CENTRE FOR ECONOMETRIC ANALYSIS
CEA@Cass

http://www.cass.city.ac.uk/cea/index.html

Cass Business School
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Co-features in Finance: Co-arrivals and Co-jumps

Jan Novotný, and Giovanni Urga

CEA@Cass Working Paper Series

WP–CEA–09-2014

Co-features in Finance: Co-arrivals and Co-jumps

Jan Novotný[a,*], Giovanni Urga[b,1]

[a]*Cass Business School, City University London, UK and CERGE-EI, CZ*
[b]*Cass Business School, City University London, UK and Bergamo University, Italy*

Abstract

In this paper, we introduce the notions of co-jumps and co-arrivals within the co-features framework. First, we formulate the limiting properties of co-jumps and discuss their discrete sample properties. In the presence of idiosyncratic price jumps, we identify the notion of weak co-jumps. We then introduce the notion of co-arrivals where the feature is the price jump arrival irrespective of the magnitude. We illustrate the proposed framework using the constituents of the DJIA30 over 01/01/2010 to 30/06/2012 and discuss the role of the multiple testing bias.

Keywords: Co-features, portfolio diversification, price jumps, price jump arrivals.

J.E.L. Classification Number: G12, C12, C32

[*]Centre for Econometric Analysis, Faculty of Finance, Cass Business School, City University London, 106 Bunhill Row, London, EC1Y 8TZ, UK and CERGE-EI, CZ. Jan.Novotny.1@city.ac.uk, Tel: +44 (0)20 7040 8089, Fax: +44 (0)20 7040 8881.

[1]Centre for Econometric Analysis, Faculty of Finance, Cass Business School, City University London, 106 Bunhill Row, London, EC1Y 8TZ, UK and Bergamo University, Italy. G.Urga@city.ac.uk, Tel: +44 (0)20 7040 8698, Fax: +44 (0)20 7040 8881.

1. Introduction

This paper proposes a novel theoretical framework to assess common price jumps in a multivariate framework using the notion of co-features (i.e. the existence of a linear combination of time series where individual features are eliminated) as originally proposed by Engle and Kozicki (1993) and recently reconsidered in the special issue of Journal of Business and Economic Statistics (2007).

There is a huge body of literature on the identification of price jumps in the univariate context. Several procedures have been proposed to test for the presence of price jumps defined as discontinuity in the price jump process. See, for example, Aït-Sahalia and Jacod (2009), Aït-Sahalia and Jacod (2011), Aït-Sahalia and Jacod (2012a), Andersen et al. (2011, 2012), Barndorff-Nielsen and Shephard (2004b, 2006), Jiang and Oomen (2008), Lee and Mykland (2008), Lee and Hannig (2010), Huang and Tauchen (2005), and Mancini (2009). Further, the analysis of the mutual performance of different price jump indicators is addressed by Dumitru and Urga (2012).

The main aim of a multivariate framework is to identify common jumps between stochastic processes. Bollerslev et al. (2008) test for the presence of portfolio-wide systemic price jumps and focus in particular on systemic common jumps without counterparts on the individual time series level. This framework is extended by Liao and Anderson (2011), who use the range-based indicators proposed by Bannouh et al. (2009). Jacod and Todorov (2009) propose a statistic to test for the joint occurrence of price jump arrivals at a pair of time series. Lahaye et al. (2011) report an empirical study and estimate the joint probabilities of common price jump arrivals. The authors also suggest a joint statistic for the estimation of common price jumps. Barndorff-Nielsen and Shephard (2004a) provide a seminal framework for testing for price jumps in the multivariate framework. Aït-Sahalia et al. (2009) use common price jumps for assets in the same sector to evaluate the optimal portfolio in the presence of jumps; Bollerslev et al. (2013) analyze the tail index, where they explicitly model the common jump term for the entire portfolio aside of the idiosyncratic terms. Linked to this issue is the concept of co-arrivals which motivates the work of Lee (2012), who proposes a model to identify the predictors of arrival times in the univariate framework. The author's results suggest the existence of common predictors for Dow Jones Index constituents. The presence of common jumps is also highlighted in Lahaye et al. (2011) where the authors map common jumps in response to specific macro-news for a broader range of assets such as USD exchange rates, US Treasury bonds futures and US equity futures. Gilder et al. (2014) analyse the contemporaneous co-jumps of US equities and link them to Federal Fund Target Rate announcements.

This paper contributes to the current literature on common price jumps as follows: First, we propose a novel notion of co-jumps identified within the co-feature framework. We provide both the limiting theory and the discrete sample properties, suitable for empirical analyses. In particular, the notion of co-jumps is linked to the diversification of price jumps in the portfolio and the co-jump can be intuitively understood as a possibility to diversify the price jumps completely out of the portfolio.[2] We further extend the notion of co-jumps to cases where each asset has idiosyncratic price jumps, implying the absence of co-jumps. The analysis of Lahaye et al. (2011) and Bollerslev et al. (2008) suggests that in many empirically relevant cases, the idiosyncratic price jumps play an important role across a portoflio. In such a case, we define the notion of weak co-jumps as a linear combination of assets with minimum contribution of price jumps and link this notion to co-jumps.

Second, we extend co-jumps and introduce the notion of co-arrivals as a co-feature based on the arrival of price jumps irrespective of their magnitude. We provide a limiting theory and discrete sample properties. The co-arrivals are linked to the signal processing, where the feature under consideration is the signal of the price jump arrival. This notion is useful to evaluate the sensitivity of the portfolio to particular news and to assess the propensity for common price jumps and contagious behavior.

[2]A portfolio wide common jumps, which cannot be diversified out of the portfolio, was discussed in Bollerslev et al. (2008).

Third, we provide an empirical illustration of co-jumps and co-arrivals using the sample of the DJIA 30 index for the period of January 1, 2010 to June 30, 2012 sampled at a 5-minute frequency. Finally, we perform a sensitivity analysis to understand the role of the spurious jump detection. In particular, we reduce the multiple testing bias by implementing the procedure suggested by Bajgrowicz and Scaillet (2011).

The paper is organized as follows: In Section 2, we introduce the model for price dynamics and the concept of co-features. Section 3 provides the definition of co-jumps. In Section 4, we define co-arrivals as an extension of co-jumps. In Section 5, we report an empirical illustration of the co-jumps and co-arrivals using the constituents of the DJIA 30 index. In Section 6, we provide a robustness check with respect to the multiple testing bias. Section 7 concludes.

2. Preliminaries

The main aim of this paper is to introduce the notions of co-jumps and co-arrivals within the co-feature framework. To this purpose, in this section, we first introduce the underlying dynamic process for N-dimensional log-prices and provide the frequency specific factorization of common price jump arrivals and their co-feature characteristics.

2.1. The Model of the Price Dynamics

Let us consider an N-dimensional vector of log-prices $Y' = \left(Y^{(1)}, \ldots, Y^{(N)}\right)$, where the vector $Y = \{Y_t\}_{0 \leq t \leq 1}$ is defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}\right)$ over the time interval $[0, 1]$. The vector of log-prices is the semi-martingale $\mathcal{F}_t$-adapted and its continuous-time dynamics can be specified by the following stochastic differential equation

$$dY_t = \mu_t dt + \sigma_t dB_t + d\mathcal{J}_t \,, \tag{1}$$

where μ_t is $(N \times 1)$-vector of drift processes, σ_t is the $(N \times N)$-dimensional covariance matrix, dB_t is the $(N \times 1)$-dimensional vector of independent standard Brownian motions, and $d\mathcal{J}_t$ is the $(N \times 1)$-dimensional vector of pure jump Levy processes.

With respect to $d\mathcal{J}_t$, we assume that the price jump process is a linear combination of

m finite activity processes, where each arrival is driven by the doubly stochastic Poisson processes of finite activity and where its magnitude given by another stochastic process. Thus, the jump term $d\mathcal{J}_t$ can be written as $U_t dJ_t$, where U_t is the magnitude of price jumps and dJ_t is the arrival of price jumps: [3]

$$\underbrace{U_t dJ_t}_{(N\times 1)} \equiv \underbrace{U_t}_{(N\times m)} \underbrace{dJ_t}_{(m\times 1)}, \tag{2}$$

where U_t is the $(N \times m)$-dimensional matrix process is càdlàg and almost surely has non-zero and finite elements, and dJ_t is the $(m \times 1)$-dimensional vector, which drives the arrivals of the price jumps with mutually independent components set as $J_t^{(j)} = \int_0^t dJ_s^{(j)}$, a doubly stochastic Poisson process with an integrated stochastic intensity $\Lambda_\theta^{(j)}(t) = \int_0^t d\Lambda_\theta^{(j)}(s)\, ds$, with θ being a vector of parameters.

In this framework, we can explicitly distinguish between idiosyncratic and common price jumps. In particular, the matrix U_t can be factorized as

$$U_t = \left(\underbrace{U_t^I}_{(N\times N)}, \underbrace{U_t^C}_{(N\times(m-N))}\right), \tag{3}$$

where $U_t^I = \operatorname{diag}\left(U_t^{I(1)}, \dots, U_t^{I(N)}\right)$ and the i-th diagonal term $U_t^{I(i)}$ is almost surely non-zero and finite if there are idiosyncratic arrivals for i-th asset, and the matrix process U_t^C has in each column at least two almost surely non-zero elements corresponding to common arrivals among assets.

The vector of arrivals can be decomposed as

[3]Moreover, the literature accounts for the presence of both small and large price jumps (see for instance Bollerslev et al., 2008), while Lee and Hannig (2010) propose a procedure to identify small price jumps. Furthermore, Aït-Sahalia and Jacod (2011) link small price jumps to the infinite activity arrival processes. In this paper, we follow Barndorff-Nielsen and Shephard (2004b, 2006), Andersen et al. (2012), Lee and Mykland (2008) and consider the large price jumps with finite activity. In addition, we distinguish between the idiosyncratic and the systemic, or common, price jumps, as defined in Bollerslev et al. (2008) and Aït-Sahalia and Jacod (2012b). Note that we adopt the notion of common price jumps rather than that of systemic price jumps since we explicitly model the price jumps occurring at the same time in market segments rather than in the entire market as modeled by Bollerslev et al. (2008).

$$dJ_t = \left[\underbrace{\left(dX_t^{(1)}, \ldots, dX_t^{(N)} \right.}_{(1\times N)}, \underbrace{\left. dZ_t^{(1)}, \ldots, dZ_t^{(m-N)} \right)}_{(1\times(m-N))} \right]' , \tag{4}$$

where we assume $m > N$. $dX_t^{(1)}, \ldots, dX_t^{(N)}$ drive the idiosyncratic shocks, while the remaining $dZ_t^{(1)}, \ldots, dZ_t^{(m-N)}$ components correspond to the common shocks captured by a respective column of the matrix U_t^C.

The presence of price jumps in (1) implies that a $(N \times N)$-dimensional quadratic variation process Σ_t can be defined as

$$\Sigma_t = \Sigma_t^{(c)} + \Sigma_t^{(d)} , \tag{5}$$

where $\Sigma_t^{(c)}$ represents the continuous part of the semi-martingale process

$$\Sigma_t^{(c)} = \int_0^t \sigma_s \sigma_s' ds \, , \text{ with } \left\{ \Sigma_t^{(c)} \right\}_{i,j} < \infty \, , \, i,j = 1, \ldots, N \, , \tag{6}$$

and $\Sigma_t^{(d)}$ represents the discontinuous part of the semi-martingale process

$$\Sigma_t^{(d)} = \sum_{j=1}^{N_t} c_j c_j' \, , \text{ with } \left\{ \Sigma_t^{(d)} \right\}_{i,j} < \infty \, , \, i,j = 1, \ldots, N \, , < \infty \, , \tag{7}$$

where c_j is N-dimensional vector $c_j \equiv U_{t_j}$ for which there exists at least one $i = 1, \ldots, N$ such that $dY_{t_j-}^{(i)} > 0$, and N_t is number of such $t_j \leq t$. We assume that $\Sigma_1^{(c)}$ is positive definite so that we rule out perfectly correlated assets. This result allows us to map the presence of price jumps in terms of quadratic variation.

The model (1)-(2) encompasses a full set of multidimensional price jump specifications with finite activity. In addition, we assume a perfect synchronicity in the price jump arrivals, as in Aït-Sahalia et al. (2009) and Bollerslev et al. (2013), and the model is able to incorporate the high-frequency stylized facts as highlighted in the recent literature; see, e.g., Aït-Sahalia and Jacod (2012b), Bollerslev et al. (2008), Lahaye et al. (2011), Jiang et al. (2011) and Jiang et al. (2011). The continuous-time framework is observationally equivalent to the discrete counterpart used in most empirical applications.

In the next section, we characterize the factorization of the log-price process Y_t within the co-feature framework.

2.2. Linear Transformation and Co-features

Consider the integrated counterpart of the process described in (1)

$$Y_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s + \sum_{j=1}^{N_t} c_j \, , \tag{8}$$

where each of the components has discontinuities in the interval $[0, t]$. The N-dimensional Brownian semi-martingale process with finite-activity jumps is closed with respect to the stochastic integration under a linear transformation given by a $(p \times N)$-dimensional matrix $\Omega^{(p\times N)}$. In general these are time-dependent; see Jacod and Shiryaev (1987). The p-dimensional process given as a linear transformation of Y_t can be written as

$$\begin{aligned} \Omega^{(p\times N)} Y_t &= \int_0^t \Omega^{(p\times N)} \mu_s ds + \int_0^t \Omega^{(p\times N)} \sigma_s dB_s + \sum_{j=1}^{N_t} \Omega^{(p\times N)} c_j \\ &= \int_0^t \mu_s^{(p)} ds + \int_0^t \sigma_s^{(p)} dB_s + \sum_{j=1}^{N_t} c_j^{(p)} \, , \end{aligned} \tag{9}$$

which is a p-dimensional Brownian semi-martingale with finite activity jumps. In general, we may find a matrix $\Omega^{(p\times N)}$ such that the jump term in one of the components $\Omega^{(p\times N)} Y_t$ disappears. This characteristic is in fact the notion of co-features, as introduced by Engle and Granger (1987) and Engle and Kozicki (1993), recently considered in a special issue of the Journal of Business and Economic Statistics (2007).

For convenience, let us summarize the concept of co-features as introduced in the econometric literature. Following Urga (2007), let us assume that the N-dimensional time series process $\mathbf{y}_t' = \left(y_t^{(1)}, \ldots, y_t^{(N)}\right)$, with $t = 1, \ldots, T$, satisfies the following set of axioms:

Axiom 1a. If each of the time series $y_t^{(j)}$ does (not) have the feature, then $\lambda y_t^{(j)}$ will (not) have the feature for any $\lambda \neq 0$ and $j = 1, \ldots, N$.

Axiom 1b. If $y_t^{(i)}$ does not have the feature and $y_t^{(j)}$ does not have the feature, then $\bar{y}_t = y_t^{(i)} + y_t^{(j)}$ will not have the feature for $\forall i \neq j,\ i, j = 1, \ldots, N$.

Axiom 1c. If $y_t^{(i)}$ does not have the feature and $y_t^{(j)}$ does have the feature, then $\tilde{y}_t = y_t^{(i)} + y_t^{(j)}$ will have the feature.

Then:

Definition 1. *A feature present in each of the time series is said to be common to the time series if there exists a non-zero linear combination of the series that does not have the feature. Such a linear combination is called the* co-feature *combination and the corresponding vector the* co-feature vector*. The set of all* co-feature vectors *spans the s-dimensional space with $s < N$.*

This definition directly implies the following decomposition of y_t:

$$\underbrace{y_t}_{(N\times 1)} = \underbrace{\mathbf{\Omega}}_{N\times(N-s)} \underbrace{F_t}_{(N-s)\times 1} + \underbrace{u_t}_{(N\times 1)}, \tag{10}$$

where vector F_t specifies the feature, $\mathbf{\Omega}$ is the loading matrix, and u_t does not have the feature.

We exploit the co-feature framework to introduce the notions of co-jumps in Section 3 and co-arrivals in Section 4.

3. Co-jumps

The notion of co-jumps is based on the disappearance of the jump term from (9) for a given time interval. We define the notion of co-jump as follows:

Definition 2. ***(Co-jumps)*** *For the N-dimensional process Y_t given by (8) with each of the components having a discontinuity in the interval $[0, t]$,* co-jumps *are defined as the existence of the N-dimensional constant vector Ω different from zero vector such that for the process $\Omega' Y_t$*

$$\Sigma_t^{(d)} = \sum_{j=1}^{N_t} \Omega' c_j c_j' \Omega = \sum_{j=1}^{N_t} \Omega' C_j \Omega = 0\,,$$

with $C_j = c_j c_j'$ being an $(N \times N)$-dimensional matrix describing the configuration and magnitudes of price jumps at a particular time t_j for which at least one component is $dY_{t_j-} > 0$. The vector Ω is called the co-jump vector *and the space of all co-jump vectors spans the* co-jump space.

The properties of the co-jump vectors are summarized in the following proposition:

Proposition 1. *(Co-jump vectors) For a given vector of price jumps c_j defined in (7), the matrix of all linear combinations, which causes the discontinuous part of the quadratic variance to disappear, is given by*

$$\Omega_{N\times N} = 1 - c_j \left(c_j' c_j\right)^{-1} c_j' \,. \tag{11}$$

If there is more than one price jump, or $N_t > 1$, the linear combinations are obtained by

$$\Omega_{N\times N} = 1 - \tilde{C} \left(\tilde{C}' \tilde{C}\right)^{-1} \tilde{C}' \,, \tag{12}$$

where $(N \times N_t)$-dimensional matrix $\tilde{C}$ is given as $\left\{\tilde{C}\right\}_{\bullet,j} = c_j$. The matrix $\Omega_{N\times N}$ is singular.

The following corollary identifies the dimension of the co-jumps space:

Corollary 1. *The dimension of the co-jumps space, s, is $s < N$, with $s \geq 1$ if and only if $rank\left(\tilde{C}\right) < N$.*

To identify co-jumps, we consider an N-dimensional log-price process Y_t defined in (8) in the interval $[0, 1]$ and define the corresponding matrix of returns as

$$r_t^{(n)} = Y_t^{(n)} - Y_{t-\Delta t}^{(n)} \,, \text{ with } t = j \cdot \Delta t \,,\ j = 1, 2, \ldots, \left\lfloor \frac{1}{\Delta t} \right\rfloor \,,$$

where the equidistant time grid is the same for all assets. For the sake of simplicity, let us denote $M \equiv \left\lfloor \frac{1}{\Delta t} \right\rfloor$ and assume that $M\Delta t = 1$.

For a univariate price process Y_t generated by (1), under the null hypothesis of no price jumps the test statistic, $\hat{G}$, has the following property

$$\hat{G} \xrightarrow{\mathcal{D}} N\left(0, \vartheta\right) \,,$$

where $\xrightarrow{\mathcal{D}}$ denotes a stable convergence in law and ϑ is some known constant depending on the estimators used. Under the alternative, we assume a one sided divergence with $\hat{G} \xrightarrow{p} -\infty$.

There is a co-jump for the N-dimensional process Y_t in the interval $[0, t]$ if a vector Ω exists such that the $\hat{G}$-statistic for the univariate process $\Omega' Y_t$ does not allow us to reject the null hypothesis. The asymptotic properties of the $\hat{G}$-statistic under the null hypothesis hold when there is no discontinuous part of the price process $\Omega' Y_t$.

The $\hat{G}$-statistic is given as

$$\hat{G} = M^{1/2} \frac{\widehat{IV}_M - \widehat{QV}_M}{\widehat{IQ}_M}\,, \tag{13}$$

where $\widehat{IV}_M$ is the estimator of the Integrated Variance, $\widehat{QV}_M$ of the Quadratic Variance, $\widehat{IQ}_M$ of the Integrated Quarticity with standard asymptotic properties:

$$\widehat{IV}_M \xrightarrow{p} \int_0^t \sigma_s^2 ds\,, \tag{14}$$

$$\widehat{QV}_M \xrightarrow{p} \int_0^t \sigma_s^2 ds + \sum_{j=1}^{N_t} c_j^2\,, \tag{15}$$

$$\widehat{IQ}_M \xrightarrow{p} \int_0^t \sigma_s^4 ds\,. \tag{16}$$

Thus:

Proposition 2. *For the univariate price process Y_t in (1) on time interval $[0,1]$ sampled into M equidistant time stamps, the $\hat{G}$-statistic defined in (13) has the following property as $M \to \infty$:*

$$\hat{G} \xrightarrow{\mathcal{D}} N\left(0, \vartheta\right)$$

if and only if

$$\Sigma_t^{(d)} = 0\,.$$

In this paper, we consider a sparse sampling approach to deal with market micro-structure noise since it provides a reasonable trade-off between accuracy and numerical feasibility at chosen sampling frequency. However, our framework can be extended to employ alternative techniques such as the pre-averaging method by Podolskij and Vetter (2009), employed by Aït-Sahalia and Jacod (2009) and Aït-Sahalia et al. (2012), or the combination of different time scales by Zhang et al. (2005), and Zhang (2011).

In the next section, we discuss the discrete sample counterpart of the asymptotic framework of co-jumps.

3.1. Discrete Sample Properties of Co-jumps

We define local co-jumps:

Definition 3. ***(Local co-jumps)*** *For the N-dimensional process Y_t given by (8) with each of the components having one price jump in the interval $[0,1]$, sampled into M equidistant time steps, a* local co-jump *is defined as the existence of the N-dimensional constant vector Ω different from zero vector such that for the process $\Omega' Y_t$*

$$\hat{G} \xrightarrow{\mathcal{D}} N(0, \vartheta) \; , \tag{17}$$

as $M \to \infty$. The vector Ω is called the local co-jump vector *and the space of all co-jump vectors spans the* local co-jump space.

The definition of local co-jumps, where each component of Y_t has exactly one price jump, allows us to link the properties of the co-jumps to the factorization of the price jump term (2).

Proposition 3. ***(Properties of local co-jumps)*** *For the N-dimensional process Y_t given by (8) with each of the components having one price jump in the interval $[0,1]$, it holds:*

1. rank $\left(\tilde{C}\right) = N_t \leq N$ with equality holding for the purely idiosyncratic price jumps.

2. The presence of almost surely non-trivial term U^C in the factorization (2) implies the existence of non-zero Ω.

3. The i-th column $\left\{\tilde{C}\right\}_{\bullet,i}$ of $\tilde{C}$ specifies the $(N-1)$-dimensional subspace $\mathcal{L}^{(i)}$. For a pair of sub-spaces $\mathcal{L}^{(i)}$ and $\mathcal{L}^{(j)}$, belonging to different columns of $\tilde{C}$, the intersection is $(N-2)$-dimensional space.

4. The local co-jump space, $\mathcal{L}$, is an $(N - N_t)$-dimensional space defined as intersection of the local co-jump spaces $\mathcal{L} = \bigcup_{i=1}^{N_t} \mathcal{L}^{(i)}$.

Let us consider the price process Y_t in the interval $[0,1]$ sampled into M equidistant steps with $\Delta t = 1/M$. By Definition 3, the local co-jump vectors Ω are all such N-dimensional vectors, such that $\hat{G}^{(\Omega)} \xrightarrow{\mathcal{D}} N(0, \vartheta)$ for $\Omega' Y_t$, as $M \to \infty$. In the discrete sampling, we say that the N-dimensional vector Ω is the local co-jump vector if

$$\hat{G} > -\frac{1}{\sqrt{\vartheta}} q_{1-\alpha} \, , \tag{18}$$

where $q_{1-\alpha}$ is the $(1-\alpha)$-th percentile of the standard normal distribution. Intuitively, the null is rejected more frequently for smaller M as the IV, QV and IQ are estimated with higher sampling error. Therefore, the local co-jump space at discrete sampling, denoted as $\mathcal{L}_M$ with M being the sampling, contains both the true local co-jump vectors obtained at $M \to \infty$, as well as all the false linear combinations due to the false rejection error caused by the discrete sampling.

The following proposition formalizes the properties of the *local co-jump space* for discrete sampling and its convergence towards the asymptotic local co-jump space, $\mathcal{L}$:

Proposition 4. ***(Local co-jump space at discrete sampling)*** *The local co-jump space at discrete sampling M, $\mathcal{L}_M$, i.e., all the N-dimensional vectors Ω solutions to (18), completed by origin, $(\{0\} \bigcup \mathcal{L}_M)$ is a closed and non-compact manifold. For any given fixed confidence level, the* $\lim_{M\to\infty} \mathcal{L}_M = \mathcal{L}$*, with $\mathcal{L}$ being the local co-jumps space, and the limit is understood as*

$$\bigcap_{M'\to\infty} \mathcal{L}_{M'} = \mathcal{L}\,, \tag{19}$$

i.e., the intersection of infinite sequence of discrete sample local co-jump spaces.

Note that Proposition 4 is based on the convergence of the confidence intervals, over which we state the hypothesis of the statistic to reject the null of no price jumps. Therefore, Proposition 4 deals with the point-wise convergence and it is not meant to substitute the proper asymptotic treatment of the test statistic. In addition, it also explicitly shows that the asymptotic solution is in the admissible region for any M. In yhe Appendix, we illustrate the discrete sample properties of local co-jumps via a Monte Carlo simulation exercise.

In the next section, we fully exploit the frequency specific factorization and discuss the case in which each component has idiosyncratic price jumps, together with any combination of common jumps.

3.2. The Full-rank Idiosyncratic Price Jumps

The notion of co-jumps, as introduced above, aims to find a linear combination which eliminates the jumps. When idiosyncratic jumps are present for every component,[4] co-jumps do not exist. To this purpose, we modify the notion of co-jumps such that we weaken the requirement for the elimination of the jump term in (9).

We define the notion of *weak co-jumps* as a linear combination which minimizes the role of jumps in the price process through the $\hat{G}$-statistic:

Definition 4. ***(Weak co-jumps)*** *For the N-dimensional process Y_t given by (8) defined in the interval $[0,1]$ such that each component $Y_t^{(i)}$, with $i = 1,\ldots,N$, has at least one idiosyncratic jump—rank$\left(U^I\right) = N$ and for each $i \in \{1,\ldots,N\}$, $dX_t^{(i)}$ has at least one*

[4]See, for instance Jiang et al. (2011), Lahaye et al. (2011), Lee (2012).

jump—the weak co-jumps are defined as the existence of the N-dimensional constant vector Ω different from zero vector such that for the process $\Omega' Y_t$, the $\hat{G}$-statistic is maximized. The vector Ω is called the weak co-jump vector.

The identification of weak co-jumps is performed in terms of pair-wise comparison of two Ω's with respect to the $\hat{G}$-statistic and it is formalized in the following proposition:

Proposition 5. ***($\hat{G}$-statistic with jumps)*** *For a pair of 1-dimensional price processes Y_t specified in (8) with $\Sigma_1^{(c)} > 0$ in the interval $[0, 1]$, such that the processes are identical, both with $N_1 = 1$, except the magnitudes of the respective price jumps $c_1^{(1)}$ and $c_1^{(2)}$ with $\left|c_1^{(1)}\right| > \left|c_1^{(2)}\right| > 0$, it holds:*

1. *$\hat{G} \xrightarrow{p} \infty$ for both processes,*
2. *$\hat{G}^{(1)} \overset{p}{>} \hat{G}^{(2)}$, with a meaning of $\frac{\hat{G}^{(1)}}{\hat{G}^{(2)}} \overset{p}{=} \left(\frac{c_1^{(1)}}{c_1^{(2)}}\right)^2 \overset{p}{>} 1$.*

Proposition 5 is defined for a price process with one price jump. However, a generalization valid for any number of price jumps, given the continuous part of the price process is the same, can be defined as follows

$$\frac{\hat{G}^{(1)}}{\hat{G}^{(2)}} \overset{p}{=} \left(\frac{\sum_{j=1}^{N_t^{(1)}} c_j^{(1)}}{\sum_{j=1}^{N_t^{(2)}} c_j^{(2)}}\right)^2 .$$

Consequently, Proposition 5 states that for an N-dimensional process Y_t, we may compare any two vectors $\Omega^{(1)}$ and $\Omega^{(2)}$ via the $\hat{G}$-statistic. As we assume that $\Sigma_1^{(c)}$ is positive definite, the $\hat{G}$-statistic is properly defined in the limit for any Ω. Therefore, we do not assume the existence of a unique minimum. Further, under the alternative, the convergence of the $\hat{G}$-statistic is unspecified and thus we search for weak co-jumps using the $\hat{G}$-statistic as an indicator function. Thus, for any Y_t considered in this paper, weak co-jumps exist.

Second, as for any $\lambda > 0$, $\hat{G}^{(\Omega)} = \hat{G}^{(\lambda\Omega)}$, we may consider only those Ω's on the hypersphere

$$\sum_{i=1}^{N} \Omega^{(i)2} = 1 . \tag{20}$$

Further, we may impose an additional constraint on the sign of $\Omega^{(i)}$ to identify $\Omega' Y_t$ to a portfolio without short selling. Third, the $\hat{G}$-statistic is diverging at rate $M^{-1/2}$ with M being the sampling frequency and we may alternatively proceed with evaluation of the weak co-jumps in terms of the re-scaled variable $\hat{D} \equiv M^{1/2}\hat{G}$.

The following corollary shows the equivalence between co-jumps and weak co-jumps in the case of rank $\left(\tilde{C}\right) < N$, with $\tilde{C}$ defined in Proposition 1:

Corollary 2. *For an N-dimensional price process Y_t defined in (8) in the interval $[0,1]$ such that the price jumps satisfy rank $\left(\tilde{C}\right) < N$, then the co-jump space and the weak co-jump space are equivalent.*

So far we have provided a framework where we define the notion of common jumps in terms of co-features implying the existence of a linear combination such that the price jumps disappear. We have also introduced a testing framework, the $\hat{G}$-statistic, to test for the presence of co-jumps. Co-jumps are thus characterized by both size and intensity of price jumps.

However, there are situations where the dominant feature is the intensity (arrivals) of price jumps, such as, for instance, in the case of evaluating the impact of macro news on asset classes or when the size is idiosyncratic and is affected by market micro-structure noise. Thus, in the next section, we introduce the notion of co-arrivals implying the existence of a linear combination such that the information about price jump arrivals disappears.

4. Co-arrivals

Definition 5. ***(Co-arrivals)*** *For an N-dimensional indicator process j_t defined as*

$$Y_t^{(i)} \to j_t^{(i)} = \begin{cases} \delta\left(t\right) & \text{if } dY_{t-}^{(i)} > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{21}$$

with $\delta\left(t\right)$ being a delta function, in the interval $[0,1]$ such that $N_1^{(i)} > 1$ for every i, co-arrival *exists if there is a non-zero linear combination Ω such that $\Omega' j_t \equiv 0$. The vector Ω is the* co-arrival vector *and the space of all co-arrival vectors forms the* co-arrival space*.*

The N-dimensional process j_t can be written as

$$j_t = \sum_{j=1}^{N_t} \overline{c}_j \delta\left(t - t_j\right) \,, \tag{22}$$

where N-dimensional vector $\overline{c}_j$ is a function of c_j such that $\overline{c}_j^{(i)} = 1_{\left(c_j^{(i)} \neq 0\right)}$. The indicator process j_t is almost always zero, being non-zero only in the presence of price jumps at a particular time t. The j_t can almost surely be identified with the price arrival process $U_t dJ_t$, with the factorization (4) holding,[5] and it is proportional to $j_t^{(i)} \propto \frac{d(U_t J_t)^{(i)}}{dt}$, where the coefficient of proportionality is inverse of the magnitude of the price jump at time t.

Such a definition allows us to construct a counting operator of price jumps in the interval $[0, t]$ for the i-th component $Y_t^{(i)}$ as

$$N_t^{(i)} = \int_0^t \sum_{j=1}^{N_t} \overline{c}_j^{(i)} \delta\left(s - t_j\right) ds\,, \tag{23}$$

as well as the counting operator for the linear transformation Ω:

$$N_t^{(\Omega)} \equiv \int_0^t \Omega' \overline{c}_j \delta\left(t - t_j\right) dt = 0\,. \tag{24}$$

This formulation allows us to apply Proposition 1 and to obtain the full set of linearly independent co-arrival vectors as

$$\Omega_{N\times N} = 1_{N\times N} - \tilde{\overline{C}} \left(\tilde{\overline{C}}' \tilde{\overline{C}}\right)^{-1} \tilde{\overline{C}}'\,, \tag{25}$$

where $(N \times N_{t=1})$-dimensional matrix $\tilde{\overline{C}}$ is given as $\left\{\tilde{\overline{C}}\right\}_{\bullet,j} = \overline{c}_j$ and the matrix $\Omega_{N\times N}$ is singular.

Then, the test for the co-arrivals, or $\Omega' j_t \equiv 0$, can be expressed as a combination of univariate tests of the form

$$\mathbf{rank}\left(\tilde{\overline{C}}\right) = N\,, \tag{26}$$

with an indicator matrix $\tilde{\overline{C}}$ obtained by application of univariate tests.

We now focus on the case in which the idiosyncratic arrivals are present for each component $j_t^{(i)}$, and therefore the rank of the matrix $\tilde{\overline{C}}$ is full. In such a case, the non-zero co-arrival vector does not exist and we proceed with weak co-arrivals defined as follows:

[5] The link between the j_t and the arrival process J_t is in probability as the magnitude is assumed to be almost surely non-zero.

Definition 6. ***(Weak co-arrivals)*** *For an N-dimensional process j_t defined in (21) in the interval* $[0,1]$ *such that $N_1^{(i)} > 1$ for every i and with at least one idiosyncratic price jump arrival for every $j_t^{(i)}$,* weak co-arrival *is defined as the linear combination Ω such that $N_1^{(\Omega)}$ is minimized. The vector Ω is the* weak co-arrival vector*.*

The notion of weak co-arrival, as in Definition 6, is equivalent to the notion of co-arrival if there exists at least one non-zero co-arrival vector with $N_1 = 0$. This property is summarized in the following corollary:

Corollary 3. ***(Weak co-arrivals vs co-arrivals)*** *For an N-dimensional process j_t defined in (21) in the interval* $[0,1]$ *such that $N_t^{(i)} > 1$ for every i, it holds that there exists a non-zero Ω such that $N_1^{(\Omega)} = 0$ iff the* weak co-arrival space *equals to the* co-arrival space *and is non empty.*

The weak co-arrival vector Ω can be identified using basic algebra tools as follows:

Corollary 4. ***(Weak co-arrivals property)*** *The* weak co-arrival vector Ω *can be obtained as a solution to the $\Omega' \tilde{\overline{C}}_{sub} = 0$, where $\tilde{\overline{C}}_{sub}$ is the sub-matrix of $\tilde{\overline{C}}$ such that:*

1. *rank* $\left(\tilde{\overline{C}}_{sub}\right) = N - 1$.
2. *Any other $\tilde{\overline{C}}'_{sub}$ which has more columns than $\tilde{\overline{C}}_{sub}$ has full rank.*

In the next section, we discuss the definition of co-arrivals in the discrete sampling case.

4.1. Discrete Sample Properties of Co-arrivals

Let us consider the discrete sampling of the interval $[0,1]$ such that it is divided into M equidistant steps $\Delta t = 1/M$. In the finite sampling, we are able to distinguish price jump arrivals belonging to different Δt_i, while price jump arrivals belonging to the same Δt_i are considered as one indistinguishable price jump arrival.

Let us define the arrival process j_t in the finite sampling as

$$j_{t_j}^{(i)} = \begin{cases} M & \text{if } \exists \tau \in \Delta t_j \equiv (t_{j-1}, t_j] = \frac{1}{M};\, dY_{\tau-}^{(i)} > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{27}$$

First, the finite sample version converges in probability to the continuous time version as $M \to \infty$, or $j_{t_j} \overset{a.s.}{\to} j_t$. Second, we assume an equidistantly sampled grid; however, j_{t_j} can be defined on any grid as long as $\max \Delta t_i \to 0$ as we increase the sampling frequency. Further, the grid has to be the same for all components of the N-dimensional price process. Third, j_{t_j} does not distinguish price jumps which arrive during the same Δt_j, the feature we have assumed in the motivation. As such, the j_{t_j} underestimates the price jump arrivals and creates a false notion common arrival when Δt_j is too coarse. Both effects, the underestimation of the number of price jump arrivals and the false common arrivals vanish as $M \to 0$.

The definition of $j_{t_j}^{(i)}$ can be used to calculate the number of price jump arrivals on the time interval $[0, t]$ analogously to (23) as

$$N_{t,M}^{(i)} = \int_0^t j_{t_j}^{(i)} ds = \sum_{j=1}^{M} 1_{\left(\exists \tau \in \Delta t_j ; dY_{\tau-}^{(i)} > 0\right)} ,$$

where the counting operator $N_{t,M}^{(i)}$ counts the number of time steps with at least one price jump arrival. Since $j_{t_j}^{(i)}$ converges $j_t^{(i)}$, the following property holds $N_{t,M}^{(i)} \to N_t^{(i)}$ as $M \to \infty$.

Further, analogously to (22), we can write an N-dimensional process j_{t_j} as

$$j_{t_j} = \sum_{j=1}^{N_t} \overline{c}_{j,M} \int_{t_{j-1}}^{t_j} M \cdot 1_{\left(\exists i, \exists \tau \in \Delta t_j ; dY_{\tau-}^{(i)} > 0\right)} dt ,$$

where the integral represents an approximation to delta function and $\overline{c}_{j,M}$ is $\overline{c}_{j,M}^{(i)} \equiv 1_{\left(\exists \tau \in \Delta t_j ; dY_{\tau-}^{(i)} > 0\right)}$. Vector $\overline{c}_{j,M}$ plays the same role as $\overline{c}_j$ and the set of all non-zero vector $\overline{c}_{j,M}$ is equivalent to $\overline{c}_j$ as $M \to \infty$. Consequently, we may proceed at finite M in the same way as for the continuous time version of co-arrivals. The continuous time co-arrival properties are recovered by the finite sampling co-arrivals as $M \to \infty$.

4.2. Co-arrivals vs Co-jumps

So far, we have introduced the notions of co-jumps and co-arrivals separately. The following proposition shows the link between the two notions.

Proposition 6. ***(Co-arrivals vs co-jumps)*** *For an N-dimensional process Y_t specified in (8) on a time interval $[0, 1]$, then it holds*

$$\exists \textit{co-jumps} \Rightarrow \exists \textit{co-arrivals} , \tag{28}$$

as $M \to \infty$.

Proposition 6 implies that in the absence of co-arrivals, co-jumps do not exist either. This Proposition has direct consequences for the testing strategy as illustrated in the next section.

5. Empirical Illustration

In this section, we illustrate the theoretical framework proposed in this paper by evaluating the presence of co-jumps and co-arrivals using the high-frequency data.

5.1. Data and Portfolio Selection

We use the individual assets of the Dow Jones Industrial 30 (DJI30) index running from 1 January 2010 to 30 June 2012 as provided by the NYSE TAQ database. We use data on trades only and utilize the appropriate cleaning mechanism by Barndorff-Nielsen et al. (2009). As a result, the data are sampled into 5-minute time steps. The trading day starts at 9:30:00 and ends at 16:00:00, which gives 79 log-prices per day. Our sample contains 621 trading days in total.

We split the DJI30 index into six portfolios, each with five equities, based on the capitalization at the beginning of the sample. Table 1 presents the composition of each of the portfolios as well as the market capitalization at the beginning of the sample. We illustrate the notion of co-jumps and co-arrivals using a representative set of six portfolios. The descriptive statistics summarized in Table 1 reveal the large kurtosis for each asset and support the deviation from normality at 5-minute frequency.

5.2. Co-jumps

We employ the notion of co-jumps as developed in Section 3, and the $\hat{G}$-statistic is calculated for each trading day using $M = 78$ log-returns. We use $\alpha = 0.05$ to test the null hypothesis that there is no price jump(s) during the given trading day. To estimate the Integrated Variance, $\widehat{QV}_M$, the Integrated Variance, $\widehat{IV}_M$, and the Integrated Quarticity, $\widehat{IQ}_M$, we follow Barndorff-Nielsen and Shephard (2006) and define the estimators as follows

Table 1: Market capitalization and descriptive statistics for DJI30.

ID	Portfolio selection		Descriptive statistics of log-returns [%]				
	Market Cap ($bn)	No.	σ	S	K	Min	Max
XOM	360,98	1	0.129	-0.089	10.970	-1.627	1.586
MSFT	266,46	1	0.147	-0.069	12.601	-2.177	2.190
WMT	211,16	1	0.102	0.101	12.735	-1.518	1.222
PG	183,81	1	0.096	-0.038	15.844	-1.529	1.630
JNJ	175,23	1	0.096	-0.208	14.461	-1.368	1.238
GE	172,17	2	0.174	0.058	11.812	-2.339	3.283
IBM	168,06	2	0.116	0.105	11.325	-1.537	1.639
JPM	166,29	2	0.199	0.174	12.746	-2.158	3.728
T	160,39	2	0.115	-0.364	14.039	-1.957	1.542
CVX	158,62	2	0.140	-0.243	10.298	-1.764	1.342
PFE	152,11	3	0146	0.263	18.125	-2.920	3.118
CSCO	137,60	3	0.159	0.097	9.474	-1.603	1.862
BAC	137,55	3	0.254	-0.436	24.734	-5.052	3.899
KO	134,58	3	0.101	-0.039	13.496	-1.405	1.368
HPQ	117,58	3	0.170	1.062	51.069	-2.624	6.150
MRK	112,65	4	0.131	0.160	13.308	-1.318	2.342
INTC	108,56	4	0.159	-0.048	8.421	-1.623	1.394
VZ	91,87	4	0.117	-0.300	21.187	-2.624	1.792
MCD	68,57	4	0.102	-0.198	14.716	-2.157	1.423
UTX	63,75	4	0.138	-0.104	17.895	-3.087	1.736
DIS	57,11	5	0.151	-0.956	38.083	-4.537	1.993
MMM	55,67	5	0.132	-0.112	9.750	-1.525	1.291
AXP	49,11	5	0.174	-0.154	11.771	-3.160	1.909
HD	47,73	5	0.149	0.080	13.045	-2.669	2.376
KFT	39,13	5	0.109	-0.126	14.168	-2.134	1.223
BA	39,03	6	0.160	-0.060	9.647	-1.737	1.658
CAT	37,16	6	0.191	-0.148	9.925	-2.282	1.949
DD	31,72	6	0.162	0.011	9.972	-2.035	1.924
TRV	28,74	6	0.129	0.077	12.942	-1.541	1.496
AA	12,47	6	0.223	-0.121	9.384	-2.722	2.178

Note: The table contains market capitalization in $bn as the markets closed on 31st December 2009, data taken from Bloomberg. The members of the DJIA are sorted in two 6 portfolios by 5 equities according to the market capitalization.

$$
\begin{aligned}
\widehat{QV}_{M,D} &= \sum_{i=1}^{M_D} r_{i,D}^2 \,, &(29)\\
\widehat{IV}_{M,D} &= \frac{M_D}{M_D-1}\mu_1^{-2}\sum_{i=2}^{M_D} |r_{i-1,D}|\,|r_{i,D}| \,, &(30)\\
\widehat{IQ}_{M,D} &= \frac{M_D}{M_D-3}\frac{1}{M_D}\mu_1^{-4}\sum_{i=4}^{M_D} |r_{i-3,D}|\,|r_{i-2,D}|\,|r_{i-1,D}|\,|r_{i,D}| \,. &(31)
\end{aligned}
$$

where $r_{i,D}$ is the i-th log-return on day D with $M_D = 78$, $\mu_1 = E[|z|] = \sqrt{2/\pi}$ with $z \sim N(0,1)$, and for which (14)-(16) hold. The $\hat{G}$-statistic for the trading day D reads

$$
\hat{G}_D = M_D^{1/2}\frac{\widehat{IV}_{M,D} - \widehat{QV}_{M,D}}{\widehat{IQ}_{M,D}} \,, \tag{32}
$$

and has the asymptotic property as

$$
\hat{G}_D \xrightarrow{\mathcal{D}} N(0,\vartheta) \,,
$$

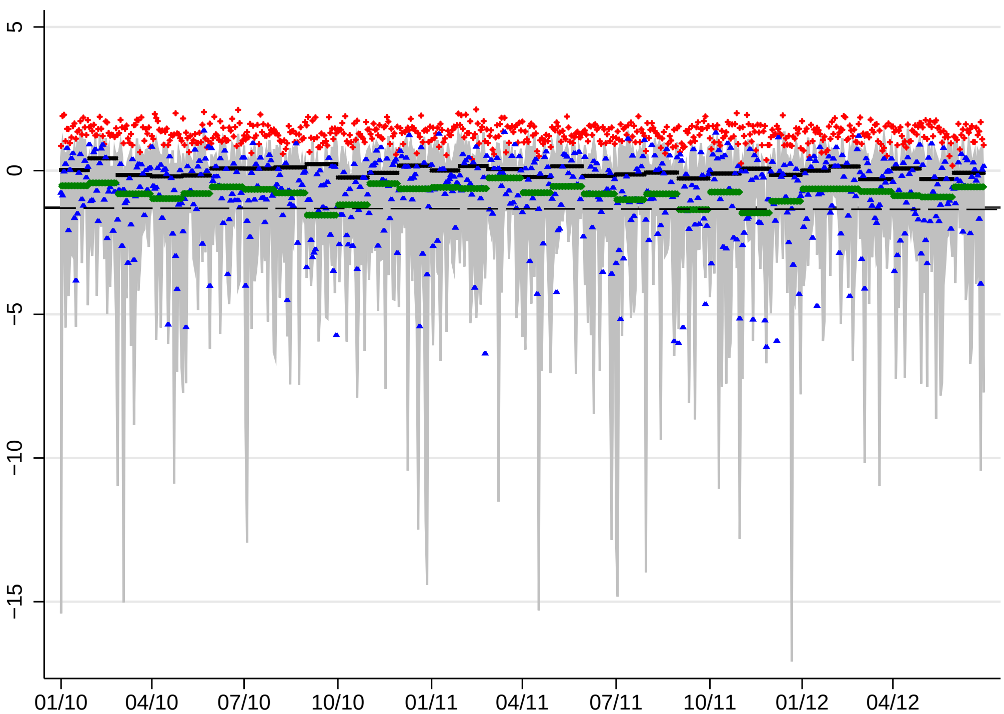
with $\vartheta = (\pi^2/4) + \pi - 5 \cong 0.609$. The test for the presence of price jumps during the trading day D at $\alpha = 0.05$ has the form

$$
\begin{array}{lll}
H_0: & \hat{G}_{M,D} \geq \sqrt{\vartheta}\Phi^{-1}(\alpha) & \text{no jump} \\
H_A: & \hat{G}_{M,D} < \sqrt{\vartheta}\Phi^{-1}(\alpha) & \text{jump(s)}
\end{array} \,,
$$

where Φ^{-1} is the inverse cumulative distribution function of the standard normal distribution giving $\sqrt{\vartheta}\Phi^{-1}(\alpha) \cong -1.284$.

Figure 1 depicts the results of the co-jumps exercise for Portfolio 1 consisting of the most capitalized set of assets in the DJIA30. In the Appendix, we report the results for the full set of portfolios. For every trading day, we find the co-jump vector Ω such that it maximizes the $\hat{G}$-statistic (red dots). For every trading day and each portfolio, we test for the presence of co-jumps and confirm the presence of co-jumps as $\hat{G}_{M,D}^{(\Omega)} \geq -1.284$, which is captured by the black long-dash line. This means that at the given sampling frequency, a linear combination of assets exists in the portfolio such that the price jumps diversifies out. Note that the

Figure 1: Co-jumps $\hat{G}$-statistic: Portfolio 1.



Note: The figure depicts the $\hat{G}$-statistic for the co-jump vector (red dots), for the equally weighted portfolio (blue dots), $\hat{G}$-statistic for the co-jump vector calculated on monthly basis (thick black solid lines), the monthly average $\hat{G}$-statistic for the equally weighted portfolio (thick green solid lines) and the gray shaded area captures the region in which lies the $\hat{G}$-statistic for each individual asset in the portfolio. The black long-dash line denotes the $\alpha = 0.05$ critical value to test the presence of price jumps, $\sqrt{\vartheta}\Phi^{-1}(\alpha) \cong -1.284$.

Table 2: Number of co-jumps vs. the individual assets.

Portfolio	1	2	3	4	5	6
$\hat{G}^{(1/N)} < \min \hat{G}^{(i)}$	60	51	47	55	60	70
$\min \hat{G}^{(i)} \leq \hat{G}^{(1/N)} \leq \max \hat{G}^{(i)}$	494	492	466	484	469	469
$\hat{G}^{(1/N)} > \max \hat{G}^{(i)}$	67	78	108	82	92	82

Note: The table presents number of instances when the $\hat{G}^{(1/N)}$-statistic for the equally weighted portfolio is smaller than any individual asset—the contribution of price jumps is amplified, the $\hat{G}^{(1/N)}$-statistic is in the range implied by the individual assets, and the $\hat{G}^{(1/N)}$-statistic is higher than any individual assets—the contribution of price jumps is suppressed.

presence of co-jumps is estimated ex post given the knowledge of the entire trading day.

Further, the figure depicts the range (gray shaded area) of the individual $\hat{G}$-statistics calculated for each asset in the portfolio. The results show that for the majority of trading days, at least one asset exists in the portfolio such that the null is rejected. At the same time, there is no case where the null would be rejected for every asset and; therefore, there is no co-jump for all five assets.

Figure 1 also reports the $\hat{G}$-statistic for equally weighted portfolio (blue dots). The results indicate that in the majority of cases, the $\hat{G}$-statistic for the equally weighted portfolio is in the range implied by the individual assets. However, a significant number of cases shows that the equally weighted portfolio may either amplify or supress the presence of price jumps. Table 2 summarizes the number of cases for each of the instances.

Figure 1 also presents the results of the co-jump analysis when estimated for every month (thick black solid lines). In particular, for every month, we maximize the daily average of the $\hat{G}$-statistic. The levels of the average $\hat{G}$-statistic are lower than the levels implied by the daily co-jumps; however, the levels are well above the critical levels. Thus, the price jumps at our sampling frequency can be completely diversified out at the monthly level as well.

Finally, we present the equally weighted portfolio for each month (thick green solid lines), where we depict the average $\hat{G}$-statistic using the daily levels. The average values are visually close to the monthly co-jump levels. The long period obliterate the effect between co-jumps and equally weighted portfolio; however, there are several instances, when the equally weighted monthly portfolio does not diversify out the price jump contribution.

To assess how much the individual assets contribute to the co-jumps, in Figure 2 we present the range of the components of each co-jump vector identified above for Portfolio 1 and in the Appendix we report the results for the full set of portfolios. We consider co-jump vectors, which are normalized as that $\sum_{i=1}^{5} \Omega^{(i)2} = 1$. First, the figure depicts the minimum (green) and maximum (red) of the magnitude of the co-jump vectors. The maximum magnitude oscillated around 0.75, while the minimum oscillated around 0.1 with the least magnitude taking the value of 0.0000202 and the largest one 0.9874796, taken from all portfolios. Therefore, despite several instances, each asset is significantly contributing to the co-jump vector and diversification of price jumps is clearly not caused by picking up an asset with fewest or no price jumps.

Second, the figure also reports the minimum (blue) and maximum (orange) of the co-jump vectors calculated monthly. The intuition suggests that the minimum and maximum magnitude of the co-jump vectors would move closer to that implied by the equally weighted portfolio as the wide range of price jump patterns would suggest. However, the empirical results do not confirm this interpretation, as in many instances, the monthly maximum or minimum magnitude is well below or above the daily levels, respectively. Finally, the solid black line corresponds to the value of the equally weighted portfolio.

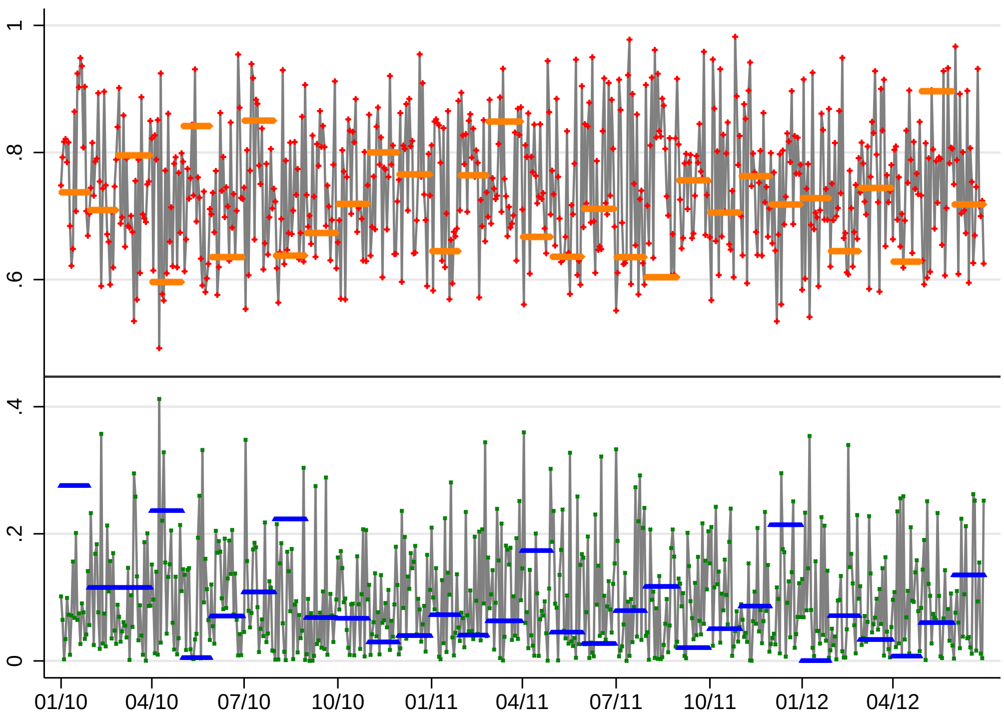
The results show the presence of co-jump vectors. From the portfolio perspective, the price jumps can be *ex post* diversified out at a 5-minute frequency, from both a daily and a monthly perspective. Further, the equally weighted portfolio is not in general sufficient to eliminate price jumps. In some cases, it actually magnifies price jumps and amplifies the deviation from Gaussianity.

In the next section, we contrast the results for co-jumps with those for co-arrivals.

5.3. Co-arrivals

First, we map each price time series on the price jump arrival indicator time series following the mapping (27). To this purpose, we identify the log-returns containing price jumps

Figure 2: Co-jump vector magnitudes: Portfolio 1.



Note: The figure depicts the minimum (green) and maximum (red) of the co-jump vectors. The figure also depicts the minimum (blue) and maximum (orange) of the co-jump vectors calculated monthly. The solid black line corresponds to the value of the equally weighted portfolio. The vectors are normalized as that $\sum_{i=1}^{5} \Omega^{(i)2} = 1$.

following the ξ-statistic of Lee and Mykland (2008).

First, we define for each time step t the absolute log-return normalized by the instantaneous Integrated Variance (iIV_t) at time t:

$$LM_t = \frac{|r_t|}{iIV_t} \, ,$$

with

$$iIV_t = \frac{n_M}{n_M - 1} \mu_1^{-2} \sum_{i=t}^{t-(n_M-2)} |r_{i-1}| \, |r_i| \, ,$$

where n_M is the length of the mowing window. Thus, the ξ-statistic to identify price jumps is given by

$$\frac{\max_{t \in A_n} |LM_t| - C_n}{S_n} \to \xi \, , \tag{33}$$

where A_n is the tested region with n observations, $C_n = (2 \ln n)^{1/2} - \frac{\ln \pi + \ln(\ln n)}{2(2 \ln n)^{1/2}}$, and $S_n = \frac{1}{(2 \ln n)^{1/2}}$. The ξ-statistic follows the standard Gumbel distribution, which is characterized by the probability distribution function $P(\xi \leq x) = \exp\left(e^{-x}\right)$. We use $n_M = 78$, corresponding to the length of the one trading day, to account for the intraday volatility pattern. We apply the ξ-statistic daily and test for the presence of price jumps using $\alpha = 0.05$ across each day.

Table 3 reports the number of identified price jump arrivals for each asset. In addition, the table provides the number of overlapping pairs of price jump arrivals within each portfolio. There is a significant number of overlapping price jump arrivals, which suggests rich co-arrival patterns.

We test for the presence of co-arrivals for each of the six portfolios on a monthly basis using (26). First, there is no evidence of co-arrivals in either of the cases as the rank of the arrival matrix $\tilde{\tilde{C}}$ is full. Therefore, we search for the presence of weak co-arrival vectors based on the results presented in Corollary 3. For each month and each portfolio, we identify the weak co-arrival vector using the findings in Corollary 4.

Table 3: Arrivals and co-arrivals for each portfolio.

P-1	XOM	MSFT	WMT	PG	JNJ	P-2	GE	IBM	JPM	T	CVX
XOM	391	65	72	74	74	GE	393	68	66	56	58
MSFT		448	91	73	82	IBM		452	68	72	69
WMT			484	107	91	JPM			426	63	57
PG				474	92	T				452	64
JNJ					466	CVX					380
No. 3	**PFE**	**CSCO**	**BAC**	**KO**	**HPQ**	**P-4**	**MRK**	**INTC**	**VZ**	**MCD**	**UTX**
PFE	445	57	59	75	69	MRK	429	60	63	69	87
CSCO		393	52	75	82	INTC		452	66	58	77
BAC			463	64	70	VZ			446	70	90
KO				505	89	MCD				437	74
HPQ					496	UTX					501
P-5	**DIS**	**MMM**	**AXP**	**HD**	**KFT**	**P-6**	**BA**	**CAT**	**DD**	**TRV**	**AA**
DIS	411	103	66	87	75	BA	481	92	97	93	74
MMM		476	86	85	87	CAT		420	95	66	85
AXP			424	70	66	DD			439	88	92
HD				476	77	TRV				519	74
KFT					477	AA					432

Note: The table presents the number of identified price jump arrivals for each asset. In addition, it captures the number of pair-wise overlapping price jump arrivals inside each portfolio.

In the majority of cases, the weak co-arrival vector chooses the asset with the least number of price jump arrivals for that particular month. In three instances, the linear combination of two assets forms the weak co-arrival vector: The *first instance* occurs in May, 2010, for General Electrics and Chevron in Portfolio 2. The *second instance* occurs in May, 2010, for Walt Disney and 3M in Portfolio 5. The *third instance* took place in August 2011, for Boeing and DuPont in Portfolio 6. Details are available in the Appendix. The weak co-arrivals are therefore driven by the idiosyncratic price jumps with a rare combination of more than asset, which carries very noisy information.

6. Robustness Check for Spurious Jump Detection

Spurious jump detection is an important issue explored in the high-frequency price jump literature (see, for instance, discussion in Lee and Mykland, 2008). Thus, the spurious jump detection may affect our results for co-jumps and co-arrivals. In this section, we report the results of a sensitivity analysis aimed to evaluate the influence of potential multiple testing

Table 4: Critical values correction for false rate detection.

window	benchmark	weekly	monthly	quarterly	full sample
α_{FD}	0.5	0.0363968	0.00718763	0.00210762	0.000167602
CV	-1.284	-1.401	-1.911	-2.234	-2.780

Note: The table reports the corrected critical values due to the multiple testing bias as the reference window increases. The benchmark value is the one employed in the co-jumps section.

biases.

In particular, in the empirical application reported above, we test for co-jumps using the G-statistic at the standard threshold $\alpha = 0.05$ and within a daily window. When we compare the co-jumps across the longer periods, we face the issue of false detection due to the repeating the statistical test many times. Thus, following Bajgrowicz and Scaillet (2011), we adjust the testing procedure for potential multiple testing bias. First, we calculate the critical number of trading days $\tilde{N}$ when the false detection starts to play a role. For $\alpha = 0.05$ a solution to

$$\alpha = \left(1 - \Phi\left(\sqrt{2 \cdot \log \tilde{N}}\right)\right) ,$$

gives $\tilde{N} = 3.87$. Bajgrowicz and Scaillet (2011) suggest that for a sample of the size larger than $\tilde{N}$, we should replace α with a more conservative α_{FD}, which suppress the false detection. In Table 4, we report the corrected critical values for the G-statistic. There is clear evidence that the length of the testing window plays a role in the correct statistical inference.

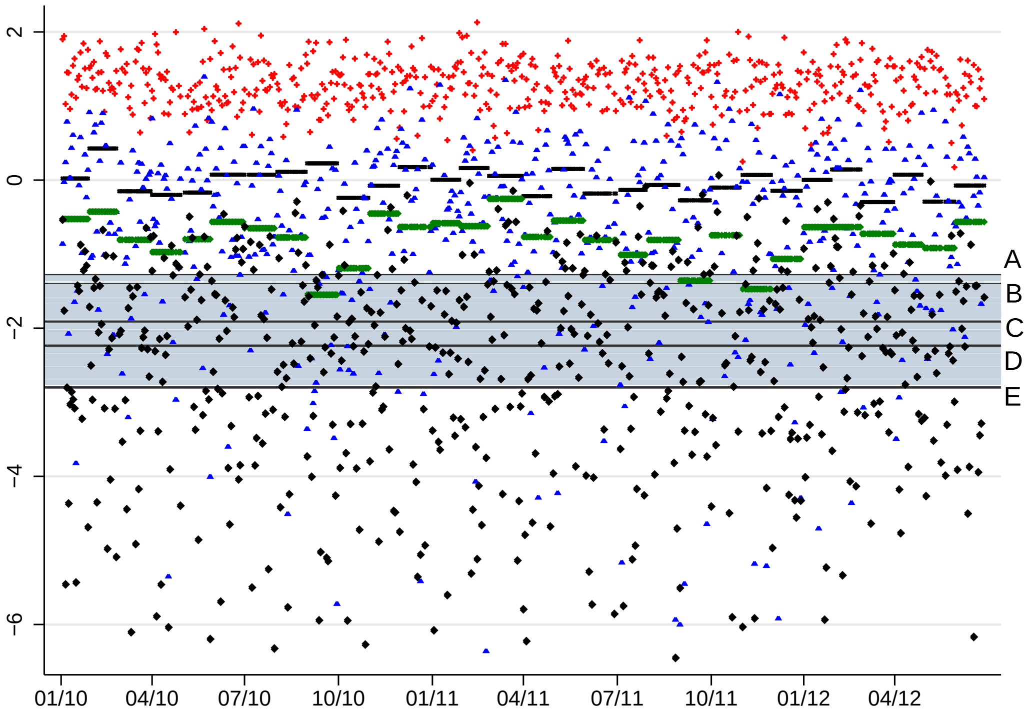
Figure 3 reports the results of testing for co-jumps using the new critical values. We plot the results for Portfolio 1 only and focus only on values of the G-statistic up to -6.5. Each black dot corresponds to the G-statistic for an asset, which reaches the lowest value; therefore, if the value is below the threshold at least one jump exists. Further, the figure contains five vertical thresholds, where each threshold value corresponds to one of the α_{FD} in Table 4. In particular, line A corresponds to the $\alpha = 0.05$ critical value to test the presence of price jumps (-1.284), line B to the weekly reference window (-1.401), line C to the monthly reference window (-1.911), line D to the quarterly reference window (-2.234), and line E to the full sample reference window (-2.780). The gray shaded area captures price jumps spuriously detected due to the multiple testing bias.

The figure shows that the number of detected price jumps decreases with longer reference window; however, a large number of price jumps remain which were not eliminated by the correction for multiple testing bias. This suggests that at our sampling frequency, there are extreme returns observationally equivalent to price jumps. In addition, the equally weighted portfolio also shows true price jumps after the multiple testing bias correction.

With respect to co-arrivals, we proceed in a similar way and adjust the critical values to account for multiple testing bias. We re-estimate the price jump arrivals for weekly, monthly, quarterly and full sample sizes, respectively. In the following, we recalculate the values reported in Table 3 using the alternative reference windows of one week, one month and one quarter. We define a week as five trading days, a month as four weeks, and a quarter as three months.

Table 5 reports the number of detected price jump arrivals for Portfolio 1 together with the overlapping pair arrivals within the portfolio for each reference window. With a paired comparison between Table 3 and 5, we see that the number of identified price jumps decreases substantially. This is in line with the fact that the increasing reference window reduces the ξ-statistic. However, the price jump arrivals are still present as are the common price jump arrivals. In addition, the ratio of common jumps is not a function of the reference window. In particular, for Portfolio 2 and a pair of CVX and JPM, the more conservative threshold in testing reveals that the ratio of mutually overlapping price jump arrivals is higher than when the multiple testing biased is not accounted for.

Figure 3: Co-jumps: $\hat{G}$-statistic – sample size check: Portfolio 1.



Note: The figure depicts the $\hat{G}$-statistic for the co-jump vector (red dots), for the equally weighted portfolio (blue dots), the minimum $\hat{G}$-statistic for an individual asset for a given day (black dots), $\hat{G}$-statistic for the co-jump vector calculated on monthly basis (thick black solid lines), and the monthly average $\hat{G}$-statistic for the equally weighted portfolio (thick green solid lines). The figure contains five vertical solid black lines corresponding subsequently to five thresholds for different reference windows. A corresponds to the $\alpha = 0.05$ critical value to test the presence of price jumps (-1.284), B to the weekly reference window (-1.401), C to the monthly reference window (-1.911), D to the quarterly reference window (-2.234), and E to the full samplereference window (-2.780). The gray shaded area captures the price jumps spuriously detected due to the multiple testing bias.

Table 5: Arrivals and co-arrivals for each portfolio – different reference windows.

Weekly reference window

P-1	XOM	MSFT	WMT	PG	JNJ	P-2	GE	IBM	JPM	T	CVX
XOM	249	34	50	43	39	GE	266	32	44	35	37
MSFT		296	54	47	52	IBM		303	41	41	39
WMT			352	74	62	JPM			295	36	33
PG				341	62	T				314	36
JNJ					311	CVX					250
P-3	PFE	CSCO	BAC	KO	HPQ	P-4	MRK	INTC	VZ	MCD	UTX
PFE	297	35	36	42	48	MRK	291	36	43	45	66
CSCO		257	30	45	52	INTC		289	36	36	53
BAC			315	37	46	VZ			303	45	58
KO				368	60	MCD				302	54
HPQ					357	UTX					360
P-5	DIS	MMM	AXP	HD	KFT	P-6	BA	CAT	DD	TRV	AA
DIS	290	68	42	52	48	BA	343	61	58	65	43
MMM		322	49	54	56	CAT		293	55	53	46
AXP			280	35	40	DD			304	58	50
HD				324	51	TRV				372	40
KFT					343	AA					275

Monthly reference window

P-1	XOM	MSFT	WMT	PG	JNJ	P-2	GE	IBM	JPM	T	CVX
XOM	168	21	31	23	30	GE	190	24	30	26	24
MSFT		218	34	29	33	IBM		214	30	30	23
WMT			274	48	40	JPM			216	25	27
PG				261	37	T				246	25
JNJ					228	CVX					187
P-3	PFE	CSCO	BAC	KO	HPQ	P-4	MRK	INTC	VZ	MCD	UTX
PFE	219	21	24	31	35	MRK	219	22	32	34	41
CSCO		192	21	28	37	INTC		212	27	24	34
BAC			237	26	38	VZ			238	31	43
KO				275	50	MCD				243	34
HPQ					287	UTX					271
P-5	DIS	MMM	AXP	HD	KFT	P-6	BA	CAT	DD	TRV	AA
DIS	204	47	28	35	28	BA	249	39	38	38	31
MMM		246	37	41	35	CAT		209	41	29	34
AXP			197	25	27	DD			230	38	31
HD				245	36	TRV				282	26
KFT					243	AA					196

Quarterly reference window

P-1	XOM	MSFT	WMT	PG	JNJ	P-2	GE	IBM	JPM	T	CVX
XOM	126	16	27	20	20	GE	143	19	25	15	17
MSFT		174	28	18	24	IBM		172	24	21	17
WMT			224	36	29	JPM			172	18	23
PG				206	29	T				189	18
JNJ					181	CVX					155
P-3	PFE	CSCO	BAC	KO	HPQ	P-4	MRK	INTC	VZ	MCD	UTX
PFE	173	13	20	15	26	MRK	180	16	22	26	33
CSCO		138	16	20	21	INTC		155	17	18	24
BAC			192	20	28	VZ			202	26	33
KO				210	36	MCD				197	26
HPQ					231	UTX					225
P-5	DIS	MMM	AXP	HD	KFT	P-6	BA	CAT	DD	TRV	AA
DIS	167	38	23	28	25	BA	203	35	29	30	23
MMM		195	32	32	27	CAT		172	30	21	31
AXP			154	18	22	DD			174	29	20
HD				202	30	TRV				238	23
KFT					198	AA					150

Full sample reference window

P-1	XOM	MSFT	WMT	PG	JNJ	P-2	GE	IBM	JPM	T	CVX
XOM	82	15	15	9	12	GE	90	8	11	9	10
MSFT		117	14	9	16	IBM		113	14	13	10
WMT			163	16	21	JPM			111	11	17
PG				140	18	T				129	8
JNJ					129	CVX					95
P-3	PFE	CSCO	BAC	KO	HPQ	P-4	MRK	INTC	VZ	MCD	UTX
PFE	114	9	14	8	13	MRK	126	11	18	12	20
CSCO		87	9	12	10	INTC		89	7	12	12
BAC			133	6	17	VZ			143	18	18
KO				138	18	MCD				137	16
HPQ					159	UTX					157
P-5	DIS	MMM	AXP	HD	KFT	P-6	BA	CAT	DD	TRV	AA
DIS	118	22	12	13	19	BA	127	21	16	19	11
MMM		128	19	21	16	CAT		104	11	15	14
AXP			112	8	13	DD			109	16	10
HD				133	19	TRV				164	10
KFT					132	AA					92

Note: The table reports the number of identified price jump arrivals for each asset and the number of pair-wise overlapping price jump arrivals within each portfolio. The test statistic is applied over four different reference windows: one week, one month, one quarter, and the full sample size, respectively.

7. Conclusion

In this paper, we employ the co-feature framework to introduce the notions of co-jumps and co-arrivals. First, the notion of co-jumps is defined as a linear combination of assets, which is free of price jumps. We provide both the limiting theory and discuss the discrete sample properties. We extend the notion of co-jumps to assets with idiosyncratic price jumps and define the weak co-jumps as a linear combination, which minimizes the price jumps. The concept is then linked to optimization of the portfolio of assets with respect to price jumps.

Second, we define the notion of co-arrival as an extension of co-jumps where the feature of interest is the price jump arrival. We provide both the limiting theory and the discrete sample properties as well. The co-arrivals are then linked to the signal processing, where the signal is represented by the arrival of price jumps.

Third, we evaluate the empirical validity of the proposed framework using assets from the Dow Jones 30 Index from 1 January 2010 to 30 June 2012 sampled at a 5-minute frequency. We form six portfolios based on the market capitalization and evaluate co-jumps and co-arrivals for each portfolio. The results show the presence of co-jumps at 5-minute frequency, while co-arrivals cease to exist. Consequently, if one wants to evaluate the impact of macro news on portfolio allocation, co-jumps may provide less accurate information than co-arrivals. This phenomena is a confirmation of the *mixed unobservability* raised by Lee (2012), where price jumps at finite sampling are observed together with the diffusion, which makes the price jump information less apparent but still asymptotically feasible.

Finally, we analyse the role of the multiple testing bias with respect to the presence of co-jumps and co-arrivals. We follow the procedure proposed by Bajgrowicz and Scaillet (2011) and eliminate the spuriously detected jumps. After the correction for this bias, the overall number of price jumps detected is suppressed; however, the qualitative properties of co-jumps and co-arrivals does not change.

The findings in this paper suggest some further developments. First, it will be interesting to extend the framework in this paper to the case of a more general price arrival process, e.g. mutually correlated self-exciting price jumps. Second, the sensitivity of the proposed framework, and in particular of the measure of commonality, can be transformed in the proper testing procedure for asynchronicity among the price jumps. Finally, the notion of co-arrivals can be used to define the information measures capturing the different features of the multivariate arrival process.

8. Acknowledgement

We wish to thank participants in the conference "Skewness, Heavy Tails, Market Crashes, and Dynamics" (University of Cambridge, 28-29 April, 2014), 13th OxMetrics User Conference (Aarhus, 5-6 September, 2013), and in the International Conference on "Systemic Risk, Contagion and Jumps" (Cass Business School, 25 January 2013), in particular Yacine Aït-Sahalia and Neil Shephard, the Finance Research Workshop (Cass Business School, 4 March 2013), in particular Richard Payne and Riccardo Pianeti, and the Economics Division Workshop (Nottingham Trent University, 6 March 2013), for useful comments. The usual disclaimer applies. Jan Novotný acknowledges funding from the European Community's Seventh Framework Program FP7-PEOPLE-2011-IEF under grant agreement number PIEF-GA-2011-302098 (Price Jump Dynamics) and GAČR grant 14-27047S.

Appendix A. Proofs

In this section, we provide the proofs of all Propositions.

Proof of Proposition 1. For any number of price jumps N_t, let us denote the matrix $\tilde{C}$ as a frame comprising of all price jump vectors c_j, with $j = 1, \ldots, N_t$. The orthogonal projection operator of the frame $\tilde{C}$ is given as

$$P_{\tilde{C}} = 1 - \tilde{C}\left(\tilde{C}'\tilde{C}\right)^{-1}\tilde{C}',$$

where the inverse of $\left(\tilde{C}'\tilde{C}\right)^{-1}$ is for $N_t > N$ understood as the Moore-Penrose pseudoinverse.

In the case of $\dim \tilde{C} < N$, then the orthogonal projection operator is non-trivial and spans the sub-space of all vectors orthogonal to every c_j. In the case of $\dim \tilde{C} = N$, the orthogonal projection is trivial as the frame spans the entire N-dimensional space. The case with $N_t = 1$ then follows from the provided formulas.

Further, $\dim \Omega_{N\times N} = N$ would imply that the dimension of $\dim \tilde{C} = 0$ and thus there are no price jumps.

□

Proof of Proposition 2. The proof of the proposition directly follows from Barndorff-Nielsen and Shephard (2006); see equation (10) on p. 8 and subsequent sections. $\Sigma_t^{(d)} = 0$ follows from (5) and

$$IV_t \xrightarrow{p} \int_0^t \sigma_s\sigma_s' ds + \sum_{j=1}^{N_t} c_j c_j'.$$

□

Proof of Proposition 3. The notion of local co-jumps assumes that every time series has exactly one price jump in $[0, 1]$. The four statements are straightforward, given that

$$E\left[dJ_t^{(i)} dJ_t^{(j)}\right] = 0, \tag{A.1}$$

for two independent counting processes i and j. In addition, this proposition is about realized price jumps irrespective of the $\hat{G}$-statistic.

Statement 1 The matrix U corresponding to purely idiosyncratic price jump arrivals is diagonal and given (A.1), rank $\left(\tilde{C}\right) = N$. In the presence of common price jumps, the matrix $\tilde{C}$ is not diagonal and thus rank $\left(\tilde{C}\right) < N$. The equality follows from the fact that each column in $\tilde{C}$ corresponds to a different column in the matrix process U, which are by construction linearly independent in probability.

Statement 2 As in the presence of common price jumps, the matrix $\tilde{C}$ is not diagonal and thus rank $\left(\tilde{C}\right) < N$, the homogenous linear system $\Omega'\tilde{C} = 0$ has non-trivial solutions.

Statement 3 In the case of local co-jumps, each column in $\tilde{C}$ corresponds to a different column in the matrix process U, which are linearly independent in probability. Further, for every vector c_j, the subspace of all orthogonal vectors is $N-1$ dimensional and the intersection of two such spaces corresponding to two linearly independent vectors c_j

Statement 4 Further, the dimension of the local co-jump space $\mathcal{L} = \bigcup_{i=1}^{N_t} \mathcal{L}^{(i)}$ is a direct consequence of the previous statement.

$\square$

Proof of Proposition 4. We prove the proposition in several steps. First, we summarize the basic properties of the $\hat{G}$-statistic (**Basic Properties of the $\hat{G}$-statistic**) and show the implications of the sequence of local alternatives in the univariate context (**Sequence of Local Alternatives**). Next, we extend the sequence of local alternatives in the multivariate framework (**Multivariate Case**). Finally, we show that the convergence of the local co-jump space is a consequence of the series of multivariate local alternatives (**Sequence of Local Alternatives in the Multivariate Case**).

Basic Properties of the $\hat{G}$-statistic. First, at a given sampling frequency M, we summarize the properties of the $\hat{G}^{(\Omega)}$-statistic as a function of $\Omega' = \left(\Omega^{(1)}, \ldots, \Omega^{(N)}\right)$, and a price jump present in any of returns $r_l^{(i)}$, with $i = 1, \ldots, N$ and $l = 1, \ldots, M$ in the following Corollary:

Corollary 5. *For the log-price process specified in (1) and for any $t \in (0,1]$ the QV_t, IV_t, and $\sqrt{IQ_t}$ are homogeneous functions of order 2 with respect to multiplying the underlying return process by constant.*[6] *In addition, for any sampling frequency M, the estimators $\widehat{RV}_M$, $\widehat{BV}_M$, and $\sqrt{\widehat{MPQ}_M}$ have the same property. As a consequence, the $\hat{G}_M$-statistic is a homogenous function of the order 0. Further, for any Ω and for any $i = 1, \ldots, M$ and $j = 1, \ldots, N$, $\lim_{\epsilon \to 0} \hat{G}_M^{(w)}\left(w'r + \epsilon \delta_i^{(j)}\right) = \hat{G}_M^{(w)}\left(w'r\right)$, where $\delta_i^{(j)}$ is a function equal to one at the i-th return of the asset j and zero otherwise.*

In addition, let us rewrite the log-price process in the interval $[0,1]$ as

$$dY_t = \mu_t dt + \sigma_t dB_t + \alpha U_t dJ_t\,, \tag{A.2}$$

where the parameter α switches the jumps. When it takes value $\alpha = 0$, the null hypothesis of no price jumps is satisfied, while for $\alpha \neq 0$, the alternative of the price jumps holds. In the case of $\alpha = 1$, the process (1) is fully recovered.

Under the fixed alternative of $\alpha \neq 0$, the $\hat{G}$-statistic is consistent

$$\lim_{M \to \infty} P\left(\hat{G}_{\alpha \neq 0} < cv\right) = 1\,,$$

where cv is the corresponding critical value for the given confidence level of the null distribution $\sqrt{\vartheta} N(0,1)$, as derived in Barndorff-Nielsen and Shephard (2006), Remark 1, v. The $\hat{G}_{\alpha \neq 0}$-statistic converges in probability under the fixed alternative as

$$\hat{G}_{\alpha \neq 0} \xrightarrow{p} \hat{G}_{\alpha = 0} - \alpha^2 M^{1/2} \frac{\sum_{i=1}^{N_t} c_i^2}{\sqrt{IQ}}\,, \tag{A.3}$$

where the convergence is actually convergence in law under the null hypothesis.[7]

[6] The function is the homogenous function of order α if $f(\lambda \cdot x) = \lambda^{\alpha} f(x)$.

[7] In terms of the distribution, the asymptotic reads $\hat{G}_{\alpha \neq 0} \xrightarrow{\mathcal{D}} \sqrt{\vartheta} N(0,1) - \delta\left(\alpha^2 \left. M^{1/2}\right|_{M=\infty} \frac{\sum_{i=1}^{N_r} c_i^2}{\sqrt{QV}}\right)$,

Sequence of Local Alternatives. Based on the asymptotic behavior of the test statistic under the alternative hypothesis, we may construct a sequence of price generating processes (A.2) with $\alpha_M \neq 0$, $M \in \mathbb{N}$, such that the sequence of price generating processes—the local alternatives—approaches asymptotically the null process (A.2) with $\alpha = 0$, for which the $\hat{G}$-statistic is not consistent. Let us index such a sequence by the changing parameter α_M defined as

$$\alpha_M = \alpha_0 + \frac{\nu}{M^{1/4}}\,, \tag{A.4}$$

where $\alpha_0 = 0$ is the parameter corresponding to the null of no price jumps and ν is some real constant $\nu \neq 0$. Such a sequence of local alternatives is approaching the α_0 process, or the null, at the Pitman's drift rate $M^{-1/4}$ as

$$\lim_{M\to\infty} \alpha_M = \lim_{M\to\infty}\left(\alpha_0 + \frac{\nu}{M^{1/4}}\right) = \alpha_0\,.$$

The asymptotic properties of the sequence of local alternatives with α_M given in (A.4) reads

$$\begin{aligned}
\hat{G}_{\alpha_M} &\xrightarrow{p} \hat{G}_{\alpha=0} - \alpha_M^2 M^{1/2}\frac{\sum_{i=1}^{N_t} c_i^2}{\sqrt{\widehat{IQ}_M}} \\
&= \hat{G}_{\alpha=0} - \left(\frac{\nu}{M^{1/4}}\right)^2 M^{1/2}\frac{\sum_{i=1}^{N_t} c_i^2}{\sqrt{\widehat{IQ}_M}} \\
&= \hat{G}_{\alpha=0} - \nu^2\frac{\sum_{i=1}^{N_t} c_i^2}{\sqrt{\widehat{IQ}_M}} \\
&= \hat{G}_{\alpha=0} - \tilde{\nu}^2\,.
\end{aligned}$$

The law of the $\hat{G}$-statistic for the sequence of local alternatives specified by (A.4) deviates

where $\delta(x)$ denotes degenerate distribution function with all the mass at point x. Therefore, the asymptotics (A.3) can be understood as $\hat{G}_{\alpha\neq 0} \xrightarrow{\mathcal{D}} \sqrt{\vartheta}N(0,1) - \delta(\infty)$, or, equally, $\hat{G}_{\alpha\neq 0} \xrightarrow{\mathcal{D}} \sqrt{\vartheta}N(-\infty,1)$.

asymptotically from the null by a non-zero constant $(-\tilde{\nu}^2)$.[8] This affects the consistency of the $\hat{G}$-statistic as follows

$$
\begin{aligned}
\lim_{M\to\infty} P\left(\hat{G}_{\alpha_M} < cv\right) &= \lim_{M\to\infty} P\left(\hat{G}_{\alpha=0} - \alpha_M^2 M^{1/2}\frac{\sum_{i=1}^{N_t} c_i^2}{\sqrt{\widehat{IQ}_M}} < cv\right) \\
&= \lim_{M\to\infty} P\left(\hat{G}_{\alpha=0} - \nu^2\frac{\sum_{i=1}^{N_t} c_i^2}{\sqrt{\widehat{IQ}_M}} < cv\right) \\
&\neq 1\,.
\end{aligned} \tag{A.5}
$$

The consistency check is evaluated at any fixed confidence level with corresponding cv of the null distribution $N(0,1)$. By choosing the proper ν, we can set the asymptotic power of the test at any value different from one. The asymptotic power of one is achieved for $\nu = 0$.

On the contrary, any sequence of local alternatives whose drift term vanishes faster than the Pitman's drift does not violate the consistency. In particular, for

$$
\alpha'_M = \alpha_0 + \frac{\nu}{M^{1/4+\epsilon}}\,, \tag{A.6}
$$

with $\epsilon > 0$, the asymptotic consistency is of the form

$$
\begin{aligned}
\lim_{M\to\infty} P\left(\hat{G}_{\alpha'_M} < cv\right) &= \lim_{M\to\infty} P\left(\hat{G}_{\alpha=0} - \left(\alpha'_M\right)^2 M^{1/2}\frac{\sum_{i=1}^{N_t} c_i^2}{\sqrt{\widehat{IQ}_M}} < cv\right) \\
&= \lim_{M\to\infty} P\left(\hat{G}_{\alpha=0} - \frac{\nu^2}{M^{2\epsilon}}\frac{\sum_{i=1}^{N_t} c_i^2}{\sqrt{\widehat{IQ}_M}} < cv\right) \\
&= 1\,.
\end{aligned} \tag{A.7}
$$

Therefore, the asymptotic properties implies that any sequence of intervals $\mathcal{A}_M = [-\alpha'_M, \alpha'_M]$, which vanishes faster than the Pitman's drift, i.e., $\alpha'_M \propto M^{-(1/4+\epsilon)}$, with $\epsilon > 0$, have the asymptotic power equal to one

[8] The constant $\tilde{\nu}^2$ should be understood in terms of the degenerate distribution.

$$\lim_{M\to\infty} P\left(\hat{G}_{\alpha'_M} < cv;\, \alpha'_M \in \mathcal{A}_M\right) = 1\,.$$

This can be interpreted as a shrinking sequence of closed intervals with asymptotics, for which the $\hat{G}$-statistic is consistent. The sequence of intervals $\mathcal{A}_M$ converges to $\mathcal{A}_\infty = 0$ in the sense of the infinite intercept of intervals

$$\mathcal{A}_M \to \mathcal{A}_\infty : \bigcap_{M\to\infty} \mathcal{A}_M = \mathcal{A}_\infty\,,$$

where the asymptotic region is $\mathcal{A}_\infty = \{\nu_0\} \equiv \{0\}$. To prove this, we have to show that $0 \subset \mathcal{A}_\infty$ and $\forall \beta \neq 0$, $\beta \not\subset \mathcal{A}_\infty$. The first part is obvious since $\alpha = 0$ is the null hypothesis. The second part of the statement follows from the consistency of the $\hat{G}$-statistic for a fixed alternative.

Multivariate Case. We extend the case to the multivariate setup specified by equation (1). We use the notation

$$dY_t^{(\Omega)} = \Omega' dY_t = \sum_{i=1}^{N} \Omega^{(i)} dY_t^{(i)}\,, \tag{A.8}$$

where $\Omega' = \left(\Omega^{(1)}, \ldots, \Omega^{(N)}\right)$ is a constant vector. The closure under stochastic integration implies that for any non-trivial vector Ω, the $\hat{G}^{(\Omega)}$-statistic has the same asymptotic properties as in the case of the univariate process.

Under the alternative, we allow for price jumps in the individual assets. The discontinuous contribution to the quadratic variance is asymptotically given as

$$\widehat{IV}_M^{(\Omega)} - \widehat{QV}_M^{(\Omega)} \xrightarrow{p} -\sum_{i=1}^{N_t} \Omega' c_i c_i' \Omega\,.$$

The discontinuous part disappears when a non-zero solution exists to

$$\sum_{i=1}^{N_t} \Omega' c_i c_i' \Omega = 0\,. \tag{A.9}$$

For every vector of price jumps c_i, it is possible to construct a subspace $\mathcal{C}_i$ of all the admissible vectors of weights $c_i^{\#}$ such that the price jump at a given time t_i disappears as

$$\mathcal{C}_i = \left\{ c_i^{\#} \in \Re^N ;\ c_i^{\#} = 1 - c_i \left(c_i' c_i \right)^{-1} c_i' \right\} ,$$

with $c_i^{\#} c_i = 0$.

The solution to (A.9) is then given by

$$\forall \Omega \in \bigcap_{j=1}^{N_t} \mathcal{C}_i \equiv \mathcal{L} \ :\ \Omega' C = 0 \, .$$

Since the trivial solution is not allowed in (A.9), the aggregate space $\mathcal{L}$ can be empty. The non-empty space defines the co-jump space.

Let us further assume that $\mathcal{L} \neq \emptyset$ and discuss the asymptotic properties of the $\hat{G}^{(\Omega)}$-statistic. If $\Omega \in \mathcal{L}$ then the

$$\hat{G}^{(\Omega)} \xrightarrow{\mathcal{D}} N\left(0, \vartheta\right) \, .$$

On the other hand, if $\Omega \notin \mathcal{L}$ then the $\hat{G}^{(\Omega)}$-statistic is consistent. This means that at any confidence level

$$\begin{aligned}
\lim_{M\to\infty} P\left(\hat{G}^{(\Omega\notin\mathcal{L})} < cv\right) &= \lim_{M\to\infty} P\left(\hat{G}^{(\Omega\notin\mathcal{L})} - M^{1/2}\frac{\sum_{i=1}^{N_t}\Omega' c_i c_i' \Omega}{\sqrt{\widehat{IQ}_M^{(\Omega)}}} < cv\right) \\
&\geq \lim_{M\to\infty} P\left(\hat{G}^{(\Omega\notin\mathcal{L})} - M^{1/2}\xi^2 < cv\right) \\
&= \lim_{M\to\infty} \Phi\left(\sqrt{\vartheta}\left(cv + \alpha^2 M^{1/2}\xi^{(w)2}\right)\right) \\
&= 1 \, ,
\end{aligned}$$

with $\Phi\left(x\right)$ being the cumulative distribution function of the $N\left(0,1\right)$, and the constant $\xi^{(\Omega)2} \equiv \liminf_{M\to\infty} \frac{\sum_{i=1}^{N_t}\Omega' c_i c_i' \Omega}{\sqrt{\widehat{IQ}_M^{(\Omega)}}}$. The asymptotic drift driving the consistency is $\propto M^{1/2}$ and has the same power as in the univariate case.

Sequence of Local Alternatives in the Multivariate Case. We assume the existence of a solution to (A.9), or $\mathcal{L} \neq \emptyset$. In the multivariate case, the parameter ν driving the deviation from locality is N-dimensional vector. We construct a sequence of local alternatives, indexed by a vector Ω_M, which converges to any $\Omega^{\#}$, $\Omega^{\#} \in \mathcal{L}$, as:

$$\Omega_M \to \Omega^{\#} + M^{1/4}\Omega_{M,\perp}\,,$$

where $\Omega'_{M,\perp}\Omega^{\#} = 0$, for all $\Omega^{\#} \in \mathcal{L}$.

Let us denote that restricting $\Omega_{M,\perp}$ to being perpendicular to all $\Omega^{\#} \in \mathcal{L}$ is not necessary. However, it allows us to form the non-overlapping sets of local alternative sequences with different $\Omega^{\#}$. This is summarized in the following corollary:

Corollary 6. *For any mutually different vectors $\Omega_1^{\#}$ and $\Omega_2^{\#}$, and for any $\Omega_{1,\perp} \in \mathcal{L}$, we cannot find another $\Omega_{2,\perp} \in \mathcal{L}$, $\Omega_{1,\perp} \neq \Omega_{2,\perp}$, such that the following equality holds*

$$\Omega_1^{\#} + M^{1/4}\Omega_{1,\perp} = \Omega_2^{\#} + M^{1/4}\Omega_{2,\perp}\,.$$

The sequence of vectors $\Omega_{M,\perp}$ is perpendicular to the solutions of (A.9) and the magnitude of the vectors $\Omega_{M,\perp}$ is not exploding or diminishing asymptotically. For that purpose, we further assume $\Omega_{M,\perp} \propto \mathcal{O}(1)$. For the sake of simplicity, we fix the sequence of $\Omega_{M,\perp} = \Omega_{\perp}$, for all M.[9]

The $\hat{G}^{(\Omega)}$-statistic under the sequence of local alternatives given by the Ω_M is not consistent

$$\begin{aligned}\lim_{M\to\infty} P\left(\hat{G}^{(\Omega_M)} < cv\right) &= \lim_{M\to\infty} P\left(\hat{G}^{\left(\Omega^{\#}+M^{1/4}\Omega_{\perp}\right)} - \frac{\sum_{i=1}^{N_t}\Omega'_{\perp}c_i c_i'\Omega_{\perp}}{\sqrt{\widehat{IQ}_M^{\left(\Omega^{\#}+M^{1/4}\Omega_{\perp}\right)}}} < cv\right)\\ &< 1\,.\end{aligned}$$

The sequence of local alternatives thus defines an envelope for local alternatives, for which the $\hat{G}$-statistic is consistent as in case of the univariate time series. The union of

[9]The same line of reasoning can be applied for slowly varying sequence $w_{M,\perp}$.

such sequences for all $\Omega^{\#}$ gives the multivariate version $\mathcal{A}_M$ with asymptotic limit $\mathcal{A}_\infty$. Analogously, in the multidimensional case we get the convergence in terms of the intercept as

$$\mathcal{L}_M \to \mathcal{L}_\infty : \bigcap_{M\to\infty} \mathcal{L}_M = \mathcal{L}_\infty \,,$$

where the asymptotic solution is $\mathcal{L}_\infty \equiv \mathcal{L}$. The proof is analogous to the univariate case. When the space of all admissible solutions to (A.9) is extended by including the origin, which is otherwise excluded, it form an s-dimensional space

$$\text{rank}\left(\{0\}\bigcup\mathcal{L}\right) = N - \text{rank}\tilde{C} = s\,.$$

By construction, $(\{0\}\bigcup\mathcal{L}_M)$ is an N-dimensional manifold, which is non-compact, where non-compactness is a direct consequence of $\hat{G}^{(\Omega)}\left(\lambda r_t\right) = \hat{G}^{(\Omega)}\left(r_t\right)$, for $\lambda \neq 0$ and $\forall M$. This concludes the proof of Proposition 4.

$\square$

Proof of Proposition 5. First, the divergence of the $\hat{G}$-statistic under the null was demonstrated in Barndorff-Nielsen et al. (2006) and Barndorff-Nielsen and Shephard (2006). The second point follows from the fact that

$$\left(\widehat{IV}_M - \widehat{QV}_M\right)\Big|_{c_1^{(1)}} \overset{p}{\to} \left(c_1^{(1)}\right)^2 \,, \tag{A.10}$$

$$\left(\widehat{IV}_M - \widehat{QV}_M\right)\Big|_{c_1^{(2)}} \overset{p}{\to} \left(c_1^{(1)}\right)^2 \,, \tag{A.11}$$

$$\left(\widehat{IQ}_M\right)\Big|_{c_1^{(1)}} \overset{p}{=} \left(\widehat{IQ}_M\right)\Big|_{c_1^{(2)}} \,, \tag{A.12}$$

where the last equality is in terms of convergence towards the same limit. Then, denoting $\hat{G}^{(i)} \equiv \hat{G}\Big|_{c_1^{(i)}}$, the limit gives

$$
\lim_{M\to\infty} \frac{\hat{G}^{(1)}}{\hat{G}^{(2)}} = \lim_{M\to\infty} \frac{M^{1/2} \left.\frac{\widehat{IV}_M - \widehat{QV}_M}{\widehat{IQ}_M}\right|_{c_1^{(1)}}}{M^{1/2} \left.\frac{\widehat{IV}_M - \widehat{QV}_M}{\widehat{IQ}_M}\right|_{c_1^{(2)}}} =
$$

$$
= \lim_{M\to\infty} \frac{\left.\widehat{IQ}_M\right|_{c_1^{(2)}}}{\left.\widehat{IQ}_M\right|_{c_1^{(1)}}} \frac{\left.\widehat{IV}_M - \widehat{QV}_M\right|_{c_1^{(1)}}}{\left.\widehat{IV}_M - \widehat{QV}_M\right|_{c_1^{(2)}}} = \frac{\left(c_1^{(1)}\right)^2}{\left(c_1^{(2)}\right)^2} .
$$

Thus, the insertion

$$
\frac{\hat{G}^{(1)}}{\hat{G}^{(2)}} \xrightarrow{p} \frac{\left(c_1^{(1)}\right)^2}{\left(c_1^{(2)}\right)^2} ,
$$

where the inequality is a consequence of the size of $\left(c_1^{(1)}\right)^2$ and $\left(c_1^{(2)}\right)^2$, respectively.

The proof can be generalized for any two price processes, given that the Integrated Variance and Integrated Quarticity are the same. In such a case

$$
\left.\left(\widehat{IV}_M - \widehat{QV}_M\right)\right|_{\sum_{j=1}^{N^{(1)}} c_j^{(1)}} \xrightarrow{p} \sum_{j=1}^{N^{(1)}} \left(c_j^{(1)}\right)^2 , \tag{A.13}
$$

$$
\left.\left(\widehat{IV}_M - \widehat{QV}_M\right)\right|_{\sum_{j=1}^{N^{(2)}} c_j^{(2)}} \xrightarrow{p} \sum_{j=1}^{N^{(2)}} \left(c_j^{(2)}\right)^2 , \tag{A.14}
$$

resulting in

$$
\frac{\hat{G}^{(1)}}{\hat{G}^{(2)}} \xrightarrow{p} \frac{\sum_{j=1}^{N^{(1)}} \left(c_j^{(1)}\right)^2}{\sum_{j=1}^{N^{(2)}} \left(c_j^{(2)}\right)^2} .
$$

□

Proof of Proposition 6. Asymptotically, co-jumps follow from Proposition 1, equation (12) using $\tilde{C}$, while co-arrivals are given by (25), using $\tilde{\tilde{C}}$.

The link between $\tilde{C}$ and $\tilde{\tilde{C}}$ is such that, $\left\{\tilde{\tilde{C}}\right\}_{i,j} = \delta_{\left(\left\{\tilde{C}\right\}_{i,j} \neq 0\right)}$. Then, it holds that

$$
\operatorname{rank}\left(\tilde{C}\right) \geq \operatorname{rank}\left(\tilde{\tilde{C}}\right) ,
$$

from which the proposition directly follows, as $\exists$co-jumps implies rank $\left(\tilde{C}\right) < N$, and as such rank $\left(\tilde{\tilde{C}}\right) < N$ as well, implying $\exists$co-arrivals.

$\square$

Appendix B. Additional Pictures

We provide a full set of pictures for all six Portfolios. Figure A.1 extends the results of Figure 1 depicting the $\hat{G}$-statistic for every portfolio.

Figure A.2 depicts the co-jump vector magnitudes for all six Portfolios, analogous to Figure 2.

Figure A.3 reports the multiple testing bias for all six Portfolios, complementing the results reported in Figure 3.

Appendix C. Illustration of the Finite Sample Properties

We illustrate the discrete sample properties of the local co-jumps in the Monte Carlo simulation exercise. We simulate a 2-dimensional vector of log-returns over the interval $t \in [0, 1]$, which are defined through the following continuous-time data generating process

$$\begin{aligned} dr_t^{(1)} &= \sigma^{(1)} dW_t^{(1)} + Y_t^{(1)} dZ_t \\ dr_t^{(2)} &= \sigma^{(2)} dW_t^{(2)} + Y_t^{(2)} dZ_t \end{aligned} , \tag{C.1}$$

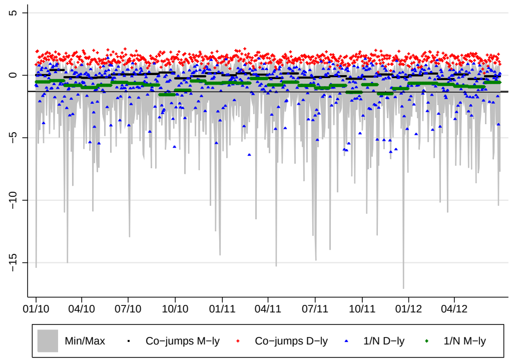
where $\sigma^{(1)}$ and $\sigma^{(2)}$ are constant volatilities, $W^{(1)}$ and $W^{(2)}$ are mutually independent Brownian motions, and $Y_t^{(1)} dZ_t$ and $Y_t^{(2)} dZ_t$ are the price jump components sharing the same arrival process. We could use a more sophisticated data generating process; however, the principle would be just illustrated equally well as with the provided model.

Let as assume that there is exactly one price jump arrival at each asset, at $t_{jump}^{(i)} \in (0, 1]$, $i = 1, 2$. Further, let us set the price jump magnitude as $Y^{(1)}_{t^{(1)}_{jump}} = -Y^{(2)}_{t^{(2)}_{jump}} \equiv Y_{jump}$.

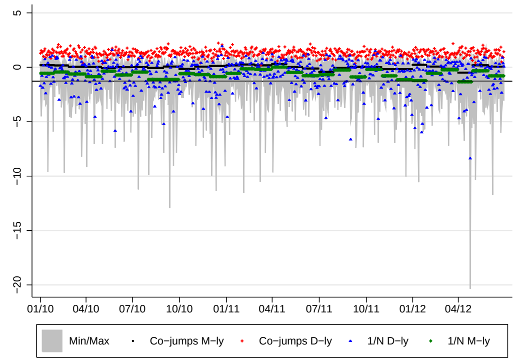
Case 1: True co-jumps.. We assume $t_{jump}^{(1)} = t_{jump}^{(2)}$, which implies that the true—the asymptotic—co-jump vector has the form $\Omega_{M\to\infty} = (\alpha, \alpha)$, where $\alpha \neq 0$. When we impose the L_2 normalization constraint, we may write any possible global co-feature vector as $\Omega_{norm,M\to\infty} =$

Figure A.1: Co-jumps: $\hat{G}$-statistic.

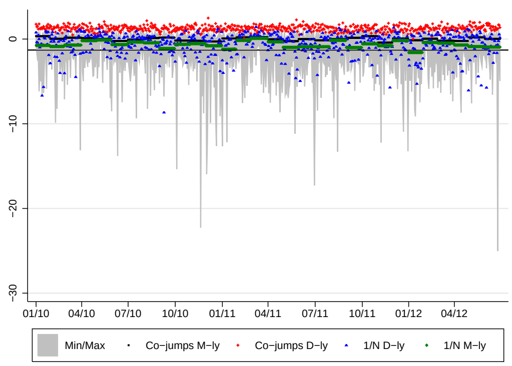
(a) Portfolio 1.



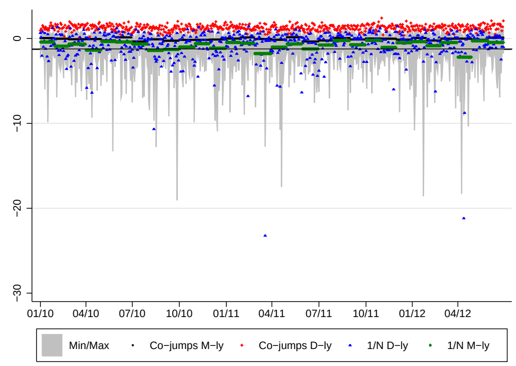
(b) Portfolio 2.



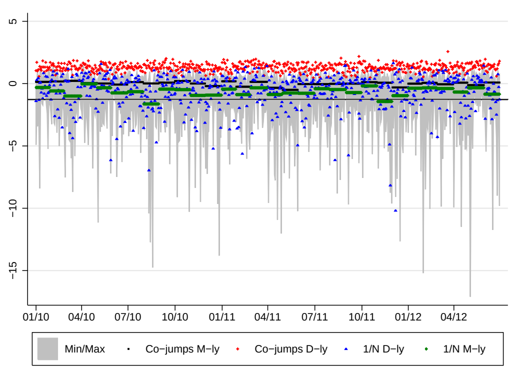
(c) Portfolio 3.



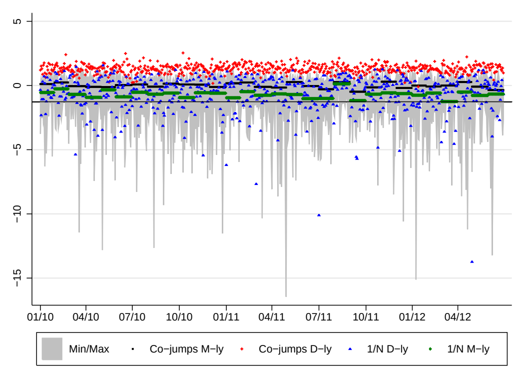
(d) Portfolio 4.



(e) Portfolio 5.



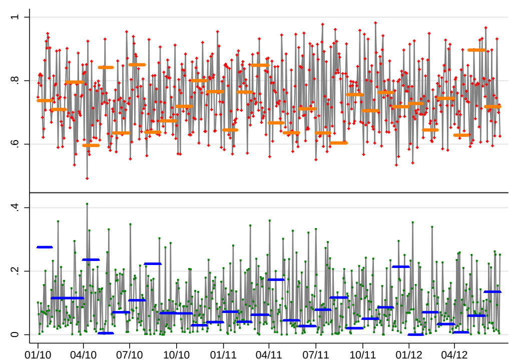
(f) Portfolio 6.



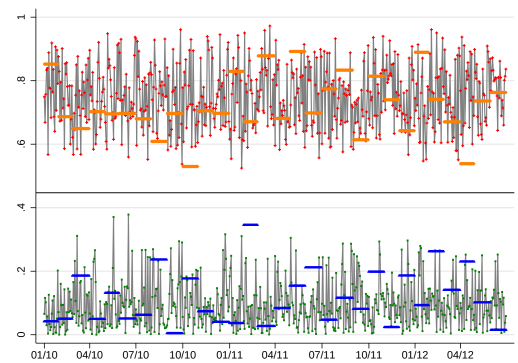
Note: The figure depicts the $\hat{G}$-statistic for the co-jump vector, for the $1/N$ portfolio, the monthly average $\hat{G}$-statistic for the $1/N$ portfolio and $\hat{G}$-statistic for the co-jump vector calculated on monthly basis. The grey area captures the region in which lies for every the $\hat{G}$-statistic for each individual asset in the portfolio. The solid black line denotes the $\alpha = 0.05$ critical value to test the presence of price jumps, $\sqrt{\vartheta}\Phi^{-1}(\alpha) \cong -1.284$.

Figure A.2: Co-jump vector magnitudes.

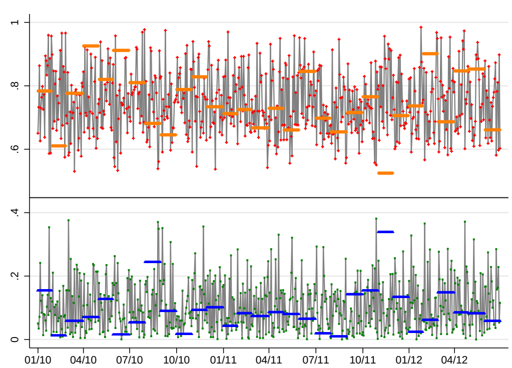
(a) Portfolio 1.



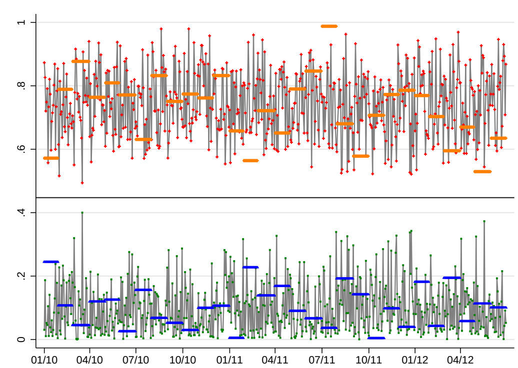
(b) Portfolio 2.



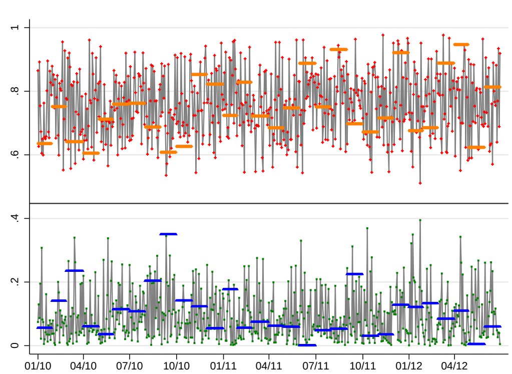
(c) Portfolio 3.



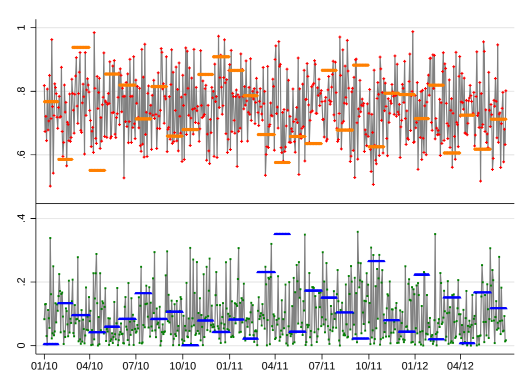
(d) Portfolio 4.



(e) Portfolio 5.



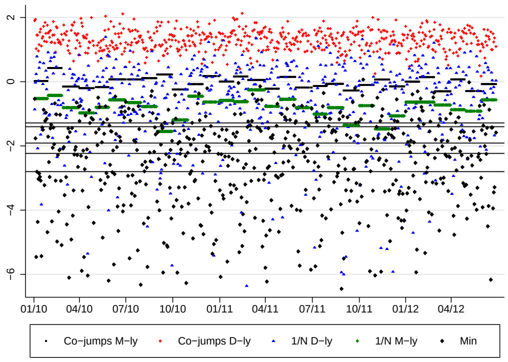
(f) Portfolio 6.



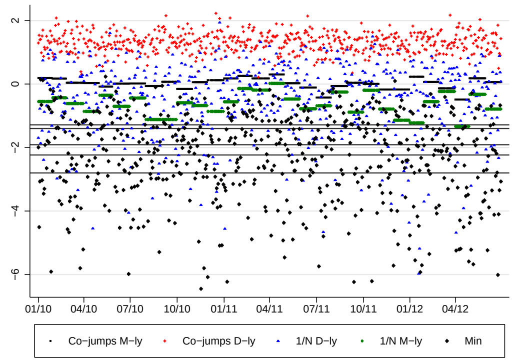
Note: The figure depicts the minimum (green) and maximum (red) of the co-jump vectors. The figure also depicts the minimum (blue) and maximum (orange) of the co-jump vectors calculated monthly. The solid black line corresponds to the value of the equally weighted portfolio. The vectors are normalized as that $\sum_{i=1}^{5} \Omega^{(i)2} = 1$.

Figure A.3: Co-jumps: $\hat{G}$-statistic – sample size check.

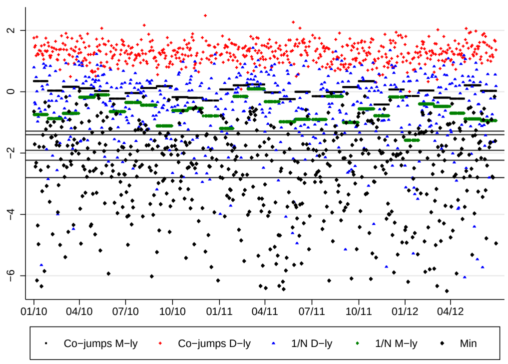
(a) Portfolio 1.



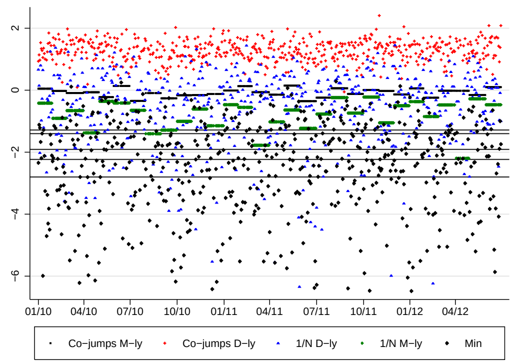
(b) Portfolio 2.



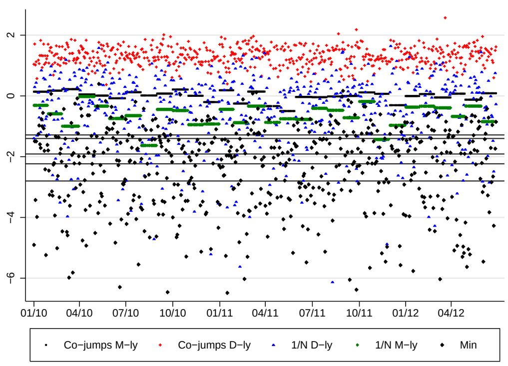
(c) Portfolio 3.



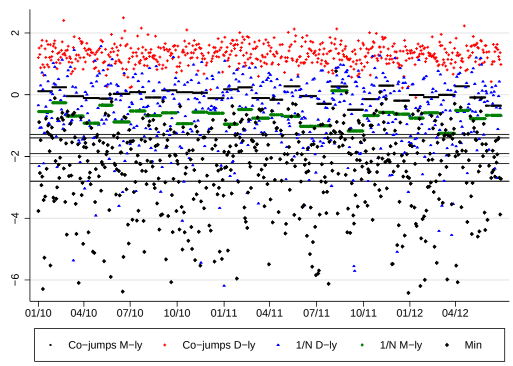
(d) Portfolio 4.



(e) Portfolio 5.



(f) Portfolio 6.



Note: The figure depicts the $\hat{G}$-statistic for the co-jump vector, for the $1/N$ portfolio, the monthly average $\hat{G}$-statistic for the $1/N$ portfolio and $\hat{G}$-statistic for the co-jump vector calculated on the monthly basis. The grey shaded area captures the region in which lies for every the $\hat{G}$-statistic for each individual asset in the portfolio. The figure contains five vertical solid black lines corresponding subsequently to five thresholds for different reference windows. The uppermost line denotes the $\alpha = 0.05$ critical value to test the presence of price jumps, $\sqrt{\vartheta}\Phi^{-1}(\alpha) \cong -1.284$. Then, the following four lines correspond to weekly, monthly, quarterly and full sample reference windows, as they are captured in α_{FD} of Table 4.

$(\cos\theta, \sin\theta)$. All admissible vectors then lie on the unit circle. Such a vector satisfies the imposed constraint by definition and any value can be uniquely described by the value of the parameter $\theta \in \langle -\pi, \pi)$. The asymptotic solution is then achieved for two values $\theta = -3/4\pi$ and $\theta = 1/4\pi$, respectively, giving the solutions $\Omega_{norm,M\to\infty} = \left(-1/\sqrt{2}, -1/\sqrt{2}\right)$ and $\Omega_{norm,M\to\infty} = \left(1/\sqrt{2}, 1/\sqrt{2}\right)$, respectively.

Case 2: No co-jumps.. We assume $t_{jump}^{(1)} \neq t_{jump}^{(2)}$ implying that no true co-jumps are present.

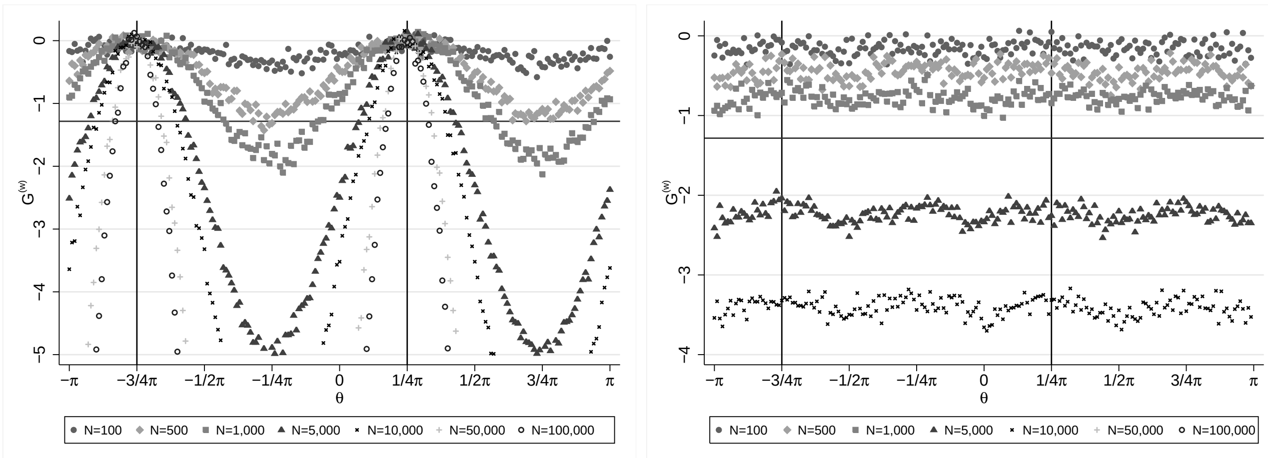
The finite sample properties of the co-jump space, as they are outlined by Proposition 4 are reported in Figure A.4, where we plot the $\hat{G}^{(\Omega)}\left(\Omega\left(\theta\right); M\right)$-statistic of the composite index constructed using the local co-jump vector $\Omega\left(\theta\right)$ for various sampling frequencies M. The data generating process is specified in (C.1) with parameters of the volatility process set as $\sigma^{(1)} = \sigma^{(2)} = 1$ and $Y_{t_{jump}} = 0.2$ with a price jump arriving in the middle of the period, i.e., $t_{jump} = 0.5$, for Case 1 (left panel), and $t_{jump}^{(1)} = 0.25$ and $t_{jump}^{(2)} = 0.75$ for Case 2 (right panel), respectively.

Whenever the $\hat{G}^{(\Omega)}\left(\Omega\left(\theta\right); M\right)$-statistic lies below the vertical line, we reject the null hypothesis that there is no price jump in the data at 95% confidence level. Further, at very low sampling frequencies, the size of the price jump is small enough in relation to the average magnitude of the noise so the null cannot be rejected. The vertical lines correspond to asymptotic values $\theta = -3/4\pi$ and $\theta = 1/4\pi$ of the normalized local co-jump vectors $\Omega_{norm,M\to\infty} = \left(-1/\sqrt{2}, -1/\sqrt{2}\right)$ and $\Omega_{norm,M\to\infty} = \left(1/\sqrt{2}, 1/\sqrt{2}\right)$, respectively. The extension from the unit circle to the entire space is than implied by using the fact that $\hat{G}^{(\Omega)} \equiv \hat{G}^{(\lambda\Omega)}$.

In particular, for Case 1 in the left panel, the figure graphically supports Proposition 4, where the region of admissible θ's, for which we reject the null hypothesis at a given sampling frequency M, is shrinking towards the limit as $M \to \infty$. The limit is represented by black vertical lines. The convergence of the intersect of the sequence of spaces $\mathcal{L}_M$ to $\mathcal{L}$ is demonstrated in the figure as well. This works as follows: As we increase the sampling frequency M, the average return is $\propto M^{-1}$, while the return with price jump is $o\left(1\right)$. Thus, the signal to noise ratio is increasing and we start to detect true price jumps. This is

Figure A.4: Local co-jump space – finite sample in 2-dimensional case.

(a) True co-jump. (b) No co-jump.



Note: The figure reports the $\hat{G}^{(\Omega)}\left(\Omega\left(\theta\right);M\right)$-statistic as a function of co-jump vector $\Omega\left(\theta\right)$ for various sampling frequencies M. The left panel reports Case 1 with true co-jump, and the right panel reports Case 2 with two idiosyncratic price jumps. Depicted are values greater than -5 for clarity. The horizontal line corresponds to 95% confidence level of the null hypothesis of no price jump in the linear combination given by $\Omega\left(\theta\right)$, at -1.284. The vertical lines correspond to $\theta = -3/4\pi$ and $\theta = 1/4\pi$, respectively, implying $\Omega^{(1)} = \Omega^{(2)}$.

further supported by Case 2 in the right panel, where there are price jumps for any linear combination.

Appendix D. Stochastic Co-jumps

Finally, we briefly introduce the notion of stochastic co-jumps. The properties of the price process under the linear transformation captured in (9) suggest the presence of such a linear transformation, which causes the jump term $\sum_{j=1}^{N_t} c_j^{(p)}$ to disappear. In the most general formulation, such a linear transformation can be a stochastic process itself. Consequently, we may define the *stochastic co-jumps* on a certain time interval $[0,1]$ is understood as the existence of the N-dimensional process Ω_s with all its elements being càdlàg and adapted to the filtration generated by the process Y_t, such that

$$Y_t^{(\Omega_t)} \equiv \Omega_t Y_t = \int_0^t \Omega_s \mu_s ds + \int_0^t \Omega_s \sigma_s dB_s\,,$$

with $\sum_{j=1}^{N_t} \Omega_{t_j-} c_j \equiv 0$, i.e., the integrated jump term in equation (8) driving the price process disappears.

A few remarks are at place: First, we want to eliminate the trivial process $\Omega_s \equiv 0$ as it satisfies the stochastic co-jump vector. For that purpose, we have to assume that at any time there almost surely exists at least one component of Ω_s, which is different from zero.

Second, analogously to the co-features literature, we may proceed with a definition of the stochastic co-jumps space. However, the properties of the stochastic co-jumps space are literarily unconstrained in most cases, constraints appear only during the moments when discontinuities take place. For that reason, we postpone analysis of the properties of stochastic co-jump space.

Lastly, analysis of the stochastic co-jumps requires anticipation of price jumps before they take place. The notion of stochastic co-jumps as a linear combination adapted to the filtration generated by the price process Y_t is a forward looking notion. In the following, we restrict ourselves to a special case of the piece-wise constant Ω_s and define the co-jumps with respect to a proper test statistic, which allows us to assess the disappearance of the jump in (8).

References

Aït-Sahalia, Y., Cacho-Diaz, J., and Hurd, T. (2009). Portfolio choice with jumps: A closed-form solution. *The Annals of Applied Probability*, 19(2):556–584.

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

Aït-Sahalia, Y. and Jacod, J. (2011). Testing whether jumps have finite or infinite activity. *The Annals of Statistics*, 39(3):1689–1719.

Aït-Sahalia, Y. and Jacod, J. (2012a). Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. *Journal of Economic Literature*, 50(4):1007–50.

Aït-Sahalia, Y. and Jacod, J. (2012b). Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data. *Journal of Econometrics*, 50:1007–1050.

Aït-Sahalia, Y., Jacod, J., and Li, J. (2012). Testing for jumps in noisy high frequency data. *Journal of Econometrics*, 168:207–222.

Andersen, T., Dobrev, D., and Schaumburg, E. (2011). A functional filtering and neighborhood truncation approach to integrated quarticity estimation. Technical report, National Bureau of Economic Research.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 169(1):75–93.

Bajgrowicz, P. and Scaillet, O. (2011). Jumps in high-frequency data: spurious detections, dynamics, and news. *Occasional Paper Series, Swiss Finance Institue*, 11(36).

Bannouh, K., Van Dijk, D., and Martens, M. (2009). Range-based covariance estimation using high-frequency data: The realized co-range. *Journal of Financial Econometrics*, 7(4):341–372.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., and Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal*, 12(3):1–32.

Barndorff-Nielsen, O. and Shephard, N. (2004a). Measuring the impact of jumps in multivariate price processes using bipower covariation. *DP, Nuffield College, Oxford University*, 2:1–37.

Barndorff-Nielsen, O. and Shephard, N. (2004b). Power and bipower variation with stochastic volatility and jumps. *Journal of financial econometrics*, 2(1):1–37.

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

Barndorff-Nielsen, O., Shephard, N., and Winkel, M. (2006). Limit theorems for multipower variation in the presence of jumps. *Stochastic processes and their applications*, 116(5):796–806.

Bollerslev, T., Law, T., and Tauchen, G. (2008). Risk, jumps, and diversification. *Journal of Econometrics*, 144(1):234–256.

Bollerslev, T., Todorov, V., and Li, S. (2013). Jump tails, extreme dependencies, and the distribution of stock returns. *Journal of Econometrics*, 172(2):307–324.

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

Engle, R. and Granger, C. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276.

Engle, R. and Kozicki, S. (1993). Testing for common features. *Journal of Business & Economic Statistics*, 11(4):369–380.

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

Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *Journal of financial econometrics*, 3(4):456–499.

Jacod, J. and Shiryaev, A. (1987). *Limit theorems for stochastic processes*, volume 288. Springer-Verlag Berlin.

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

Jiang, G. J., Lo, I., and Verdelhan, A. (2011). Information shocks, liquidity shocks, jumps, and price discovery: Evidence from the us treasury market. *Journal of Financial and Quantitative Analysis*, 46(02):527–551.

Jiang, G. J. and Oomen, R. C. (2008). Testing for jumps when asset prices are observed with noise: a swap variance approach. *Journal of Econometrics*, 144(2):352 – 370.

Lahaye, J., Laurent, S., and Neely, C. (2011). Jumps, cojumps and macro announcements. *Journal of Applied Econometrics*, 26(6):893–921.

Lee, S. (2012). Jumps and information flow in financial markets. *Review of Financial Studies*, 25(2):439–479.

Lee, S. and Hannig, J. (2010). Detecting jumps from lévy jump diffusion processes. *Journal of Financial Economics*, 96(2):271–290.

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

Liao, Y. and Anderson, H. M. (2011). Testing for co-jumps with high-frequency financial data: an approach based on first-high-low-last prices. *Monash Econometrics and Business Statistics Working Paper*, 11(9).

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2):270–296.

Podolskij, M. and Vetter, M. (2009). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3):634–658.

Urga, G. (2007). Common features in economics and finance. *Journal of Business & Economic Statistics*, 25(1):2–11.

Zhang, L. (2011). Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics*, 160(1):33–47.

Zhang, L., Mykland, P., and Ait-Sahalia, Y. (2005). A tale of two time scales. *Journal of the American Statistical Association*, 100(472):1394–1411.