Article

Data-Driven Jump Detection Thresholds for Application in Jump Regressions

Robert Davies [1] and George Tauchen [2],*

[1]   Amazon.com, 399 Fairview Ave N, Seattle, WA 98109, USA; reddavie@amazon.com
[2]   Department of Economics, Duke University, Durham, NC 27708, USA
*   Correspondence: george.tauchen@duke.edu; Tel.: +1-919-660-1800

check for
updates

Abstract: This paper develops a method to select the threshold in threshold-based jump detection methods. The method is motivated by an analysis of threshold-based jump detection methods in the context of jump-diffusion models. We show that over the range of sampling frequencies a researcher is most likely to encounter that the usual in-fill asymptotics provide a poor guide for selecting the jump threshold. Because of this we develop a sample-based method. Our method estimates the number of jumps over a grid of thresholds and selects the optimal threshold at what we term the 'take-off' point in the estimated number of jumps. We show that this method consistently estimates the jumps and their indices as the sampling interval goes to zero. In several Monte Carlo studies we evaluate the performance of our method based on its ability to accurately locate jumps and its ability to distinguish between true jumps and large diffusive moves. In one of these Monte Carlo studies we evaluate the performance of our method in a jump regression context. Finally, we apply our method in two empirical studies. In one we estimate the number of jumps and report the jump threshold our method selects for three commonly used market indices. In the other empirical application we perform a series of jump regressions using our method to select the jump threshold.

Keywords: efficient estimation; high-frequency data; jumps; semimartingale; specification test; stochastic volatility

JEL Classification: C5; C52; G12

1. Introduction

Modeling asset prices with jumps has proven to be successful, both empirically and theoretically. Because of this a method to accurately and reliably estimate the timing and magnitude of jumps in asset pricing models would greatly aid the existing literature.

Since being introduced in Mancini (2001, 2004) and threshold-based methods have become popular ways to estimate the jumps in time series data. The essential idea of these methods is that if an observed return is sufficiently large in absolute value then it is likely that the interval in which that return was taken contained a jump. To think about such an idea consider a standard jump-diffusion process for a log-asset price:

$$X_t = \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t \tag{1}$$

where b_s is thought of as the drift of the process, σ_s is a time-varying volatility process, and J_t is some finite activity jump process. (See Section 2 for a more rigorous definition of the jump-diffusion processes we consider in this paper.) Defining the returns of the observed process X as

$$\Delta_i^n X \equiv X_{i\Delta_n} - X_{(i-1)\Delta_n}, \quad i = 1, \ldots, n. \tag{2}$$

where $\Delta_n = 1/n$ is the sampling interval, n is the number of high frequency increments per day, and $n \to \infty$ for the asymptotic approximations. Note that $\Delta_i^n X$ is the geometric return in the asset price over the interval $[(i-1)\Delta_n, i\Delta_n]$. Given a sequence of thresholds, v_n, a threshold technique would label a return interval as containing a jump if $|\Delta_i^n X| > v_n$. While Mancini (2001) originally set $v_n = \sqrt{\Delta_n \log(1/\Delta_n)}$ a common practice has emerged to use $v_n = \alpha \sigma \Delta_n^\varpi$ where α and ϖ are parameters selected by the researcher and σ is the level of the local volatility around each return interval.[1] Typically, values are $\varpi = 0.49$ or $\varpi = 0.45$ and α is left as a tuning parameter.

If $\varpi = 0.49$ or $\varpi = 0.45$ the parameter α has a convenient interpretation. Since the diffusive moves in X_t are on the order $\sigma \Delta_n^{1/2}$ using $\varpi = 0.49$ or $\varpi = 0.45$ we see that the tuning parameter α has the interpretation of being essentially the number of local standard deviations of the process. A threshold-based jump selection scheme of this form then has the convenient interpretation of labeling returns as containing a jump (or multiple jumps) if the return is larger in absolute value than α local standard deviations. While this provides a nice interpretation of the method, unfortunately the literature leaves the choice α up to the researcher. The goal of the current paper is to provide a method for the selection of α. (We leave $\varpi = 0.49$ or $\varpi = 0.45$ since what is important in practice is the relative size of a 'typical' increment $\Delta_i^n X$ and v_n. See Jacod and Protter (2012, p. 248) for a discussion.)

Our primary focus in this paper is effective jump detection for the jump regression context of Li et al. (2017a) and Li et al. (2017b) where $X = (Z, Y)$. We can think of such a setting as estimating β in the following model

$$\Delta_{i_p}^n Y = \beta \Delta_{i_p}^n Z + \epsilon_{i_p}, \tag{3}$$

where $\{i_p\}_{p=1}^{N_n}$ are the N_n return intervals in Z thought to contain a jump, where N_n is the total number of identified jumps. In finance applications, Z is the log of a market index and Y is the log of a stock price. The underlying theoretical model is

$$\Delta Y_t = \beta \Delta Z_t + \Delta E_t \tag{4}$$

where Δ is the instantaneous jump operator (i.e., $\Delta X_t = X_t - X_{t-}$), and the orthogonality condition that identifies β is $\Delta Z_t \Delta E_t = 0$.[2] Equation (3) is the empirical counterpart of (4). Li et al. (2017a) contains more explanation of the theoretical model and the identifying orthogonality condition.

With a truncation threshold of the form $v_n = \alpha \sigma \Delta_n^\varpi$ we would estimate $\{i_p\}_{p=1}^{N_n}$ as

$$\{i_p\}_{p=1}^{N_n} = \{i : 1 \le i \le nT \text{ and } |\Delta_i^n Z| > \alpha \sigma \Delta_n^\varpi\}. \tag{5}$$

Notice how crucially $\{i_p\}_{p=1}^{N_n}$ and thereby the estimated $\hat{\beta}$ in (3) depends on the choice of α. There are two types of jump classification errors that can be made: (i) incorrectly labeling a particular interval as containing a jump when it does not, and (ii) omitting an interval that actually contains a jump. If we set too low a threshold, we make type-(i) errors and include return intervals in $\{i_p\}_{p=1}^{N_n}$ that do not actually contain jumps, and that could potentially badly bias the estimated 'jump beta'.

To see why we do not want to set too high a threshold and make a type-(ii) errors, we need to think about the variance of our estimated jump beta. To do so consider a heuristic model where at the jumps times $\{i_p\}_{p \ge 1}$ we have

$$\Delta_{i_p}^n Y = \beta_0 \Delta_{i_p}^n Z + \epsilon_{i_p} \tag{6}$$

---

[1]　In discussions on estimating asset pricing jumps, the volatility σ is typically treated as being known and locally constant. In practice it needs to be estimated.

[2]　In the financial econometrics literature, the symbol Δ is used three different ways: Δ_n is the sampling interval; $\Delta_i^n X \equiv X_{i\Delta_n} - X_{(i-1)\Delta_n}$ is the first difference operator over the interval of width Δ_n, and ΔX_t means the instantaneous jump in X at time t. Note that $\Delta X_t = 0$ if X is continuous at t.

where the ϵ_{i_p} are independent and identically distributed with common variances σ_ϵ^2. In addition, assume that the continuous returns are sufficiently small so that regardless of the truncation threshold used no continuous returns are included in the set of estimated jump returns. In this simplified setting

$$\mathrm{var}(\hat{\beta}) = \frac{\sigma_{\epsilon^2}}{\sum_{p \geq 1} (\Delta_{i_p}^n Z)^2}. \tag{7}$$

Notice that in this heuristic model that the variance of the estimated jump beta decreases as the number of jumps included in regression increases, i.e., as the set $\{i_p\}_{p \geq 1}$ grows. If a truncation threshold of the form $v_n = \alpha\sigma\Delta_n^\omega$ were used the variance of the estimator would be decreasing in α.

Figure 1 illustrates these ideas using some empirical data. The left panel plots the jump beta for the jump regression of the SDPR S&P500 ETF (SPY) against the SDPR utilities ETF (XLU) for the years 2007 to 2014 using five-minute returns over a grid of α jump threshold parameters. Notice that going down from $\alpha = 7$ to about $\alpha = 3.75$ that the hypothesis of a constant jump beta might be supported, i.e., that $\Delta Y_{\tau_p} = \beta \Delta Z_{\tau_p}$ where $\{\tau_p\}_{p \geq 1}$ are the true jump times in Z. The plotted jump beta is obviously noisy, but the estimated jump betas might very well be centered around a true and constant value. However after about $\alpha = 3.75$ the estimated jump betas for this asset begin to rapidly decline. It is not hard to imagine that after about $\alpha = 3.75$ the jump regressions became wildly corrupted by the addition of return intervals containing only diffusive moves. The right panel plots the reciprocal of the variance of the estimated jump betas along the same grid of α jump threshold parameters as the reciprocal of the variance of an estimator is often thought of as a measure of the 'precision' of the estimator. Notice how the precision increases as α decreases and we add return intervals to the jump regression. As in the left panel we plot a line at $\alpha = 3.75$. If after around $\alpha = 3.75$ our jump regression begins to be rapidly corrupted by the addition of return intervals that only contain diffusive moves then, even though the precision of our estimator is increasing, our estimates of the jump beta are likely to be significantly biased. These panels illustrate the trade-off mentioned earlier in selecting a jump threshold. To decrease the variance of our estimated jump beta (or increase its precision) we would like a low jump threshold, but too low a jump threshold will likely bias our jump beta since we will likely include many returns that only contain diffusive moves.

In this paper, we develop a new method to balance the trade-off of setting too low a threshold and potentially including return intervals that only contain diffusive moves versus setting too high a threshold and potentially excluding return intervals that actually contain true jumps. The main idea is to find the value of α for which the jump count function (defined below) 'bends' most sharply. Intuitively this could be thought of as the 'take-off' point of the jump count function. Selecting a threshold at this 'take-off' point should greatly reduce the number of misclassifications while maintaining many of the true jumps. We implement this idea by computing the point of maximum curvature to a smooth sieve-type estimator applied to the jump count function.

A related paper Figueroa-López and Nisen (2013) derives an optimal rate for the threshold in a threshold-based jump detection scheme with the goal of estimating the integrated variance. Using a loss function that equally penalizes jump misclassifications and missed jumps, Figueroa-López and Nisen (2013) find that the optimal threshold should be on the order of $v_n^* = \sqrt{3\sigma^2\Delta_n \log(1/\Delta_n)} + o(\sqrt{\Delta_n \log(1/\Delta_n)})$ similar to the threshold originally proposed in Mancini (2001). Since $\sqrt{3\sigma^2\Delta_n \log(1/\Delta_n)}$ is of order $\sqrt{\Delta_n \log(1/\Delta_n)}$ this result does not provide any guidance on the scale of the threshold to choose. Any threshold of the form Av_n for any $A > 0$ would be just as optimal in their setting. This presents a major challenge for practitioners. Because of this Figueroa-López and Nisen (2013) provide an iterative method for selecting the scale of the truncation threshold. This iterative method however is not motivated by a theory of the jumps or the returns and adds an additional estimation step for any researcher hoping to use their method.
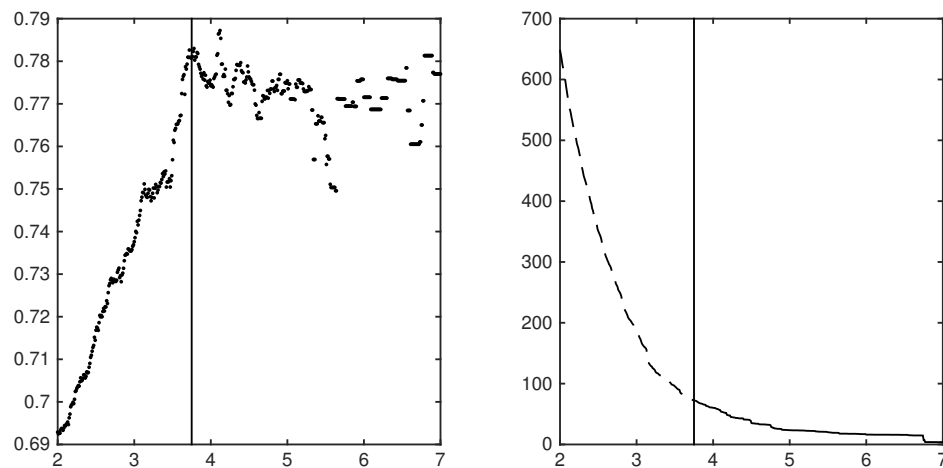
Figure 1. A jump regression illustration of the importance of the α threshold parameter. NOTE: Along the horizontal axes is the jump threshold parameter α. The left panel plots the jump beta for the jump regression of the SDPR S&P500 ETF (SPY) on the utilities sector ETF (XLU) across a grid of α threshold parameters used to estimate the jumps in SPY. The right panel plots the inverse variance of the estimated jump betas in these regressions. A vertical line has been plotted in both panels at $\alpha = 3.75$ where it appears the estimated jump beta begins to rapidly decrease. The estimates for these plots are based on five minute return data for both series spanning the years 2007 to 2014.

While we could have used truncation thresholds of the form $A\sqrt{3\sigma^2\Delta_n \log(1/\Delta_n)}$ for some $A > 0$ in our paper and investigated the choice of the scale of the threshold, i.e., the choice of A we did not for two reasons. First, the difference in the relative convergence rates of $\sqrt{\Delta_n \log(1/\Delta_n)}$ and $\alpha\Delta_n^\varpi$ are tiny when $\varpi = 0.49$ or $\varpi = 0.45$ (see the discussion in Jacod and Protter (2012, p. 248)). Second, we feel using $v_n = \alpha\sigma\Delta_n^\varpi$ provides a convenient interpretation for the tuning parameter α and therefore using $v_n = \alpha\sigma\Delta_n^\varpi$ is preferable.

The rest of the paper is organized as follows. Section 2 presents the setting. Our methodology and the main theory about its consistency are developed in Sections 3–4. Sections 5 and 6 present the results from a series of Monte Carlo studies and two empirical applications respectively. Finally, Section 7 provides a conclusion. All proofs are in the Appendix A.

2. The Setting

We start with introducing the formal setup for our analysis. The following notations are used throughout. We denote the transpose of a matrix A by $A^\top$. The adjoint matrix of a square matrix A is denoted $A^\#$. For two vectors a and b, we write $a \leq b$ if the inequality holds component-wise. The functions $vec\,(\cdot)$, $\det\,(\cdot)$ and $\mathrm{Tr}(\cdot)$ denote matrix vectorization, determinant and trace, respectively. The Euclidean norm of a linear space is denoted $\|\cdot\|$. We use $\mathbb{R}_*$ to denote the set of nonzero real numbers, that is, $\mathbb{R}_* \equiv \mathbb{R} \setminus \{0\}$. The cardinality of a (possibly random) set $\mathcal{P}$ is denoted $|\mathcal{P}|$. For any random variable ξ, we use the standard shorthand notation $\{\xi$ satisfies some property$\}$ for $\{\omega \in \Omega : \xi(\omega)$ satisfies some property$\}$. The largest smaller integer function is denoted by $\lfloor\cdot\rfloor$. For two sequences of positive real numbers a_n and b_n, we write $a_n \asymp b_n$ if $b_n/c \leq a_n \leq cb_n$ for some constant $c \geq 1$ and all n. All limits are for $n \to \infty$. We use $\xrightarrow{\mathbb{P}}$, $\xrightarrow{\mathcal{L}}$ and $\xrightarrow{\mathcal{L}\text{-}s}$ to denote convergence in probability, convergence in law, and stable convergence in law, respectively.

2.1. The Underlying Processes

The object of study of the paper is the optimal selecting of the cutoff level for a threshold-style jump detection scheme. Let X be the process under consideration and, for simplicity of exposition,

assume that X is one-dimensional. (The results can be trivially generalized to settings where X is multidimensional, but doing so would unnecessarily burden the notation.)

We proceed with the formal setup. Let X be defined on a filtered probability space represented as $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Throughout the paper, all processes are assumed to be càdlàg adapted. Our basic assumption is that X is an Itô semimartingale (see, e.g., Jacod and Protter 2012, sct 2.1.4) with the form

$$X_t = x_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t, \quad J_t = \int_0^t \int_{\mathbb{R}} \delta(s, u) \, \mu(ds, du), \tag{8}$$

where the drift b_t takes value in $\mathbb{R}$; the volatility process σ_t takes value in $\mathbb{R}_+$, the set of positive real numbers; W is a standard Brownian motion; $\delta : \Omega \times \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}$ is a predictable function; μ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with its compensator $\nu(dt, du) = dt \otimes \lambda(du)$ for some measure λ on $\mathbb{R}$. The jump of X at time t is denoted by $\Delta X_t \equiv X_t - X_{t-}$, where $X_{t-} \equiv \lim_{s \uparrow t} X_s$. Finally, the spot volatility of X at time t is denoted by σ_t. Our basic regularity condition for X is given by the following assumption.

Assumption 1. *(a) The process b is locally bounded; (b) σ_t is nonsingular for $t \in [0, T]$; (c) $\nu([0, T] \times \mathbb{R}) < \infty$.*

The only nontrivial restriction in Assumption 1 is the assumption of finite activity jumps in X. This assumption is used mainly for simplicity as our focus in the paper are 'big' jumps, i.e., jumps that are not 'sufficiently' close to zero. Alternatively, we can drop Assumption 1(c) and focus on jumps with sizes bounded away from zero.[3]

Turning to the sampling scheme, we assume that X is observed at discrete times $i\Delta_n$, for $0 \leq i \leq n \equiv \lfloor T/\Delta_n \rfloor$, within the *fixed* time interval $[0, T]$. Following standard notation as discussed in the Introduction, the increments of X are denoted by $\Delta_i^n X \equiv X_{i\Delta_n} - X_{(i-1)\Delta_n}$, $i = 1, \ldots, n$. Below, we consider an infill asymptotic setting, that is, $\Delta_n \to 0$ as $n \to \infty$.

3. Limits

Here we present some initial results needed to develop the data-drive method described in Section 4. To do so we first discuss how to think about inference for the jumps; next, we introduce the jump count function, and then we proceed to discuss jump misclassifications.

3.1. Inference for the Jump Marks

As was discussed in the introduction, in order to disentangle jumps from the diffusive component of asset returns, we choose a sequence v_n of truncation threshold values which satisfy the following condition:

$$v_n \asymp \Delta_n^{\varpi} \text{ for some constant } \varpi \in (0, 1/2). \tag{9}$$

In order to analyze the jumps of the process X it is helpful to introduce some notation. First, define $\{\tau_p\}_{p \geq 1}$ to be the successive jump times of the process X. Next, define two random sets $\mathcal{P} \equiv \{p \geq 1 : \tau_p \leq T\}$ and $\mathcal{T} \equiv \{\tau_p : p \in \mathcal{P}\}$ which collect respectively the indices of the jumps times in the interval $[0, T]$ and the jump times themselves. Since the jumps in X are assumed to be of finite activity, these two sets are almost surely finite as well. For the jump in X that occurs at time $\tau \in \mathcal{T}$, we call $(\tau, \Delta X_\tau)$ its *mark*. Finally, define a Borel measurable subset $\mathcal{D} \subset [0, T] \times \mathbb{R}_*$ as a (temporal-spatial) *region*. We do so in order to think about restricting our observation set to only those jumps that fall within a given region. To do so define the set $\mathcal{P}_\mathcal{D} \equiv \{p \geq 1 : (\tau_p, \Delta X_{\tau_p}) \in \mathcal{D}\}$.

---

[3] Yet another strategy, that can allow for studying dependence in infinite activity jumps, is to use higher order powers in the statistics that we develop henceforth. This, however, comes at the price of losing some efficiency for the analysis of the 'big' jumps.

With these definitions we can think about the true and estimated sets that index the jumps in a given sample. For each $p \in \mathcal{P}$, we denote by $i(p)$ the unique random index i such that $\tau_p \in ((i-1)\Delta_n, i\Delta_n]$. We set

$$
\begin{aligned}
\mathcal{I}_n(\mathcal{D}) &\equiv \{i : 1 \le i \le n, ((i-1)\Delta_n, \Delta_i^n X) \in \mathcal{D}, |\Delta_i^n X| > v_n\}, \text{ and} \\
\mathcal{I}(\mathcal{D}) &\equiv \{i(p) : p \in \mathcal{P}_\mathcal{D}\}.
\end{aligned}
\tag{10}
$$

The set-valued statistic $\mathcal{I}_n(\mathcal{D})$ collects the indices of returns whose 'marks' $((i-1)\Delta_n, \Delta_i^n X)$ are in the region $\mathcal{D}$, where the truncation criterion $|\Delta_i^n X| > v_n$ eliminates diffusive returns asymptotically. The set $\mathcal{I}(\mathcal{D})$ collects the indices of sampling intervals that contain the jumps with marks in $\mathcal{D}$. Clearly, the set $\mathcal{I}(\mathcal{D})$ is random and unobservable. We also impose the following mild regularity condition on $\mathcal{D}$, which amounts to requiring that the jump marks of X almost surely do not fall on the boundary of $\mathcal{D}$.

Assumption 2. $\nu(\{(s, u) \in [0, T] \times \mathbb{R} : (s, \delta_X(s, u)) \in \partial\mathcal{D}\}) = 0$, *where $\partial\mathcal{D}$ denotes the boundary of $\mathcal{D}$.*

Under Assumptions 1 and 2 it can be shown that for a fixed $v_n \asymp \Delta_n^\varpi$ that $\mathcal{I}_n(\mathcal{D})$ consistently estimates the jumps, i.e., $\mathcal{I}_n(\mathcal{D}) \xrightarrow{\mathbb{P}} \mathcal{I}(\mathcal{D})$. (See, for example, Li et al. 2017a.) The goal of the current paper is to make v_n dependent on the sample and the sampling frequency.

3.2. The Jump Count Function

The now-standard method to define the truncation level is

$$
v_n = \alpha\sigma\Delta_n^\varpi \text{ for some constant } \varpi \in (0, 1/2).
\tag{11}
$$

where σ is an estimate of the general level of local volatility, typical settings are $\varpi = 0.49$ or $\varpi = 0.45$, and α is a tuning parameter. Since the diffusive moves in X are on the order $\sigma\Delta_n^{1/2}$ and ϖ is just under $1/2$, the tuning parameter α has the convenient interpretation of essentially being the number of local standard deviations. This definition of v_n motivates a definition of the sample index of the jumps that depends on the truncation threshold α. With this in mind define

$$
\mathcal{I}_n(\alpha, \mathcal{D}) \equiv \{i : 1 \le i \le n, ((i-1)\Delta_n, \Delta_i^n X) \in \mathcal{D}, |\Delta_i^n X| > \alpha\sigma\Delta_n^\varpi\}.
\tag{12}
$$

By the presumed finite activity of the jump process in X there are only a (random) finite number of jumps and we wish to identify the set $\mathcal{I}_n(\alpha, \mathcal{D})$.

In order to do so it proves convenient to define the jump count function

$$
N_n(\alpha) = \sum_{i=1}^{nT} \mathbb{1}\left(|\Delta_i^n X| > \alpha\sigma\Delta_n^\varpi\right), \qquad \alpha \in [0, \infty).
\tag{13}
$$

Evidently, $N_n(\alpha)$ is non-increasing, piecewise flat with discontinuities at the order statistics of $|\Delta_i^n X|$. Notice $N_n(\alpha)$ decreases to zero as $\alpha \to \infty$. For each fixed α (and for any $\varpi \in (0, 1/2)$), it can be shown that for a large enough n, i.e., for a small enough Δ_n that

$$
N_n(\alpha) = |\mathcal{I}(\mathcal{D})|
\tag{14}
$$

since $N_n(\alpha) = |\mathcal{I}_n(\alpha, \mathcal{D})|$ and $\mathcal{I}_n(\alpha, \mathcal{D})$ converges to $\mathcal{I}(\mathcal{D})$. (See (Li et al. 2017a) for the details and a more thorough discussion.)

3.3. Jump Misclassifications

We think of a jump selection procedure as having a 'misclassification' if, for some return interval $\Delta_i^n X$, we $|\Delta_i^n X| > \alpha\sigma\Delta_n^\varpi$ yet over the region $((i-1)\Delta_n, i\Delta_n]$ we have $\Delta X_t = 0$. That is, if we label the return interval as containing a jump when no true jump occurred.

In order to think about jump misclassifications consider the jump count function solely for the diffusive moves. Defining the continuous moves of the process as $X_t^c = X_t - \sum_{s \leq t} \Delta X_s$ for $t \geq 0$ and $\Delta_i^n X^c \equiv X_{i\Delta_n}^c - X_{(i-1)\Delta_n}^c$ we can define the jump count function of the continuous moves as

$$N_n^c(\alpha) = \sum_{i=1}^{nT} \mathbb{1}\left(|\Delta_i^n X^c| > \alpha\sigma\Delta_n^\varpi\right), \qquad \alpha \in [0, \infty). \tag{15}$$

For a given jump threshold α and sampling frequency Δ_n, the function $N_n^c(\alpha)$ counts the diffusive moves that are 'incorrectly' labeled as jumps. Since the diffusive moves are locally Gaussian we see that $\mathbb{1}\left(|\Delta_i^n X^c| > \alpha\sigma\Delta_n^\varpi\right)$ is simply a Bernoulli random variables with probability of success equal to $2Q(\alpha\Delta_n^{\varpi-1/2})$ where $Q(\cdot) = 1 - \Phi(\cdot)$ and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal density. This is because

$$
\begin{aligned}
\mathbb{P}\left(|\Delta_i^n X^c| > \alpha\sigma\Delta_n^\varpi\right) &= \mathbb{P}\left(\left|\Delta_n^{-1/2}\frac{\Delta_i^n X^c}{\sigma}\right| > \alpha\Delta_n^{\varpi-1/2}\right) \\
&= \mathbb{P}\left(|Z| > \alpha\Delta_n^{\varpi-1/2}\right) \\
&= 2Q(\alpha\Delta_n^{\varpi-1/2})
\end{aligned}
\tag{16}
$$

where Z is a standard normal random variable. Because of this $N_n^c(\alpha)$ is simply a binomial random variable with the same probability and n draws. This implies

$$\mathbb{E}[N_n^c(\alpha)] = n2Q(\alpha\Delta_n^{\varpi-1/2}). \tag{17}$$

For a fixed $\alpha > 0$ it is fairly straight forward to show that $\mathbb{E}[N_n^c(\alpha)] = n2Q(\alpha\Delta_n^{\varpi-1/2}) \to 0$ as $\Delta_n \to 0$, which implies in the limit that the number of misclassifications goes to zero. This result however turns out not to be a good guide for the range of sampling frequencies most often encountered in practice. Table 1 reports the expected number of yearly misclassification, i.e., $T \times \mathbb{E}[N_n^c(\alpha)]$ with $T = 252$, using $\varpi = 0.49$ for $n = \{39, 78, 390, 390 \times 60\}$ over a range of α threshold parameter values. The range of sampling frequencies corresponds to ten minute, five minute, one minute, and one second sampling in a typical trading day. Notice that for each selected threshold α that the number of expected misclassifications is always increasing in the table.[4] This is in stark contrast to what one might expect given the asymptotic theory. Since the timing and magnitude of the jumps do not vary with the sampling frequency and the diffusive moves vanish as Δ_n shrinks to zero one might be led to conclude that the truncation thresholds could be decreased as the sampling frequency increases. The result in Table 1 shows that for the range of frequencies a researcher is most likely to encounter that this is not the case. Because of this result, while we remain alert to the asymptotic theory, we seek to find an optimal threshold parameter α that is sample driven, not one based solely on the asymptotic theory.

---

[4]   The intuition behind the result in Table 1 is that when $\varpi = 0.49$ we have $\Delta_n^{\varpi-1/2} \approx 1$ for $n = \{39, 78, 390, 390 \times 60\}$ resulting in $n2Q(\alpha\Delta_n^{1/2-\varpi})$ actually increasing as Δ_n decreases (or n increases).

Table 1. Expected Number of Yearly Jump Misclassifications.

Freq.	Threshold Parameter							
	$\alpha = 3.5$	$\alpha = 4$	$\alpha = 4.5$	$\alpha = 5$	$\alpha = 5.5$	$\alpha = 6$	$\alpha = 6.5$	$\alpha = 7$
39	2.78	0.328	0.030	0.0021	0.0001	4.8×10^{-6}	1.5×10^{-7}	3.8×10^{-9}
78	5.04	0.578	0.051	0.0035	0.0002	7.2×10^{-6}	2.2×10^{-7}	5.2×10^{-9}
390	19.96	2.140	0.175	0.0109	0.0005	1.9×10^{-5}	5.1×10^{-7}	1.1×10^{-8}
390×60	640.63	57.304	3.822	0.1897	0.0070	0.0002	3.9×10^{-6}	5.8×10^{-8}

NOTE: Table reports the expected number of yearly ($n \times 252$) diffusive returns that would be misclassified as jumps for each fixed α using the result that $\mathbb{E}[N_n^c(\alpha)] = n2Q(\alpha \Delta_n^{\varpi-1/2})$ where $N_n^c(\alpha)$ is the jump count function of the diffusive moves. We set $\varpi = 0.49$.

4. The Curvature Method

As briefly discussed above in the introduction, the selection of a jump threshold, i.e., the selection of α in Equation (11), involves a trade off between setting too high a threshold and failing to include all of the jumps against setting too low a threshold and erroneously labeling diffusive moves as jumps. For example, setting $\alpha = 0$ would correctly identify every jump but would also include every diffusive move. Similarly, setting $\alpha > \max_i |\Delta_i^n X|$ would guarantee that no diffusive moves were incorrectly labeled as jumps, but would fail to identify any of the jumps.

We can use the results of Section 3 to guide the selection of a suitable α. Under the modeling assumptions of Section 2, there are a finite (but random) number of jumps N^* on the interval $[0, T]$. From the theory (Jacod and Protter 2012; Li et al. 2017a) we know that for any fixed α the truncation scheme correctly classifies all N^* jumps when n is sufficiently large. Thus, for any fixed α the jump count Function (13) satisfies $\lim_{n \to \infty} N_n(\alpha) = N^*$ almost surely. Furthermore, for a fixed n and for higher values of α we should expect the jump count function to have a long flat region that is level at about N^*, but we should also expect the jump count function to rise sharply at lower values of α where many diffusive moves start getting erroneously classified as jumps. So the task is to determine from the jump count function that value of α where the jump count function starts to increase sharply as α declines. We think of this point as the point at which the jump count function begins to 'take-off'. Our solution to find this 'take-off' point is to look for the value of α at which the jump count function $N_n(\alpha)$ 'kinks' or 'bends' most sharply.

The way to mathematically define a 'kink' or sharpest 'bend' in a smooth function is the point of maximum curvature. The curvature of a smooth function $f : \mathbb{R} \to \mathbb{R}$ is defined as

$$\kappa(f) \equiv \frac{|f''|}{(1 + [f']^2)^{3/2}}. \tag{18}$$

Intuitively, if we think of the function f as lying in a two-dimensional plane and representing the direction of travel of some object, the curvature of f represents the rate at which the direction of travel is changing. (Or more rigorously the magnitude of the rate of change of the unit tangent vector to the curve.) The point of maximum curvature then is the point at which the direction of travel changes the most.

However, being piece-wise flat, the raw jump count function itself is ill-suited for this purpose, as evident in the top two panels of Figure 2. These panels plot the jump count function for the five minute returns of the SDPR S&P500 ETF during the year 2014. In the top left panel the domain $\alpha \in [2, 10]$ is too wide making the steps barely noticeable. Zooming in however on the domain $\alpha \in [6, 7.5]$ the top right panel clearly shows the jump count function to be piece-wise flat.
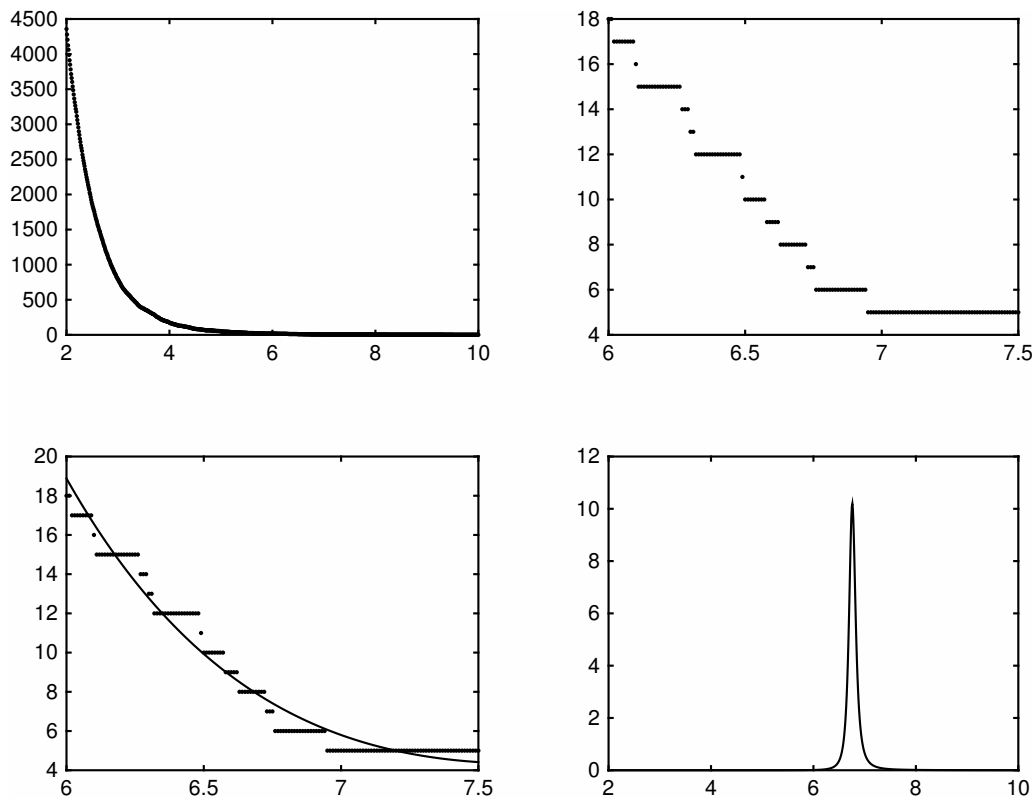
Figure 2. An illustration of the jump threshold selection method. NOTE: Along the horizontal axes is the jump threshold parameter α. The top left panel plots the estimated number of jumps over a grid of $\alpha \in [2, 10]$. The top right zooms in and plots the estimated number of jumps over $\alpha \in [6, 7.5]$. The bottom left panel adds a plot of the fitted basis function. The bottom right plots the curvature of the fitted basis function. The estimates for these plots are based on five minute return data from the SDPR S&P500 ETF (SPY) spanning the year 2014. We set $\varpi = 0.49$. See Sections 3.3 and 4 for details on the estimated number of jumps and the fitted basis functions.

Given this problem, we work with a smoother sieve estimator fitted to the jump count function. A natural choice might seem to be kernels or splines but these turn out to be ineffective due to the small wiggles and discontinuities that these functions have in their higher order derivatives. These wiggles and discontinuities in turn significantly affect the curvature of these functions making the point of maximum curvature often more dependent on the particular choice of which kernels or splines was chosen rather than the data. A far better approach is to do a least-squares projection of the observed jump count function onto a set of smooth basis functions. Given the shape of the jump count function, we use basis functions

$$g(\alpha, \gamma) \equiv \{\alpha \in [\underline{a}, \overline{a}], \gamma \in \mathbb{R}^p : \gamma_0 + \gamma_1 \alpha^{-1} + \cdots + \gamma_p \alpha^{-p}\} \tag{19}$$

where we need $p \geq 1$ for the point of maximum curvature to be well-defined. Using these basis functions we can define the projection of the jump count function onto $g(\alpha, \gamma)$ as

$$g_n(\alpha) \equiv \text{proj}_{g(\alpha, \gamma)} N_n(\alpha). \tag{20}$$

In practice we find that these basis functions result in projections with extremely tight[5] fits that have very high R^2s for low values of $p = 3$ or $p = 4$. Because of this the projection itself amounts to a compact numerical representation of approximately the same information as in the raw jump count function itself.

With this idea in mind we select α_n^* as the value that maximizes the curvature of the appropriately smoothed jump count function, i.e.,

$$\alpha_n^* = \max_{\alpha} \kappa[g_n(\alpha)] \ \text{ for } \alpha \in \left[\underline{\alpha}, \bar{\alpha}\right] \tag{21}$$

where $\alpha \in [\underline{\alpha}, \bar{\alpha}] \subset \mathbb{R}_+$. We refer to such a selection method in what follows as the 'curvature method'.

Setting the threshold right at this point of maximum curvature or 'kink' point then allows for a great many of the true jumps to be located, but guards against overly misclassify diffusive moves. Because of this, the procedure is evidently very conservative in that it lets through only a small number of diffusive moves. However, in a jump regression setting a very conservative jump selection procedure is to be preferred as the loss from including diffusive moves is very high because doing so potentially biases the estimates whereas incorrectly missing a true jump only entails a small loss of efficiency.

Though conservative, the curvature method is asymptotically accurate. We show this in Theorem 1 below. The theorem shows that in the limit the curvature method correctly identifies all of the jumps and excludes any returns that contain only diffusive moves. The theorem relies on the following definition for the convergence of random vectors with possibly different length: for a sequence N_n of random integers and a sequence $((A_{j,n})_{1 \leq j \leq N_n})_{n \geq 1}$ of random elements, we write $(A_{j,n})_{1 \leq j \leq N_n} \xrightarrow{\mathbb{P}} (A_j)_{1 \leq j \leq N}$ if we have both $\mathbb{P}(N_n = N) \longrightarrow 1$ and $(A_{j,n})_{1 \leq j \leq N} 1_{\{N_n = N\}} \xrightarrow{\mathbb{P}} (A_j)_{1 \leq j \leq N}$

Theorem 1. *Under Assumptions 1 and 2 and with $\alpha \in [\underline{\alpha}, \bar{\alpha}] \subset \mathbb{R}_+$ we have that*

(a) $\mathbb{P}[\mathcal{I}_n(\alpha_n^*, \mathcal{D}) = \mathcal{I}(\mathcal{D})] \to 1$, *and*
(b) $((i-1)\Delta_n, \Delta_n^i X)_{i \in \mathcal{I}_n(\alpha_n^*, \mathcal{D})} \xrightarrow{\mathbb{P}} (\tau_p, \Delta X_{\tau_p})_{p \in \mathcal{P}_{\mathcal{D}}}.$

The theorem above shows that as $\Delta_n \to 0$ that the jump count function $N_n(\alpha_n^*)$ using our procedure will converge in probability to the true number of jumps and that the estimated index of the jumps $\mathcal{I}_n(\alpha_n^*, \mathcal{D})$ over a region $\mathcal{D}$ will converge in probability to the true index $\mathcal{I}(\mathcal{D})$ over that region.

5. Monte Carlo Studies

We evaluate the performance of our threshold selection method on simulated data in three Monte Carlo studies. The first study compares our method with a method that simply chooses a fixed value of the truncation parameter α. The second study evaluates how our method does at recovering jumps of varying magnitudes. Finally, the third study shows the performance of our method in a jump regression setting.

5.1. Comparing Our Method with Choosing a Fixed Truncation Constant

In the first Monte Carlo study we evaluate the performance of our threshold selection method against a method that simply chooses a fixed value of α. (Where recall we label a return interval as containing a jump if $|\Delta_i^n X| > \alpha \Delta_n^\omega$.) The sample span is one year, containing $T = 252$ trading days. Each day we simulate data using $N = 390 \times 60$ high-frequency returns to match what would

---

5   The projection minimizes $\int_{\underline{\alpha}}^{\bar{\alpha}} (N_n(\alpha) - g(\alpha, \gamma))^2 \, d\alpha$ with respect to γ.

correspond to one second sampling and consider return intervals of one second, one minute, five minutes, and ten minutes. There are 1000 Monte Carlo replications and we set $\varpi = 0.49$.

The data generating process in this Monte Carlo study follows the model below. The model is taken from Li et al. (2017b) and accommodates features such as the leverage effect, price-volatility co-jumps, and heteroskedasticity in jump sizes. Let W and B be independent Brownian motions. We generate prices according to

$$dX_t = \sqrt{V_t}(\rho dB_t + \sqrt{1 - \rho^2}dW_t) + \varphi_t dN_t$$
$$d \log V_t = \mu_V dt + 0.5(dB_t + J_{V,t}dN_t) \tag{22}$$

The first displayed equation shows the price dynamics: X is the log asset price, V the local diffusive variance, ρ is parameter that captures correlation between the continuous parts of X and V (the leverage effect), φ_t is a mean-zero Gaussian price jump, and N_t is a Poisson counting process with intensity λ. The second displayed equation shows the variance (V) dynamics: μ_V is the drift in V, $J_{V,t}$ is the log-variance jump, which occurs at the same time as price jumps and is exponential with parameter η_{JV}. The parameters calibrated (realistically) are given by

$$V_0 = (18)^2, \ \rho = -0.7$$
$$J_{V,t} \overset{i.i.d.}{\sim} \text{Exp}(\eta_{JV}), \ \eta_{JV} = 0.1$$
$$\varphi_t | V_t \overset{i.i.d.}{\sim} N(0, \phi^2 V_t), \ \phi = 0.055 \tag{23}$$
$$\mu_V = -2$$
$$N_t \text{ is a Poisson process with intensity } \lambda = 20.$$

The negative value for the variance drift μ_V is needed to offset the positive upward drift generated by variance jumps with positive mean, and thereby keep $\log V_t$ from increasing off to infinity.

In addition to the selected threshold α_n^*, we report two statistics for the Monte Carlo study. The first we term the jump 'recovery rate'. This is the number of correctly identified jumps divided by the number of true jumps. A recovery rate of 100% means every true jump was correctly identified whereas a recovery rate of 0% means no true jumps were identified. The second statistic we term the 'accuracy' of the jump detection procedure. This is the number of correctly matched jumps divided by the number of estimated jumps. An accuracy of 100% means that every return interval we estimated to include a jump actually contained a true jump whereas an accuracy of 0% means that none of the return intervals we estimated to include a jump actually contained a true jump. Table 2 below reports the results of the first Monte Carlo study. All the statistics in the table are averages across the 1000 Monte Carlo replications. First notice that while the average selected value of α_n^* decreases from ten minute sampling down to one second sampling. Such a result is to be hoped for since over the sampling range of ten minutes to one second the number of jump misclassifications, as was shown in Section 3.3, is actually increasing at higher sampling frequencies. A method that attempted to minimize jump misclassifications would ideally increase the jump threshold over this range to guard against such misclassifications. Our method appears to make some effort to do so.

Notice that the average recovery rates of the curvature procedure are generally as good as and sometimes better (rarely worse) than those using a fixed α. At the same time, the average accuracy of the procedure is above 90% for all sampling frequencies, unlike the fixed $\alpha = 4$ case. The curvature method can achieve substantially increased recovery rates with little sacrifice in accuracy. As for the other values of α, the accuracy remains high but at the expense of a lower recovery rate than that of the curvature method.

Table 2. Comparison against a fixed truncation scheme: Monte Carlo averages (%).

| | Curvature | | | Fixed Truncated Parameter | | | | | | | |
| | Method | | | $\alpha = 4$ | | $\alpha = 5$ | | $\alpha = 6$ | | $\alpha = 7$ | |
Freq.	α_n^*	REC	ACC	REC	ACC	REC	ACC	REC	ACC	REC	ACC
10 min	4.83	29.35	98.61	39.22	89.36	27.75	99.63	17.17	100.00	11.08	100.00
5 min	4.65	49.49	98.95	56.09	92.63	46.21	99.91	36.83	100.00	28.65	100.00
1 min	4.22	79.61	94.85	80.71	87.09	76.08	99.90	71.42	100.00	65.89	100.00
1 s	5.95	96.70	91.48	97.46	25.74	97.03	98.95	96.54	100.00	95.73	100.00

NOTE: REC is average jump recovery rate, ACC is the average accuracy of estimated jumps, and α_n^* is the average selected threshold parameter α across the Monte Carlo replications. The jump recovery rate is defined as the number of correctly matched jumps divided by the number of true jumps. The jump accuracy is defined as the number of correctly matched jumps divided by the number of estimated jumps. The jump accuracy and recovery rate are in percentage terms. The results are based on 1000 replications following the data generating process outlined in (22) and (23).

5.2. Recovering Jumps of Varying Magnitudes

For the second Monte Carlo study, we use modification of a standard setup to examine how our method performed in recovering jumps of differing magnitudes. To this end we simulated jumps that, with equal probability, took sizes varying from one to ten unit standard deviations of the local volatility.[6] To do this, we modeled the jumps as following a compound Poisson process, that once scaled for the local volatility, had a jump size density that followed a discrete uniform distribution taking values in the range $\{1, 2, \ldots 10\}$. Using such a jump density allows us to observe how well our method can and cannot detect jumps of various magnitudes.

Letting (W_t, B_t) be a vector of Brownian motions with $\mathrm{Corr}(W_t, B_t) = 0.5$, the model is defined as

$$
\begin{aligned}
dX_t &= \sqrt{V_t} dW_t + \sqrt{V_t} u_t N_t, \\
dV_t &= 0.03(1.0 - V_t)dt + 0.1\sqrt{V_t} dB_t, \\
&\quad N_t \text{ is a Poisson process with intensity } \lambda = 1/12 \times 252, \text{ and}
\end{aligned}
\tag{24}
$$

where u_t is an i.i.d. discrete uniform distribution that takes values in $\{1, \ldots, 10\}$. (Setting $\mathrm{Corr}(W_t, B_t) = 0.5$ allows for a dependence between X_t and V_t, i.e., a leverage effect.) We set $\lambda = 1/12 \times 252$ so there should be on average a one-twelfth chance of a jump occurring each day. This is consistent with previous studies on market jumps. The data generating process for the diffusive moves and the volatility process is similar and based on that found in Li et al. (2017a).

We perform the study using 1000 replications and set $T = 3 \times 252$, which corresponds to three years' worth of simulated data. We use an Euler scheme to simulate the high-frequency data doing an initial simulation with $N = 390 \times 60 \times 10$ which corresponds to sampling once every tenth of a second. We then sample these high-frequency returns at one second, one minute, five minute, and ten minute frequencies.

Table 3 reports the results of this Monte Carlo study. The table lists the averages across all 1000 Monte Carlo replications. Consistent with a theory of vanishing diffusive moves the recovery rates increase significantly with each increase in the sampling frequency. At a 10 minute frequency we recovery most jumps greater than eight local standard deviations, a few jumps between five and seven local standard deviations, and virtually no jumps of sizes one to four local standard deviations. Sampling at a five minute frequency we make significant gains in recovering jumps of five to seven local standard deviations. At a one minute sampling frequency we can uncover nearly all jumps except those of one local standard deviation. Finally, at a one second sampling all of the jumps are

---

[6] Where to preserve the jump sizes across sampling frequencies we used the local volatility in terms of return intervals at the coarsest sampling frequency, which here corresponded to sampling at a ten minute frequency.

recovered. (Note though that the increase in the sampling frequency going from one minute to one second sampling is significantly greater than going from ten to five to one minute sampling so the stark contrast between the one minute and the one second sampling should not be exaggerated.)

The average selected threshold parameter α_n^* appears to decrease somewhat from ten minute to five minute to one minute sampling, but increases quite significantly going from one minute to one second sampling. Following the discussion in Section 3.3 the large increase in the selected threshold from one minute to one second sampling is to be hoped for as the number of expected jump misclassifications increases greatly going from one minute to one second sampling. The slight decrease in the average selected threshold parameter going from ten minute to one minute sampling, while not ideal in terms of the arguments of Section 3.3, does not appear to drastically change the accuracy of the estimated jumps. The accuracy over these three sampling frequencies is always above 98% and only decreases to 94.28% at one second sampling.

Table 3. Recovering jumps of varying magnitudes: Monte Carlo averages.

| Freq. | α_n^* | Accuracy (%) | Recovery Rates (%) by Jump Size | | | | | | | | | |
			1σ	2σ	3σ	4σ	5σ	6σ	7σ	8σ	9σ	10σ
10 min	5.27	99.71	0.32	0.15	1.27	5.51	18.53	41.48	65.93	83.83	94.07	97.79
5 min	5.07	99.83	0.31	0.91	11.68	46.62	84.69	97.67	99.86	99.96	99.94	99.97
1 min	4.41	98.22	6.38	93.90	99.99	100.00	99.95	99.97	99.99	100.00	99.98	100.00
1 s	5.72	94.28	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

NOTE: The recovery rates are for jumps of sizes equal to 1–10 unit local standard deviations in terms of return intervals sampled at a ten minute frequency. The σ above indicates a unit of local standard deviation. The average selected threshold parameter α across Monte Carlo replications is denoted α_n^*. The jump recovery rate is defined as the number of correctly matched jumps divided by the number of true jumps. The jump accuracy is defined as the number of correctly matched jumps divided by the number of estimated jumps. The results are based on 1000 replications following the data generating process outlined in Section 5.2.

5.3. Jump Regression Setting

The third Monte Carlo study examines how our procedure performs in a jump regression context. A thorough overview of jump regressions can be found in Li et al. (2017a). Below we only give a brief overview of jump regressions and the results we use in our Monte Carlo study. Given two series of returns $\Delta_i^n Z$ (often a proxy for the market) and $\Delta_i^n Y$ (often the return on an asset price) a jump regression considers a regression of $\Delta_i^n Z$ on $\Delta_i^n Y$ only over the return intervals in which Z is thought to contain a jump. The null in many jump regression settings is that the jump regression coefficient, termed the jump beta, is constant at every jump time, i.e.,

$$\Delta Y_{\tau_p} = \beta \Delta Z_{\tau_p}, \text{ where } \tau_p \text{ are the jump times of } Z. \tag{25}$$

For this Monte Carlo study we perform a test of a constant jump beta under both a simulated model that has a constant jump beta and a model with a time varying jump beta. We report rejection rates for the test as well as the average selected thresholds α_n^* and the accuracy and recovery rates of the estimated jumps. For the test of a constant jump beta we use a bootstrap version of the determinant test of Li et al. (2017a).

We simulate data using a model adapted from Li et al. (2017a). The model takes the form

$$
\begin{aligned}
dZ_t &= \sigma_t dW_t + \sigma_t dJ_t \\
dY_t &= \beta^c \sigma_t dW_t + \beta_t^J \sigma_t dJ_t + \frac{\sigma_t}{\sqrt{2}} d\tilde{W}_t + \sigma_t d\tilde{J}_t \\
d\sigma_t^2 &= 0.03(1 - \sigma_t^2)dt + 0.15\sigma_t dB_t
\end{aligned}
\tag{26}
$$

where W, $\tilde{W}$, and B are three independent Brownian motions. J_t and $\tilde{J}_t$ are compound Poisson jump processes where the jump size densities follow double-exponential (or Laplacian) distributions and the jump intensities are $\lambda = 1/12 \times 252$ and $\tilde{\lambda} = 1/48 \times 252$ respectively. We set $\beta^c = 0.89$.

The jump beta process β_t^J follows the following specifications under the null and the alternative

$$\beta_t^J = 1, \text{ for all } t \in [0, T], \text{ under } H_0 \text{ (null hypothesis)}$$
$$d\beta_t^J = 0.005(1 - \beta_t^J)dt + 0.005\sqrt{\beta_t^J}d\tilde{B}_t, \text{ under } H_a \text{ (alternative hypothesis)} \tag{27}$$

where $\tilde{B}$ is a Brownian motion independent of W, $\tilde{W}$, and B. The unconditional mean of β_t^J under the alternative is 1. The model differs from Li et al. (2017a) only in the specification of different jump and diffusive betas.

We perform the study using 1000 replications and set $T = 3 \times 252$, which corresponds to three years' worth of simulated data. We use an Euler scheme to simulate the high-frequency data doing an initial simulation with $N = 390 \times 60 \times 10$ which corresponds to sampling once every tenth of a second. We then sample these high-frequency returns at one second, one minute, five minute, and ten minute frequencies. These parameters were chosen to match the Monte Carlo study in Li et al. (2017a) as closely as possible.

Table 4 below reports the results of our study. Notice how differently the size of the test is affected by the choice of the jump threshold parameter. Using the curvature method the test is only moderately over-sized at ten and five minute sampling and not terribly over-sized at one minute sampling. (This is perhaps to be somewhat expected as Li et al. (2017a) found the test of a constant jump beta to be moderately over-sized.) These fairly mild over rejections using the curvature method however are in stark contrast to using a fixed $\alpha = 4$. Notice that using a fixed $\alpha = 4$ how the size of the tests becomes progressively worse and worse as the sampling frequency increases. Even at ten and five minute sampling the test is quite over-sized. This result is due to the inclusion of return intervals in the jump regression that only contained diffusive moves thereby biasing the estimated jump beta. To see this notice that using $\alpha = 4$ the accuracy of the jump detection procedure deteriorates significantly as the sampling frequency increases. At one minute sampling the average accuracy is 88.1% and at one second sampling the average accuracy is a very low 26.5%. This means that in the respective jump regressions on average fully 11.9% and 84.5% of the respectively estimated returns did not actually contain a jump. In contrast using the curvature method the accuracy of the estimated jumps remains high at all sampling frequencies. Finally, note that the power of the test using both methods is consistent with the results in Li et al. (2017a).

Table 4. Monte Carlo rejection rates (%) for tests of a constant jump beta.

Curvature Method											
		Under H_0							**Under H_a**		
		Average		**Nominal Level**				**Average**		**Nominal Level**	
Freq.	α_n^*	**ACC**	**REC**	**10%**	**5%**	**1%**	α_n^*	**ACC**	**REC**	**10%**	**5%**	**1%**
10 min	5.64	99.8	34.4	11.6	6.5	1.8	5.65	99.8	34.6	52.1	42.6	22.8
5 min	4.42	99.7	52.4	10.6	5.8	1.8	4.42	99.8	52.9	71.0	62.0	45.4
1 min	5.04	98.1	78.6	13.5	8.3	2.2	5.01	98.1	78.6	98.6	96.9	94.1
1 s	5.57	92.6	96.0	19.7	11.6	4.1	5.59	92.7	96.1	100.0	100.0	100.0

Table 4. *Cont.*

Fixed $\alpha = 4$ as in Li et al. (2017a)											
	Under H_0						Under H_a				
	Average			Nominal Level			Average			Nominal Level		
Freq.	α	ACC	REC	10%	5%	1%	α	ACC	REC	10%	5%	1%
10 min	4.00	92.8	48.3	12.7	7.1	3.0	4.00	92.9	48.6	52.0	41.4	23.1
5 min	4.00	93.6	61.0	12.7	7.5	2.6	4.00	93.7	61.0	70.9	61.5	43.1
1 min	4.00	88.1	80.5	18.8	10.5	4.2	4.00	88.2	80.6	98.7	96.9	94.5
1 s	4.00	26.5	97.2	99.1	96.6	86.5	4.00	26.3	97.3	100.0	100.0	100.0

NOTE: The rejection rates are for the null of a constant jump beta using a bootstrap version of the determinant test of Li et al. (2017a). REC is average jump recovery rate, ACC is the average accuracy of estimated jumps, and α_n^* is the average selected threshold parameter α across the Monte Carlo replications. The jump recovery rate is defined as the number of correctly matched jumps divided by the number of true jumps. The jump accuracy is defined as the number of correctly matched jumps divided by the number of estimated jumps. The jump accuracy and recovery rate are in percentage terms. The results are based on 1000 replications following the data generating processes for the null and alternative as outlined in Section 5.3.

6. Empirical Application

We considered two empirical applications. The first estimates the jumps and reports the jump threshold selected by our method for three commonly used and high liquid market indices. The second application reports the results of jump regressions of the nine SDPR sector ETFs against the SDPR S&P500 ETF using our method to select the jump threshold.

6.1. Estimating Jumps in Market Indices

For the E-mini S&P500 index futures (ES), the SPDR S&P500 ETF (SPY), and the VIX futures (VIX) we use the tools developed in this study to estimate the optimal jump thresholds for each series over a range of dates and a range of sampling frequencies. We report both the jump threshold selected by our method as well as the estimated number of jumps at each selected jump threshold. The SPY and ES series span the dates 3 January 2007 to 12 December 2014. The VIX series spans the dates 2 July 2012 to 30 April 2015. Only the more recent VIX futures data are used because Bollen and Whaley (2015) provide evidence that the VIX futures market was highly illiquidity and immature in prior periods. For each series we remove market holidays and partial trading days; and, to guard against possible adverse microstructure effects, we discard the first five minutes and the last five minutes of each trading day.

For each series we performed the estimation over both the entire sample and each complete calendar year within each sample. In addition, we performed the estimation using one minute, five minute, and ten minute intraday returns. Tables 5 and 6 report the selected jump threshold α_n^* and the estimated number of jumps at the selected jump threshold.

In Table 6 which reports the selected jump threshold α_n^*, notice that for the E-mini S&P500 index futures (ES) and the SPDR S&P500 ETF (SPY) there appears to be somewhat of an increase in the selected threshold as the sampling frequency increases from ten minute to five minute to one minute sampling. As was discussed in Section 3.3 this is to be hoped for as the number of jump misclassifications is actually increasing over this range of sampling frequencies. For the VIX futures (VIX) we do not see much of a pattern in the selected jump threshold α_n^*. This however should not be seen as evidence against our threshold selection procedure since Andersen et al. (2015) provide evidence that the high-frequency returns of the VIX futures might be well modeled as following an α-stable distribution with $\alpha \approx 1.8$. If this were true then not only would we not expect the same misclassification dynamics as in the diffusive case, but the correct scaling of the returns would be on the order of Δ_n^α rather than $\Delta_n^{1/2}$.

For Table 5, which reports the estimated number of jumps, notice that the number of estimated jumps is always increasing as the sampling frequency increases. Note also that the number of jumps

detected at the 5-min and 10-min frequencies is very small, reflecting the inherent conservative nature of the curvature method. In practice, common sense suggests that at these coarser frequencies the practitioner might elect to experiment a bit with slightly lower values of α than those produced directly by the curvature method, which does define a sensible baseline however.

Table 5. Estimated Number of Jumps.

SPDR S&P500 ETF (SPY)									
Frequency	**Full Sample**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
10 min	6	3	1	1	0	0	0	0	1
5 min	20	3	4	1	3	1	5	3	1
1 min	81	6	13	13	10	10	12	5	6
E-mini S&P500 Futures (ES)									
Frequency	**Full Sample**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
10 min	6	1	0	1	0	0	1	0	1
5 min	20	2	3	2	3	1	4	3	1
1 min	100	12	9	9	13	14	14	11	12
VIX Futures									
Frequency	**Full Sample**	**2013**	**2014**						
10 min	3	2	1						
5 min	9	6	2						
1 min	43	17	11						

NOTE: The table reports the estimated number of jumps for each sample at the chosen α_n^* jump threshold given in Table 6. The frequency refers to the sampling frequency of the returns. The SPY and ES series span the dates 3 January 2007 to 12 December 2014. The VIX series spans the dates 2 July 2012 to 30 April 2015.

Table 6. Selected Jump Threshold (α_n^*).

SPDR S&P500 ETF (SPY)									
Frequency	**Full Sample**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
10 min	5.61	5.81	5.20	5.25	6.07	5.38	6.13	5.26	5.71
5 min	6.13	5.98	5.54	5.81	6.30	5.80	6.60	6.17	7.01
1 min	6.67	7.12	6.13	6.24	6.49	6.43	7.59	6.79	6.76
E-mini S&P500 Futures (ES)									
Frequency	**Full Sample**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
10 min	5.60	5.97	5.25	5.39	5.81	5.37	6.06	5.24	5.52
5 min	6.15	6.20	5.66	5.86	6.26	5.90	6.52	6.04	6.72
1 min	6.11	6.45	6.08	5.94	5.80	5.95	6.69	5.93	6.13
VIX Futures									
Frequency	**Full Sample**	**2013**	**2014**						
10 min	5.78	5.65	5.67						
5 min	5.52	5.38	5.41						
1 min	5.49	5.18	5.30						

NOTE: The table reports the selected jump threshold for each sample based on the procedure in Section 4. The frequency refers to the sampling frequency of the returns. The SPY and ES series span the dates 3 January 2007 to 12 December 2014. The VIX series spans the dates 2 July 2012 to 30 April 2015.

6.2. Jump Regressions

Using the nine SDPR sector ETFs we perform a series of jump regressions of the sector ETFs against the SDPR S&P500 market ETF (SPY). We determine the jumps in the SPY series via the jump threshold parameter α_n^* based on the curvature method developed Section 4 above. Then to examine

how sensitive these jump regressions are to different jump thresholds we consider two other thresholds α_n^+ and α_n^- which are equal to α_n^* plus and minus 15% respectively. The reason for basing the jump threshold on the SPY series is that a jump regression only considers the beta for the regression of the specific asset return on the market return for intervals in which the market (SPY) is thought to have jumped. Note that the data are for the year 2009 and that we use one-minute returns to estimate the jumps but five-minute returns to perform the jump regressions.[7] We chose the year 2009 because it was a representative year and one for which there appeared empirical support for a constant jump beta for each asset over the year.[8]

Table 7. Jump regression results for the nine SDPR sector ETFs against the SDPR S&P500 ETF.

Jump Thresholds					
$\alpha_n^- = 5.18$	$\alpha_n^* = 6.09$	$\alpha_n^+ = 7.00$			
Asset	α	$\hat{\beta}$	**S.E.**	R^2	*p*-**Value**
XLB	α_n^-	1.049	0.060	0.894	0.021
	α_n^*	0.996	0.102	0.905	0.045
	α_n^+	0.988	0.112	0.922	0.015
XLE	α_n^-	0.948	0.065	0.877	0.029
	α_n^*	0.902	0.101	0.984	0.679
	α_n^+	0.934	0.117	0.994	0.471
XLF	α_n^-	1.182	0.144	0.866	0.360
	α_n^*	1.687	0.577	0.984	0.956
	α_n^+	1.691	0.661	0.998	0.872
XLI	α_n^-	1.046	0.048	0.955	0.151
	α_n^*	0.967	0.078	0.991	0.648
	α_n^+	0.973	0.086	0.994	0.271
XLK	α_n^-	0.697	0.057	0.916	0.196
	α_n^*	0.711	0.105	0.922	0.086
	α_n^+	0.741	0.114	0.988	0.350
XLP	α_n^-	0.684	0.059	0.908	0.363
	α_n^*	0.712	0.133	0.935	0.275
	α_n^+	0.754	0.146	0.962	0.173
XLU	α_n^-	0.890	0.079	0.865	0.118
	α_n^*	1.127	0.128	0.914	0.183
	α_n^+	1.192	0.135	0.974	0.073
XLV	α_n^-	0.764	0.060	0.905	0.339
	α_n^*	0.885	0.102	0.903	0.071
	α_n^+	0.973	0.105	0.999	0.819
XLY	α_n^-	0.956	0.049	0.932	0.023
	α_n^*	1.032	0.073	0.955	0.073
	α_n^+	0.998	0.077	0.962	0.018

NOTE: This table reports the results from jump regression of the listed asset against the SDPR S&P500 ETF (SPY). The data are returns on the nine SDPR sector ETFs for the year 2009. The selected jump threshold parameters are based on the estimated jumps in the SPY returns series as this is the left-hand side variable in the jump regressions. The jumps were located using one-minute returns and the jump regressions where performed using five-minute returns. The threshold α_n^* is the estimated threshold using the curvature method and α_n^+ and α_n^- are α_n^* plus and minus 15% respectively. The standard errors are calculated under the simplifying assumption that the volatility is continuous over the day. The p-values are from a bootstrap version of the determinant test in Li et al. (2017a).

---

[7] The SPY asset is sufficiently liquid to use to identify jump intervals at the 1-min level; the subsequent aggregation to 5-min returns is a correction for possible trading friction noise in the returns of the less liquid sector-specific assets.

[8] Not all years showed such evidence of constant jump betas. For the sake of exposition we do not report the results from these years since there is not as much to learn from examining the jump regression results using different jump thresholds if the jump beta is time-varying. Results for all years are available on request, however.

Table 7 reports the jump beta, the standard error of the jump beta, the R^2 of the regression, and the *p*-value of the null of a constant jump beta over the year. The *p*-values are calculated using a bootstrap version of the determinant test in Li et al. (2017a). The standard errors are calculated under a simplifying assumption that the volatility process of the diffusive moves is continuous across the market jump times; otherwise, inference becomes far more complicated but the conclusions barely changed in the end. For some of the portfolios the estimated beta seen in the table seems relatively insensitive to the 15% perturbations to α_n^*, but there are some notable exceptions. In particular, the jump beta for the XLF (Finance) portfolio is quite lower (1.182 vs. 1.687) using α_n^- versus α_n^*. The same is also true but to a lesser degree for XLK (Technology), XLU (Utilities), and XLV (Health Care). These four are economically important portfolios where the beta value matters, and one does not want a misleading estimate obtained by letting in too many diffusive moves and thereby throwing off the jump regression. At the same time, note that for all nine portfolios the estimation precision obtained with α_n^* is higher (lower standard error) than with α_n^+, which reflects of course the inclusion of the more jumps, i.e., data points.

7. Conclusions

This paper introduced a method for selecting the threshold in threshold-based jump detection schemes. Previously the selection of the threshold in such schemes has been left to each researcher in each project to choose. This creates a problem because the number of estimated jumps in a series of observed returns can vary substantially depending on which threshold a researcher selects. Our method therefore advances the existing literature on asset price jumps because it provides a method for the selection of the jump threshold. Even further, we believe researchers will find our method intuitive and easy to implement in practice.

In developing our method, we first showed that over the range of sampling frequencies a researcher is most likely to encounter that the standard in-fill asymptotics provide a poor guide for the selection of the jump threshold. Because of this we developed a sample-based method. Our method is developed as follows. Given a series of observed returns, our method relies on first estimating the number of jumps in this series over a grid of possible thresholds. Doing so results in a jump count function where the value of the function is the number of estimated jumps in the series of returns at each value of the threshold in the grid. Our method then selects the chosen threshold as the threshold for which the curvature of a suitably smoothed version of the jump count function is maximized. We think of this point as being the point were the estimated number of jumps begins to 'take-off'. We argue that selecting the threshold at this point should include many of the true jumps in the process and should guard against overly including returns that only contain diffusive moves. As the sampling size of the returns goes to zero we show that such a methodology will consistently estimate the jumps of a jump-diffusion model and asymptotically will exclude returns that only contain diffusive moves.

Having developed a methodology for selecting the threshold in threshold-based jump detection schemes we show its performance in several Monte Carlo studies and an empirical application. The Monte Carlo studies showed our method was able to recovery many of the true jumps in the data generating processes considered and maintained a high degree of accuracy in the returns it labeled as containing jumps. Further, one Monte Carlo study showed the improvement our method gave in a jump regression context. Finally, in two empirical studies we applied the method discussed to real world data. In the first empirical study we estimated the number jumps and provided the jump threshold selected by our method over a range of dates and sampling frequencies for three commonly used series in finance: the SPDR S&P500 ETF, the S&P500 E-mini futures, and the VIX futures. In the second empirical study we performed a series of jump regressions where we regressed the return intervals thought to contain jumps in the SDPR S&P500 ETF (SPY) on the corresponding return intervals in the SDPR sector ETFs using our method to select the jump times.

Appendix A. Proofs

To notational brevity below we refer to the optimally selected jump threshold as α_n, that is $\alpha_n = \arg\max \kappa(g_n(\alpha))$, rather than α_n^* as in the main text. For two positive sequences of real numbers a_n and b_n we use the following set of notations below: we write $a_n \overset{a}{\sim} b_n$ if $a_n/b_n \to 1$ as $n \to \infty$, we write $a_n = O(b_n)$ if for some $N^* \in \mathbb{N}$ there exists $c > 0$ such that for all $n > N^*$ we have $a_n \le cb_n$, we write $a_n = \Omega(b_n)$ if for some N^* there exists $c > 0$ such that for all $n > N^*$ we have $a_n \ge cb_n$, and finally we write $a_n = \Theta(b_n)$ if for some N^* there exists $c > 0$ and $d > 0$ such that for all $n > N^*$ we have $cb_n \le a_n \le db_n$.

We also drop the dependence of $\mathcal{I}_n(\alpha, \mathcal{D})$ and $\mathcal{I}(\mathcal{D})$ on the region $\mathcal{D}$ and simply write $\mathcal{I}_n(\alpha)$ and $\mathcal{I}$ instead.

Proof of Theorem 1. (a) Since the jumps of X have finite activity, we can assume without any loss of generality that each interval $((i-1)\Delta_n, i\Delta_n]$ contains at most one jump. (If not we can restrict the focus to the w.p.a.1 set of the sample paths upon which this condition holds.) We denote the continuous part of X by

$$X_t^c = X_t - \sum_{s \le t} \Delta X_s, \ t \ge 0. \tag{A1}$$

Following Li, Todorov, and Tauchen (2014) notice that $\mathcal{I}_n(\alpha)$ can be broken into two disjoint sets $\mathcal{I}_{1n}(\alpha)$ and $\mathcal{I}_{2n}(\alpha)$ defined as

$$\mathcal{I}_{1n}(\alpha) = \mathcal{I} \cap \mathcal{I}_n(\alpha) \ \text{and} \ \mathcal{I}_{2n}(\alpha) = \mathcal{I}_n(\alpha) \setminus \mathcal{I}(\alpha). \tag{A2}$$

The proof proceeds by showing

(i) $\mathbb{P}(\mathcal{I}_{1n}(\alpha_n) = \mathcal{I}) \to 1$, and
(ii) $\mathbb{P}(\mathcal{I}_{2n}(\alpha_n) = \varnothing) \to 1$.

Part (i):
Recall that $\alpha_n \in [\underline{\alpha}, \bar{\alpha}]$ with $\underline{\alpha} > 0$. By Lemma A1 below we have $\alpha_n \to \underline{\alpha}$. This implies $\alpha_n \sigma \Delta_n^\varpi \to 0$ as $n \to \infty$ (or equivalently as $\Delta_n \to 0$) since $\Delta_n^\varpi \to 0$. Following Li, Todorov, and Tauchen (2014), we notice that $\alpha_n \sigma \Delta_n^\varpi \to 0$ implies that for any $p \in \mathcal{P}$ and an $n \in \mathbb{N}$ sufficiently large that we will have $|\Delta_{i(p)}^n X| > \alpha_n \sigma \Delta_n^\varpi$. These results imply

$$\mathcal{I}_{1n}(\alpha_n) = \left\{ i(p) : p \in \mathcal{P}, ((i(p)-1)\Delta_n, \Delta_{i(p)}^n X) \right\} \ \text{w.p.a. 1}. \tag{A3}$$

Next notice that

$$\sup_{p \in \mathcal{P}} \left\| \left((i(p)-1)\Delta_n, \Delta_{i(p)}^n X\right) - (\tau_p, \Delta X_{\tau_p}) \right\| \to 0 \ \text{a.s..} \tag{A4}$$

To see this notice that, almost surely,

$$\begin{aligned}
\sup_{p \in \mathcal{P}} &\left\| \left((i(p)-1)\Delta_n, \Delta_{i(p)}^n X\right) - (\tau_p, \Delta X_{\tau_p}) \right\| \\
&= \sup_{p \in \mathcal{P}} \left\| \left((i(p)-1)\Delta_n - \tau_p, \Delta_{i(p)}^n X\right) \right\| \\
&\le \Delta_n + \sup_{s,t \le T, |s-t| \le \Delta_n} |X_t^c - X_s^c| \\
&\to 0.
\end{aligned} \tag{A5}$$

Therefore w.p.a.1 the sets $\mathcal{I}_{1n}(\alpha_n)$ and $\mathcal{I}$ will coincide.

Part (ii):

We first show a result concerning the distribution of the diffusive moves. Notice that because the diffusive moves are locally Gaussian that for a fixed $\alpha > 0$ that we have $\mathbb{P}(|\Delta X_i^c| > \alpha\sigma\Delta_n^\varpi) = 2Q(\alpha\Delta_n^{\varpi-1/2})$ where $Q(\cdot) = 1 - \Phi(\cdot)$ and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal density.

Recalling that $\alpha_n \in [\underline{\alpha}, \bar{\alpha}]$ with $\underline{\alpha} > 0$ notice that

$$
\begin{aligned}
0 &\leq \mathbb{P}(\mathcal{I}_{2n}(\alpha_n) \neq \varnothing) \\
&\leq \mathbb{P}\left(\bigcup_{i=1}^n \{|\Delta_i^n X^c| > \alpha_n\sigma\Delta_n^\varpi\}\right) \\
&\leq \mathbb{P}\left(\bigcup_{i=1}^n \{|\Delta_i^n X^c| > \underline{\alpha}\Delta_n^\varpi\}\right) \\
&\leq n\mathbb{P}(|\Delta_i^n X^c| > \underline{\alpha}\sigma\Delta_n^\varpi).
\end{aligned}
\tag{A6}
$$

Consider $n\mathbb{P}(|\Delta_i^n X| > \underline{\alpha}\sigma\Delta_n^\varpi)$. Since $\underline{\alpha}$ is non-random we know $\mathbb{P}(|\Delta_i^n X| > \underline{\alpha}\sigma\Delta_n^\varpi) = 2Q(\underline{\alpha}\Delta_n^{\varpi-1/2})$ and therefore $n\mathbb{P}(|\Delta_i^n X^c| > \underline{\alpha}\sigma\Delta_n^\varpi) = n2Q(\underline{\alpha}\Delta_n^{\varpi-1/2})$. Next, we will use the result that $Q(z) \leq \phi(z)/z$ for any $z > 0$ where $\phi(z)$ is the standard normal density. Using this result and the fact that $Q(z) > 0$ for any $z > 0$ we see

$$
\begin{aligned}
0 < n2Q(\underline{\alpha}\Delta_n^{\varpi-1/2}) &\leq n2\frac{\phi(\underline{\alpha}\Delta_n^{\varpi-1/2})}{\underline{\alpha}\Delta_n^{\varpi-1/2}} \\
&= \frac{n2}{\sqrt{2\pi}}\frac{\exp\{-(\underline{\alpha}\Delta_n^{\varpi-1/2})^2/2\}}{\underline{\alpha}\Delta_n^{\varpi-1/2}} \\
&= \frac{2}{\sqrt{2\pi}}\frac{\exp\{-(\underline{\alpha}\Delta_n^{\varpi-1/2})^2/2\}}{\underline{\alpha}\Delta_n^{\varpi+1/2}} \\
&\to 0
\end{aligned}
\tag{A7}
$$

where the convergence above follows since $\exp\{-(\underline{\alpha}\Delta_n^{\varpi-1/2})^2/2\} \to 0$ at a faster rate than $\underline{\alpha}\Delta_n^{\varpi+1/2} \to 0$. This shows $\mathbb{P}(\mathcal{I}_{2n}(\alpha_n) \neq \varnothing) \to 0$. □

(b) By part (a), it suffices to show that

$$
((i-1)\Delta_n, \Delta_i^n X)_{i\in\mathcal{I}} - (\tau_p, \Delta X_{\tau_p})_{p\in\mathcal{P}} = o_p(1).
\tag{A8}
$$

Observe that $((i-1)\Delta_n, \Delta_i^n X)_{i\in\mathcal{I}}$ is simply $((i(p)-1)\Delta_n, \Delta_{i(p)}^n X)_{p\in\mathcal{P}}$. We deduce the desired convergence by the same arguments as in part (a).

Lemma A1. *For $\alpha_n \in [\underline{\alpha}, \bar{\alpha}]$ with $\underline{\alpha} > 0$ and α_n as defined in Section 3.3, we have that $\alpha_n \to \underline{\alpha}$ at rate $n^{\varpi+1/2}e^{-n^{1-2\varpi}}$.*

Proof. Recall that $\alpha_n = \arg\max_{\alpha\in[\underline{\alpha},\bar{\alpha}]}\kappa[g_n(\alpha)]$. Where $\kappa(\cdot)$ is the curvature of a function and where $g_n(\alpha)$ is the projection of $N_n(\alpha)$ onto the set of basis functions $g(\alpha,\gamma) = \{\alpha \in [\underline{a},\bar{a}], \gamma \in \mathbb{R}^{p+1} : \gamma_0 + \gamma_1\alpha^{-1} + \cdots + \gamma_p\alpha^{-p}\}$.

Since the approximation of $N_n(\alpha)$ onto the set of basis functions $g(\alpha,\gamma)$ is found by a least squares projection of $N_n(\alpha)$ onto g we can solve for $g_n(\alpha) = \text{proj}_{N_n(\alpha)}g(\alpha,\gamma)$ analytically. To do so define $a \equiv (1, \alpha^{-1}, \ldots, \alpha^{-p})'$ so that $\gamma_n = \langle a, a'\rangle^{-1}\langle a, N_n(\alpha)\rangle$ where $\langle f, g\rangle = \int_{\underline{a}}^{\bar{a}} f(x)'g(x)dx$ is the inner

product on the space of real-valued functions on the domain $[\underline{\alpha}, \overline{\alpha}]$. Defining $A \equiv \langle a, a' \rangle^{-1}$ we can express the $(p+1) \times 1$ vector of coefficients

$$
\begin{aligned}
\gamma_n &= \langle a, a' \rangle^{-1} \langle a, N_n(\alpha) \rangle \\
&= \left[\int_{\underline{a}}^{\overline{a}} a' a \, d\alpha \right]^{-1} \left[\int_{\underline{a}}^{\overline{a}} a' N_n(\alpha) \, d\alpha \right] \\
&= A \left[\int_{\underline{a}}^{\overline{a}} a' N_n(\alpha) \, d\alpha \right].
\end{aligned}
\tag{A9}
$$

which implies each coefficient

$$
\gamma_{j,n} = \sum_{k=0}^{p} A_{j,k+1} \langle \alpha^{-k}, N_n(\alpha) \rangle.
\tag{A10}
$$

Recall from the proof of Theorem 1 that $\mathbb{P}(|\Delta X_i^c| > \alpha \sigma \Delta_n^{\varpi}) = 2Q(\alpha \Delta_n^{\varpi - 1/2})$ for a fixed $\alpha > 0$. Since the jump process in X is assumed to be of finite activity and $N_n(\alpha)$ increases without bound as $\alpha \to 0$ we can see that $N_n(\alpha) \overset{a}{\sim} N_n^c(\alpha)$. (Note the limit here is for $\alpha \to 0$ holding n fixed.) This result implies that in a neighborhood around $\alpha = 0$ that $N_n(\alpha) \overset{a}{\sim} n2Q(\alpha \Delta^{\varpi - 1/2})$ since $\mathbb{E}[N_n^c(\alpha)] = n2Q(\alpha \Delta_n^{\varpi - 1/2})$. (Where the limit is now for $n \to \infty$ holding α fixed in a neighborhood around $\alpha = 0$.) We can use this result to derive an expression for each $\langle \alpha^{-k}, N_n(\alpha) \rangle$ when α is small and n is large.[9]
For $k \geq 2$ some algebra shows

$$
\begin{aligned}
\langle \alpha^{-k}, N_n(\alpha) \rangle &= \int_{\underline{\alpha}}^{\overline{\alpha}} \alpha^{-k} N_n(\alpha) d\alpha \\
&= \frac{1}{2(k-1)} n \\
&\left[\underline{\alpha}^{1-k} - \underline{\alpha}^{1-k} \left(\mathrm{erf} \left(\frac{\underline{\alpha} n^{\frac{1}{2} - \varpi}}{\sqrt{2}} \right) \right) - \frac{2^{\frac{1}{2} - \frac{k}{2}} n^{\frac{1}{2}(k-1)(1-2\varpi)} \Gamma \left(1 - \frac{k}{2}, \frac{1}{2} \underline{\alpha}^2 n^{1-2\varpi} \right)}{\sqrt{\pi}} \right. \\
&\left. + \overline{\alpha}^{1-k} \mathrm{erf} \left(\frac{\overline{\alpha} n^{\frac{1}{2} - \varpi}}{\sqrt{2}} \right) - \overline{\alpha}^{1-k} + \frac{2^{\frac{1}{2} - \frac{k}{2}} n^{\frac{1}{2}(k-1)(1-2\varpi)} \Gamma \left(1 - \frac{k}{2}, \frac{1}{2} n^{1-2\varpi} \overline{\alpha}^2 \right)}{\sqrt{\pi}} \right]
\end{aligned}
\tag{A11}
$$

where the function $\Gamma(a, z)$ is the incomplete Gamma function. (When $k = 0$ or $k = 1$ similar expressions can be deduced.) We proceed by examining each component of (A11) in the limit in order to derive a bound on $\langle \alpha^{-k}, N_n(\alpha) \rangle$ in the limit. For any fixed $m \in \mathbb{N}$ we can express $\Gamma(a, z)$ as

$$
\Gamma(a, z) = z^{a-1} e^{-z} \left(\sum_{k=0}^{m-1} \frac{u_k}{z^k} + R_m(a, z) \right)
\tag{A12}
$$

where $u_k = (-1)^k (1-a)_k$ and $R_m(a, z) = O(z^{-m})$. (DLMF, Section 8.11(i)) If we think of $z \to \infty$ in (A12) we see $\Gamma(a, z) = \Theta(z^{a-1} e^{-z})$. This implies that $\Gamma \left(1 - \frac{k}{2}, \frac{1}{2} \alpha^2 n^{1-2\varpi} \right) = \Theta(n^{-\frac{1}{2}k(1-2\varpi)} e^{-n^{1-2\varpi}})$ for $\alpha = \underline{\alpha}, \overline{\alpha}$ in (A11) and therefore that

$$
\frac{2^{\frac{1}{2} - \frac{k}{2}} n^{\frac{1}{2}(k-1)(1-2\varpi)} \Gamma \left(1 - \frac{k}{2}, \frac{1}{2} n^{1-2\varpi} \alpha^2 \right)}{\sqrt{\pi}} = \Theta(n^{\varpi - 1/2} e^{-n^{1-2\varpi}})
\tag{A13}
$$

---

[9] We limit our scope to the case when α is small because we are primarily interested in limiting the number of misclassifications that might occur in the jump count function and, as was shown in Section 3.2, these increase exponentially as $\alpha \to 0$.

for $\alpha = \underline{\alpha}, \overline{\alpha}$ in (A11). The error function $\mathrm{erf}(z)$ can be written as (DLMF, Section 7.12)

$$\mathrm{erf}(z) = 1 - \frac{e^{-z^2}}{z\sqrt{\pi}} \sum_{m=0}^{\infty} (-1)^m \frac{(2m-1)!!}{(2z^2)^m} \tag{A14}$$

which shows $\mathrm{erf}(z) - 1 = \Theta(z^{-1}e^{-z^2})$. This implies that

$$\alpha^{1-k}\mathrm{erf}\left(\frac{\alpha n^{\frac{1}{2}-\varpi}}{\sqrt{2}}\right) - \alpha^{1-k} = \Theta(n^{\varpi-1/2}e^{-n^{1-2\varpi}}) \tag{A15}$$

as well for $\alpha = \underline{\alpha}, \overline{\alpha}$ in (A11). Combing these results we see

$$\langle \alpha^{-k}, N_n(\alpha) \rangle = \Theta(n)\left[\Theta(n^{\varpi-1/2}e^{-n^{1-2\varpi}}) + \Theta(n^{\varpi-1/2}e^{-n^{1-2\varpi}})\right.$$
$$\left. + \Theta(n^{\varpi-1/2}e^{-n^{1-2\varpi}}) + \Theta(n^{\varpi-1/2}e^{-n^{1-2\varpi}})\right] \tag{A16}$$
$$= \Theta(n^{\varpi+1/2}e^{-n^{1-2\varpi}})$$

for $k \geq 2$. Using a similar derivation as in (A11) and the arguments above it can be shown that both $\langle \alpha^0, N_n(\alpha) \rangle = \Theta(n^{\varpi+1/2}e^{-n^{1-2\varpi}})$ and $\langle \alpha^{-1}, N_n(\alpha) \rangle = \Theta(n^{\varpi+1/2}e^{-n^{1-2\varpi}})$ as well. Recall $\gamma_{j,n} = \sum_{k=0}^{p} A_{j,k+1}\langle \alpha^{-k}, N_n(\alpha) \rangle$ since $\langle \alpha^{-k}, N_n(\alpha) \rangle = \Theta(n^{\varpi+1/2}e^{-n^{1-2\varpi}})$ for all $k \geq 0$ we see $\gamma_{j,n} = \Theta(n^{\varpi+1/2}e^{-n^{1-2\varpi}})$.

Having provided rates for how the coefficients of $g_n(\alpha)$ go to zero we will use these results to describe the behavior of the the curvature function of $g_n(\alpha)$ in the limit as well. Doing so will allow us to think about how $\alpha_n = \arg\max \kappa(g_n(\alpha))$ will behave in the limit. The curvature[10] of $g_n(\alpha)$ is

$$\kappa(g_n(\alpha)) = \frac{g_n''(\alpha)}{(1 + [g_n'(\alpha)]^2)^{3/2}}. \tag{A17}$$

Since $g_n(\alpha) = \sum_{j=1}^{p} \gamma_{j,n}\alpha^{-j}$ notice that

$$g_n'(\alpha) = \sum_{j=1}^{p} -j\gamma_{j,n}\alpha^{-j-1} \tag{A18}$$

and

$$g_n''(\alpha) = \sum_{j=1}^{p} j(j+1)\gamma_{j,n}\alpha^{-j-2}. \tag{A19}$$

We showed above that for $j = 0, \ldots, p$ that each coefficient

$$\gamma_{j,n} = \Theta(n^{\varpi+1/2}e^{-n^{1-2\varpi}}) \tag{A20}$$

which implies that both

$$g_n'(\alpha) = \Theta(n^{\varpi+1/2}e^{-n^{1-2\varpi}}) \tag{A21}$$

and

$$g_n''(\alpha) = \Theta(n^{\varpi+1/2}e^{-n^{1-2\varpi}}). \tag{A22}$$

---

[10]　The equation in (A17) is actually for the signed curvature. However the basis functions $g(\alpha, \gamma)$ used here always have a positive signed curvature so that the curvature and the signed curvature coincide.

Looking at the denominator of $\kappa(g_n(\alpha))$ in (A17) we see that

$$\left(1 + [g_n'(\alpha)]^2\right)^{3/2} - 1 = \Theta(n^{3\omega+3/2}e^{-3n^{1-2\omega}}) \tag{A23}$$

and therefore that for an n sufficiently large that $\kappa(g_n(\alpha)) \overset{a}{\sim} g_n''(\alpha)$ since $\left(1 + [g_n'(\alpha)]^2\right)^{3/2} \to 1$ at a faster rate than $g_n''(\alpha) \to 0$. To see this, note that

$$\frac{g_n''(\alpha)}{\kappa(g_n(\alpha))} = \frac{\kappa(g_n(\alpha))\left(1 + [g_n'(\alpha)]^2\right)^{3/2}}{\kappa(g_n(\alpha))} = \left(1 + [g_n'(\alpha)]^2\right)^{3/2} \to 1. \tag{A24}$$

Note as well that since $\kappa(g_n(\alpha)) = g_n''(\alpha)(1 + [g_n'(\alpha)]^2)^{-3/2}$ and $g_n''(\alpha) \to 0$ at slower rate than $(1 + [g_n'(\alpha)]^2)^{-3/2} \to 1$ that $\kappa(g_n(\alpha)) \to 0$ and therefore $|\kappa(g_n(\alpha)) - g_n''(\alpha)| \to 0$.

Having concluded that $\kappa(g_n(\alpha)) \overset{a}{\sim} g_n''(\alpha)$ and $|\kappa(g_n(\alpha)) - g_n''(\alpha)| \to 0$ we can think about how α_n might behave in the limit as well. Recall that $\alpha_n \in [\underline{\alpha}, \overline{\alpha}]$ with $\underline{\alpha} > 0$. Fixing an $n > 0$ and thinking about the function $g_n''(\alpha)$ on the domain $[\underline{\alpha}, \overline{\alpha}]$ notice that for any $n > 0$ that $\arg\max_\alpha g_n''(\alpha) = \underline{\alpha}$ since $g_n''(\alpha)$ is a monotonically decreasing function of α. Since $|\kappa(g_n(\alpha)) - g_n''(\alpha)| \to 0$ and $\arg\max_\alpha g_n''(\alpha) = \underline{\alpha}$ for all $n > 0$ we see that $\alpha_n = \arg\max \kappa(g_n(\alpha)) \to \underline{\alpha}$ as well.

Having shown that $\alpha_n \to \underline{\alpha}$ we will find its rate. First, however, we need to establish a result concerning linear functions. Note that for any non-zero linear function $f(x) : \mathbb{R} \to \mathbb{R}$ the rate at which $f(x) \to f(c)$ when $x \to c$ for a constant $c \in \mathbb{R}$ will be the same as the rate that $x \to c$ because we express $f(x) = mx + b$ for some $m \neq 0$ and $b \in \mathbb{R}$. With this in mind define the function $h_n(\alpha_n) = \kappa(g_n(\alpha_n)) - g_n''(\alpha_n)$ and consider its Taylor approximation around $\alpha_n = \underline{\alpha}$ when $\underline{\alpha}$ is 'small'. That is

$$h_n(\alpha_n) = h_n(\underline{\alpha}) + h_n'(\underline{\alpha})\alpha_n + O(\alpha_n^2). \tag{A25}$$

Since $|\kappa(g_n(\alpha)) - g_n''(\alpha)| \to 0$ as $\Delta_n \to 0$, for a sufficiently large n, we will have $h_n(\underline{\alpha}) = \kappa(g_n(\underline{\alpha})) - g_n''(\underline{\alpha}) \approx 0$. Since we took the approximation around $\alpha_n = \underline{\alpha}$ and assumed $\underline{\alpha}$ was 'small' we see that the Taylor approximation error in (A25) will also be negligible compared to $h_n'(\underline{\alpha})\alpha_n$. This shows that in a sufficiently small neighborhood around $\alpha_n = \underline{\alpha}$ when $\underline{\alpha}$ is 'small' that $h_n(\alpha_n)$ will be approximately linear in α_n and therefore the rate that $\alpha_n \to \underline{\alpha}$ will be the same as the rate that $h_n(\alpha_n) \to 0$. Since $h_n(\alpha_n) = \kappa(g_n(\alpha_n)) - g_n''(\alpha_n)$ the rate that $h_n(\alpha_n) \to 0$ is given by the rate that $|\kappa(g_n(\alpha)) - g_n''(\alpha)| \to 0$. We showed earlier that $\kappa(g_n(\alpha)) \overset{a}{\sim} g_n''(\alpha)$ this implies that $\kappa(g_n(\alpha)) = \Theta(g_n''(\alpha))$ and since $g_n''(\alpha) = \Theta(n^{\omega+1/2}e^{-n^{1-2\omega}})$ we see $|\kappa(g_n(\alpha)) - g_n''(\alpha)| \to 0$ at rate $n^{\omega+1/2}e^{-n^{1-2\omega}}$. This implies finally that $\alpha_n \to \underline{\alpha}$ at rate $n^{\omega+1/2}e^{-n^{1-2\omega}}$. $\square$

Appendix B. NOTES

The entire interval $[0, nT]$ comprises nT intervals of width Δ_n. Let are J_n denote indexes (labels) of the intervals that actually contain Z jumps. Note that the labels in J_n vary with n but the cardinality $|J_n|$ does not vary with n since that is the actual (finite) number of jumps.

We need to characterize the asymptotic behavior of the increments in the Y and Z processes across both the non-jump and jump intervals. For the diffusive (non-jump) intervals, we have from Jacod and Protter (2012)

$$\begin{pmatrix} \Delta_r^n Z \\ \Delta_r^n Y^c \end{pmatrix} = \begin{pmatrix} \Delta_n^{1/2}\sigma_{z,r}\zeta_{z,r} \\ \Delta_n^{1/2}\sigma_{y,r}\zeta_{y,r} \end{pmatrix} \qquad r \notin J_n \tag{A26}$$

where $\sigma_{z,r}, \sigma_{y,r}$ are local volatilities, and $\zeta_{z,r}, \zeta_{y,r}$ are conditionally correlated Gaussian random variables with unit variance each. For the intervals that actually Z jumps we have from LTT

$$\Delta_i^n Y = \Delta Z_{t_i}\beta + \Delta_n^{1/2}\eta_i \qquad i \in J_n, \tag{A27}$$

where t_i are the jump times, ΔZ_{t_i} the actual jumps, β is the (constant) jump beta, and η_i are "mixed-normal" random variables with a relatively simple but non-stand distribution defined by the diffusive variation in Y and Z across the jump interval.

As per the main text, let $\mathcal{I}_n$ denote the intervals labeled as jumps by the jump detection scheme, where we suppress the dependence on α for now. If $J_n = \mathcal{I}_n$ then all jumps are perfectly detected, and we are in the setup of LTT, which has been covered. The following is interesting only if $J_n \cap \mathcal{I}_n^c \neq \phi$, i.e., there are diffusive intervals erroneously miss-classified as jump intervals.

Suppose across the entire jump interval

$$\Delta_{i_p}^n Y = \Delta_{i_p}^n Z \beta + \Delta_{i_p}^n e_{i_p}^c, \qquad p = 1, \ldots, \mathcal{P} \tag{A28}$$

which implicitly assumes equal beta at jump intervals and by construction e_t^c is a continuous process.

Suppose we incorrectly include R_n extra diffusive terms $\Delta_{i_r}^n Z^c$, $\Delta_{i_r}^n Y^c$, $r = 1, 2, \ldots, R_n$ into the regression. From [Jacod and Protter (2012)](#) we have that

$$\begin{aligned}
\Delta_{i_r}^n Z^c &= \Delta_n^{1/2} \sigma_{z,t_r} \zeta_{z,r} \\
\Delta_{i_r}^n Y^c &= \Delta_n^{1/2} \rho_{zy,t_r} \sigma_{y,t_r} \zeta_{z,r} + \Delta_n^{1/2} \sigma_{y,t_r} \sqrt{1 - \rho_{zy,t_r}^2} \zeta_{y,r}
\end{aligned} \tag{A29}$$

In what follows it matters how R_n grows (if at all) with n. Thus we write

$$\frac{R_n}{n} \to B, \qquad 0 \leq B \leq \infty \tag{A30}$$

The most interesting and relevant case is when $0 < B < \infty$, but considering the cases $B = 0$, $B = \infty$ provide further insights.

The jump regression estimator is

$$\widehat{\beta} = \frac{\sum_{p=1}^P \Delta_{i_p}^n Z \Delta_{i_p}^n Y + \sum_{r=1}^{R_n} \Delta_{i_r}^n Z^c \Delta_{i_r}^n Y^c}{\sum_{p=1}^P (\Delta_{i_p}^n Z)^2 + \sum_{r=1}^{R_n} (\Delta_{i_r}^n Z^c)} \tag{A31}$$

Suppose B is positive and finite. Then

$$\begin{aligned}
\sum_{r=1}^{R_n} \Delta_{i_r}^n Z^c \Delta_{i_r}^n Y^c &= \\
\sum_{r=1}^{R_n} (\Delta_{i_r}^n Z^c)^2 &= \Delta_n \sum_{r=1}^{R_n} \sigma_{z,t_r}^2 \zeta_{z,r}^2 \\
&= \frac{B}{\tilde{n}} \sum_{r=1}^{\tilde{n}} \sigma_{z,t_r}^2 \zeta_{z,r}^2
\end{aligned} \tag{A32}$$

The above is just the mean of $\tilde{n}$ random variables, where the r for the $\zeta_{z,r}^2$ is drawn from whatever probability density $\eta(r)$ governs the (incorrect) inclusion of the diffusive terms in $[0,1]$; hence,

$$\frac{B}{\tilde{n}} \sum_{r=1}^{\tilde{n}} \sigma_{z,t_r}^2 \zeta_{z,r}^2 \to BK \int_0^1 \sigma_{z,s}^2 \eta(s) ds, \tag{A33}$$

where $K = E(\zeta_s^2)$. Note that the support of $\eta(s)$ could be a strict subset of $[0,1]$, and that $\eta(s)$ will put zero mass points at the jump times t_p, $p = 1, \ldots, P$. By similar reasoning we have that

$$\sum_{r=1}^{R_n} \Delta_{i_r}^n Z^c \Delta_{i_r}^n Y^c \to BK \int_0^1 \sigma_{zy,s}^2 \eta(s) ds \tag{A34}$$

Using familiar jump regression arguments we have that

$$\sum_{p=1}^{P} \Delta_{i_p}^n Z \Delta_{i_p}^n Y = \beta \sum_{p=1}^{P} (\Delta_{i_p}^n Z)^2 + \Delta_n^{1/2} \sum_{p=1}^{P} \Delta_{i_p}^n Z \sigma_{e,t_p} \zeta_{e,r} \tag{A35}$$

Putting everything together we have that asymptotically (A31) acts as

$$\widehat{\beta} = w_{JJ}\beta + w_{dd}b + w_{JJ} \frac{\Delta_n^{1/2}}{\sum_{p=1}^{P}(\Delta_{i_p}^n Z)^2} \sum_{p=1}^{P} \Delta_{i_p}^n Z \tag{A36}$$

where $b = \sum_{r=1}^{R_n} \Delta_{i_r}^n Z^c \Delta_{i_r}^n Y^c / \sum_{r=1}^{R_n} (\Delta_{i_r}^n Z^c)^2$ is the now classical "realized beta" for diffusive regression, and

$$w_{JJ} = \frac{\sum_{p=1}^{P}(\Delta_{i_p}^n Z)^2}{\sum_{p=1}^{P}(\Delta_{i_p}^n Z)^2 + BK \int_0^1 \sigma_{z,s}^2 \eta(s) ds} \tag{A37}$$

$$w_{dd} = 1 - w_{JJ}$$

References

Andersen, Torben G., Oleg Bondarenko, Viktor Todorov, and George Tauchen. 2015. The fine structure of equity-index option dynamics. *Journal of Econometrics* 187: 532–46.

Bollen, Nicolas P. B., and Robert F. Whaley. 2015. On the Supply of and Demand for Volatility. Working paper, Vanderbilt University, Nashville, USA.

Figueroa-López, José E., and Jeffrey Nisen. 2013. Optimally thresholded realized power variations for Lévy jump diffusion models. *Stochastic Processes and their Applications* 123: 2648–77.

Jacod, Jean, and Philip E. Protter. 2012. *Discretization of Processes.* Berlin: Springer.

Li, Jia, Viktor Todorov, and George Tauchen. 2017a. Jump Regressions. *Econometrica* 85: 173–95.

Li, Jia, Viktor Todorov, and George Tauchen. 2017b. Robust Jump Regressions. *Journal of the American Statistical Association* 112: 332–41.

Mancini, Cecilia. 2001. Disentangling the Jumps of the Diffusion in a Geometric Jumping Brownian Motion. *Giornale dell'Istituto Italiano degli Attuari* LXIV: 19–47.

Mancini, Cecilia. 2004. Estimation of the characteristics of the jumps of a general Poisson-diffusion model. *Scandinavian Actuarial Journal* 1: 42–52.