=1}^{d} \sum_{t_l^m<t_j^i} \alpha e_{im}^{\beta(t_j^i - t_l^m)}\right) + \sum_{m=1}^{d} \sum_{t_j^m<T} \frac{\alpha_{im}}{\beta_{im}}(1 - e^{\beta_{im}(T-t_j^m)}). \tag{31}$$

Ogata (1978) showed that maximum likelihood estimation for this class of processes is consistent and asymptotically normal as $T \to \infty$. However in our case the jump times are only imperfectly identified through the jump detection procedure. Fortunately, the associated jump detection error vanishes as $\delta \to 0$ such that the parameters in (28) or (29) can be consistently estimated as $T \to \infty$ and $\delta T \to 0$. This is convenient as now, when comparing the number of jumps over intervals $(t_{j-1}, t_j], j = 1, ..., K$ with their expected number under the null, the jump processes can be considered a non-homogeneous Poisson processes with deterministic intensity functions $\lambda_t^{(i)}$ and $\overline{\lambda_j^{(i)}} := \int_{t_{j-1}}^{t_j} \lambda_s^{(i)}(s)ds$. Thus the intensities can be treated as locally constant over discrete observation intervals.

Our general cojump testing procedure can be applied in the same way as for homogeneous jump intensities using $p_{i,j} = \overline{\lambda_j^{(i)}}\delta$ as the probability of a jump in asset i in observation interval j. Then we have under the null, that g follows the distribution function

$$g \sim F = \frac{1}{\widetilde{M}} \sum_{j=K+1}^{M} F_j,$$

where F is an equally weighted multinomial mixture distribution and F_j being the null distribution function of g in $(t_{j-1}, t_j]$. A point worth noting is that because of (20) and (26), there is moderate computational burden involved when when calculating F. This is because the number

of times, the derivative function must be evaluated with its respective λ's grows linearly with the number of intervals $(t_{j-1}, t_j]$ over which we have distinct distribution functions F_j.

3 Monte Carlo simulations

The aim of this section is to study the finite sample properties of our tests and to decide, which of the competing forms is to be preferred. Our simulation design follows Gilder et al. (2014) and the DGP is given by the following equations

$$dY_t^{(i)} = a_t dt + \theta\sigma_t^{(i)} dB_t^{(i)} + \sqrt{(1-\theta^2)}\sigma_t^{(i)} dW_t + \sum_{k=1}^{N_t^{(i)}} c_k^{(i)} \tag{32}$$

$$\sigma_t^{(i)} = \exp(\alpha_0 + \alpha_1 \zeta_t^{(i)}) \tag{33}$$

$$d\zeta_t^{(i)} = \alpha_2 \zeta_t^{(i)} dt + dB_t^{(i)}, \tag{34}$$

$$f_{c^{(i)}}(x) = \frac{1}{2}\phi\left(x, -\kappa q, \frac{\kappa q}{3}\right) + \frac{1}{2}\phi\left(x, +\kappa q, \frac{\kappa q}{3}\right) \tag{35}$$

where W and B are independent standard Brownian Motions modeling a common diffusive factor and idiosyncratic innovations respectively and $B^{(i_1)} \perp\!\!\!\perp B^{(i_2)}, i_1 \neq i_2$. Equation (33) ensures positivity of SV, α_2 captures the persistence of typical market volatility, while α_0 and α_1 adjust the level and variance of the SV process.[5] The jump size $c^{(i)}$ follows a normal mixture with modes $\mp\kappa q$ and variances $\frac{\kappa q}{3}$, which mimics our empirical findings in section 4. Herein, q is the 0.841-quantile of the unconditional distribution of the diffusion process and κ is a scaling factor which regulates the separability of the jump and diffusive process.

We employ an Euler discretization, where prices are simulated every minute; i.e., $\delta = 1/390$. The parameters are assumed to be the same for all assets and are given by $a_t = 0.03$, $\theta = 0.4$, $\alpha_0 = -1.6094$ implying a starting value for the volatility of 0.2, $\alpha_1 = 0.125$, and $\alpha_2 = 0.9$. The scaling factor is set to $\kappa = 32$ to ensure reasonable separability. Finally, the starting value for the log-price process Y_0 is arbitrarily set to 3.4.

In order to obtain dependent jump processes we proceed as follows. We simulate

$$\boldsymbol{z}_t \sim N(\mathbf{0}, R)$$

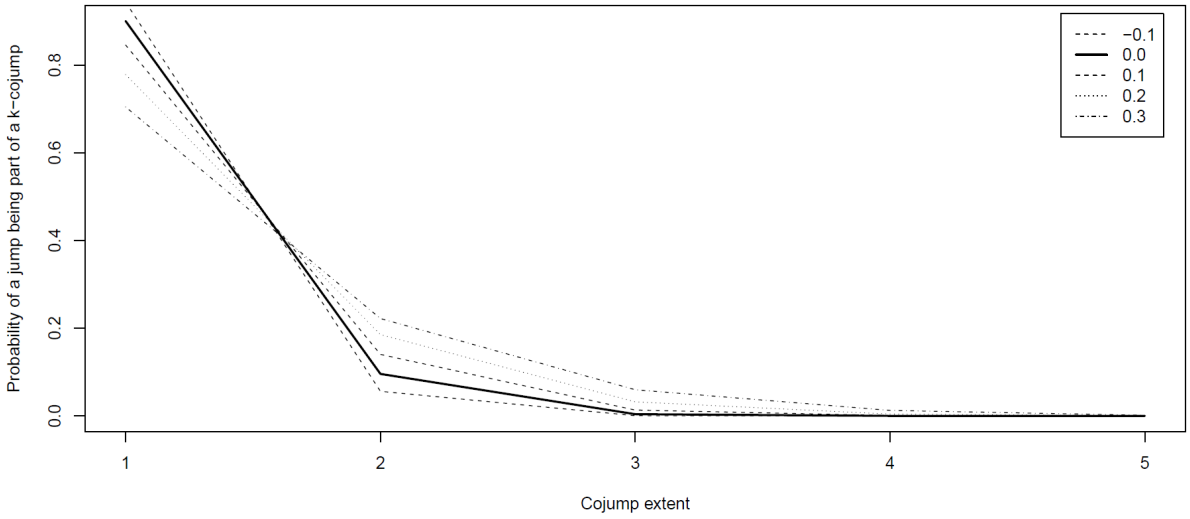
$$R = \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & & \vdots \\ \vdots & & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix}$$

where R is a $d \times d$ equicorrelation matrix with correlation $\rho \in [-1, 1]$. The jump process for asset i then follows as

$$\Delta N_t^{(i)} = \begin{cases} 1 \text{ if } \Phi(z_t^{(i)}) > 1 - \lambda_t^{(i)}\delta \\ 0 \text{ otherwise} \end{cases}, \tag{36}$$

[5] Gilder et al. (2014) additionally introduce a deterministic intraday volatlity pattern in their Monte Carlo simulation, but we refrained from that for the sake of simplicity. We nevertheless consider this relevant empirical fact in the application section 4.

Figure 1: Cojump probability for different ρ



with $\Phi(\cdot)$ the standard normal CDF. Thus, a jump occurs whenever $z_t^{(i)}$ exceeds the $1 - \lambda_t^{(i)}\delta$ quantile of the standard normal distribution. When $\rho = 0$ the jump processes are independent and cojumps only occur by coincidence, which coincides with the situation under $H_0^{(2)}$ described in Section 2.3: Cojumps may occasionally happen in discrete time even though they are independent. Whenever $\rho > 0$ the $z_t^{(i)}$ are correlated and cojumps becomes more likely. To study the power of our tests we consider $\rho \in \{0.1, 0.2, 0.3\}$ corresponding to an increasing probability of cojumps. Figure 1 plots different values of ρ against the probability of a jump being part of a k-cojump in a five-dimensional data set to illustrate how far the alternative hypotheses, i.e. values of $\rho \neq 0$, are from the null.

Regarding the intensity, the case $\lambda_t^{(i)} = \lambda$ corresponds to the iid case, whereas for the non-iid case $\lambda_t^{(i)}$ is generated by the Hawkes process in (28) with $\alpha = 0.04$ and $\beta = 0.07$. For the iid case we consider the values $\lambda \in \{0.5, 1, 2, 3, 5, 10\}$. For the Hawkes model the unconditional mean of the intensity is given by $\mathbb{E}[\lambda_t] = \frac{\lambda_0}{1-\alpha/\beta}$ and we vary the baseline intensity $\lambda_0 \in \{0.21, 0.43, 0.86, 1.29, 2.14, 4.29\}$ to match the intensities in the iid case. Figure 2 plot a typical Hawkes process based corresponding to $\mathbb{E}[\lambda_t] = 10$. The clustering of events and the persistence in the intensity is apparent. Finally, for the dimension of the problem we consider $d \in (2, 5)$ and the number of Monte Carlo simulations is $M = 1024$ for the tests based on asymptotic results and $M = 512$ for the tests using bootstrapped critical value. The number of bootstrap replications is also chosen as $B = 512$.

Table 3 reports the size of the tests based on 1-minute returns and after aggregating the data to 5-minute returns for the following cases. First, we consider iid jump processes and rely on the asymptotic distributions in Proposition 3. For 1-minute returns the size properties are only acceptable for the statistic Z. The other test statistics are all undersized, more severely so for $d = 5$. Here Z^2 still has the most acceptable size. For 5-minute returns all the tests are severely oversized except for the case $\lambda = 10$, i.e. when jumps occur rather frequently. For smaller value of λ, however, the size can be as high as 100%. Here we need to recall that there are a lot of sources of error in finite samples, most notably the detection of jumps which appears to work much poorer when data are aggregated to 5-minute returns. The bootstrapped versions of the tests, on the other hand, are all close to the nominal level of 5% across all settings.

Table 3: Size

		1-minute returns														
		Asymptotic					Bootstrapped					Hawkes				
	λ	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}
$d=2$	10	0.068	0.000	0.013	0.000	0.047	0.047	0.047	0.053	0.053	0.049	0.080	0.086	0.061	0.045	0.082
	5	0.075	0.000	0.014	0.000	0.049	0.072	0.074	0.072	0.072	0.068	0.049	0.068	0.023	0.023	0.035
	3	0.061	0.000	0.013	0.000	0.030	0.041	0.041	0.041	0.041	0.039	0.055	0.063	0.039	0.037	0.041
	2	0.080	0.000	0.014	0.000	0.043	0.043	0.043	0.043	0.043	0.043	0.057	0.059	0.051	0.045	0.051
	1	0.067	0.000	0.027	0.000	0.044	0.023	0.023	0.023	0.023	0.023	0.047	0.049	0.047	0.047	0.041
	0.5	0.077	0.000	0.077	0.000	0.077	0.074	0.074	0.074	0.074	0.074	0.039	0.041	0.041	0.041	0.035
$d=5$	10	0.054	0.001	0.076	0.009	0.000	0.047	0.053	0.031	0.031	0.037	0.061	0.064	0.045	0.039	0.053
	5	0.045	0.000	0.015	0.004	0.000	0.053	0.057	0.059	0.059	0.057	0.055	0.063	0.049	0.043	0.049
	3	0.062	0.000	0.016	0.003	0.002	0.051	0.049	0.055	0.057	0.023	0.057	0.072	0.055	0.051	0.051
	2	0.060	0.000	0.046	0.005	0.002	0.063	0.063	0.059	0.059	0.059	0.033	0.041	0.049	0.049	0.053
	1	0.064	0.000	0.006	0.005	0.000	0.047	0.047	0.051	0.051	0.051	0.061	0.072	0.068	0.064	0.064
	0.5	0.094	0.000	0.008	0.002	0.000	0.039	0.043	0.031	0.031	0.031	0.043	0.045	0.043	0.045	0.045
		5-minute returns														
$d=2$	10	0.048	0.002	0.018	0.003	0.041	0.055	0.055	0.070	0.070	0.066					
	5	0.173	0.004	0.063	0.007	0.097	0.063	0.063	0.045	0.045	0.045					
	3	0.806	0.030	0.631	0.171	0.684	0.051	0.053	0.045	0.045	0.045					
	2	0.983	0.195	0.951	0.659	0.956	0.053	0.053	0.057	0.057	0.057					
	1	1.000	0.485	1.000	0.933	0.999	0.053	0.053	0.053	0.053	0.053					
	0.5	0.999	0.553	0.998	0.960	0.996	0.031	0.031	0.031	0.031	0.031					
$d=5$	10	0.014	0.060	0.050	0.035	0.162	0.047	0.049	0.041	0.043	0.039					
	5	0.216	0.286	0.662	0.552	0.211	0.055	0.057	0.061	0.061	0.041					
	3	0.867	0.945	1.000	1.000	0.867	0.055	0.057	0.066	0.064	0.076					
	2	0.995	0.997	1.000	1.000	0.988	0.057	0.057	0.057	0.059	0.047					
	1	1.000	1.000	1.000	1.000	0.998	0.061	0.063	0.063	0.064	0.047					
	0.5	1.000	1.000	1.000	1.000	0.999	0.063	0.061	0.059	0.059	0.061					

Notes: Table 3 reports the size of the tests defined in equations (21) to (25) in Section 2.3. The data is generated by the model in equations (32) to (35) with $\mu = 0$, $\theta = 0.4$, $\alpha_0 = -1.6094$, $\alpha_1 = 0.125$, $\alpha_2 = 0.9$ and $\kappa = 32$. The jump process is generated by (36) with $\rho = 0$ and intensity λ either constant or driven by the Hawkes process in (28). Data over $T = 100$ days are generated at the 1-minute frequency ($\delta = 1/390$) with an Euler discretization and are aggregated to 5-minute returns for the lower panel. The size is based on the asymptotic results provided in Proposition 3 for the columns labeled 'Asymptotic' and on the block bootstrap described in Section 2.4 for the other two cases. The number of simulations is $M = 1024$ for the asymptotic tests and $M = 512$ for the bootstrapped tests with $B = 512$ bootstrap replications.

Table 4: Power iid case

		1-minute returns														
		$\rho = 0.1$					$\rho = 0.2$					$\rho = 0.3$				
	λ	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}
$d = 2$	10	0.922	0.922	0.895	0.895	0.891	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	5	0.596	0.596	0.553	0.553	0.541	0.988	0.990	0.986	0.986	0.986	1.000	1.000	1.000	1.000	1.000
	3	0.385	0.385	0.381	0.381	0.381	0.877	0.877	0.877	0.877	0.875	0.994	0.994	0.994	0.994	0.994
	2	0.313	0.313	0.313	0.313	0.316	0.715	0.719	0.715	0.715	0.719	0.965	0.965	0.965	0.965	0.965
	1	0.109	0.109	0.109	0.109	0.109	0.406	0.406	0.406	0.406	0.406	0.699	0.699	0.699	0.699	0.699
	0.5	0.188	0.191	0.188	0.188	0.188	0.328	0.332	0.328	0.328	0.328	0.539	0.539	0.539	0.539	0.539
$d = 5$	10	1.000	1.000	1.000	1.000	0.570	1.000	1.000	1.000	1.000	0.990	1.000	1.000	1.000	1.000	1.000
	5	0.998	0.998	0.998	0.998	0.781	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000
	3	0.984	0.984	0.979	0.979	0.332	1.000	1.000	1.000	1.000	0.971	1.000	1.000	1.000	1.000	1.000
	2	0.898	0.906	0.738	0.738	0.625	1.000	1.000	1.000	1.000	0.922	1.000	1.000	1.000	1.000	1.000
	1	0.582	0.586	0.570	0.570	0.578	0.977	0.984	0.980	0.980	0.980	1.000	1.000	1.000	1.000	1.000
	0.5	0.273	0.277	0.277	0.277	0.277	0.781	0.785	0.789	0.789	0.789	0.984	0.984	0.984	0.984	0.984
		5-minute returns														
$d = 2$	10	0.285	0.287	0.191	0.193	0.184	0.654	0.654	0.557	0.557	0.549	0.977	0.977	0.951	0.951	0.951
	5	0.203	0.199	0.148	0.148	0.145	0.617	0.625	0.527	0.527	0.514	0.955	0.953	0.930	0.930	0.928
	3	0.162	0.168	0.127	0.127	0.125	0.443	0.445	0.367	0.367	0.359	0.857	0.857	0.809	0.809	0.803
	2	0.146	0.145	0.133	0.133	0.131	0.377	0.379	0.348	0.348	0.344	0.703	0.705	0.666	0.666	0.662
	1	0.098	0.098	0.098	0.098	0.098	0.205	0.205	0.205	0.205	0.207	0.484	0.484	0.484	0.484	0.482
	0.5	0.076	0.076	0.076	0.076	0.076	0.129	0.129	0.129	0.129	0.129	0.293	0.293	0.293	0.293	0.293
$d = 5$	10	0.684	0.789	0.588	0.621	0.195	0.996	1.000	1.000	1.000	0.732	1.000	1.000	1.000	1.000	0.986
	5	0.680	0.748	0.529	0.570	0.205	1.000	1.000	0.998	1.000	0.574	1.000	1.000	1.000	1.000	0.982
	3	0.518	0.580	0.248	0.271	0.178	0.988	0.992	0.936	0.945	0.518	1.000	1.000	1.000	1.000	0.879
	2	0.398	0.408	0.281	0.289	0.178	0.949	0.957	0.914	0.918	0.617	1.000	1.000	1.000	1.000	0.984
	1	0.262	0.260	0.201	0.201	0.086	0.773	0.781	0.674	0.678	0.248	0.994	0.994	0.977	0.977	0.656
	0.5	0.164	0.162	0.162	0.162	0.160	0.467	0.477	0.482	0.482	0.475	0.895	0.914	0.896	0.896	0.893

Notes: Table 4 reports the power of the tests defined in equations (21) to (25) in Section 2.3. The data is generated by the model in equations (32) to (35) with $\mu = 0$, $\theta = 0.4$, $\alpha_0 = -1.6094$, $\alpha_1 = 0.125$, $\alpha_2 = 0.9$ and $\kappa = 32$. The jump process is generated by (36) with $\rho \in \{0.1, 0.2, 0.3\}$ and constant intensity $\lambda \in \{0.5, 1, 2, 3, 5, 10\}$. Data over $T = 100$ days are generated at the 1-minute frequency ($\delta = 1/390$) with an Euler discretization and are aggregated to 5-minute returns for the lower panel. The power is based on the block bootstrap described in Section 2.4. The number of simulations is $M = 512$ with $B = 512$ bootstrap replications.

Table 5: Power Hawkes intensity

		1-minute returns														
		$\rho = 0.1$					$\rho = 0.2$					$\rho = 0.3$				
	λ	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}	Z	Z_1	Z^2	Z^2_{sd}	Z^2_{var}
$d = 2$	10	0.998	0.998	0.984	0.959	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	5	0.814	0.840	0.762	0.729	0.807	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	3	0.645	0.648	0.570	0.547	0.615	0.990	0.990	0.984	0.982	0.988	1.000	1.000	1.000	1.000	1.000
	2	0.449	0.455	0.418	0.402	0.428	0.910	0.912	0.893	0.889	0.900	0.994	0.994	0.992	0.992	0.994
	1	0.264	0.270	0.262	0.258	0.273	0.578	0.570	0.578	0.578	0.582	0.896	0.887	0.896	0.896	0.904
	0.5	0.117	0.119	0.117	0.117	0.109	0.309	0.305	0.309	0.307	0.311	0.566	0.564	0.564	0.564	0.547
$d = 5$	10	1.000	1.000	1.000	1.000	0.918	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	5	1.000	1.000	1.000	1.000	0.494	1.000	1.000	1.000	1.000	0.969	1.000	1.000	1.000	1.000	1.000
	3	1.000	1.000	1.000	1.000	0.826	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	2	0.998	0.998	0.979	0.975	0.541	1.000	1.000	1.000	1.000	0.984	1.000	1.000	1.000	1.000	1.000
	1	0.797	0.830	0.523	0.508	0.313	1.000	1.000	0.998	0.996	0.756	1.000	1.000	1.000	1.000	0.980
	0.5	0.430	0.447	0.432	0.426	0.445	0.920	0.938	0.926	0.926	0.930	0.998	0.998	0.998	0.998	0.998

Notes: Table 5 reports the power of the tests defined in equations (21) to (25) in Section 2.3. The data is generated by the model in equations (32) to (35) with $\mu = 0$, $\theta = 0.4$, $\alpha_0 = -1.6094$, $\alpha_1 = 0.125$, $\alpha_2 = 0.9$ and $\kappa = 32$. The jump process is generated by (36) with $\rho \in \{0.1, 0.2, 0.3\}$ and time-varying intensity λ_t driven by the Hawkes process in (28) with expected value $\mathbb{E}[\lambda_t] \in (0.5, 1, 2, 3, 5, 10)$. Data over $T = 100$ days are generated at the 1-minute frequency ($\delta = 1/390$) with an Euler discretization and are aggregated to 5-minute returns for the lower panel. The power is based on the block bootstrap described in Section 2.4. The number of simulations is $M = 512$ with $B = 512$ bootstrap replications.

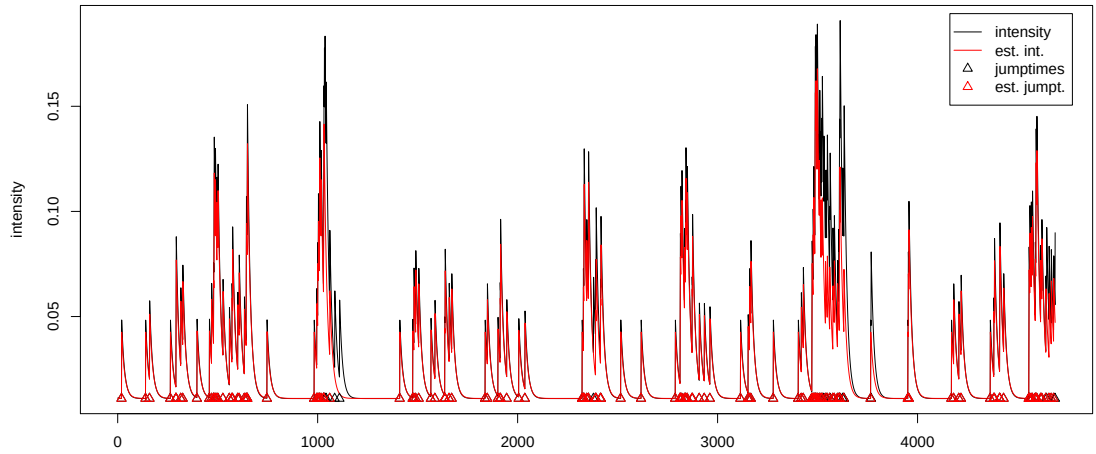


Figure 2: Typical Hawkes process, simulated under (32)-(36) and (28) with $\alpha = 0.04, \beta = 0.07$ and λ_0 such that $\mathbb{E}[\lambda_t] = 10$.

The power of the tests can be found in Table 4 for iid jumps and in Table 5 for the jump processes driven by Hawkes processes. As to be expected, the power increases with λ, i.e. with the frequency of jumps, with the dimension of the problem, and with the strength of the dependence of the jumps ρ. Aggregating the data to 5-minute returns greatly decreases the power, which is not surprising, as the effective sample size decreases accordingly (recall that the number of days is fixed at $T = 100$). All proposed tests have good power and in many cases the power of the different tests is identical or nearly identical. This can be explained by the fact that these tests are very similar in that they compare empirical and theoretical cojumps and therefore tend to reject for the same samples. However, the tests Z and Z_1 dominate the other tests in several cases and are to be preferred over the other tests. Between the two, Z_1 is to be slightly preferred over Z. These findings hold both for the iid jumps and for jumps following a Hawkes process.[6]

4 Empirical application

In this section we apply our tests to stock prices data from the S&P500 from January 2$^{\text{nd}}$ 2014 to September 4$^{\text{th}}$ 2015.[7] Prices are recorded each minute in trading hours of the New York Stock Exchange from 9:30am to 16:00pm. To avoid trivial cojumps that are not due to coincidental news, but rather due to accumulated overnight information to which prices adjust as soon as markets open, we exclude overnight returns. This leaves 164547 observations per asset on a total of 423 trading days. Furthermore, we exclude assets with more than 5000 missing values (roughly 3%) from the analysis. From the complete S&P500 this leaves us with a total of 354 assets from which 1-minute log-returns are computed.[8] Remaining missing returns are imputed by mean zero normal random variables with the average variance of the respective series.

The rest of this section is structure as follows. In Section 4.1 we discuss how we deal with the intraday volatility pattern and look at first results for an arbitrarily chosen 10-dimensional portfolio and at some descriptive statistics for cojumps. In Section 4.4 we discuss how our test

[6] Unreported simulation results for different value of T and based on a bivariate Hawkes process with spillovers are available from the authors upon request.

[7] The data were purchased from the provider QuantData.com.

[8] See table B.1 in the appendix for a list of the included assets.

can be applied to find portfolios that minimize the cojump risk and we investigate, whether cojumps risk can be diversified away.

4.1 Intraday volatility pattern

An obvious problem of the LM procedure is the stylized fact that many financial time series exhibit a persistent intraday volatility pattern. Typically in equity markets, we find that volatility is highest in the earliest hours of the trading day, lowest mid-day and again rises slightly before markets close. Therefore the LM test may underestimate jumps in the middle of the day and mistakes high volatility returns for jumps in the beginning and end of the day. We solve this problem by estimating the intraday volatility pattern in a jump-robust way and account for it in the normalization of returns; see e.g. Boudt et al. (2011). For simplicity, assume that we want to estimate the periodic volatility component only based on the time of *day*, that is we ignore possible day-of-the-week effects. In our analysis, we employ the Shortest-Half scale estimator by Rousseeuw and Leroy (1988). For its construction consider the discrete time version of our baseline model (1) under $H_0^{(1)}$, i.e. without any jumps, and introduce a periodic volatility component f_t. This results is the model

$$y_j = f_t s_j u_j, \tag{37}$$

where u_j is standard normal, s_j is the stochastic volatility at period j that is considered constant over the interval $\big((j-1)\delta, j\delta\big]$ and t, $t \in \{2, ..., 390\}$ indicates the trading minute in the trading hours 9:30am-16:00pm of j-th return by the relation $t = \delta^{-1}(j \mod \delta^{-1}) + 1$. The periodic volatility component f_t must sum in squares to one over a trading day, that is $\delta \sum_i^{\delta^{-1}} f_i^2 = 1$. This standardization condition ensures that the decomposition in (37) is unique. Suppose the local volatility-standardized return $\mathcal{L}_j$ was recorded in trading minute t. Then we denote $\mathcal{L}_{1;t}, ..., \mathcal{L}_{n;t}$ the returns that were observed at the same time of day as $\mathcal{L}_j$, and by $\mathcal{L}_{(1);t}, ..., \mathcal{L}_{(n);t}$ their corresponding order statistics such that $\mathcal{L}_{(1);t} \leq \mathcal{L}_{(2);t} \leq ... \leq \mathcal{L}_{(n);t}$. The variable n counts the observations in the sample made at a specific time of day. The Shortest Half scale (SH) is given by the shortest distance between $h_t = \lfloor n/2 \rfloor + 1$ contiguous order statistics, that is

$$SH_t := 0.741 \min\{\mathcal{L}_{(h_t);t} - \mathcal{L}_{(1);t}, ..., \mathcal{L}_{(n);t} - \mathcal{L}_{(n-h_t);t}\}.$$

The SH estimator of the periodicity of $\mathcal{L}_j$ reads as

$$\hat{f}_t^{\,SH} = \frac{SH_t}{\sqrt{\delta \sum_{i=1}^{\delta^{-1}} SH_i^2}},$$

such that the periodicity-corrected returns can be constructed as

$$\widetilde{\mathcal{L}}_j = \frac{\mathcal{L}_j}{\hat{f}_t^{\,SH}}.$$

In Figure 3 we present the intraday volatility pattern estimated from our data. In the remainder of this section we work with periodicity-corrected returns.

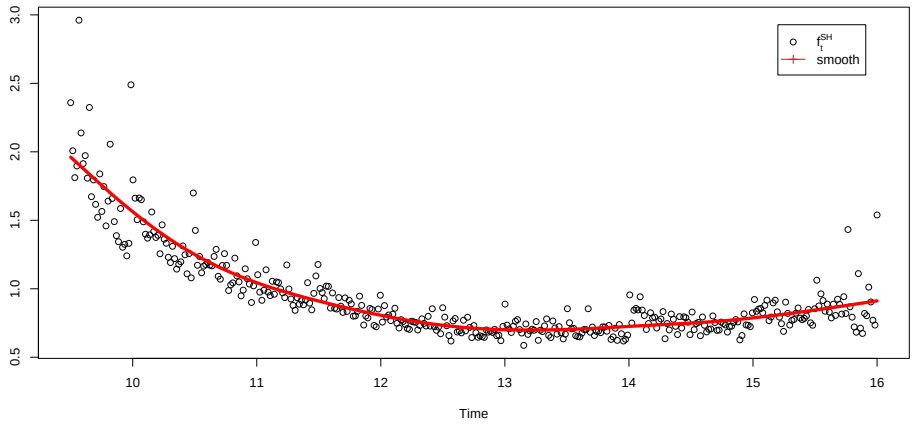


Figure 3: Intraday Volatility Pattern

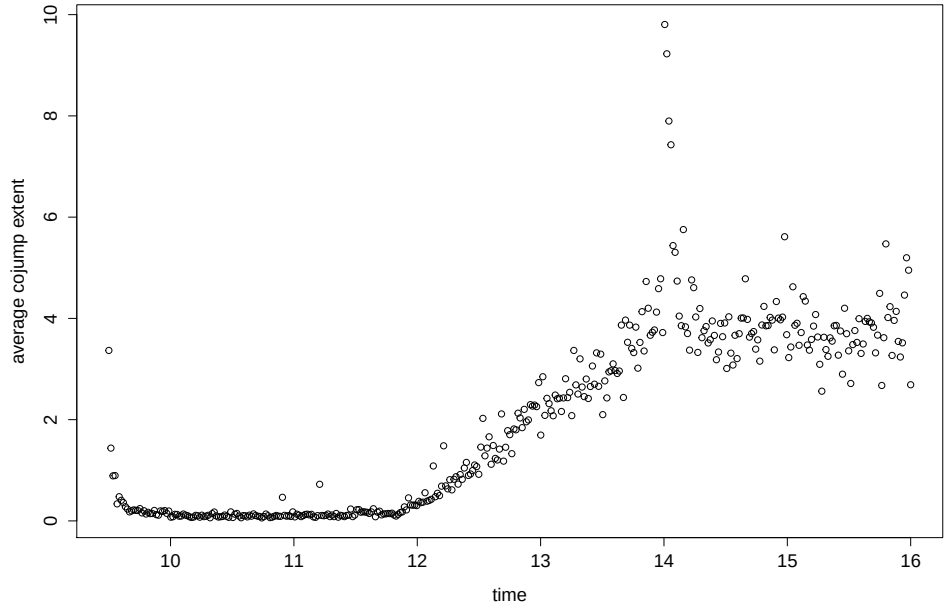


Figure 4: Average cojump extent per trading minute

4.2 Descriptive statistics

Next, we look at the complete market and study the behavior of cojumps descriptively. Figure 4 show the average cojump extent for each minute of the trading day for our sample of 354 assets. Most cojumps occur after noon with a peak at 2:00pm, which is the time the Federal Open Market Committee (FOMC) releases news. This is probably the single most important source of information causing jumps and cojumps.

Table 6 summarizes the distribution of the estimated jump intesities $\hat{\lambda}$ in the full sample. We find to empirical λs to be in line with our simulations, ranging from roughly one jump every three days to more than four jumps per day. Figure 5 illustrates the distribution of cojumps extents in the market on a log-scale. Large extents tend to more dispersed and cojumps involving more than 150 assets are mostly limited to one or two occurances. Nevertheless, the heavy right tail in the distribution suggests a significant amount of cojumps in the data.

Table 6: Summary statistics for $\hat{\lambda}$

	Min	1st Q.	Median	Mean	3rd Q.	Max
$\hat{\lambda}$	0.2748	1.4642	1.7556	1.8054	2.0844	4.3499

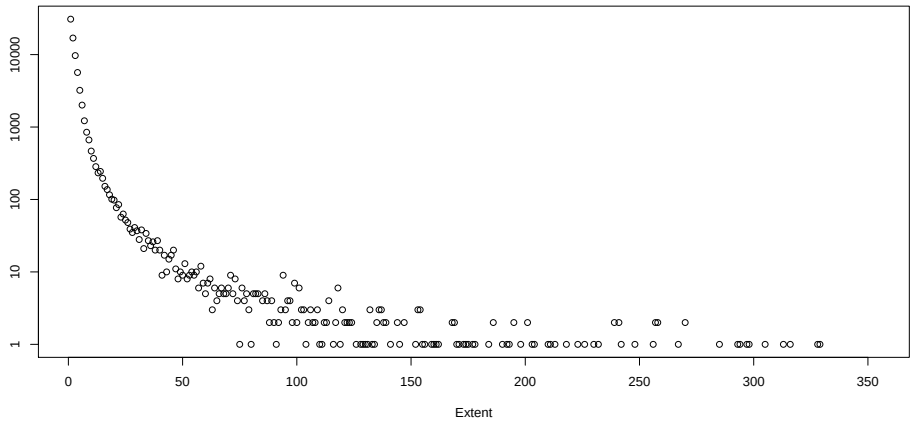


Figure 5: Market-wide extent distribution

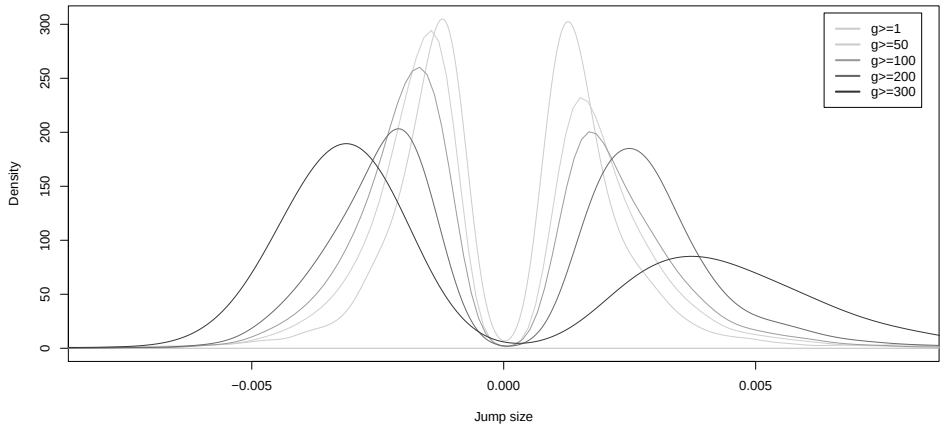


Figure 6: Conditional jump size density in the market.

An interesting picture emerges when considering the *conditional* jump size distribution in the market as a whole. Figure 6 plots the densities of jump sizes of all assets in our sample, conditional on the cojump extent g being larger than some number. Three patterns are apparent. First, the conditional jump size density is bimodal. That is the least surprising finding, as small jumps are blurred by diffusive increments and are thus hardly detected. Second, jump sizes grow in magnitude with their extent, meaning large-extent cojumps contain larger jumps on average. However their skewness seems to disappear for larger extents. Compare the case of $g \leq 1$, which is the jump size density for all jumps. This particular density has much wider tails both in the negative and positive direction. Third, and probably most informative, large-extent cojumps seem to be dominated by negative jumps and the asymmetry is becoming more pronounced the larger the extent. Together with our results from table 7, we conclude that large-extent cojumps predominantly constitute bad news for the market.

4.3 Portfolio tests

Next, we applied the jump detection procedure as described in Section 2.2. Here we also distinguish between positive and negative jumps, as negative jumps are much more relevant for a typical investor. Before looking at the full data set, Table 7 summarizes empirical findings and test results for a typical portfolio consisting of the stocks of Amazon, Berkshire Hathaway, Google, J.P. Morgan, Morgan Stanley, Microsoft, Netflix, Procter&Gamble, Starbucks and UPS. The difference in individual jump frequency is striking with Google exhibiting the lowest (201 jumps) and Procter&Gamble the highest jump frequency (778). We find about 2.18% jumps in total. Furthermore, positive and negative jumps occur in roughly equal proportions of 1.08% and 1.11%, respectively. Concerning cojumps, their frequency declines in the extent k as to be expected. However, a considerable amount of cojumps can also be found for medium to large extents. The greatest extent detected is, considering all, positive and negative jumps respectively, 10, 9 and 10. Interestingly, there also appear to be some co-jumps of mixed sign as the numbers of positive and negative cojumps do not add up to the number of total cojumps for $k = 2, 3$. Cojumps of opposite signs are, of course, desirable as part of the cojumps risk can be diversified away. For $k \geq 4$, however, all cojumps are of the same sign.[9] Based on this finding one may hypothesize, that news that is suprising enough to cause market disruptions, but which carries mixed informational content is exceedingly rare (or maybe does not even exist at all), implying that big news is either good or bad for all affected. In other words, there is strong dependence in the multivariate tail risk.

Table 7 also contains a first set of test results for cojumps. For each k, we test the null hypothesis of cojumps of extent larger or equal to k. For example, for $k = 3$ bivariate cojumps are contained in the null hypothesis and we test against the alternative that cojumps of extent 3 or larger occur. Here we rely only the test Z_m for $m \geq 2$, as this test statistic performed best in the simulations and is applicable for any m. The test rejects for all cases. Looking at the theoretical probabilities and the empirical frequencies of the different cojump extents it stands out that cojump of a large degree are highly unlikely to occur by chance alone. Therefore, it is reasonable that a single cojump of extend 10 is enough to warrant rejection of the null

[9]Strictly speaking, an accurate decomposition of jump size signs and their corresponding diffusive increments is difficult, when only their superposition is observed. So the underlying assumption here is that jumps dominate and we identify the jump direction just by the sign of the associated return.

hypothesis. The existence of cojumps is also not surprising economically, as news to the market, e.g. concerning the state of the economy, tends to affect all, or most, assets in the market.

Table 7: Test results for 10-dimensional portfolio

					all jumps						
		amzn	brk-b	googl	jpm	ms	msft	nflx	pg	sbux	ups
count		302	464	201	603	458	460	382	778	469	631
rel. frequency		0.0018	0.0028	0.0012	0.0037	0.0028	0.0028	0.0023	0.0047	0.0029	0.0038
extent k	0	1	2	3	4	5	6	7	8	9	10
k-Cojumps	160654	3010	323	112	49	35	18	13	13	7	1
rel. frequency	0.978196	0.018327	0.001967	0.000682	0.000298	0.000213	0.000110	7.92E-05	7.92E-05	4.26E-05	6.09E-06
probability	0.971488	0.028147	0.000362	2.73E-06	1.33E-08	4.40E-11	9.96E-14	1.52E-16	1.50E-19	8.61E-23	2.18E-26
rej. Z_m	NA	NA	•	•	•	•	•	•	•	•	•
					positive jumps only						
		amzn	brk-b	googl	jpm	ms	msft	nflx	pg	sbux	ups
count		148	226	101	290	221	231	213	362	239	292
rel. frequency		0.0009	0.0014	0.0006	0.0018	0.0013	0.0014	0.0013	0.0022	0.0015	0.0018
extent k	0	1	2	3	4	5	6	7	8	9	10
k-Cojumps	162461	1528	133	50	15	16	11	6	10	5	0
frequency	0.989198	0.009304	0.000810	0.000304	9.13E-05	9.74E-05	6.70E-05	3.65E-05	6.09E-05	3.04E-05	0
probability	0.985959	0.013953	8.80E-05	3.25E-07	7.80E-10	1.27E-12	1.42E-15	1.07E-18	5.22E-22	1.48E-25	1.87E-29
rej. Z_m	NA	NA	•	•	•	•	•	•	•	•	
					negative jumps only						
		amzn	brk-b	googl	jpm	ms	msft	nflx	pg	sbux	ups
count		154	238	100	313	237	229	169	416	230	339
rel. frequency		0.0009	0.0014	0.0006	0.0019	0.0014	0.0014	0.0010	0.0025	0.0014	0.0021
extent k	0	1	2	3	4	5	6	7	8	9	10
k-Cojumps	162408	1519	176	59	34	19	7	7	3	2	1
rel. frequency	0.988876	0.009249	0.001072	0.000359	0.000207	0.000116	4.26E-05	4.26E-05	1.83E-05	1.22E-05	6.09E-06
probability	0.985346	0.014558	9.54E-05	3.65E-07	9.01E-10	1.50E-12	1.71E-15	1.31E-18	6.50E-22	1.87E-25	2.37E-29
rej. Z_m	NA	NA	•	•	•	•	•	•	•	•	•

Note: This table contains count and frequency of jumps and cojumps for all jumps (top panel), only positive jumps (middle panel) and only negative jumps (lower panel). Furthermore, it contains the relative frequency of cojumps and the corresponding theoretical probability of the event under independent jump processes. The row 'rej. Z_k' indicates whether the null of independent cojump processes was rejected, based on Z_k and asymptotic critical values for given extent k. A "•" indicates rejection.

4.4 Finding cojump-diversified portfolios

So far, we have found that our test rejects the null of independent jump processes in a typical portfolio. This finding continues to hold when generating a number of arbitrary portfolios, working on subsamples and varying the portfolio size. Cojumps are a typical feature of the data and we always reject the null hypothesis of no cojumps. Therefore, we ask the following question: From all 354 assets in the sample, can we find a portfolio of 10 assets that does not contain significant cojumps?[10] If yes, this implied that systematic risk could be evaded, at least partially in a small portfolio. A brute-force approach would be to test all possible portfolio with 10 out of 354 assets. Unfortunately the $\binom{354}{10} = 7.49 \times 10^{18}$ arising possible portfolios are hard to test in their entirety even on modern computers. Therefore, we opted for another approach. In the spirit of minimum variance portfolio from classical Portfolio Theory (see e.g. Fleming et al., 2001), we construct a minimum cojump portfolio (MCP). For the MVP the portfolio weights

[10]Strictly speaking, one would need to select a portfolio on a training sample and test the alleged cojump-freeness out of sample. However this objection is rendered somewhat irrelevant since we do not even find a jump-independent portfolio in sample.

Table 8: Minumum cojump portfolio

Typ. Port.		amzn	brk-b	googl	jpm	ms	msft	nflx	pg	sbux	ups
	amzn	302	37	23	43	36	28	17	37	34	36
	brk-b	37	464	39	109	83	73	22	102	67	89
	googl	23	39	201	42	34	37	17	49	32	38
	jpm	43	109	42	603	122	80	33	114	78	100
	ms	36	83	34	122	458	65	24	74	57	76
	msft	28	73	37	80	65	460	19	78	52	69
	nflx	17	22	17	33	24	19	382	26	24	24
	pg	37	102	49	114	74	78	26	778	89	114
	sbux	34	67	32	78	57	52	24	89	469	66
	ups	36	89	38	100	76	69	24	114	66	631
QP Port.		alxn	amzn	bac	biib	blk	ftr	googl	hst	nflx	pcln
	alxn	376	18	14	20	20	6	12	13	15	6
	amzn	18	302	18	19	24	5	23	17	17	9
	bac	14	18	298	11	21	6	13	19	12	7
	biib	20	19	11	187	17	1	18	13	11	11
	blk	20	24	21	17	218	5	24	23	19	8
	ftr	6	5	6	1	5	460	6	6	3	1
	googl	12	23	13	18	24	6	201	18	17	12
	hst	13	17	19	13	23	6	18	351	6	4
	nflx	15	17	12	11	19	3	17	6	382	9
	pcln	6	9	7	11	8	1	12	4	9	116
extent		1	2	3	4	5	6	7	8	9	10
count typ.	160958	3017	324	112	49	35	18	13	13	7	1
count QP	161973	2394	112	34	13	12	4	5	0	0	0
Rej. typ.	NA	NA	•	•	•	•	•	•	•	•	•
Rej. QP	NA	NA	•	•	•	•	•	•			

Note: The top two panels show the number of detected bivariate cojump for each pair of the typical portfolio and the 'minimum cojump portfolio', respectively. The respective portfolio constituents can be read from the first row/column in the matrices. The lower panel contains the the number of cojumps by extent. The row 'rejections' indicates whether the null of independent cojump processes was rejected, based on Z_k and asymptotic critical values for given extent k. A "•" indicates rejection.

are given by, disregarding potential time variation,

$$w = \frac{\widehat{\Sigma}^{-1}\iota}{\iota'\widehat{\Sigma}^{-1}\iota},$$

where ι is a $(d \times 1)$-vector of ones and $\widehat{\Sigma}$ is the sample covariance matrix of returns. We modify this approach and define the sample *cojump covariance matrix* $\widehat{\Sigma}_{CJ} := \frac{1}{M}\widehat{\boldsymbol{J}}'\widehat{\boldsymbol{J}}$, where $\widehat{\boldsymbol{J}}$ was defined in Section 2.3. This is, in a way, a restricted covariance matrix, as all non-cojump events are filtered out. With this modification we define the MCP weights

$$w_{CJ} = \frac{\widehat{\Sigma}_{CJ}^{-1}\iota}{\iota'\widehat{\Sigma}_{CJ}^{-1}\iota}.$$

We find all elements in w_{CJ} to be nonzero, so to construct a 10-dimensional MCP we select the assets with 10 largest weights in absolute value. In Table 8 we compare the arbitrary portfolio studied in the last section with the portfolio constructed in this way. The table shows the pairwise cojump frequency for the included pairs with jump frequencies on the main diagonal.

Table 9: Minimum cojump portfolio for negative jumps only

Typ. Port		amzn	brk-b	googl	jpm	ms	msft	nflx	pg	sbux	ups
	amzn	154	17	11	19	18	13	7	17	16	18
	brk-b	17	238	18	53	49	38	6	48	28	49
	googl	11	18	100	17	18	17	3	17	11	19
	jpm	19	53	17	313	65	41	10	59	35	51
	ms	18	49	18	65	237	35	9	36	23	44
	msft	13	38	17	41	35	229	4	39	22	38
	nflx	7	6	3	10	9	4	169	8	9	9
	pg	17	48	17	59	36	39	8	416	40	65
	sbux	16	28	11	35	23	22	9	40	230	31
	ups	18	49	19	51	44	38	9	65	31	339
QP Port.		amzn	bac	biib	blk	ftr	ge	googl	hban	nflx	pcln
	amzn	154	6	8	11	2	13	11	6	7	3
	bac	6	160	5	10	4	13	4	7	5	2
	biib	8	5	94	10	0	10	7	4	3	4
	blk	11	10	10	115	2	13	10	10	6	2
	ftr	2	4	0	2	235	2	2	4	0	0
	ge	13	13	10	13	2	134	9	10	4	2
	googl	11	4	7	10	2	9	100	5	3	4
	hban	6	7	4	10	4	10	5	196	7	1
	nflx	7	5	3	6	0	4	3	7	169	2
	pcln	3	2	4	2	0	2	4	1	2	54
extent	0	1	2	3	4	5	6	7	8	9	10
count typ.	162717	1522	176	59	34	19	7	7	3	2	1
count QP	163276	1185	62	9	8	3	2	0	2	0	0
Rej. typ.	NA	NA	•	•	•	•	•	•	•	•	•
Rej. QP	NA	NA	•	•	•	•	•	•	•		

Note: The top two panels show the number of detected bivariate cojump for each pair of the typical portfolio and the 'minimum cojump portfolio', respectively. The respective portfolio constituents can be read from the first row/column in the matrices. The lower panel contains the the number of cojumps by extent. The row 'rejections' indicates whether the null of independent cojump processes was rejected, based on Z_k and asymptotic critical values for given extent k. A "•" indicates rejection.

The bottom panel shows the frequency of higher order cojumps and the test outcome for testing against cojumps of extent k or higher. The same analysis has been done considering negative jumps only and the results can be found in Table 9.

Interestingly, when restricting our attention to negative jumps, the diversification potential seems to be larger than for unrestricted jumps. We find a 78.2% decrease in the restricted MCP compared to the typical portfolio. For the unrestricted MCP the cojump decrease amounts to 73.4%. We conclude that there is some moderate cojump diversification potential, which may still be improved upon by recognizing potential time variation in the multivariate dependence structure of jump events. Nevertheless, large cojumps are hard to get rid of, as our test still rejects both portfolios with maximum significant extent of 7 and 8 respectively. Relatively few cojumps of a large extent will always lead to the rejection of the null hypothesis.

5 Conclusion

In this paper we have presented a set of formal test statistics for the null hypothesis of absence of cojumps in high frequency asset prices against the alternative that cojumps of arbitrary degree are present. The basic idea underlying the tests is that in discretely observed prices a certain number of cojumps occur simply by chance. Another source of observed cojumps are type I errors in the (univariate) jump detection procedure on which the cojumps tests are based. We provide an iterative updating algorithm that improves the finite sample properties of the jump detection procedure proposed by Lee and Mykland (2008) and the improved properties carry over to the cojump tests. The test can easily be adapted to test against the alternative of cojumps of a certain degree m, i.e., of at least m assets cojumping. Furthermore, non iid'ness of the jumps can be taken into account in our approach by assuming that jumps are driven by a Hawkes process, either univariate or even allowing for spillovers, i.e., jumps in one asset leading jumps in other assets. The asymptotic distributions of our statistics are derived, but simulations showed that in general they provide poor approximations to the finite sample behavior of the test. A block bootstrap procedure mitigates these problems. Our Monte Carlo simulations show that the tests have good power against processes with even moderate cojumps probabilities and that the competing test statistics have very similar power, which implies that it suffices to consider only the most basic version, the statistic Z. Our empirical analysis of 1-minute returns of a large part of the constituents of the S&P 500 shows that cojumps are a general feature of the data, both when looking a cojumps of any type, as well as when considering only jumps of positive or negative sign. In fact, any attempts to construct portfolios of a subset of assets that does not feature significant cojumps was unsuccessful. Nevertheless, we suggest constructing the minimum cojump portfolio, similar to the well known minimum variance portfolio, but based on the covariance of the jump indicators instead of the standard covariance matrix. Even though such a portfolio still is characterized by the significant presence of cojumps we observe that the cojump risk can be reduced significantly with this approach. The extended model based on Hawkes processes suggests that the jump intensities are not predictable and do not cluster significantly.

Future research may apply our proposed tests in different context such as a foreign exchange markets, commodity prices or to stock market indices instead of individual stocks. Furthermore, currently the application to very large dimensional portfolios is restricted due to the fact that the number of derivatives of the probability generating function that is needed to compute the expected probabilities of cojumps grows with the dimension of the problem. Efficient tools to approximate these probabilities or justified truncations of these probabilities may be considered. Finally, an alternative approach to test for cojumps may be based on the parametric multivariate peaks-over-threshold Hawkes model by Grothe et al. (2014), which explicitly assigns positive probability to co-exceedances, but may be restricted to the case of independent tails. A likelihood ratio test may thus be derived as an alternative to the approach presented in this paper.

References

Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117:585–606.

Aït-Sahalia, Y. and Jacod, J. (2009). Testing for jumps in a discrete observed process. *The Annals of Statistics*, 37(1):184–222.

Andersen, T. G., Bollerslev, T., and Dobrev, D. (2007). No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics*, 138:125–180.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(75-93).

Barndorff-Nielsen, O. E., Graversen, S. E., Jacod, J., Podolskij, M., and Shephard, N. (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In Yu, K., Liptser, R., and Stoyanov, J., editors, *From Stochastic Calculus to Mathematical Finance: The Shiryaev Festschrift*, chapter 3, pages 33–68. Springer.

Barndorff-Nielsen, O. E. and Shepard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37.

Barndorff-Nielsen, O. E. and Shepard, N. (2006). Econometrics of testing for jumps in financial econometrics using bipower variation. *Journal of Financial Econometrics*, 4(1):1–30.

Bibinger, M. and Reiß, M. (2014). Spectral estimation of covolatility from noisy observations using local weights. *Scandinavian Journal of Statistics*, 41(1):23–50.

Bibinger, M. and Winkelmann, L. (2015). Econometrics of co-jumps in high-frequency data with noise. *Journal of Econometrics*, 184(2):361–378.

Bollerslev, T., Law, T. H., and Tauchen, G. (2008). Risk, jumps and diversification. *Journal of Econometrics*, 144:234–256.

Bollerslev, T. and Todorov, V. (2011). Tails, fear and risk premia. *Journal of Finance*, 66(6):2165–2211.

Bollerslev, T., Todorov, V., and Li, S. Z. (2013). Jump tails, extreme dependencies, and the distribution of stock returns. *Journal of Econometrics*, 172:307–324.

Bormann, C. and Schaumburg, M. S. J. (2016). Beyond dimension two: A test for higher-order tail risk. *Journal of Financial Econometrics*, page forthcoming.

Bormetti, G., Calcagnile, L. M., Treccani, M., Corsia, F., Marmi, S., and Lillo, F. (2015). Modelling systemic price cojumps with hawkes factor models. *Quantitative Finance*, 15(7):1137–1156.

Bos, C. S. and Janus, P. (2013). A quantile-based realized measure of variation: New tests for outlying observations in financial data. Technical report, Tinbergen Institute Discussion Paper.

Boudt, K., Croux, C., and Laurent, S. (2011). Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Empirical Finance*, 18:353–367.

Bryc, A. (2002). A uniform approximation to the right normal tail integral. *Applied Mathematics and Computation*, 127(2-3):365–374.

Caporin, M., Kolokolov, A., and Renò, R. (2015). Multi-jumps. Technical report, Stryto Working Paper Series No. 3.

Chavez-Demoulin, V., Davison, A. C., and McNeil, A. J. (2005). Estimating Value-at-Risk: A point process approach. *Quantitative Finance*, 5(2):227–234.

Chavez-Demoulin, V. and McGill, J. (2012). High-frequency financial data modeling using Hawkes processes. *Journal of Banking & Finance*, 36:34153426.

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC.

Corsi, F., Pirino, D., and Ren, R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159:276–288.

Dumitru, A.-M. and Urga, G. (2012). Identifying jumps in financial assets: A comparison between nonparametric jump tests. *Journal of Business & Economic Statistics*, 30(2):242–255.

Embrechts, P., Liniger, T., and Lin, L. (2011). Multivariate Hawkes Processes: an Application to Financial data. *Journal of Applied Probability*, 48(A):367–378.

Eraker, B., Johannes, M., and Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance*, 58(3):1269–1300.

Fleming, J., Kirby, C., and Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, 56:329–352.

Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics*. Robert E. Krieger Publishing Co., Inc.

Gilder, D., Shackleton, M. B., and Taylor, S. J. (2014). Cojumps in stock prices: Empirical evidence. *Journal of Banking & Finance*, 40:443–459.

Grothe, O., Korniichuk, V., and Manner, H. (2014). Modeling multivariate extreme events using self-exciting point processes. *Journal of Econometrics*, 182:269–289.

Hawkes, A. G. (1971). Point Spectra of Some Mutually Exciting Point Processes. *Journal of the Royal Statistical Society B*, 33(3):438–443.

Jacod, J. and Todorov, V. (2009). Testing for common arrival of jumps for didiscrete observed multidimensional processes. *The Annals of Statistics*, 37(4):1792–1838.

Jiang, G. J. and Oomen, R. C. A. (2008). Testing for jumps when asset prices are observed with noise - a swap variance approach. *Journal of Econometrics*, 144:352–370.

Lee, S. S. and Mykland, P. A. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Studies*, 21(6):2535–2563.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36:270–296.

Mancini, C. and Gobbi, F. (2012). Identifying the Brownian covariation from the co-jumps given discrete observations. *Econometric Theory*, 28:249–273.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton University Press.

Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 30(Part A):243–261.

Pan, J. (2002). The jump-risk premia implicit in options: evidence from and integrated time-series study. *Journal of Financial Economics*, 63:3–50.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89:1303–1313.

Rousseeuw, P. and Leroy, A. (1988). A robust scale estimator based on the shortest half. *Statistica Neerlandica*, 42(2):103–116.

Todorov, V. (2010). Variance risk-premium dynamics: the role of jumps. *The Review of Financial Studies*, 23(1):345–383.

van der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.

Appendix A Proofs

Proof of Lemma 1. We will first establish the usefull equivalence in law $\int_{\delta(j-1)}^{\delta j} \sigma_s dW_s \overset{law}{=} \sigma_{\delta(j-1)}\sqrt{\delta}U_j$. For simplicity write $a := \delta(j-1)$, $b := \delta j$, $\delta := b - a$ and $a = t_0 < \cdots < t_n = b$ with $t_{i+1} - t_i = \delta/n$. Let $U_j(\cdot) \sim \mathcal{N}(0, \cdot)$ and $U_j \sim \mathcal{N}(0,1)$. By definition of the Itō integral

$$\begin{aligned}
\int_a^b \sigma_s dW_s &= \lim_{n\to\infty} \left(\sum_{j=1}^{n} \sigma_{t_{i-1}} \Delta W_{t_i} \right) \\
&\overset{law}{=} \lim_{n\to\infty} \left(\sum_{j=1}^{n} U_j \left(\frac{\delta}{n} \sigma^2_{t_{i-1}} \right) \right) \\
&= \lim_{n\to\infty} U_j \left(\sum_{j=1}^{n} \frac{\delta}{n} \sigma^2_{t_{i-1}} \right) \\
&= U_j \left(\int_a^b \sigma_s^2 ds \right) \\
&= \sqrt{\overline{\sigma^2_{s^*}}}\sqrt{c} U_j \\
&= \sigma_a \sqrt{c} U_j + O_p(\delta^{3/2-\epsilon})
\end{aligned}$$

where $\Delta W_{t_i} \sim \mathcal{N}(0, \frac{\delta}{n})$, which implies $\sigma_{t_{i-1}} \Delta W_{t_i} \sim \mathcal{N}(0, \frac{\delta}{n}\sigma^2_{t_{i-1}})$. The third equality holds as long as the σ_{t_i} are square summable, which is ensured by the integrability of the stochastic volatility process. By the Mean Value Theorem and continuity of σ_t, there exists $\delta\overline{\sigma^2_{s^*}} = \int_a^b \sigma_s^2 ds$, with $s^* \in [a,b]$. Note that, as $b \to a$, $\overline{\sigma^2_{s^*}} \to \sigma_a^2$. Therefore also $\sqrt{\overline{\sigma^2_{s^*}}} \to \sigma_a$, which yields the last equality.

Now consider the denominator of (5). By Proposition 1 in Andersen et al. (2012), as long as $K = O_p(\delta^\alpha)$, $\alpha \in (-1, -\frac{1}{2})$, we have for our robust local volatility estimate over the window $(t_{j-k}, t_j]$,

$$\text{Med-RV}(t_{j-k}, t_j) := \vartheta \frac{K}{K-1} \sum_{i=0}^{K-2} \text{med}(|y_{j-i}|, |y_{j-i-1}|, |y_{j-i-2}|)^2 \xrightarrow{p} \int_{t_{j-K}}^{t_j} \sigma_s^2 ds = O_p(\delta^{1+\alpha}) \tag{38}$$

This implies

$$\sqrt{\frac{1}{K}\text{Med-RV}(t_{j-k}, t_j)} = O_p(\delta^{-1/2}).$$

Now, if there are no jumps in $[a,b]$. Then, by using assumption 1 in the third equality yields

$$\begin{aligned}
\mathcal{L}_j = \frac{y_j}{\hat{\sigma}_j} &= \frac{\int_{\delta(j-1)}^{\delta j} a_s ds + \int_{\delta(j-1)}^{\delta j} \sigma_s dW_s}{\sqrt{\vartheta \frac{K}{(K-1)(K-2)} \sum_{i=0}^{K-2} \text{med}(|y_{j-i}|, |y_{j-i-1}|, |y_{j-i-2}|)^2}} \\
&= \frac{a_{\delta(j-1)}\delta + O_p(\delta^{3/2-\epsilon})}{O_p(\delta^{1/2})} + \frac{\sigma_a\sqrt{c}U_j + O_p(\delta^{3/2-\epsilon})}{\sqrt{\vartheta \frac{K}{(K-1)(K-2)} \sum_{i=0}^{K-2} \text{med}(|y_{j-i}|, |y_{j-i-1}|, |y_{j-i-2}|)^2}} \\
&\xrightarrow{p} U_j
\end{aligned} \tag{39}$$

as long as our local volatility estimator converges to $\sigma_a\sqrt{\delta}$, which we will show in turn. To see

this, apply the continuous mapping theorem to the denominator to conclude that

$$\sqrt{\frac{1}{K}\text{Med-RV}(t_{j-k},t_j)} \xrightarrow{p} \sqrt{\int_{t_{j-1}}^{t_j}\sigma_u^2 du} = \sigma_{t_{j-1}}\sqrt{\delta} + O_p(\delta^{3/2-\epsilon})$$

On the other hand, if there is a jump in $(t_{j-1},t_j]$ (there can be at most one jump in any infinitesimal interval), analogously to (39)

$$\mathcal{L}_j = \frac{a_{\delta(j-1)}\delta + O_p(\delta^{3/2-\epsilon})}{O_p(\delta^{1/2})} + \frac{\sigma_a\sqrt{c}U_j + O_p(\delta^{3/2-\epsilon}) + c_{\tilde{s}}}{\sqrt{\vartheta\frac{K}{(K-1)(K-2)}\sum_{i=0}^{K-2}\text{med}(|y_{j-i}|,|y_{j-i-1}|,|y_{j-i-2}|)^2}} \tag{40}$$

$$\xrightarrow{p} U_j + \frac{c_{\tilde{s}}}{\sigma_{t_{j-1}}\sqrt{\delta}} \longrightarrow \infty, \tag{41}$$

where $\tilde{s} \in (t_{j-1},t_j]$. □

Proof of Proposition 1. We drop the subindex i in this proof, as it only concerns univariate results. The idea of the proof is as follows. First we consider a continuous version of the problem, wherein the iterative updating procedure can be shown to converge to a unique fixed-point. To show equivalence of the fixed point and the minimizer of the joint error probabilities, we obtain an equation which must hold if they are equal and a second equation which characterizes the optimum. These two equations are shown to be equivalent asymptotically in κ. The last step of the proof shows that the "discretization" error that incurs when going from the continuous version back to the original formulation becomes negligible in $\widetilde{M}$.

Using Lemma 1, we can derive the distribution of LM_j conditional on N jumps in the sample. Let $w := \frac{\widetilde{M}-N}{\widetilde{M}}$ denote the fraction of returns which do not contain jumps. Then, LM follows the mixture folded distribution

$$\begin{aligned} F_m(\beta) &= wF_d(S_{\widetilde{M}}\beta + C_{\widetilde{M}}) + (1-w)F_{d+j}(S_{\widetilde{M}}\beta + C_{\widetilde{M}}) \\ &\approx wF_d(S_{\widetilde{M}}\beta + C_{\widetilde{M}}) + (1-w)F_j(S_{\widetilde{M}}\beta + C_{\widetilde{M}}) \\ &= w\,\text{erf}\left(\frac{S_{\widetilde{M}}\beta + C_{\widetilde{M}}}{\sqrt{2}}\right) + (1-w)F_j(S_{\widetilde{M}}\beta + C_{\widetilde{M}}) \end{aligned} \tag{42}$$

where $F_d(x) = \text{erf}\left(\frac{x}{\sqrt{2}}\right)$ is the folded normal CDF of the diffusive increments (i.e. returns without jumps) and F_{d+j} is the distribution of the sum of the diffusive increment and the jump for interval in which jumps occur. We approximate $F_{d+j} \approx F_j$ as usually $c_j \gg \int_{(j-1)\delta}^{j\delta}\sigma_s dW_s$. With this distribution the overall size of the univariate jump detection is bounded above by α as $1-\alpha = F_\xi(\beta) = F_m(\beta)^{\widetilde{M}} \Rightarrow F_m(\beta) = (1-\alpha)^{1/\widetilde{M}} > 1-\alpha$.

F_m as a sum of the folded diffusion density and the folded bimodal normal mixture is everywhere differentiable. The function $T(x) : \mathbb{R} \to \mathbb{R}, x \mapsto -\log\left(\frac{\hat{\sigma}\sqrt{T}\widetilde{M}(1-F_m(x))}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right)$ is monotone and a contraction on $\mathbb{R}$ as for $x = \beta, y = \beta+\mu,\ \mu > 0$ such that w.l.o.g. $x < y$ and

$$d(T(x),T(y)) = \log\left(\frac{\hat{\sigma}\sqrt{T}\widetilde{M}(1-F_m(\beta))}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right) - \log\left(\frac{\hat{\sigma}\sqrt{T}\widetilde{M}(1-F_m(\beta+\mu))}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right) \tag{43}$$

$$= \log\left(\frac{1-F_m(\beta)}{1-F_m(\beta+\mu)}\right) < \mu = |y-x|,\ \forall\mu > 0 \tag{44}$$

By the compactness of $\mathbb{R}$ it follows from the Banach Fixed-Point Theorem that β_n converges to the unique fixed-point β_{fix}

$$\beta_{fix} = -\log\left(\frac{\hat{\sigma}\sqrt{T}\widetilde{M}(1-F_m(\beta_{fix})}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right).$$

By equating β^* and β_{fix} we obtain the CDF equation

$$-\log\left(\frac{\sigma\mu_1^2\sqrt{T}N}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right) = -\log\left(\frac{\hat{\sigma}\sqrt{T}\widetilde{M}(1-F_m(\beta_{fix}))}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right) \tag{45}$$

$$\Leftrightarrow -\log\left(\frac{\sigma\mu_1^2\sqrt{T}N}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right) = -\log\left(\frac{\sigma\sqrt{T}\widetilde{M}(1-F_m(\beta_{fix}))}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right) + O_p\big(-\log(1+\widetilde{M}^{-1/2})\big) \tag{46}$$

$$\Leftrightarrow N = \widetilde{M}(1-F_m(\beta_{fix})) + O_p\big(-\log(1+\widetilde{M}^{-1/2})\big) \tag{47}$$

$$\Leftrightarrow F_m(\beta_{fix}) = \underbrace{\frac{\widetilde{M}-N}{\widetilde{M}}}_{=:w} + O_p\big(-\widetilde{M}\log(1+\widetilde{M}^{-1/2})\big) \tag{48}$$

$$\Leftrightarrow wF_d(\beta_{fix}) + (1-w)F_j(\beta^*) = w + O_p\big(-\widetilde{M}\log(1+\widetilde{M}^{-1/2})\big) \tag{49}$$

$$\textbf{(CDF)} \qquad \Leftrightarrow (1-w)F_j(\beta_{fix}) = w(1-F_d(\beta_{fix})) + O_p\big(-\widetilde{M}\log(1+\widetilde{M}^{-1/2})\big) \tag{50}$$

For a given sample the optimization problem reads slightly different, as jumps and volatility have realized

$$\beta^* = \text{argmin}_\beta\,(\#\text{Spurious Detections} + \#\text{Not Detected Jumps}) \tag{51}$$

$$= \text{argmin}_\beta\,((1-w)\mathbb{F}_j(\beta) + w(1-\mathbb{F}_d(\beta^*))) \tag{52}$$

By again considering the continuous version, we can apply first-order conditions to get as conditions for optimality

$$\textbf{(PDF)} \qquad (1-w)f_j(\beta^*) = wf_d(\beta^*) \tag{53}$$

We proof the equivalence of (50) and (53) for a simple case of bimodal normal mixture jump sizes with unit variance and modes $\pm\kappa$. Without loss of generality we set the mean of the diffusion density to zero, (53) then reads as

$$(1-w)\frac{1}{\sqrt{2\pi}}e^{-\frac{(\beta-\kappa)^2}{2}} = w\frac{2}{\sqrt{2\pi}}e^{-\frac{\beta^2}{2}} \tag{54}$$

$$\beta = \frac{\kappa^2 - 2\log\left(\frac{1-w}{2w}\right)}{2\kappa} \tag{55}$$

By an appropriate approximation to the normal tail integral we can transform the CDFs in (50) to functions of their PDFs times some factor which yields

$$(1-w)F_j(\beta^*) = w(1-F_d(\beta^*)) \tag{56}$$

$$(1-w)\phi(\beta-\kappa)g(\beta-\kappa) = 2w\phi(\beta)g(x) \tag{57}$$

$$\beta_{fix} = \frac{\kappa^2 - 2\log\left(\frac{(1-w)\eta}{2w}\right)}{2\kappa} \tag{58}$$

where $\eta = \frac{g(x-\kappa)}{g(x)}$. Using Bryc (2002)'s uniform approximation,

$$\eta(x) = \frac{\frac{\sqrt{2\pi}(x-\kappa+3.\bar{3})}{\sqrt{2\pi}(x-\kappa)^2+7.32(x-\kappa)+6.\bar{6}}}{\frac{\sqrt{2\pi}(x+3.\bar{3})}{\sqrt{2\pi x^2+7.32x+6.\bar{6}}}} = O(\kappa). \tag{59}$$

Therefore,

$$\beta^* - \beta_{fix} = \frac{\log\eta}{\kappa} = O\left(\frac{\log\kappa}{\kappa}\right). \tag{60}$$

It remains to be shown that the error made by considering the discrete version of the problem becomes negligible in $\widetilde{M}$. Recall that the iterative procedure will stop converging at value $\hat{\beta}^*$ as soon as $\beta_{n+1} = \beta_n$ for some n, which is the case when $\mathbb{F}(\beta_{n+1}) = \mathbb{F}(\beta_n)$, i.e. when both β_{n+1}, β_n fall into the interval $[q_m, q_{m+1})$ of two adjacent empirical quantiles q_m, q_{m+1} for some $m \in \{1, ..., \widetilde{M} - 1\}$. Assume w.l.o.g. that $\beta^* \in [q_m, q_{m+1})$ for some m. Therefore the discretization error is

$$\left|\beta^* - \hat{\beta}^*\right| < \max\left(\left|\beta^* + \log\left(\frac{\hat{\sigma}\sqrt{T}\widetilde{M}(1-\mathbb{F}_m(q_m))}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right)\right|, \left|\beta^* + \log\left(\frac{\hat{\sigma}\sqrt{T}\widetilde{M}(1-\mathbb{F}_m(q_{m+1}))}{\sqrt{2\widetilde{M}\log\widetilde{M}}}\right)\right|\right) \tag{61}$$

$$< \left|\log\left(\frac{1-\mathbb{F}(q_{m+1})}{1-\mathbb{F}(q_m)}\right)\right| \xrightarrow{a.s.} 0 \text{ as } \widetilde{M} \to \infty, \tag{62}$$

where the last step follows from Glivenko-Cantelli and the Continuous Mapping Theorem. □

Proof of Corollary 2. We drop the subindex i in this proof, as it only concerns univariate results. As mentioned earlier, we have perfect jump identification for fixed T under the δ-asymtotics

$$\hat{N}_t \xrightarrow{p} N_t,$$

such that the jump detection can δ-asymtotically be ignored. By definition of the Poisson process, $N_\lambda(t) - N_\lambda(s), s < t$, is Poisson distributed with parameter $\lambda(t-s)$. This implies that the number of jumps over the whole sample $N_t \sim Poisson(\lambda T)$, such that, by the strong law of large numbers and Continuous Mapping, as $T \longrightarrow \infty$

$$\hat{\lambda} = \frac{\hat{N}_T}{T} \xrightarrow{a.s.} \lambda.$$

Now consider the polynomial in (20). It's simple to see that the lowest order terms when expanding the polynomial is of order λ_i to the power of one. By the Continuous Mapping Theorem, the order of the λ-estimating error goes through and we can establish

$$P_k = \hat{P}_k + O_p(T^{-1/2}).$$

□

Proof of Proposition 3. By the Poisson property and the assumed independence of the individual jump processes under the null, we have that g_j are independent and identically distributed with distribution function F. By Donsker's Theorem (van der Vaart, 1998, Thm. 19.3), as $\delta \downarrow 0$,

$$\sqrt{\widetilde{M}} \begin{pmatrix} \mathbb{F}_M(0) - F(0) \\ \vdots \\ \mathbb{F}_M(d) - F(d) \end{pmatrix} \rightsquigarrow \mathcal{N}_d \left(\mathbf{0}, \begin{pmatrix} F(0)(1-F(0)) & \cdots & F(0) - F(0)F(d) \\ \vdots & F(a \wedge b) - F(a)F(b) & \ddots & \vdots \\ F(0) - F(0)F(d) & \cdots & F(d)(1-F(d)) \end{pmatrix} \right), \tag{63}$$

where $F(a \wedge b) = \min[F(a), F(b)]$.

1) Distribution of Z

$$\begin{aligned} Z &= \sqrt{\widetilde{M}} \Big[(\mathbb{P}_2 - P_2) + ... + (\mathbb{P}_d - P_d) \Big] && (64) \\ &= \sqrt{\widetilde{M}} \Big[\sum_{i=2}^{d} \mathbb{P}_i - \sum_{i=2}^{d} P_i \Big] && (65) \\ &= \sqrt{\widetilde{M}} \Big[(\mathbb{F}(d) - \mathbb{F}(1)) - (F(d) - F(1)) \Big] && (66) \\ &= -\sqrt{\widetilde{M}} \Big[\mathbb{F}(1) - F(1) \Big], && (67) \end{aligned}$$

where $\mathbb{F}(k) = \frac{1}{d} \sum_{i=0}^{k} \mathbb{P}_i$ is the empirical CDF of g_j and $\mathbb{F}(d) = F(d) = 1$ by construction. Now, by Donsker's Theorem, $\sqrt{\widetilde{M}} \left(\mathbb{F}(x) - F(x) \right) \xrightarrow{d} \mathcal{N}(0, F(x)(1-F(x))$, and by symmetry of the standard normal distribution, the assertion follows for Z.

2) Distribution of Z_1 For Z_1 observe

$$Z_1 = \sqrt{(n)}(\mathbb{F}_n(1) - F(1)) - \sqrt{(n)}(\mathbb{F}_n(0) - F(0)),$$

which by Donsker's Theorem and a standard result for sums of correlated normal variates is distributed as stated.

3) Distribution of Z^2 For all versions of the squared test statistics, their stated asymptotic distributions follow straightforwardly from Lemma 17.1 and the proof of Theorem 17.2 in van der Vaart (1998, p. 243). □

Appendix B List of assets in the empirical analysis

Table B.1: List of assets in our sample. Assets marked with * are excluded in our analysis due to missing values.

	A	Agilent Technologies Inc		DE	Deere & Co.		JNPR	Juniper Networks		PSX	Phillips 66
	AA	Alcoa Inc		DFS	Discover Financial Services	*		Joy Global	*	PVH	PVH Corp.
	AAL	American Airlines Group		DG	Dollar General		JPM	JPMorgan Chase & Co.	*	PWR	Quanta Services Inc.
*	AAP	Advance Auto Parts		DGX	Quest Diagnostics		JWN	Nordstrom		PX	Praxair Inc.
	AAPL	Apple Inc.		DHI	D. R. Horton		K	Kellogg Co.		PXD	Pioneer Natural Resources
	ABBV	AbbVie		DHR	Danaher Corp.		KEY	KeyCorp	*	PYPL	PayPal
	ABC	AmerisourceBergen Corp		DIS	The Walt Disney Company		KIM	Kimco Realty		QCOM	QUALCOMM Inc.
	ABT	Abbott Laboratories		DISCA	Discovery Communications-A		KLAC	KLA-Tencor Corp.	*	QRVO	Qorvo
		ACE Limited		DLPH	Delphi Automotive		KMB	Kimberly-Clark	*	R	Ryder System
	ACN	Accenture plc		DLTR	Dollar Tree		KMI	Kinder Morgan		RAI	Reynolds American Inc.
	ADBE	Adobe Systems Inc	*	DNB	Dun & Bradstreet		KMX	Carmax Inc		RCL	Royal Caribbean Cruises Ltd
	ADI	Analog Devices, Inc.		DO	Diamond Offshore Drilling		KO	The Coca Cola Company	*	REGN	Regeneron
	ADM	Archer-Daniels-Midland Co		DOV	Dover Corp.		KORS	Michael Kors Holdings		RF	Regions Financial Corp.
	ADP	Automatic Data Processing		DOW	Dow Chemical		KR	Kroger Co.	*	RHI	Robert Half International
*	ADS	Alliance Data Systems		DPS	Dr Pepper Snapple Group		KSS	Kohl's Corp.		RHT	Red Hat Inc.
	ADSK	Autodesk Inc	*	DRI	Darden Restaurants	*	KSU	Kansas City Southern		RIG	Transocean
	ADT	ADT Corp	*	DTE	DTE Energy Co.	*	L	Loews Corp.	*	RL	Polo Ralph Lauren Corp.
	AEE	Ameren Corp		DUK	Duke Energy		LB	L Brands Inc.	*	ROK	Rockwell Automation Inc.
	AEP	American Electric Power	*	DVA	DaVita Inc.	*	LEG	Leggett & Platt	*	ROP	Roper Industries
	AES	AES Corp		DVN	Devon Energy Corp.		LEN	Lennar Corp.		ROST	Ross Stores
	AET	Aetna Inc		EA	Electronic Arts	*	LH	Laboratory Corp. of America Holding		RRC	Range Resources Corp.
	AFL	AFLAC Inc		EBAY	eBay Inc.	*	LLL	L-3 Communications Holdings	*	RSG	Republic Services Inc
*	AGN	Allergan plc		ECL	Ecolab Inc.		LLTC	Linear Technology Corp.		RTN	Raytheon Co.
	AIG	American International Group, Inc.		ED	Consolidated Edison		LLY	Lilly (Eli) & Co.		SBUX	Starbucks Corp.
*	AIV	Apartment Investment & Mgmt	*	EFX	Equifax Inc.	*	LM	Legg Mason	*	SCG	SCANA Corp
*	AIZ	Assurant Inc		EIX	Edison Int'l		LMT	Lockheed Martin Corp.		SCHW	Charles Schwab Corporation
	AKAM	Akamai Technologies Inc		EL	Estee Lauder Cos.		LNC	Lincoln National		SE	Spectra Energy Corp.
	ALL	Allstate Corp		EMC	EMC Corp.		LOW	Lowe's Cos.	*	SEE	Sealed Air Corp.(New)
*	ALLE	Allegion		EMN	Eastman Chemical		LRCX	Lam Research	*	SHW	Sherwin-Williams
		Altera		EMR	Emerson Electric Company	*	LUK	Leucadia National Corp.	*	SIG	Signet Jewelers
	ALXN	Alexion Pharmaceuticals		ENDP	Endo International		LUV	Southwest Airlines	*	SJM	Smucker (J.M.)
	AMAT	Applied Materials Inc		EOG	EOG Resources		LVLT	Level 3 Communications		SLB	Schlumberger Ltd.
*	AME	Ametek	*	EQIX	Equinix		LYB	LyondellBasell	*	SLG	SL Green Realty
*	AMG	Affiliated Managers Group Inc		EQR	Equity Residential		M	Macy's Inc.	*	SNA	Snap-On Inc.
	AMGN	Amgen Inc		EQT	EQT Corporation		MA	Mastercard Inc.		SNDK	SanDisk Corporation
	AMP	Ameriprise Financial		ES	Eversource Energy	*	MAC	Macerich	*	SNI	Scripps Networks Interactive Inc.
	AMT	American Tower Corp A		ESRX	Express Scripts		MAR	Marriott Int'l.		SO	Southern Co.
	AMZN	Amazon.com Inc	*	ESS	Essex Property Trust Inc		MAS	Masco Corp.		SPG	Simon Property Group Inc
*	AN	AutoNation Inc		ESV	Ensco plc		MAT	Mattel Inc.		SPLS	Staples Inc.
	ANTM	Anthem Inc.		ETFC	E*Trade		MCD	McDonald's Corp.	*	SRCL	Stericycle Inc
	AON	Aon plc		ETN	Eaton Corporation		MCHP	Microchip Technology		SRE	Sempra Energy
	APA	Apache Corporation		ETR	Entergy Corp.		MCK	McKesson Corp.		STI	SunTrust Banks
	APC	Anadarko Petroleum Corp	*	EW	Edwards Lifesciences	*	MCO	Moody's Corp		STJ	St Jude Medical
*	APD	Air Products & Chemicals Inc		EXC	Exelon Corp.		MDLZ	Mondelez International		STT	State Street Corp.
*	APH	Amphenol Corp A	*	EXPD	Expeditors Int'l		MDT	Medtronic plc		STX	Seagate Technology
*	ARG	Airgas Inc		EXPE	Expedia Inc.		MET	MetLife Inc.		STZ	Constellation Brands
*	AVB	AvalonBay Communities, Inc.		F	Ford Motor	*	MHFI	McGraw Hill Financial		SWK	Stanley Black & Decker
	AVGO	Avago Technologies		FAST	Fastenal Co	*	MHK	Mohawk Industries		SWKS	Skyworks Solutions
*	AVY	Avery Dennison Corp		FB	Facebook		MJN	Mead Johnson		SWN	Southwestern Energy
	AXP	American Express Co		FCX	Freeport-McMoran Cp & Gld	*	MKC	McCormick & Co.		SYK	Stryker Corp.
*	AZO	AutoZone Inc		FDX	FedEx Corporation	*	MLM	Martin Marietta Materials		SYMC	Symantec Corp.
	BA	Boeing Company		FE	FirstEnergy Corp		MMC	Marsh & McLennan		SYY	Sysco Corp.
	BAC	Bank of America Corp	*	FFIV	F5 Networks		MMM	3M Company		T	AT&T Inc
	BAX	Baxter International Inc.	*	FIS	Fidelity National Information Services	*	MNK	Mallinckrodt Plc	*	TAP	Molson Coors Brewing Company
	BBBY	Bed Bath & Beyond	*	FISV	Fiserv Inc		MNST	Monster Beverage	*	TDC	Teradata Corp.
	BBT	BB&T Corporation		FITB	Fifth Third Bancorp		MO	Altria Group Inc	*	TE	TECO Energy
	BBY	Best Buy Co. Inc.	*	FLIR	FLIR Systems		MON	Monsanto Co.		TEL	TE Connectivity Ltd.
*	BCR	Bard (C.R.) Inc.		FLR	Fluor Corp.		MOS	The Mosaic Company	*	TGNA	Tegna
	BDX	Becton Dickinson	*	FLS	Flowserve Corporation		MPC	Marathon Petroleum		TGT	Target Corp.
	BEN	Franklin Resources	*	FMC	FMC Corporation		MRK	Merck & Co.	*	THC	Tenet Healthcare Corp.
*	BF-B	Brown-Forman Corporation	*	FOX	Twenty-First Century Fox Class B		MRO	Marathon Oil Corp.	*	TIF	Tiffany & Co.
	BHI	Baker Hughes Inc		FOXA	Twenty-First Century Fox Class A		MS	Morgan Stanley		TJX	TJX Companies Inc.
	BIIB	BIOGEN IDEC Inc.		FSLR	First Solar Inc		MSFT	Microsoft Corp.	*	TMK	Torchmark Corp.
	BK	The Bank of New York Mellon Corp.		FTI	FMC Technologies Inc.		MSI	Motorola Solutions Inc.		TMO	Thermo Fisher Scientific
	BLK	BlackRock		FTR	Frontier Communications	*	MTB	M&T Bank Corp.		TRIP	TripAdvisor
*	BLL	Ball Corp	*	GAS	AGL Resources Inc.		MU	Micron Technology		TROW	T. Rowe Price Group
	BMY	Bristol-Myers Squibb		GD	General Dynamics		MUR	Murphy Oil		TRV	The Travelers Companies Inc.
	BRCM	Broadcom Corporation		GE	General Electric		MYL	Mylan N.V.	*	TSCO	Tractor Supply Company
	BRK-B	Berkshire Hathaway		GGP	General Growth Properties Inc.	*	NAVI	Navient		TSN	Tyson Foods
	BSX	Boston Scientific		GILD	Gilead Sciences		NBL	Noble Energy Inc		TSO	Tesoro Petroleum Co.
	BWA	BorgWarner		GIS	General Mills	*	NDAQ	NASDAQ OMX Group	*	TSS	Total System Services
*	BXLT	Baxalta		GLW	Corning Inc.		NEE	NextEra Energy		TWC	Time Warner Cable Inc.
*	BXP	Boston Properties		GM	General Motors		NEM	Newmont Mining Corp. (Hldg. Co.)		TWX	Time Warner Inc.
	C	Citigroup Inc.		GMCR	Keurig Green Mountain		NFLX	Netflix Inc.		TXN	Texas Instruments
	CA	CA, Inc.		GME	GameStop Corp.		NFX	Newfield Exploration Co		TXT	Textron Inc.
	CAG	ConAgra Foods Inc.			Genworth Financial		NI	NiSource Inc.		TYC	Tyco International
	CAH	Cardinal Health Inc.	*	GOOG	Alphabet Inc Class C		NKE	Nike		UA	Under Armour
	CAM	Cameron International Corp.		GOOGL	Alphabet Inc Class A		NLSN	Nielsen Holdings	*	UHS	Universal Health Services, Inc.
	CAT	Caterpillar Inc.	*	GPC	Genuine Parts		NOC	Northrop Grumman Corp.		UNH	United Health Group Inc.
	CB	Chubb Limited		GPS	Gap (The)		NOV	National Oilwell Varco Inc.	*	UNM	Unum Group
*	CBG	CBRE Group	*	GRMN	Garmin Ltd.		NRG	NRG Energy		UNP	Union Pacific
	CBS	CBS Corp.		GS	Goldman Sachs Group		NSC	Norfolk Southern Corp.		UPS	United Parcel Service
	CCE	Coca-Cola Enterprises		GT	Goodyear Tire & Rubber		NTAP	NetApp	*	URBN	Urban Outfitters
	CCI	Crown Castle International Corp.	*	GWW	Grainger (W.W.) Inc.	*	NTRS	Northern Trust Corp.		URI	United Rentals, Inc.
	CCL	Carnival Corp.		HAL	Halliburton Co.		NUE	Nucor Corp.		USB	U.S. Bancorp
	CELG	Celgene Corp.	*	HAR	Harman Int'l Industries		NVDA	Nvidia Corporation		UTX	United Technologies
	CERN	Cerner	*	HAS	Hasbro Inc.	*	NWL	Newell Rubbermaid Co.		V	Visa Inc.
*	CF	CF Industries Holdings Inc		HBAN	Huntington Bancshares	*	NWSA	News Corp. Class A	*	VAR	Varian Medical Systems
	CHK	Chesapeake Energy	*	HBI	Hanesbrands Inc		O	Realty Income Corporation		VFC	V.F. Corp.
	CHRW	C. H. Robinson Worldwide		HCA	HCA Holdings	*	OI	Owens-Illinois Inc		VIAB	Viacom Inc.
	CI	CIGNA Corp.	*		Hudson City Bancorp		OKE	ONEOK		VLO	Valero Energy
*	CINF	Cincinnati Financial		HCN	Welltower Inc.		OMC	Omnicom Group	*	VMC	Vulcan Materials
	CL	Colgate-Palmolive		HCP	HCP Inc.		ORCL	Oracle Corp.	*	VNO	Vornado Realty Trust
	CLX	The Clorox Company		HD	Home Depot	*	ORLY	O'Reilly Automotive	*	VRSN	Verisign Inc.
*	CMA	Comerica Inc.		HES	Hess Corporation		OXY	Occidental Petroleum		VRTX	Vertex Pharmaceuticals Inc
	CMCSA	Comcast A Corp		HIG	Hartford Financial Svc.Gp.		PAYX	Paychex Inc.		VTR	Ventas Inc
	CME	CME Group Inc.		HOG	Harley-Davidson		PBCT	People's United Financial		VZ	Verizon Communications
*	CMG	Chipotle Mexican Grill		HON	Honeywell Int'l Inc.	*	PBI	Pitney-Bowes	*	WAT	Waters Corporation
	CMI	Cummins Inc.		HOT	Starwood Hotels & Resorts		PCAR	PACCAR Inc.		WBA	Walgreens Boots Alliance
	CMS	CMS Energy		HP	Helmerich & Payne		PCG	PG&E Corp.		WDC	Western Digital
	CNP	CenterPoint Energy		HPQ	HP Inc.	*	PCL	Plum Creek Timber Co.		WEC	Wisconsin Energy Corporation
	CNX	CONSOL Energy Inc.	*	HRB	Block H&R		PCLN	Priceline.com Inc		WFC	Wells Fargo
	COF	Capital One Financial	*	HRL	Hormel Foods Corp.		PCP	Precision Castparts		WFM	Whole Foods Market
	COG	Cabot Oil & Gas	*	HRS	Harris Corporation	*	PDCO	Patterson Companies	*	WHR	Whirlpool Corp.
	COH	Coach Inc.	*	HSIC	Henry Schein		PEG	Public Serv. Enterprise Inc.		WM	Waste Management Inc.
*	COL	Rockwell Collins		HST	Host Hotels & Resorts		PEP	PepsiCo Inc.		WMB	Williams Cos.
	COP	ConocoPhillips		HSY	The Hershey Company		PFE	Pfizer Inc.		WMT	Wal-Mart Stores
	COST	Costco Co.	*	HUM	Humana Inc.	*	PFG	Principal Financial Group		WU	Western Union Co
*	CPB	Campbell Soup		IBM	International Bus. Machines		PG	Procter & Gamble		WY	Weyerhaeuser Corp.
*	CPGX	Columbia Pipeline Group Inc	*	ICE	Intercontinental Exchange		PGR	Progressive Corp.	*	WYN	Wyndham Worldwide
	CRM	Salesforce.com	*	IFF	Intl Flavors & Fragrances		PH	Parker-Hannifin		WYNN	Wynn Resorts Ltd
*		Computer Sciences Corp.		INTC	Intel Corp.		PHM	Pulte Homes Inc.	*	XEC	Cimarex Energy
	CSCO	Cisco Systems		INTU	Intuit Inc.	*	PKI	PerkinElmer		XEL	Xcel Energy Inc
	CSX	CSX Corp.		IP	International Paper		PLD	Prologis		XL	XL Capital
*	CTAS	Cintas Corporation		IPG	Interpublic Group		PM	Philip Morris International		XLNX	Xilinx Inc
	CTL	CenturyLink Inc		IR	Ingersoll-Rand PLC		PNC	PNC Financial Services		XOM	Exxon Mobil Corp.
	CTSH	Cognizant Technology Solutions	*	IRM	Iron Mountain Incorporated		PNR	Pentair Ltd.	*	XRAY	Dentsply International
	CTXS	Citrix Systems	*	ISRG	Intuitive Surgical Inc.	*	PNW	Pinnacle West Capital		XRX	Xerox Corp.
	CVC	Cablevision Systems Corp.		ITW	Illinois Tool Works	*	POM	Pepco Holdings Inc.	*	XYL	Xylem Inc.
	CVS	CVS Caremark Corp.		IVZ	Invesco Ltd.	*	PPG	PPG Industries		YHOO	Yahoo Inc.
	CVX	Chevron Corp.	*	JBHT	J. B. Hunt Transport Services		PPL	PPL Corp.		YUM	Yum! Brands Inc
	D	Dominion Resources		JCI	Johnson Controls	*	PRGO	Perrigo	*	ZBH	Zimmer Biomet Holdings
	DAL	Delta Air Lines	*	JEC	Jacobs Engineering Group		PRU	Prudential Financial		ZION	Zions Bancorp
	DD	Du Pont (E.I.)		JNJ	Johnson & Johnson	*	PSA	Public Storage		ZTS	Zoetis