Warnings about Future Jumps: Properties of the Exponential Hawkes Model

Rachele Foschi
University of Pisa, Department of Management and Economics,
Via Ridolfi 10, 56124 Pisa, Italy.
rachele.foschi@unipi.it.

Francesca Lilla
Bank of Italy, Directorate General for Economics, Statistics and Research
Via Nazionale 91, 00184 Rome, Italy.
francesca.lilla@bancaditalia.it.

Cecilia Mancini
University of Verona, Department of Economics
Via Cantarane 24, 37129 Verona, Italy.
cecilia.mancini@univr.it.

Abstract

We analyze jump risks in financial asset prices modeled by Ito semimartingales with an exponential Hawkes process as the jump counter. First, using little information, we estimate the probability that an observed jump cluster is not yet exhausted. Second, we make explicit the conditional density of consecutive jump durations and prove that durations stochastically increase. Third, we provide bounds for jump probabilities in consecutive time intervals. Application to 5-minute US returns shows that cluster depletion probabilities strongly correlate with the expected yearly jump count, and improve forecasts of jumps and realized variance.

Keywords: jumps cluster, cluster residual length, truncation, jumps risk, jumps and realized variance forecasting.
JEL-Classification: C18, C41, C58.

1 Introduction*

Many financial asset price models are based on semimartingales, which include a stochastic volatility Brownian component and a jump component. The first element accounts for the risk of continuous, predictable and usually relatively small price adjustments. The jump component, on the other hand, is related to the risk of abrupt, abnormal and possibly large fluctuations. In many cases, the latter is modelled as a Compound Poisson Process, where the durations between consecutive jumps are independent and identically distributed (iid). However, as shown for example in Maheu and McCurdy (2004) and Christensen et al. (2022), large movements may accumulate over time, forming jump clusters. For example, a market crash or crisis may be preceded by several substantial price changes and may materialize in further subsequent jumps.

Whenever a cluster of jumps is detected, a trader or market participant may wonder whether the circumstance that triggered it will generate further jumps, boosting the cluster, or, on the contrary, the price will return to its 'usual' fluctuations. Moreover, market participants may be interested in understanding how likely it is that a cluster will occur in the near future, given the past price realizations. In this framework, our paper aims to obtain finite sample quantifications of the probability of events related to the jump risk.

In order to allow for serial dependence of jumps over time, and inspired by Aït-Sahalia et al. (2015), we assume an *Exponential Hawkes* model (EHM) for the times of occurrence of jumps in the price of a financial security, and we provide closed-form formulas.

The first probability we provide is intended to inform whether a cluster is depleted. Specifically, having observed a cluster of jumps, we compute the probability that it is not yet finished but will continue, conditionally on the estimated value of the jump intensity in a past instant of time. This result is much more informative than the generic probability that a given jump is triggered by one of the previously observed jumps (see Zhuang et al., 2002). The crucial quantity is the *decay instant* of the jump intensity, which represents the residual length of a partially observed cluster of jumps.

Clearly, the probabilities we obtain are closely related to how consecutive jump durations (also called *interarrival times*) are linked. The second result is a property of strong positive dependence of the interarrival times for an EHM, called *stochastic increasingness* (Purcaru and Denuit, 2002). This property is related to the decay instant and describes precisely how a longer preceding duration is related to a longer subsequent duration.

Moreover, to obtain a qualitative insight into the possibility that observed jumps can be interpreted as the propagation of one previous event, we compute the probability that a given number of jumps occurs in small consecutive time intervals. This information can also be used to check the accuracy of the process chosen for modelling jump arrival times.

Li et al. (2014) study the *length* of a cluster of jumps within a framework of topics search tasks using an extended Hawkes model. However when we observe the occurrence of a cluster we do not know when the triggering event occurred, so information about the length does not tell us when the cluster finishes. But precisely when the cluster ends is crucial.

Reynaud-Bouret and Roy (2007) determine an upper bound for the tail of the random length of a cluster produced by a Hawkes process, and estimate the number of points of the cluster in an interval. Their results cannot be used in our framework, because their assumptions exclude the case of exponential kernel. Moreover, we are interested in determining the probabilities of more specific events, e.g., the probability that the next jump will still belong to the previous cluster and the probability of jumps occurring in consecutive 5-minute intervals.

In the insurance risk framework, the Compound Hawkes process is often used to extend the classical Compound Poisson model. For instance Swishchuk et al. (2021) give a closed-form formula for the probability that bankruptcy occurs within a fixed time horizon. While this formula can be used in our framework to compute the probability that a given number of jumps occur in a time interval, it does not allow us to compute the probability of a sequence of jumps in consecutive intervals.[1]

The above results are obtained in continuous time, a framework in which it is possible to precisely identify the position of the jumps. However, in many empirical applications (e.g. Ferriani and Zoi, 2020) the jump times are estimated from discrete price observations. In the discrete-time literature, the jump time series is often obtained by thresholding the high frequency asset returns. In this case the exact location of the jumps is not fully identified, which prevents the correct identification of the jump intensity even if the EHM parameters are known. In turn, the error in the estimation of the intensity due to the lack of information on the exact jump times can lead to biased probabilities (Hawkes, 2018).

In the context of foreign exchange markets, Rambaldi et al. (2018) consider a Hawkes model driven by a multi-exponential kernel, where some kernels give rise to intensity bursts, i.e. episodes of particularly high intensity. Given discrete observations, within each estimated jump interval the authors place the jump time randomly. In this way, they reconstruct the path realized by the jump intensity and estimate the time of occurrence of all the bursts. Unlike them, for each probability provided we measure the bias due to the lack of information on the exact jump times.

We apply our results to the analysis of 5-minute price observations of financial assets from the US stock market. For each stock, we identify the jump times by applying the threshold method proposed in Mancini (2001); Mancini (2009), and apply our formulas only to those assets for which the estimated jump times are well described by an Exponential Hawkes model (20 out of 96).[2]

If we estimated the model parameters knowing the true jump times we would be uncertain only about the reliability of the estimator, but since the jump times are estimations, there is a further source of uncertainty, which can amplify the bias in the measurement of the probabilities. However, in a simple simulation exercise, we show that the overall error is reasonable and allows us to have well-measured probabilities.

Applying our formulas to the selected stocks, we find different degrees of risk between different securities. Basically, a high jump risk is directly related to a high annual jump frequency characterizing each asset. We also find that the probability that an observed cluster of jumps has not yet ended is homogeneous within the Financial sector but varies significantly in the Healthcare sector. The stock of General Electric (GE, industrial sector) achieves the highest probability, which is close to 1. The theoretical probability of consecutive jumps fits strikingly well, especially for BHP Billiton Ltd. stock (BHP, basic material sector) and GE.

These findings are clearly useful for portfolio risk management. Consider the case in which we have to choose which of two financial securities to buy. If we observed several close jumps in the previous prices of both assets and we are worried that some further jumps will occur, we can compare the probabilities that the cluster in each asset is not yet finished, and choose to buy the less risky.

Or, if pertinent, we could make a comparison of two portfolios, portfolio 1 obtained if we include the first asset, against portfolio 2 obtained if we include the second asset: provided both portfolios displayed some previous jumps and the jump times for each portfolio are consistent with an EHM, we could use our formulas to compare the risk of a further jump in each case. More generally, we could look for an asset, within a basket of our interest, that minimizes the risk.

We then implement the proposed probabilities that observed clusters of jumps are yet to be exhausted, to improve forecasting performance for asset jumps and realized variance. In particular, we extend the HAR-J-CJ model of Corsi and Renò (2012) by introducing the interaction between these probabilities and jump terms as an additional source of information on future jumps. Inspired by the HAR-J-CJ specification, we impose a heterogeneous structure on probabilities as well, by computing the risk that a cluster of jumps observed over the previous day, week and month is not yet finished. The empirical findings indicate that jumps are predictable, as we obtain a remarkable jump forecasting performance for most of the assets considered, especially when one-step ahead forecasts are analyzed. This suggests that the probability that observed clusters of jumps are not yet exhausted incorporate important information about future jumps.

In light of these empirical results, we investigate the forecasting performance due to the interaction of the proposed risk measures with jumps, when the realized variance is used as dependent variable. We find that the probabilities we proposed also work as supplementary information for predicting realized variance.

The paper is organized as follows. Section 2 illustrates the theoretical framework, the crucial role of the jump intensity and the properties which are fundamental to our analysis. Sections 3 and 4 present our new theoretical results on the probabilistic properties of the Exponential Hawkes process and the measurements of jump risk that we propose. In Section 5 we explain how we estimate the jumps and the Hawkes model, verify the accuracy of the probabilities provided on simulated data, and implement the proposed formulas on empirical data. Section 6 presents our empirical contribution, namely the evidence for the advantage of using the probability that observed clusters of jumps are not yet exhausted in predicting asset jumps and realized variance. Section 7 concludes. The proofs of the mathematical results are set out in Appendix A and supplementary explanatory material is given in the Web Appendix *Supplementary explanatory material.*

2 Hawkes processes

A càdlàg counting process $N = \{N_t\}_{t\geq 0}$, with natural filtration $\mathcal{F} \doteq (\mathcal{F}_t)_{t\geq 0}, \mathcal{F}_t = \sigma(N_s, 0 \leq s \leq t)$, is of Hawkes type when it is *simple*[3] and has a *self exciting* stochastic intensity $\lambda = \{\lambda_t\}_{t\geq 0}$, which, for any $t \in \mathbb{R}_+$, is influenced, in a deterministic way, by the occurrence of its jumps prior to t. Namely, the intensity is a stochastic process such that for any $t \geq 0$,

$$\lambda_t = \lambda_0 + \int_0^t \Phi(t-s)dN_s = \lambda_0 + \sum_{\ell:0 \leq T_\ell < t} \Phi(t - T_\ell), \tag{1}$$

$\lambda_0 \in (0, +\infty)$ being a constant, $\Phi \geq 0$ a deterministic Lebesgue-integrable *kernel* function defined on $\mathbb{R}_+$ (Bacry et al., 2015 p.3), while T_ℓ are the random times when both N and λ jump, and are named *jump times*. Following the literature dealing with Hawkes processes, we take $\int_0^t \Phi(t-s)dN_s \doteq \int_{[0,t)} \Phi(t-s)dN_s$, thus λ is càglàd and adapted to the filtration $(\mathcal{F}_{t-})_{t>0}$, where $\mathcal{F}_{t-}$ represents the information available up to t excluding t.

We assume here $N_0 = 0$: for each path, at the first jump time T_1 of N, the intensity passes from the starting level λ_0 to $\lambda_0 + \Phi(0)$, and at the $k-th$ jump time $t = T_k$ the intensity passes from λ_t to $\lambda_t + \Phi(0)$. As soon as $\Phi(0) > 0$, at any jump time of N, the intensity path undergoes an increase and, since the paths of λ until t allow the probability of the occurrence of jumps from t onwards to be measured increased paths of λ involve a higher jump risk. This generates the self-excitation mechanism, which in fact would be guaranteed by the conditions $\Phi \geq 0, \Phi \not\equiv 0$.[4] The term *cluster of jumps*, sometimes simply abbreviated to *cluster*, indicates a generic group of close jumps. When $\Phi \equiv 0$, the self excitation is absent and the Hawkes process is reduced to a Poisson process.

Under a *branching processes* perspective, N is the superposition of a Poisson process with parameter λ_0 and the process counting the remaining jumps, which are triggered by the fact that after any previous jump the intensity increased. The jumps generated by the Poisson process are called *fathers* and, in modeling asset prices, they represent sporadic and totally independent jumps having non-persistent causes. The others are called *descendants* and represent jumps generated as consequences of more persistent phenomena.

The existence and uniqueness of a point process with intensity as in equation (1) for any t, is guaranteed, for example, in Bowsher (2007) (Theorem 3.1). The process $N = \{N_t\}_{t \geq 0}$ has asymptotically stationary increments and λ_t is asymptotically stationary, as $t \to +\infty$, if the kernel function Φ satisfies the *stability condition* $\int_0^{+\infty} \Phi(x)dx < 1$ (Clinet and Yoshida, 2017, Proposition 4.4).

In empirical applications, the kernel function Φ is generally specified either in exponential form, which is expressed as a finite sum of exponential terms, or in power form. Thanks to its mathematical tractability, the most popular Hawkes model is the *Exponential Hawkes* which has a one-term exponential kernel with the form $\Phi(x) = \alpha e^{-\beta x}$, with $\alpha, \beta > 0$. We focus on this latter kernel, which fits our samples quite well

If we were able to observe the process N continuously until a point in time t, we would know exactly when the jump times before t occurred. If we also knew the parameters of the process we would know the path realized by λ up to $t-$. In our application we are not able to directly observe N, but we are able to estimate the values it takes on at specified instants of time. Under the assumption of an exponential kernel, this allows us to estimate the model parameters. In Sections 4 and 5, we work under different information sets on the realized path of process N. In any case, we assume that the model parameters λ_0, α and β are known.

At each time t, the jump intensity has the property $\frac{1}{\delta}P\{N_{t+\delta} - N_t = 1|\mathcal{F}_{t-}\} \xrightarrow{P} \lambda_t$, as $\delta \to 0$ (Aït-Sahalia et al., 2015, second formula in (2.1)). Thus, for fixed $\delta > 0$, $\lambda_t \delta$ approximates, all the better the smaller δ, the probability $P\{N_{t+\delta} - N_t = 1|\mathcal{F}_{t-}\}$ that, given the state of the system up to time $t-$, a jump may occur within the time interval $(t, t+\delta]$.[5] This approximation quantifies the risk of a jump in terms of λ_t, and shows how crucial knowing the magnitude of the intensity is, in order to measure the jump risk.

Since in our empirical application we do not know the precise value λ_t, we start showing bounds for the possible levels it can reach. Note that the contribution to λ_t from any jump occurred at a time $T_\ell \leq t$ is $\alpha e^{-\beta(t-T_\ell)}$. If we know that $T_\ell \in [a, b]$, then we obtain the following simple but very useful bounds:

$$\alpha e^{-\beta(t-a)} \leq \alpha e^{-\beta(t-T_\ell)} \leq \alpha e^{-\beta(t-b)} \tag{2}$$

For instance, the first consequence is that, if we have at our disposal only the values of λ_0, α, β and the *number* of jumps of N occurred on $[0, t)$, without knowing the exact jump times locations, the best possible bounds for λ_t are as follows: a.s. $\lambda_0 + \alpha n e^{-\beta t} < \lambda_t < \lambda_0 + \alpha n$.

The value of λ_t crucially depends on how the jump times are distributed on $[0, t)$, thus if we want narrower bounds we need more information about the jumps locations. The next lemma concerns two specific cases: the case in which the jump times all accumulate close to 0 and the case in which all accumulate close to t.

Lemma 1. For any $\eta > 0$,

1. conditionally on $\{N_{t-} = n, T_i < \eta \ \ \forall i = 1, ..., n\}$, then a.s. $\lambda_0 + \alpha n e^{-\beta t} < \lambda_t < \lambda_0 + \alpha n e^{-\beta(t-\eta)}$;
2. conditionally on $\{N_{t-} = n, T_i > t - \eta \ \ \forall i = 1, ..., n\}$, we have a.s. $\lambda_0 + \alpha n e^{-\beta\eta} < \lambda_t < \lambda_0 + \alpha n$.

We conclude this section by showing a recursive formula which relates the value λ_t to the value λ_s at a previous time $s < t$, with the model parameters and with the jump times T_ℓ between s and t. This property and (2) are basic but crucial tools for all our results.

Lemma 2. For the Exponential Hawkes model with kernel $\Phi(x) = \alpha e^{-\beta x}$, for any t and s with $0 \leq s < t$ we have

$$\lambda_t = \lambda_0 + [\lambda_s - \lambda_0]e^{-\beta(t-s)} + \sum_{T_\ell \in [s,t)} \alpha e^{-\beta(t-T_\ell)}. \tag{3}$$

In practice, we are able to estimate λ_s, the parameters of the model and the number of jumps that occurred on $[s, t)$. With this information, in Section 3, jointly using (2) and Lemma 2, we produce finer bounds for the values of λ_t. From this we obtain a measure of the probability that a cluster of jumps has not yet finished at time t. Using additional, still partial, information about the jump locations within specified time intervals, in Section 4 we further refine the bounds for λ_t and obtain a measure of the probability of a jump in the near future and of the probability that k *consecutive* jumps will occur. Proofs of all theoretical results are provided in Appendix A.

3 Probability that a cluster of jumps is going to produce a further jump

In Section 3.1 we assume that up to time $\bar{s}$ we detected the occurrence of a cluster of jumps from process N, and we want to assess whether it has finished or is going to continue. We characterize the occurrence of a cluster, within a time interval, in terms of the values assumed by the intensity in that interval and define the *decay instant,* essentially representing the residual length of the cluster starting from $\bar{s}$. Specifically, we assume what follows: to know the kernel parameters; to have a realistic idea of the value that the intensity has taken in a specified past moment, called $\underline{s}$; and to know that k jumps have occurred in the time interval $[\underline{s}, \bar{s}]$, whereas we do not know the exact jump time locations within $[\underline{s}, \bar{s}]$. With this information we show bounds for the possible values of $\lambda_{\bar{s}}$. The bounds allow us to measure the probability that the next jump will arrive before the end of the cluster, which is measured through the residual length. That is, we estimate the probability that the cluster has not yet finished.

The residual length of the cluster is related to the duration, or *interarrival time*, between two consecutive jumps of the process N. In Section 3.2 we show a specific property of the interarrival times for an Exponential Hawkes model allowing us to better understand the decay instant.

3.1 Decay instant and cluster size

We begin by formalizing the notion of a cluster of jumps in terms of the values assumed by the intensity. Note that, from (1), for any $t > 0$ we have $\lambda_t \geq \lambda_0$, in fact at each jump arrival any path of process λ jumps up by α and then starts to decrease at the exponential rate β. Even if no other jumps occurred, after the first jump the intensity could never again reach the value λ_0, whereas as soon as a new jump occurs, λ jumps up by another α and goes even further away from λ_0. Thus the occurrence of several close jumps causes the intensity to take on high values.

We identify as a *father* a jump occurring when the value of λ_t is close to λ_0 with an accuracy of ϵ, i.e. $\lambda_t < \lambda_0(1+\epsilon)$, where $\epsilon > 0$ is small. We define a *cluster* as the set of all the jumps occurring after a father and before the subsequent time t when λ_t drops below $\lambda_0(1+\epsilon)$ again. We thus consider the jumps within a cluster as (direct or indirect) descendants of a father jump that occurred at some previous point in time, say u. We are interested in understanding *for how long* the event that occurred at u can produce consequences. If a jump, occurring at a time $T > u$, belongs to the cluster then its descendants are also descendants of the father event that occurred at u and the cluster triggered by the father event is further extended, in other words the consequences of the event in u have not yet been exhausted. When on the other hand after time u the intensity is again close to λ_0 (i.e. below $\lambda_0(1+\epsilon)$) then no events of the cluster produced descendants anymore, meaning that the father event has already been absorbed by the market.

Let us now consider an observation time $\bar{s}$. We first illustrate the case in which $\bar{s}$ is not a time of jump for the specific path at hand, and recall that $\lambda_0 > 0$. If $\lambda_{\bar{s}} > \lambda_0(1+\epsilon)$ we define as *decay instant* the quantity

$$t_\epsilon(\bar{s}) = \min\{v > 0 |\ \lambda_0 + (\lambda_{\bar{s}} - \lambda_0)e^{-\beta v} < \lambda_0(1+\epsilon)\},$$

which measures the amount of time we would have to wait after $\bar{s}$ for the intensity to again fall below $\lambda_0(1+\epsilon)$, *if* no jumps occurred after $\bar{s}$. Indeed, by the recursive formula given in Lemma 2 (apply (28) in Appendix A with $s = \bar{s}$ and $T_i \to +\infty$), for all times $t = \bar{s} + v$, with $v \geq 0$, we have $\lambda_0 + (\lambda_{\bar{s}} - \lambda_0)e^{-\beta v} =$

$\lambda_{\bar{s}+v}$. It turns out that

$$t_\epsilon(\bar{s}) = \frac{1}{\beta} \log\left(\frac{\lambda_{\bar{s}} - \lambda_0}{\epsilon\lambda_0}\right). \tag{4}$$

The definition of decay instant when $\bar{s}$ is a jump time is similar, as follows. If $\lambda_{\bar{s}}+\alpha > \lambda_0(1+\epsilon)$ the *decay instant* is defined by $t_\epsilon(\bar{s}) := \min\{v > 0 | \ \lambda_0 + (\lambda_{\bar{s}} - \lambda_0 + \alpha)e^{-\beta v} < \lambda_0(1+\epsilon)\}$, and has an interpretation similar to that of the previous case, since (apply (27) with $\bar{s} = T_{i-1}$ and $T_i \to +\infty$), for all the times $t = \bar{s} + v$, with $v \geq 0$, we have $\lambda_0 + (\lambda_{\bar{s}} - \lambda_0 + \alpha)e^{-\beta v} = \lambda_{\bar{s}+v}$. In this case $t_\epsilon(\bar{s}) = \frac{1}{\beta}\log\left(\frac{\lambda_{\bar{s}} - \lambda_0 + \alpha}{\epsilon\lambda_0}\right)$.

Consider again a no jump time $\bar{s}$ and assume that before it we have detected the arrival of a cluster of k jumps from N, say $T_i, ..., T_{i+k-1}$ are the last k jump times before $\bar{s}$, and assume $\lambda_{\bar{s}} > \lambda_0(1+\epsilon)$. The k jumps already took place, however we do not know when. By our definition of cluster, regardless of whether the father occurred at or before T_i, the intensity is larger than $\lambda_0(1+\epsilon)$ during the whole period $[T_i, \bar{s}]$. Since T_{i+k} denotes the unknown time of the first jump after $\bar{s}$, we know that T_{i+k} belongs to the same cluster if and only if for all $t = \bar{s} + v \in [\bar{s}, T_{i+k}]$ we still have $\lambda_t > \lambda_0(1+\epsilon)$, that is $\lambda_0 + (\lambda_{\bar{s}} - \lambda_0)e^{-\beta v} = \lambda_{\bar{s}+v} > \lambda_0(1+\epsilon)$, for all $v \in [0, T_{i+k} - \bar{s}]$. This happens if and only if up to the time T_{i+k} the intensity never falls below the level $\lambda_0(1+\epsilon)$, that is $T_{i+k} < \bar{s} + t_\epsilon(\bar{s})$. This allows us to interpret $t_\epsilon(\bar{s})$ as the *residual length* (after $\bar{s}$) of the cluster. More generally, in both cases when $\bar{s}$ is a jump instant and when it is not, the probability that the subsequent jump is still a direct or indirect effect of the father event is $P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s})\}$.

Note that the event 'the $(i+k)$-th jump belongs to the previous cluster' is radically different from the event 'the jump occurred at T_{i+k} is a descendant'. See a detailed explanation in Remark 1 in the Web Appendix.

It is particularly useful to exploit the information available, or estimatable, at the moment in which we want to compute $P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s})\}$. One possibility is to look at $P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s}) \mid \lambda_{\bar{s}}\}$, because $\lambda_{\bar{s}}$ contains the entire past history of the jump counter, including the timing of the last k jumps. This is consistent with the fact that we can characterize the occurrence of a cluster through the magnitude of the jump intensity. The conditional probability $P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s}) | \lambda_{\bar{s}}\}$ is available in closed form. For the case $\bar{s} \in (T_{i+k-1}, T_{i+k})$ it is stated below.

Proposition 1. Let $\bar{s} \in (T_{i+k-1}, T_{i+k})$ and $\epsilon > 0$ such that $\lambda_{\bar{s}} > \lambda_0(1+\epsilon)$, then

$$P \doteq P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s}) \mid \lambda_{\bar{s}}\} = 1 - e^{-\frac{\lambda_{\bar{s}} - \lambda_0(1+\epsilon)}{\beta}} \left(\frac{\lambda_0 \epsilon}{\lambda_{\bar{s}} - \lambda_0}\right)^{\frac{\lambda_0}{\beta}}. \tag{5}$$

Note that, under the hypothesis $\lambda_{\bar{s}} > \lambda_0(1+\epsilon)$, we have $\lambda_{\bar{s}} > \lambda_0$, so that $P < 1$. Further, we have both $e^{-\frac{\lambda_{\bar{s}}-\lambda_0(1+\epsilon)}{\beta}} < 1$ and $\frac{\lambda_0\epsilon}{\lambda_{\bar{s}}-\lambda_0} < 1$, so $P > 0$.

From (5) we see that, if we knew the value $\lambda_{\bar{s}}$, as well as the parameters (λ_0, β), we could compute the probability of our interest, i.e. the probability that the cluster observed up to $\bar{s}$ is not yet finished. However, we assumed we only knew the value $\lambda_{\underline{s}}$ of the intensity at a time $\underline{s} < \bar{s}$ and the number of jumps occurred within $[\underline{s}, \bar{s}]$, rather than knowing $\lambda_{\bar{s}}$. This assumption is justified by the following practical reason. Since $\bar{s}$ comes after several jumps, in practice the estimation of the value of λ at $\bar{s}$ is less reliable than the estimation of the value at a properly chosen previous time belonging to a more tranquil period. For instance we can choose a time $\underline{s}$ when no turbulence affected the market, so that we can attribute to $\lambda_{\underline{s}}$ a value close to λ_0 (however greater than $(1+\epsilon)\lambda_0$). When $\underline{s}$ has been chosen we can estimate the number of jumps that occurred in the time interval $[\underline{s}, \bar{s}]$ as explained in our empirical application. For the sake of simplicity we once again call this number k.

Even without knowing the jump time locations within $[\underline{s}, \bar{s}]$, the information assumed is sufficient to provide bounds for $t_\epsilon(\bar{s})$ allowing for an approximation of $P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s}) \mid \lambda_{\bar{s}}\}$.

Recall that $\bar{s} \in (T_{i+k-1}, T_{i+k})$ (the case $\bar{s} = T_{i+k-1}$ being analogous). Denoting $\bar{s} - \underline{s} = d$ (as *distance*), using Lemma 2 and (2), we obtain (see the details in Appendix A) that $t_\epsilon(\bar{s}) \in (\underline{t}_\epsilon, \bar{t}_\epsilon)$, where

$$\underline{t}_\epsilon = \max\left\{\frac{1}{\beta}\log\left(\frac{(\lambda_{\underline{s}} - \lambda_0 + k\alpha)e^{-\beta d}}{\lambda_0\epsilon}\right), 0\right\},\ \bar{t}_\epsilon = \frac{1}{\beta}\log\left(\frac{(\lambda_{\underline{s}} - \lambda_0)e^{-\beta d} + k\alpha}{\lambda_0\epsilon}\right) \tag{6}$$

The decay instant $t_\epsilon(\bar{s})$ coincides with the lower bound $\underline{t}_\epsilon$ in the limit case where all the k jumps of the cluster occurred at $\underline{s}$, whereas $t_\epsilon(\bar{s})$ coincides with the upper bound $\bar{t}_\epsilon$ in the limit case in which all the k jumps occurred at $\bar{s}$. It follows that, conditioning on (the sigma-algebra generated by) any information set $\mathcal{G}$, almost surely

$$P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s})|\mathcal{G}\} \in \Big(P\{T_{i+k} < \bar{s} + \underline{t}_\epsilon|\mathcal{G}\}\ ,\ P\{T_{i+k} < \bar{s} + \bar{t}_\epsilon|\mathcal{G}\}\Big) \tag{7}$$

The lower bound for the conditional probability corresponds to attributing to the cluster a shorter residual length compared to the theoretical length, and this involves considering that the cluster will produce a lower number of jumps after $\bar{s}$. A symmetric consideration holds for the upper bound. The next Theorem makes the bounds in (7) explicit when the information set is $\mathcal{G} \doteq \{N_{\bar{s}} - N_{\underline{s}} = k, \lambda_{\underline{s}}\}$.

Theorem 1. Let $\bar{s} \in (T_{i+k-1}, T_{i+k})$ such that $\lambda_{\bar{s}} > \lambda_0(1+\epsilon)$ then, with $\mathcal{G} \doteq \{N_{\bar{s}} - N_{\underline{s}} = k, \lambda_{\underline{s}}\}$, we have

$$LB \doteq P\{T_{i+k} < \bar{s} + \underline{t}_\epsilon | \mathcal{G}\} = 1 - e^{-\frac{1}{\beta}\left(e^{-\beta d}(\alpha k + \lambda_{\underline{s}} - \lambda_0) - \lambda_0 \epsilon\right)} \left((\alpha k + \lambda_{\underline{s}} - \lambda_0)\frac{e^{-\beta d}}{\lambda_0 \epsilon}\right)^{-\frac{\lambda_0}{\beta}} \tag{8}$$

$$UB \doteq P\{T_{i+k} < \bar{s} + \bar{t}_\epsilon | \mathcal{G}\} = 1 - e^{-\frac{1}{\beta}\left(\alpha k + e^{-\beta d}(\lambda_{\underline{s}} - \lambda_0) - \lambda_0 \epsilon\right)} \left((e^{\beta d}\alpha k + \lambda_{\underline{s}} - \lambda_0)\frac{e^{-\beta d}}{\lambda_0 \epsilon}\right)^{-\frac{\lambda_0}{\beta}} \tag{9}$$

and the two bounds LB and UB are strictly decreasing in d and increasing in k.

Theorem 1 makes the dependence of LB and UB on $d = \bar{s} - \underline{s}$ and k explicit. For instance, given k, for a fixed jump time configuration on $[\underline{s}, \bar{s}]$, if $d = \bar{s} - \underline{s}$ increases then $[\underline{s}, \bar{s}]$ is wider, and the contribution to λ from the k jumps of the cluster has more time to decay. This means that the level reached by λ at time $\bar{s}$ is lower, and, from (4), the residual length $t_\epsilon(\bar{s})$ of the cluster is shorter. Also the bounds $\underline{t}_\epsilon$ and $\bar{t}_\epsilon$ for the residual cluster length are lower, so $(\bar{s}, \bar{s} + \underline{t}_\epsilon)$ and $(\bar{s}, \bar{s} + \bar{t}_\epsilon)$ are shorter, and the probabilities of a jump on $(\bar{s}, \bar{s} + \underline{t}_\epsilon)$ or $(\bar{s}, \bar{s} + \bar{t}_\epsilon)$ are lower.

By contrast, for a fixed d, an increase in the number k of elements in the cluster involves further contributions α to the intensity process within the period $[\underline{s}, \bar{s}]$. In particular $\lambda_{\bar{s}}$ is higher, so the bounds $\underline{t}_\epsilon, \bar{t}_\epsilon$ are larger and the probabilities of a further jump in $(\bar{s}, \bar{s} + \underline{t}_\epsilon)$ or $(\bar{s}, \bar{s} + \bar{t}_\epsilon)$ are higher.
Note that by taking $\bar{s}$ sufficiently close to the estimated time $\hat{T}_{i+k-1}$ of the last jump in the detected cluster before $\bar{s}$, and taking sufficiently small ϵ, the condition $\lambda_{\bar{s}} > \lambda_0(1+\epsilon)$ is satisfied.

3.2 Durations between consecutive jump times

For the Exponential Hawkes process, the decay instant, and thus the probability that an observed cluster is yet to be exhausted, is linked to a property that we illustrate in this subsection, the *stochastic increasingness* between consecutive jump interarrival times. Intuitively, the longer we have to wait for the arrival of the n-th jump, the longer we will have to wait for the arrival of the $(n+1)$-th jump. Theorem 2 establishes the property and Proposition 2 is crucial for its proof. We conclude the subsection by showing the relation to the decay instant.

Definition 1. (Joe, 1997) Let X, Y be two random variables. Y is stochastically increasing (resp. decreasing) in X if, for any y, the conditional law $P\{Y > y | X = x\}$ is increasing (resp. decreasing) in x.

We abbreviate the term stochastically increasing with SI.

The next result clarifies how the interarrival time $T_{n+1} - T_n$ depends on the past history through λ_{T_n}.

Proposition 2. For any $\ell \geq \lambda_0$, conditionally on $\lambda_{T_n} = \ell$, the density and the distribution function of the interarrival time $T_{n+1} - T_n$ are, respectively, given by

$$P\{T_{n+1} - T_n = t | \lambda_{T_n} = \ell\} = \Big[\lambda_0 + (\ell - \lambda_0 + \alpha) e^{-\beta t}\Big] e^{(e^{-\beta t}-1)(\ell-\lambda_0+\alpha)/\beta - \lambda_0 t} \tag{10}$$

$$P\{T_{n+1} - T_n > \tau | \lambda_{T_n} = \ell\} = e^{\left[e^{-\beta\tau}(\ell+\alpha-\lambda_0) - \beta\lambda_0\tau - \ell - \alpha + \lambda_0\right]/\beta} \tag{11}$$

A non-obvious consequence of the Proposition is that $T_{n+1} - T_n$ is stochastically decreasing in λ_{T_n}, see Section 1.2 in the Web Appendix. The following Theorem is the main result of this subsection.

Theorem 2. *Conditionally on $\lambda_{T_{n-1}}$, $T_{n+1} - T_n$ is stochastically increasing in $T_n - T_{n-1}$.*

In particular, as displayed in (36) in Appendix A, this result formalizes how the interarrival jump times are, conditionally, dependent on each other. As is well-known, if we specify λ by taking either $\alpha = 0$ or $\beta \to +\infty$, we describe a simple Poisson process. Consistently, starting from the conditional distributions of the interarrival times displayed in (37), we reach the same conclusion.

The connection of the SI property with the decay instant is that, given s and τ, then the probability $P\{T_{n+1} - T_n > \tau | T_n - T_{n-1} = s, \lambda_{T_{n-1}}\}$, given in (37), is decreasing in $t_\epsilon(T_{n-1})$. In fact, by abbreviating $t_\epsilon(T_{n-1}) = \frac{1}{\beta}\log\left(\frac{\lambda_{T_{n-1}} - \lambda_0 + \alpha}{\lambda_0 \epsilon}\right)$ by t_ϵ, we obtain

$$P\{T_{n+1} - T_n > \tau | T_n - T_{n-1} = s, \lambda_{T_{n-1}}\} = e^{-\frac{1-e^{-\beta\tau}}{\beta}(\lambda_0\epsilon \cdot e^{\beta\, t_\epsilon}\, e^{-\beta s} + \alpha) - \lambda_0 \tau}.$$

This means that, if up to T_{n-1} we observed a cluster, a bigger $t_\epsilon(T_{n-1})$ corresponds to a furthest end, and thus to a higher probability that T_n and T_{n+1} belong to the same cluster as T_{n-1}. Since the jumps of a cluster tend to be close, the conditional probability that $T_{n+1} - T_n > \tau$ has to be smaller.

4 Evenly spaced high frequency observations

In this section we assume that we have access to richer information about the path realized by process N. For a fixed $\delta > 0$ we consider a discrete time grid $t_0 = 0 < t_1 ... < t_{m-1} < t_m$ of evenly spaced times

$t_i = i\delta$ for all $i = 1, ..., m$, and assume that we know in which small intervals $(t_{i-1}, t_i]$ the process N made a jump and the parameters of the exponential Hawkes process. Since δ is fixed, knowing that N jumped within $(t_{i-1}, t_i]$ does not give us the exact jump location, and in our risk measures we explicitly account for this uncertainty. By contrast we neglect other possible, and less probable, sources of error in estimating the times of the jumps that have occurred. For instance, after identifying a jump interval, we register the occurrence of only one event, whereas two or more distinct jumps could have occurred in that interval. Since the Hawkes process is simple and with finite activity jumps (meaning that a.s. on any finite time interval only finitely many jumps occur), then almost surely, for sufficiently small δ, in each interval $(t_{i-1}, t_i]$ at most one jump can occur, so there is in fact no error.

In this framework, given the information assumed in this section up to an instant of time t_n, we get bounds for the possible values of λ_{t_n}. Now the bounds are more complicated but more precise, and allow us to measure the probability of occurrence for the specific type of cluster consisting of *consecutive jumps,* i.e. a set of k jumps that fall in different, but consecutive, k intervals $[t_{i-1}, t_i)$.
Note that, since the compensator $\left\{ \int_0^t \lambda_s ds \right\}_{t \geq 0}$ of $\{N_t\}_{t\geq 0}$ is absolutely continuous with respect to the Lebesgue measure, then the process N has no fixed times of discontinuity, i.e. $P\{N_t - N_{t-} = 1 | \mathcal{F}_{t-}\} = 0$, thus $P\{N_{t+\delta} - N_t = 1 | \mathcal{F}_{t-}\} = P\{\text{one jump in } (t, t+\delta] | \mathcal{F}_{t-}\}$ coincides with $P\{\text{one jump in } [t, t+\delta) | \mathcal{F}_{t-}\}$. Since we are going to condition on $\mathcal{F}_{t_n -}$, it is more natural to wonder whether a jump will occur in $[t_n, t_{n+1})$ or k jumps occur in consecutive $[t_n, t_{n+1}), ..., [t_{n+k}, t_{n+k+1})$.

In Section 4.1 we assess the conditional probability, or $\tilde{\lambda}_n$, of the occurrence of a jump in the next interval $[t_n, t_{n+1})$. We obtain Theorem 3 and (15), which deliver feasible recursive formulas for $\tilde{\lambda}_n$, given either $\tilde{\lambda}_{n-1}$ or $\tilde{\lambda}_{n-\ell}$ for some integer $\ell > 1$. In Section 4.2 (through (12) below and the first of the two recursive formulas) we obtain, in (18), a measure of the conditional probability $\tilde{\lambda}_n^{(k)}$ that a cluster of consecutive jumps occurs in the near future. An immediate consequence is to assess the probability of the occurrence of a further consecutive jump after the cluster.

4.1 Probability of a jump in the next time interval

The quantity we want to estimate in this subsection is, for fixed δ and n,

$$\tilde{\lambda}_n \doteq P\{ \text{ one jump occurs in } [n\delta, (n+1)\delta) \,|\, \mathcal{F}_{n\delta-} \ \}.$$

Consistently with the Markov property of the couple (N, λ) for an Exponential Hawkes process, we obtain the following exact expression for $\tilde{\lambda}_n$, whose randomness only depends on the state $\lambda_{n\delta}$ of the intensity:

$$\tilde{\lambda}_n = \int_{n\delta}^{(n+1)\delta} \lambda_{s|\mathcal{F}_{n\delta-},NJ_{[n\delta,s)}} e^{-\int_{n\delta}^{s} \lambda_{r|\mathcal{F}_{n\delta-},NJ_{[n\delta,r)}} dr} ds = 1 - \exp\left(\frac{1}{\beta}(\lambda_{n\delta} - \lambda_0)(e^{-\beta\delta} - 1) - \delta\lambda_0\right) \quad (12)$$

where for any two points in time r, s we denote $NJ_{[s,r)} \doteq \{\text{no jumps in } [s,r)\}$, and for any set A the notation $\lambda_{r|\mathcal{F}_{n\delta-},A}$ means $E[\lambda_r|\,\mathcal{F}_{n\delta-}, A]$. Expression (12) makes the relation $\tilde{\lambda}_n \approx \delta\lambda_{n\delta}$ precise. In particular, $\tilde{\lambda}_0 = 1 - e^{-\delta\lambda_0}$. However the computation of the right-hand term is not feasible, because we do not know the value of $\lambda_{n\delta}$, due to the uncertainty regarding the past jump locations. Thus Theorem 3 provides a recursive formula where $\tilde{\lambda}_n$ is approximated in terms of $\tilde{\lambda}_{n-1}$. Given that we know the exact expression for $\tilde{\lambda}_0$, once we have estimated the parameters of the model, by iterating the formula we can obtain a clear quantification of $\tilde{\lambda}_n$ for any n.

To obtain the result of Theorem 3 we proceed as follows: through relation (12) we express $\lambda_{n\delta}$ in terms of $\tilde{\lambda}_n$, then we relate $\lambda_{n\delta}$ to $\lambda_{(n-1)\delta}$ using (13) below, and then use again (12) to connect $\lambda_{(n-1)\delta}$ and $\tilde{\lambda}_{n-1}$. In detail, defined

$$\omega_i = \mathbf{1}_{\{\text{one jump occurs in } [i\delta,(i+1)\delta)\}}, \; i = n, n+1, ...,$$

we have

$$\tilde{\lambda}_n = P\{\omega_n = 1|\mathcal{F}_{n\delta-}\},$$

and relation (13) is obtained using (3) and (2), as follows. Given $\lambda_{(n-1)\delta}$: if $\omega_{n-1} = 1$ then the recursive formula (3) with $s = (n-1)\delta$ and $t = n\delta$ delivers $\lambda_{n\delta} = \lambda_0 + [\lambda_{(n-1)\delta} - \lambda_0]e^{-\beta\delta} + \alpha e^{-\beta(n\delta - T_\ell)}$, where $T_\ell \in [(n-1)\delta, n\delta)$. If instead $\omega_{n-1} = 0$ then we simply have $\lambda_{n\delta} = \lambda_0 + [\lambda_{(n-1)\delta} - \lambda_0]e^{-\beta\delta}$. That is,

$$\lambda_{n\delta} \begin{cases} \in \big[\lambda_0 + (\lambda_{(n-1)\delta} - \lambda_0 + \alpha)e^{-\beta\delta}, \lambda_0 + (\lambda_{(n-1)\delta} - \lambda_0)e^{-\beta\delta} + \alpha\big) & \text{if } \omega_{n-1} = 1 \\ = \lambda_0 + (\lambda_{(n-1)\delta} - \lambda_0)e^{-\beta\delta} & \text{if } \omega_{n-1} = 0 \end{cases} \quad (13)$$

In particular, if within $\big[(n-1)\delta, n\delta\big)$ no jumps occurred, then we have an exact expression for $\lambda_{n\delta}$, given $\lambda_{(n-1)\delta}$. On the other hand, when a jump occurred then the infimum of the interval in the first row of (13) is the value that the intensity would have achieved if the jump had occurred at $(n-1)\delta$, whereas the supremum would be the value if the jump occurred at time $n\delta$. When $\omega_{n-1} = 1$ the approximation of $\lambda_{n\delta}$ is given with an absolute error smaller than $\alpha(1 - e^{-\beta\delta}) \approx \alpha\beta\delta$.

Theorem 3 is the main result of this subsection (see the proof in Section A.4): given $\tilde{\lambda}_{n-1}$, it provides forecasting of a jump in $[n\delta,(n+1)\delta)$. If we know that the model is an Exponential Hawkes process, then given discrete observations of N this forecast is feasible.

Theorem 3. Given $\tilde{\lambda}_{n-1}$ and ω_{n-1},

$$\begin{aligned}\tilde{\lambda}_n \quad \in \quad & \Big[\ 1-(1-\tilde{\lambda}_{n-1})^{e^{-\beta\delta}} \exp\Big((e^{-\beta\delta}-1)\Big(\delta\lambda_0+\omega_{n-1}\cdot\tfrac{\alpha e^{-\beta\delta}}{\beta}\Big)\Big)\ , \\ & 1-(1-\tilde{\lambda}_{n-1})^{e^{-\beta\delta}} \exp\Big((e^{-\beta\delta}-1)\Big(\delta\lambda_0+\omega_{n-1}\cdot\tfrac{\alpha}{\beta}\Big)\Big)\ \Big)\end{aligned} \tag{14}$$

If $\omega_{n-1}=0$ then the interval is reduced to the sole value $1-(1-\tilde{\lambda}_{n-1})^{e^{-\beta\delta}}\cdot e^{(e^{-\beta\delta}-1)\delta\lambda_0}$ and $\tilde{\lambda}_n$ is determined without error. Otherwise, if $\omega_{n-1}=1$ we have some uncertainty about the exact value of $\tilde{\lambda}_n$, due to the fact that we do not know precisely the location of the jump that occurred in $[(n-1)\delta,n\delta)$. We can estimate $\tilde{\lambda}_n$ using any value within the interval given in (14), and we know the size of the estimation error. If we recursively estimate $\tilde{\lambda}_j$, for $j=n,...,1$, the approximation error is amplified at each step, if the number of jumps that occurred within $[0,n\delta)$ is higher and/or if the jumps are closer to each other. Conversely, the error is null if no jumps occurred. [6]

By applying ℓ times the recursion given in Theorem 3, we obtain the following closed expression of $\tilde{\lambda}_n$ in terms of $\tilde{\lambda}_{n-\ell}$; $\omega_{n-\ell},\omega_{n-\ell-1},...,\omega_{n-1}$; and the model parameters:

$$\begin{aligned}\tilde{\lambda}_n \in \Big[\ & 1-(1-\tilde{\lambda}_{n-\ell})^{e^{-\ell\beta\delta}} e^{(e^{-\beta\delta}-1)(\delta\lambda_0+\frac{\alpha e^{-\beta\delta}}{\beta})\sum_{i=0}^{\ell-1}\omega_{n-\ell+i}e^{-i\beta\delta}} \cdot e^{(e^{-\beta\delta}-1)\delta\lambda_0\sum_{i=0}^{\ell-1}(1-\omega_{n-\ell+i})e^{-i\beta\delta}}, \\ & 1-(1-\tilde{\lambda}_{n-\ell})^{e^{-\ell\beta\delta}} e^{(e^{-\beta\delta}-1)(\delta\lambda_0+\frac{\alpha}{\beta})\sum_{i=0}^{\ell-1}\omega_{n-\ell+i}e^{-i\beta\delta}} \cdot e^{(e^{-\beta\delta}-1)\delta\lambda_0\sum_{i=0}^{\ell-1}(1-\omega_{n-\ell+i})e^{-i\beta\delta}}\ \Big)\end{aligned} \tag{15}$$

4.2 Probability of a cluster of consecutive jumps and of a jump immediately afterwards

The aim of this subsection is to obtain an estimate of the conditional probability of the occurrence of k consecutive jumps. Using this result, we are then also able to estimate the conditional probability of the arrival of a further consecutive jump after the cluster. These results, in addition to giving us practical indications on the risk of future jumps, can be used to check whether the statistical properties of our data match those of the model, as shown for instance in Table 2 in the empirical section.

The first conditional probability we want to estimate is denoted by

$$\tilde{\lambda}_n^{(k)} \doteq P\{ \text{ exactly one jump in each } [i\delta,(i+1)\delta), i = n, ..., n+k-1 \mid \mathcal{F}_{n\delta-} \}. \tag{16}$$

By the Markov property, it coincides with the probability of $\{\omega_i = 1,\ i = n, \ldots, n+k-1\}$ conditional on $\lambda_{n\delta}$, or, equivalently (see (12)), on $\tilde{\lambda}_n$. For a Hawkes process the sets $\{\omega_i = 1\}$ are not independent events, but each $\{\omega_i = 1\}$ depends on the outcomes of the previous random variables $\omega_j = 1,\ j = n, ..., n+i-1$, and we want to quantify (16) in a feasible way, given the information at our disposal.

Hence we factorize $\tilde{\lambda}_n^{(k)}$ by subsequently conditioning, and are able to express the j-th factor in terms of a conditional expectation of $\tilde{\lambda}_{n+j},\ j = 1, ..., k-1$. Via (12) the term $\tilde{\lambda}_{n+j}$ is written using $\lambda_{(n+j)\delta}$. Then from the following Proposition 3, which generalizes (13), we implement the bounds for $\lambda_{(n+j)\delta}$ in terms of $\lambda_{n\delta}$. Finally, the sought approximation for the probability of occurrence of k consecutive jumps follows in Theorem 4. Both Proposition 3 and Theorem 4 are proved in Section A.5.

Proposition 3. Conditionally on $\{\omega_i = 1,\ \forall\ i = n, \ldots, n+j-1\}$ and $\lambda_{n\delta}$, for any integer $j \geq 1$ we have

$$\lambda_{(n+j)\delta} \in [A_{j,n}\ ,\ B_{j,n}) \tag{17}$$

where

$$A_{j,n} := \lambda_0 + e^{-\beta\delta j}(\lambda_{n\delta} - \lambda_0) + \alpha\frac{1 - e^{-\beta\delta j}}{e^{\beta\delta} - 1}\ , \qquad B_{j,n} := \lambda_0 + e^{-\beta\delta j}(\lambda_{n\delta} - \lambda_0) + \alpha\frac{1 - e^{-\beta\delta j}}{1 - e^{-\beta\delta}}$$

Theorem 4. *Conditionally on $\lambda_{n\delta}$, the probability $\tilde{\lambda}_n^{(k)}$ of occurrence of k consecutive jumps belongs to the interval*

$$\left[\ \tilde{\lambda}_n \cdot \prod_{j=1}^{k-1}\left[1 - \exp\left(\frac{1}{\beta}(A_{j,n} - \lambda_0)(e^{-\beta\delta} - 1) - \delta\lambda_0\right)\right], \right. \tag{18}$$

$$\left. \tilde{\lambda}_n \cdot \prod_{j=1}^{k-1}\left[1 - \exp\left(\frac{1}{\beta}(B_{j,n} - \lambda_0)(e^{-\beta\delta} - 1) - \delta\lambda_0\right)\right]\ \right)$$

Both extremes of the latter interval contain exponents which are deterministic expressions of $\lambda_{n\delta}$, so through (12) the interval can also be written only in terms of $\tilde{\lambda}_n$. Once the estimates of the kernel parameters have been obtained, $\lambda_{n\delta}$ can be reconstructed using the jump times estimated up to $n\delta-$, while $\tilde{\lambda}_n$ is estimated using Theorem 3.

A direct consequence of (18) is that we obtain an estimate of the conditional probability of a further jump in the next small interval, after the occurrence of a cluster of consecutive k jumps, for any integer $k \geq 1$. It is sufficient to apply the following relation

$$P\{\omega_n = 1|\omega_{n-1} = 1, ..., \omega_{n-k} = 1, \mathcal{F}_{(n-k)\delta-}\} = \frac{P\{\omega_n = 1, \omega_{n-1} = 1, ..., \omega_{n-k} = 1|\mathcal{F}_{(n-k)\delta-}\}}{P\{\omega_{n-1} = 1, ..., \omega_{n-k} = 1|\mathcal{F}_{(n-k)\delta-}\}} \quad (19)$$

Note that the bounds presented in Theorem 1 differ from those obtained from (19) and (18). In fact, Theorem 1 does not explicitly account for the discreteness of the observations and considers a cluster with potentially arbitrarily sparse jumps.

5 Implementation

In this section we implement the previously obtained theoretical results to empirical data consisting of high-frequency returns on evenly spaced time intervals for 20 financial assets. We measure the jump risk of the assets, and find that it varies, but is directly related to the average yearly number of jumps and partially related to the sector to which the asset belongs.

We first explain how we model the asset prices, and how we estimate the jump times and fit an Exponential Hawkes model to each asset (Section 5.1). We then illustrate a simulation study (Section 5.2) which proves the satisfactory reliability of our measure of the probability, even where the Hawkes model parameters are obtained on estimated jump times, rather than true jump times. In Section 5.3 we briefly illustrate the data employed and show the estimated models. In Section 5.4 we proceed to estimate the probability that a cluster of jumps observed at a given date is unfinished. Further, we calculate the risk of the occurrence of a cluster of consecutive jumps, which allows us also to check the fit of our models.

5.1 Asset price model, identification of the jumps and fit of the Hawkes model.

We assume the following data generating process for the log-prices of a financial asset:

$$dX_t = a_t dt + \sigma_t dW_t + J_t, \quad t \geq 0 \quad (20)$$

where the *drift* coefficient a and the *volatility* coefficient σ are stochastic processes with càdlàg paths, W is a Wiener process, and J a pure jump semimartingale, which (as for instance in Aït-Sahalia et al., 2015) is independent of a, σ and W. More precisely, $J_t = \sum_{i=1}^{N_t} Y_i$ is a finite activity jump process where the counter $\{N_t\}_{t\geq 0}$ of the jump times follows an Exponential Hawkes model. We assume $\sigma_t(\omega) \geq \underline{\sigma}$ a.s., $\forall t \in [0, +\infty)$, for some constant level $\underline{\sigma} > 0$, but note that we do not need to assume that σ is a semimartingale. Our data comprises discrete high frequency and evenly spaced observations X_{t_1}, X_{t_2}, ..., X_{t_m} of X with $t_i = i\delta$, for fixed $\delta > 0$. To apply our formulas we need to estimate both the jump times and the parameters of the jump intensity. The first step requires infill asymptotics and the second long-span asymptotics. Thus we assume that our data is a finite sample within a theoretical framework where $\delta \to 0$ and $m \to \infty$ in such a way that $\delta m \to \infty$.

Under model (20), by the *threshold method* a jump is detected when a return has size above a suitable level, and for this reason it is also called *large variation*, or *jump over the threshold*. Namely, the time T_ℓ of the occurrence of a jump is estimated as the left extreme t_{i-1} of an interval $[t_{i-1}, t_i)$ where $(\Delta_i X)^2 = (X_{t_i} - X_{t_{i-1}})^2$ is above the threshold $r_i(\delta) = 2\hat{\sigma}^2_{t_{i-1}} \delta \log \frac{1}{\delta}$. The quantity $\hat{\sigma}_{t_{i-1}}$ is a recursive kernel-truncated estimator of spot σ at t_{i-1}, based on Mancini et al. (2015) and Figueroa-López and Li (2020), and the function $2\delta \ln \frac{1}{\delta}$ is the squared modulus of continuity of the Brownian motion paths. See more details in Section 1.4.1 in the Web Appendix.

We denote by $\{\hat{T}_\ell, \ell = 1, 2, ..., \hat{N}_{\delta m}\}$ the estimated jump times, and by $\{\hat{N}_t\}_t$ the process where $\hat{N}_t$ counts the number of jumps estimated up to t. To check for possible correlation of the jump intensity with the spot volatility in the empirical data considered, we also verify the fit with the alternative model for $\{\hat{N}_t\}_t$, where the jump times form a point process with intensity $\check{\lambda}_t = q + b\sigma_t$ and constant numbers q, b. Running maximum likelihood we find that for 17 out of the 20 selected assets the correlation coefficient b is not significant. For the remaining 3 assets we find that the Hawkes model better fits the data (see some details in the Web Appendix, Section 1.4.2). Thus we maintain the independence assumption of N on σ. The estimated parameters λ_0, α, β of an EHM are obtained by applying maximum likelihood (ML) to the estimated jump times. The maximization is made by imposing on the parameters the positivity condition and the condition ensuring asymptotic stationarity of the point process. Details of the ML estimation are given in the Web Appendix (Section 1.4.3).

5.2 Reliability check on simulations.

The theoretical results presented in Sections. 3 and 4 require the parameters qualifying the Hawkes process to be known. Actually, we do not know the true values of λ_0, α and β but estimate them from the jump times which in turn are identified starting from the observed asset prices. These two estimation steps lead to a sum of measurement errors and may deliver a significant bias in the estimate of our conditional probabilities. As for the probability of future consecutive jumps, the good results of a reliability check on simulations are set out, for two of the considered assets, in the last column of Table 2 below. We now focus on the conditional probability that a cluster of k jumps, which at time $\bar{s}$ we know to have occurred on an interval $[\underline{s}, \bar{s}]$, has not yet ended. As mentioned before Proposition 1, the best way to measure this risk is to compute $P \doteq P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s}) \mid \lambda_{\bar{s}}\}$, however the approximation $P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s}) | N_{\bar{s}} - N_{\underline{s}} = k, \lambda_{\underline{s}}\}$, with appropriately chosen $\underline{s}$, can be estimated more reliably. To check for the magnitude of the accumulated estimation error on P we implement the following simulation exercise.

A dataset of log prices $X_{t_i},\ t_i = i\delta,\ i = 1, ..., m$, is obtained through discretization (Euler scheme) of the path of a jump-diffusion process having leverage, and evolving in continuous time by

$$dX_t = -\frac{\sigma_t^2}{2} dt + \sigma_t dW_t + dJ_t \tag{21}$$

$$\sigma_t = e^{v_t}, \quad dv_t = 0.09\big(\log(0.25) - v_t\big)\ dt + 0.05\ dW_t^v, \quad v_0 = \log(0.3)$$

with $d[W, W^v]_t = -0.7dt$. All the parameters are annual. The continuous martingale part of X is as in Cont and Mancini (2011) (p.792), and similar to Huang and Tauchen (2005) (model SV1FJ). The jump part $J_t = \sum_{j=1}^{N_t} Z_j$ is independent on (W, W^v); N is an Exponential Hawkes process with parameters $\lambda_0 = 22, \alpha = 50, \beta = 80$ such that X makes on average 59 jumps per year (similarly as in Jacod et al., 2023). The jump sizes Z_j are iid and distributed as a truncated double exponential random variable with density

$$f(x) = \frac{e^{\frac{a}{b}}}{2b} e^{-\frac{|x|}{b}} \mathbf{1}_{\{(-\infty,-a]\}} + \frac{e^{\frac{a}{b}}}{2b} e^{-\frac{|x|}{b}} \mathbf{1}_{\{[a,+\infty)\}}$$

where $a = \bar{\sigma}_t \sqrt{\delta}$; $\bar{\sigma}_t$ is the average of the values realized by the simulated path of σ_t and b=0.012 is such that the standard deviation of ΔJ_s is 0.018 and the proportion of the Quadratic Variation of X due to the jumps is about 30%. In this framework we are interested in the identification of large jumps of asset prices, thus with the Hawkes process we want to model the arrival times only for the jumps with size larger than volatility. Finally, we chose $\delta = 1/(252 \times 77) = 5.15 \times 10^{-5}$, corresponding to 77 observations per day for 4815 days, as in our datasets, so m=370'755 is the number of returns at our disposal and $T = m\delta$ is about 19 years.

We simulate 500 paths, for each path we estimate spot σ and the jump times as described in the previous subsection. Then, on the estimated jump times we implement MLE for the EHM parameters. With $\epsilon = 0.01$ and two different choices for the couple $(\underline{s}, \bar{s})$, we substitute the estimated parameters for the true ones in the expressions (8) for LB and (9) for UB and compare the estimated bounds with the true probability P given by (5).

We find that the average percentage estimation error on λ_0, α, β is about 23%, 30% and 18%, respectively, and $\hat{LB}$ and $\hat{UB}$ are very good indications of the true probability P (see the details in the Web Appendix, Section 1.4.4).

5.3 Data and model fitting

Our empirical analysis is based on the intraday returns of 96 equities among the most liquid assets traded in the U.S. market and representing industrial sectors with high market capitalization. The sample is composed as follows: 34 equities from the Basic Materials sector, 15 from Financials, 13 from Healthcare, 8 from Industrial Goods, and 26 from Technology (see Section 2 in the Web Appendix for the complete list of all stocks considered).[7] The data are sampled at 5-minute frequency in the period from 2/1/2003 to 10/3/2022. We exclude the overnight returns and include possible zero returns, thus for each asset we have 77 observations per day, with $\delta = 1/(252 \cdot 77)$ years. The largest sample covers 4815 days and includes a total of 370'755 observations. The smallest sample also spans 4815 days, while the average number of days covered across all samples is 4582.

The formulas presented in this paper can only be applied to stocks whose data-generating process for estimated jump times is represented by the Exponential Hawkes Model (EHM), as they are specifically designed for this type of model. To assess this, as is commonly done with the Hawkes process, we evaluate the goodness-of-fit for each stock by examining the quantile-quantile (QQ) plot of the transformed durations against the exponential distribution (see Section 1.4.3 in the Web Appendix for technical details). If the points in the QQ-plot, specifically the quantiles of the appropriately transformed durations, lie close to the straight line (i.e., the bisector, which represents the quantiles of the exponential distribution with parameter one), then the model fits the estimated jump times well. More precisely, we cannot reject the null hypothesis that the transformed durations follow a unit exponential distribution, which allows us to assume that the underlying data-generating process for the jump times is the EHM. Conversely, if the points in the QQ-plot lie slightly above the straight line, this indicates that the transformed durations exhibit heavier tails than the exponential distribution. In such cases, the null hypothesis that the empirical transformed durations follow a unit exponential distribution is rejected at a typical significance level, suggesting that the observed jump times cannot be well described by the EHM. Using this approach, we estimate the jump times and fit the Exponential Hawkes model for each stock. We then evaluate all 96 QQ-plots and analyze only the stocks that exhibit the best goodness-of-fit, i.e., those for which the quantiles align with those of the exponential law up to at least the third quantile (a common threshold in empirical studies, see Figure 1.4.3 in the Web Appendix). As a result, we end up with 20 stocks out of 96, and we apply the proposed methodology only to these assets.[8]

Table 1 shows the estimated parameters for the 20 assets analyzed together with some relevant information. The estimated parameters are significantly different from zero for all the assets. In particular, all assets exhibit self-excitation, as indicated by the strong significance of the parameter α. The very high values of the estimated β suggest a rapid decay in the effect of one jump on the process intensity λ. We also verified that the estimation results and the Hawkes model fit is stable over time, in the following way. For each selected asset, we performed a rolling window estimation where the length of each window is one decade and each window was moved forward by one year.

5.4 Measuring the jump risk

We now apply our formulas. For each of the 20 selected assets, based on the estimated EHM model, we compute: the probability for the occurrence of a cluster of at least k consecutive jumps in the next k consecutive 5-minute intervals (Table 2); and the probability that an observed cluster of jumps is not yet finished (Tables 3 and 4). To produce Table 2 we use (18) and (17), for Table 3 (8) and (9), and for Table 4 (5). We account for all the information obtained until the end of the sample.

The results given in Table 2 also serve as a further check, after the QQ-plot, for the goodness of fit. The Hawkes process fits the data extremely well, for instance, for General Electric (GE) and BHP Billiton (BHP), as we can see in the Table.

For each asset the third and fourth columns provide the lower and the upper bounds delivered by our formulas. The fifth column of the Table allows a comparison of the bounds with the empirical probabilities (P_{emp}) computed from the data, and the sixth column flags that the P_{emp} falls within the bounds. The following column gives P_{sim}, the mean empirical probabilities obtained on 1000 simulated paths of the Hawkes process with the parameters estimated on the whole sample for the asset considered. The last column flags that the P_{sim} falls within the bounds. We can see that for almost all values of k the empirical probabilities fall within the bounds. Note that for BHP the empirical probabilities for large values of k are zero due to the relatively small number of estimated jumps.

We now implement the formulas devised to quantify the probability that an observed cluster of jumps is not yet finished. Based on the Hawkes model fitting the data of an asset, we focus on an observation time $\bar{s}$ where $\hat{\lambda}_{\bar{s}}$ is high (namely, $\hat{\lambda}_{\bar{s}} > \hat{\lambda}_0 + \hat{\alpha}/4$), because a high level of the jump intensity indicates that a cluster of jumps occurred recently. We then look for the first time $\underline{s}$ backwards, such that $\hat{\lambda}_{\underline{s}}$ drops below the level $\hat{\lambda}_0 + \hat{\alpha}/4$. This means that $\underline{s}$ is situated at the end of a less turbulent period than the period before $\bar{s}$, so we can rely more on our estimate $\hat{\lambda}_{\underline{s}}$ and use this estimate to implement the formulas for the probability. On the interval $[\underline{s}, \bar{s}]$ some jumps occurred, namely the k jumps at times $T_i, ..., T_{i+k-1}$. We take ϵ such that $\hat{\lambda}_0(1+\epsilon) < \hat{\lambda}_0 + \hat{\alpha}/4$ so that the intensity remains above $\hat{\lambda}_0(1+\epsilon)$ during the whole time interval, and therefore those jumps form a cluster. We do not know where the jumps are precisely located, because we only estimated the 5-minute intervals containing them. By applying Theorem 1 we compute the probability that at $\bar{s}$ such a cluster has not yet finished, that is, the next jump, at time T_{i+k}, is still a consequence of the same causes that produced the cluster. In other words, we deliver the probability that the next jump will occur before the cluster decays, i.e. the probability of $T_{i+k} < \bar{s} + t_\epsilon(\bar{s})$. The results with $\epsilon = 0.01$ are illustrated in Table 3.

The table shows the lower bound $LB = P\{T_{i+k} < \bar{s} + \underline{t}_\epsilon | N_{\bar{s}} - N_{\underline{s}} = k, \lambda_{\underline{s}}\}$ for $P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s}) | \lambda_{\bar{s}}\}$ and the upper bound $UB = P\{T_{i+k} < \bar{s} + \bar{t}_\epsilon | N_{\bar{s}} - N_{\underline{s}} = k, \lambda_{\underline{s}}\}$, where $\bar{s}$ is 4/6/2021, for those assets for which $\hat{\lambda}_{\bar{s}}$ is high. In addition to the number $\hat{N}_T$ of estimated jumps on the whole sample of each asset, the table provides the following information: the *half-life* defined by $\log(2)/\hat{\beta}$ and representing the time needed for the effect α on the intensity, due to the arrival of a jump, to decrease and become $\alpha/2$; the (asymptotically) expected number of jumps in one year, given by the formula $E[N_1] = \lambda_0\beta/(\beta - \alpha)$; the market sector to which the asset belongs; the average trading volume in 5 minutes; the time $\underline{s}$; the levels $\lambda_{\bar{s}}$ and $\lambda_{\underline{s}}$; the number of 5-minute intervals partitioning $[\underline{s}, \bar{s})$ and given by $(\bar{s} - \underline{s})/\delta$; the number k of jumps detected within that interval; LB and UB.

First, note that, except in two cases, the bounds are quite close to each other, which allows us to have quite a good indication of the probability that the cluster is still active. For instance we see that for WMB the probability is low, i.e. between 12% and 22%, for UNH it is quite high, i.e. between 55% and 64%, while for GE it is very high, between 91% and 94%.

Second, also for other choices of $\bar{s}$, it turns out that UNH and GE have particularly high probabilities, with GE always close to 1. This means that their jumps are highly self-exciting, and the occurrence of a further jump of the cluster is easily predictable.

Third, for all assets, and over the different days $\bar{s}$, we show that, consistently with theory, the higher the number k of jumps in the cluster, the higher the probability that by time $\bar{s}$ the cluster has not finished.

For a given k, the narrower the interval $[\underline{s}, \bar{s}]$ the higher the probability: a cluster made up of closely spaced jumps generates a strong increase in the jump intensity. However, in practice in a large time interval $[\underline{s}, \bar{s}]$ the number of jumps tends to be high and the range $UB - LB$ tends to be large. A wider range corresponds to a greater uncertainty about the probability that we can deduce from the bounds. As an extreme example we mention that BHP displays LB=0 and UB=1, which are not useful at all, because the only time $\underline{s}$ before $\bar{s}$ when $\lambda_{\underline{s}}$ is low is far, and 188 jumps fall within the interval $[\underline{s}, \bar{s}]$. Since LB is the probability when all the 188 jumps occur close to $\underline{s}$, then at time $\bar{s}$ they are too far to be considered an unfinished cluster. Analogously, UB is the probability when all the 188 jumps occur close to $\bar{s}$, which makes it almost certain that the next jump still belongs to the cluster.

The bounds shown in Table 3 indicate that the probability that an observed cluster has not yet finished is typically higher with a higher expected number of jumps $E[N_1]$.

The reason is clear, as for an asset with frequently jumping prices many jumps fall in $[\underline{s}, \bar{s}]$, increasing the level of the jump intensity $\lambda_{\bar{s}}$, and consequently the probability that the cluster was not yet exhausted. However this direct relation was not obvious, because a higher yearly jump frequency does not necessarily mean a uniform increase in the number of jumps in each time interval, in particular it does not signify an increase in the number of jumps in the interval preceding $\bar{s}$ for all the displayed assets.

Moreover, we see that within the Financial sector the jump frequencies $E[N_1]$ are homogeneous, and the same holds for the probabilities. Within the Technology sector there is less homogeneity, and in the Healthcare there is a significant variability.

Table 2 shows the probability of at least k consecutive jumps and gives the results for only two assets, however taking into account the other 18 selected assets we again observed the direct relation between P_{emp} and $E[N_1]$, for all k. Note that $E[N_1] = \lambda_0/(1 - \alpha/\beta)$ is increasing in α/β (as well as in λ_0, or decreasing in β). Thus in order to show the direct relation of P_{emp} with $E[N_1]$, in Figure 1 we plotted P_{emp} and $\hat{\alpha}/\hat{\beta}$, as the asset varies, for all k.

The direct relation is mathematically explained by the fact that when β is small and/or α is large the effect of any jump on the intensity is persistent and/or strong, and this makes the accumulation of consecutive jumps easier.

Finally, Table 4 shows the probability P of Proposition 1 for those assets that on $\bar{s} = 16/12/2021$ display a level of $\hat{\lambda}_{\bar{s}}$ belonging to $(\hat{\lambda}_0(1+\epsilon), \hat{\lambda}_0 + \hat{\alpha}/4)$. In this case $\hat{\lambda}_{\bar{s}}$ is small enough to indicate that $\lambda_{\bar{s}}$ does not fall in a turbulent period containing many close jumps, and we can rely on the precision of its estimated value. Thus we implement the formula in the proposition rather than settle for the bounds. Clearly, since the level of $\lambda_{\bar{s}}$ is not high, the probability that at time $\bar{s}$ the cluster has not yet finished is rather small, not exceeding 17% in all cases.

6 Probabilities and forecasting of jumps and realized variance

The purpose of this section is to empirically use the probabilities presented in the previous sections to improve the forecasting performance for asset jumps and realized variance (RV). We follow the well-known literature on realized variance forecasting, which uses a heterogeneous structure (Corsi, 2009) for two empirically main determinants of the RV dynamics, namely lagged continuous integrated variance and lagged jumps (see, among others, Corsi and Renò, 2012). We add the jump risk measures proposed in this paper as a possible source of information on future jumps and RV, thus imposing a heterogeneous structure on these risk measures too.

In particular, in section 6.1 we extend the reduced-form model HAR-J-CJ put forward by Corsi and Renò (2012) for forecasting jumps, by including the probability that a cluster observed over the previous day, week and month has not yet finished. At odds with the conclusions drawn in Corsi and Renò (2012), we find that jumps are predictable and jump forecasting performance is remarkable for most of the assets considered, especially for 1 day ahead. In section 6.2, we investigate the forecasting performance when the realized variance is used as dependent variable and the interaction of jumps with the proposed risk measures is included. In this case we use as benchmark the HAR-RV-CJ model of Corsi and Renò (2012) and record a significant improvement for some assets.

6.1 Forecasting Jumps with probabilities

Busch et al. (2011) and Corsi and Renò (2012) use the jump component of the total quadratic variation as dependent variable in order to forecast jumps. They find that jumps are almost unpredictable. Following Corsi and Renò (2012), in addition to the heterogeneous structure of the continuous and jump component, we use the probability that, having observed a cluster of jumps over the past day, week, and month, the cluster is not yet exhausted. Combining heterogeneity for the continuous part of the quadratic variation, jumps and probabilities, we extend the following HAR-J-CJ model by Corsi and Renò (2012):

$$\begin{aligned} \hat{JV}_{t+h}^{(h)} = c &+ \beta^{(d)}\hat{C}_t + \beta^{(w)}\hat{C}_t^{(5)} + \beta^{(m)}\hat{C}_t^{(22)} \\ &+ \alpha^{(d)}\hat{JV}_t + \alpha^{(w)}\hat{JV}_t^{(5)} + \alpha^{(m)}\hat{JV}_t^{(22)} + \epsilon_{t+h}^{(h)}, \end{aligned} \tag{22}$$

where $\hat{C}_t$ and $\hat{JV}_t$ are identified according to the threshold method explained in Section 5.1.

As commonly in the literature, we construct daily variables as:

$$\hat{C}_t = \sum_{i \in \text{day } t} \hat{C}_{i,t}, \quad \hat{JV}_t = \sum_{i \in \text{day } t} \hat{JV}_{i,t},$$

where $\hat{C}_{i,t} = (\Delta_i X)^2 \mathbf{1}_{\{|\Delta_i X| \leq r_i(\delta)\}}$ and $\hat{JV}_{i,t} = (\Delta_i X)^2 \mathbf{1}_{\{|\Delta_i X|^2 > r_i(\delta)\}}$. The weekly ($h = 5$) and monthly ($h = 22$) variables are obtained by averaging the continuous component and aggregated jumps over integer number of h days:

$$\hat{C}_t^{(h)} = \frac{1}{h}\sum_{j=1}^{h} \hat{C}_{t-j+1}, \quad \hat{JV}_t^{(h)} = \sum_{j=1}^{h} \hat{JV}_{t-j+1}, \quad \text{with} \quad h = 5, 22. \tag{23}$$

We introduce the probabilities provided in Section 3 and use these additional variables to inform the jump contribution to the total quadratic variation. The proposed model reads HAR-J-CJP and is specified as follows:

$$
\begin{aligned}
\hat{JV}_{t+h}^{(h)} = c + &\beta^{(d)}\hat{C}_t + \beta^{(w)}\hat{C}_t^{(5)} + \beta^{(m)}\hat{C}_t^{(22)} + \alpha^{(d)}\hat{JV}_t + \alpha^{(w)}\hat{JV}_t^{(5)} + \alpha^{(m)}\hat{JV}_t^{(22)} \\
&+ \gamma^{(d)}\hat{JV}_t \times P_t + \gamma^{(w)}\hat{JV}_t^{(5)} \times P_t^{(5)} + \gamma^{(m)}\hat{JV}_t^{(22)} \times P_t^{(22)} + \epsilon_{t+h}^{(h)},
\end{aligned}
\tag{24}
$$

where $P_t^{(h)}$ is the probability that the cluster observed over the interval $[t-h, t]$ is not exhausted at time t, with $h = 1, 5, 22$, and $P_t = P_t^{(1)}$. Model (24) nests the HAR-J-CJ, when all probabilities are zero.

We estimate models (22) and (24) for all the assets considered in our application, with $h = 1, 5, 22$, by using ordinary least squares (OLS) with Newey-West covariance correction for serial correlation.

The forecasts of the different models over all assets are evaluated on the basis of the adjusted R^2 of the regressions, the root mean square error (RMSE) and the Diebold-Mariano (DM) test, as shown in Table 5.[9] To make the relative gains from our probabilities stand out more clearly, we show the ratio between the RMSE of the HAR-J-CJP model and that of the benchmark model HAR-J-CJ. Note that this indicator is a ratio between two loss functions, indeed a value less than 1 indicates an outperformance of the model set as the numerator, that is, HAR-J-CJP model.

We can see that jumps are predictable, with an adjusted R^2 that increases considerably for many assets when the model HAR-J-CJP, as in equation (24), is evaluated at $h = 1$. In this case the adjusted R^2 increase varies from 1% to 37%. The HAR-J-CJP model outperforms the alternative model, the RMSE always being smaller than one. Specifically, the HAR-J-CJP model improvement ranges from about 1% to 23%. Moreover, the probability contribution to the model performance is significant for more than 50% of assets, as suggested by the DM test at the 10% confidence level.

When weekly and monthly horizons are considered, the predictability and the performance improvements are reduced, even if the adjusted R^2 for the HAR-J-CJP remains larger with respect to the HAR-J-CJ and the RMSE ratio is always smaller than one, with the exception of three cases for $h = 22$, where the ratio is approximately one. The negative sign of the DM test indicates that the HAR-J-CJ produced a larger average loss than the HAR-J-CJP forecast for all assets, however the improved performance of the model is not statistically different in the majority of cases, for $h = 5, 22$.

To sum up, the results of this section show that the probability that observed clusters of jumps are not yet exhausted, incorporates information about future jumps and can be used to improve their predictability for all assets considered in this paper, especially when one-step ahead jumps forecasts are evaluated.

6.2 Forecasting Realized Variance with probabilities

In light of the above section results, we empirically assess if the ability of forecasting jumps using the probabilities proposed in this paper carries over to the prediction of realized variance. We achieve this goal estimating both specifications in equations (22)-(24) using the realized variance for each asset as the dependent variable. Specifically, we use as benchmark model the HAR-RV-CJ specification as in Corsi and Renò (2012):

$$\begin{aligned} \hat{RV}^{(h)}_{t+h} = c &+ \beta^{(d)}\hat{C}_t + \beta^{(w)}\hat{C}^{(5)}_t + \beta^{(m)}\hat{C}^{(22)}_t \\ &+\alpha^{(d)}\hat{JV}_t + \alpha^{(w)}\hat{JV}^{(5)}_t + \alpha^{(m)}\hat{JV}^{(22)}_t + \epsilon^{(h)}_{t+h}. \end{aligned} \tag{25}$$

Following the extension proposed in Section 6.1, we define the HAR-RV-CJP model as:

$$\begin{aligned} \hat{RV}^{(h)}_{t+h} = c &+ \beta^{(d)}\hat{C}_t + \beta^{(w)}\hat{C}^{(5)}_t + \beta^{(m)}\hat{C}^{(22)}_t + \alpha^{(d)}\hat{JV}_t + \alpha^{(w)}\hat{JV}^{(5)}_t + \alpha^{(m)}\hat{JV}^{(22)}_t \\ &+ \gamma^{(d)}\hat{JV}_t \times P_t + \gamma^{(w)}\hat{JV}^{(5)}_t \times P^{(5)}_t + \gamma^{(m)}\hat{JV}^{(22)}_t \times P^{(22)}_t + \epsilon^{(h)}_{t+h}. \end{aligned} \tag{26}$$

As for the jump analysis, the forecast performance for all assets is evaluated on the basis of the adjusted R^2 of the regressions, the root mean square error (RMSE) and the Diebold-Mariano (DM) test, as shown in Table 6. Looking at the adjusted R^2 for the model HAR-RV-CJP, we observe an increase across all assets and forecast horizons h. In some cases, the improvement is substantial, although the gains are slightly smaller in magnitude compared to those reported in Table 5. The HAR-RV-CJP model outperforms the alternative, for all the tickers with the exception of JNPR and ACN when $h = 1$, and GE when $h = 22$, since the RMSE is greater than one. However, according to the DM test at the 10% confidence level, the model improvement when the probabilities are considered is significant only in the case of seven assets. Hence, the probability that a cluster of jumps observed over previous time intervals is not yet exhausted, works as supplementary information for predicting realized variance.

7 Conclusions

Many large changes in financial asset prices are not isolated episodes, but are preceded by exceptional price instability. Indeed a cluster of jumps can be interpreted as an early-warning sign of a future extreme event. Supposing a jump-diffusion model for the log-price of an asset, the price variations are classified as jumps if they are above a proper threshold making them incompatible with a Brownian semimartingale.

We focus on the jump times, and assume that their counting process follows a Hawkes model. We investigate the probabilistic properties of the Exponential Hawkes process, characterized by one exponential kernel.

We study the dependence relation between contiguous jump interarrival times, proving their stochastic increasingness. This property is stronger than the positive correlation. Since in the empirical applications we usually have discrete observations, we aim to measure specific jump risks explicitly accounting for the lack of precise information on the jump locations. Through the *decay instant* of the jump intensity, we study the residual length of a partially observed cluster, which enables the computation of the bounds for the probability P that the cluster has yet to be exhausted. A high P means that the cluster is expected to produce a further jump soon. A low P means that the next jump is completely random and unpredictable. The decay instant is connected with the stochastic increasingness of the jump interarrival times. Moreover, we provide bounds for the probability of occurrence of jumps in consecutive observation intervals.

We check the robustness of our risk measures on simulations and then apply our formulas to records of discrete prices of US financial stocks. We find that the estimated jump times of 20 selected assets are well modeled by Exponential Hawkes processes and display significant dependence and clustering.

We find that the magnitude of the probability that an observed cluster was yet to be exhausted by a given date, varies among assets and market sectors, but is quite homogeneous within the Financial sector. By contrast, it is systematically related to the expected number of jumps in the unit period.

Since this probability represents an additional source of information on future jumps, we use it to improve the forecasting of asset jumps and realized variance. We extend the HAR-J-CJ model by considering a heterogeneous structure that also takes into account the conditional probabilities of having observed jumps in the previous day, week and month. According to our results, jumps are predictable and provide a remarkable forecasting performance for jumps of most of the assets considered, especially when one-step ahead forecasts are analyzed. The probability we propose in this paper works as supplementary information for predicting realized variance.

A Proof of the mathematical results.

A.1 For Section 2: Hawkes processes

Proof of Lemma 1. We have $\lambda_t = \lambda_0 + \alpha \sum_{k=1}^{n} e^{-\beta(t-T_i)}$. Consider the case where the n jumps on $(0,t]$ all occurred within $(0,\eta)$. Using (2) with $a = 0+, b = \eta-$, we have

$$\alpha e^{-\beta t} < \alpha e^{-\beta(t-T_i)} < \alpha e^{-\beta(t-\eta)}$$

for each i, and claim 1. follows.

Analogously, if the n jumps on $(0,t]$ all occurred within $(t-\eta, t)$, then for each jump we apply (2) with $a = (t-\eta)+, b = t-$, and obtain claim 2. from the relations

$$\alpha e^{-\beta\eta} < \alpha e^{-\beta(t-T_i)} < \alpha \qquad \square$$

Proof of the recursive formula, Lemma 2. We split the self-exciting term of the intensity into two parts: the first part is due to the history before s, which gave to λ a contribution that at time t decayed, and the second part is due to the jumps within $[s,t)$. Namely,

$$\lambda_t = \lambda_0 + \alpha \sum_{T_\ell < t} e^{-\beta(t-T_\ell)} = \lambda_0 + \alpha \sum_{T_\ell < s} e^{-\beta(s-T_\ell+t-s)} + \alpha \sum_{T_\ell \in [s,t)} e^{-\beta(t-T_\ell)}$$

$$= \lambda_0 + e^{-\beta(t-s)} \alpha \sum_{T_\ell < s} e^{-\beta(s-T_\ell)} + \alpha \sum_{T_\ell \in [s,t)} e^{-\beta(t-T_\ell)}$$

$$= \lambda_0 + e^{-\beta(t-s)}(\lambda_s - \lambda_0) + \alpha \sum_{T_\ell \in [s,t)} e^{-\beta(t-T_\ell)} \qquad \square$$

We show some consequences of the lemma that are used in the rest of the paper.

First consequence: we display two distinct formulas that will be used in Section 3.1: in (27) below we know the value $\lambda_{T_{i-1}}$ of the intensity at a time of jump, while in (28) we know the value $\lambda_{\bar{s}}$ at a time $\bar{s}$ of no jump.

For $t \in (T_{i-1}, T_i]$, by applying (3) with $s = T_{i-1}$, we have that on (T_{i-1}, t) there are no jumps, and we obtain

$$\lambda_t = \lambda_0 + [\lambda_{T_{i-1}} - \lambda_0 + \alpha] e^{-\beta(t-T_{i-1})}. \tag{27}$$

If $T_{i-1} < \bar{s} \le t \le T_i$, on $(\bar{s}, t)$ there are no jumps, and using(3) with $s = \bar{s}$ we obtain

$$\lambda_t = \lambda_0 + (\lambda_{\bar{s}} - \lambda_0)e^{-\beta(t-\bar{s})}. \tag{28}$$

The second consequence of Lemma 2 is used in Section 5.3 to verify the fit of the exponential Hawkes model to the considered empirical data. Defined $\Lambda_s = \int_0^s \lambda_t dt$ the *integrated intensity,* from (27) it follows that the increments $\Delta_{T_i}\Lambda := \int_{T_{i-1}}^{T_i} \lambda_t dt$ are given by

$$\Delta_{T_i}\Lambda = \lambda_0(T_i - T_{i-1}) + \frac{1}{\beta}[\lambda_{T_{i-1}} - \lambda_0 + \alpha]\left[1 - e^{-\beta(T_i - T_{i-1})}\right]. \tag{29}$$

A.2 For Section 3.1: Decay instant and cluster size

Recall the following notation: given two instants of time r, s, let us define the sets

$$NJ_{[r,s)} \doteq \{\text{no jumps in } [r,s)\}, NJ_{(r,s)} \doteq \{\text{no jumps in } (r,s)\}.$$

Proof of Proposition 1. Conditionally on $\lambda_{\bar{s}}$, and following Rangan and Grace (1988), we have

$$P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s})|\lambda_{\bar{s}}\} = \int_0^{t_\epsilon(\bar{s})} \lambda_{\bar{s}+t|NJ_{[\bar{s},\bar{s}+t)}} \exp\left(-\int_0^t \lambda_{\bar{s}+\tau|NJ_{[\bar{s},\bar{s}+\tau)}} d\tau\right) dt. \tag{30}$$

Since by definition of $NJ_{[\bar{s},\bar{s}+t)}$ there are no jumps within $[\bar{s}, \bar{s}+t)$, then $\lambda_{\bar{s}+t|NJ_{[\bar{s},\bar{s}+\tau)}} = \lambda_0 + (\lambda_{\bar{s}} - \lambda_0)e^{-\beta t}$, and the above becomes

$$\int_0^{t_\epsilon(\bar{s})} \left[\lambda_0 + (\lambda_{\bar{s}} - \lambda_0)e^{-\beta t}\right] e^{-\int_0^t \lambda_0 + (\lambda_{\bar{s}} - \lambda_0)e^{-\beta\tau} d\tau} dt = \int_0^{t_\epsilon(\bar{s})} \left[\lambda_0 + (\lambda_{\bar{s}} - \lambda_0)e^{-\beta t}\right] e^{-\lambda_0 t + (\lambda_{\bar{s}} - \lambda_0)\frac{e^{-\beta t} - 1}{\beta}} dt.$$

With $c = \lambda_{\bar{s}} - \lambda_0$ and $b = t_\epsilon(\bar{s})$ the above display equals

$$1 - e^{\frac{ce^{-b\beta}}{\beta}} e^{-b\lambda_0 - \frac{c}{\beta}}.$$

Since $e^{-b\beta} = \frac{\lambda_0\epsilon}{\lambda_{\bar{s}} - \lambda_0}$ and $e^{-b\lambda_0} = \left(\frac{\lambda_0\epsilon}{\lambda_{\bar{s}} - \lambda_0}\right)^{\frac{\lambda_0}{\beta}}$ then

$$P = P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s})|\lambda_{\bar{s}}\} = 1 - e^{\frac{\lambda_{\bar{s}} - \lambda_0}{\beta} \frac{\lambda_0\epsilon}{\lambda_{\bar{s}} - \lambda_0}} \left(\frac{\lambda_0\epsilon}{\lambda_{\bar{s}} - \lambda_0}\right)^{\frac{\lambda_0}{\beta}} e^{-\frac{\lambda_{\bar{s}} - \lambda_0}{\beta}} = 1 - e^{-\frac{\lambda_{\bar{s}} - \lambda_0(1+\epsilon)}{\beta}} \left(\frac{\lambda_0\epsilon}{\lambda_{\bar{s}} - \lambda_0}\right)^{\frac{\lambda_0}{\beta}}$$

□

Proof that $t_\epsilon(\bar{s}) \in (\underline{t}_\epsilon, \bar{t}_\epsilon)$, with $\underline{t}_\epsilon, \bar{t}_\epsilon$ as in (6). By the recursive relation in (3) we have

$$\lambda_{\bar{s}} = \lambda_0 + e^{-\beta d}(\lambda_{\underline{s}} - \lambda_0) + \alpha \sum_{\underline{s} \le T_\ell < \bar{s}} e^{-\beta(\bar{s} - T_\ell)}. \tag{31}$$

The cluster observed on $[\underline{s}, \bar{s})$ has k jumps, and for any jump of the cluster we have $T_\ell \ge \underline{s}$, so the third term above satisfies

$$k\alpha e^{-\beta d} = k\alpha e^{-\beta(\bar{s}-\underline{s})} \le \alpha \sum_{\underline{s} \le T_\ell < \bar{s}} e^{-\beta(\bar{s}-T_\ell)} \le k\alpha.$$

Note that the right extreme $k\alpha$ would be obtained if all the jumps of the cluster were concentrated at $\bar{s}$ ($T_\ell = \bar{s}$ for any T_ℓ of the cluster), as in the limit for $\eta \to 0$ of case 2 in Lemma 1. Analogously, the left extreme $k\alpha e^{-\beta(\bar{s}-\underline{s})}$ would be obtained if all the jumps were concentrated at $\underline{s}$, as in the limit of case 1 in the Lemma. It follows that

$$\lambda_{\bar{s}} - \lambda_0 \in \Big(e^{-\beta d}(\lambda_{\underline{s}} - \lambda_0) + k\alpha e^{-\beta d}, \quad e^{-\beta d}(\lambda_{\underline{s}} - \lambda_0) + k\alpha\Big).$$

Because $\bar{s}$ is assumed not to be a jump time then $t_\epsilon(\bar{s}) = \frac{1}{\beta}\log\left(\frac{\lambda_{\bar{s}}-\lambda_0}{\epsilon\lambda_0}\right)$, and the stated inequality is straightforward. □

Proof of Theorem 1. Conditionally on $N_{\bar{s}} - N_{\underline{s}} = k$ and on $\lambda_{\underline{s}}$, similarly as for Proposition 1, we have

$$P\{T_{i+k} < \bar{s} + t_\epsilon(\bar{s}) | \mathcal{G}\} = E\left[\int_0^{t_\epsilon(\bar{s})} \lambda_{\bar{s}+t|NJ_{[\bar{s},\bar{s}+t)}} \exp\left(-\int_0^t \lambda_{\bar{s}+\tau|NJ_{[\bar{s},\ \bar{s}+\tau)}} d\tau\right) dt \,\Big|\, \mathcal{G}\right]. \tag{32}$$

Similarly as in (31), with $\bar{s}+t$ in place of $\bar{s}$, and recalling that $\bar{s} - \underline{s} = d$, we have

$$\lambda_{\bar{s}+t} = \lambda_0 + (\lambda_{\underline{s}} - \lambda_0)e^{-\beta(d+t)} + \alpha \sum_{T_\ell \in [\underline{s}, \bar{s}+t)} e^{-\beta(\bar{s}+t-T_\ell)},$$

and, since by definition of $NJ_{[\bar{s},\bar{s}+t)}$ there are no jumps within $[\bar{s}, \bar{s}+t)$ then within $[\underline{s}, \bar{s}+t)$ we still have k jumps, and the last sum has k terms.

In the limit case where all the k jumps in the cluster within $[\underline{s}, \bar{s})$ happen at $\underline{s}$, then $T_\ell = \underline{s}$ for all ℓ and $t_\epsilon(\bar{s}) = \underline{t}_\epsilon$, thus $\lambda_{\bar{s}+t|NJ_{[\bar{s},\bar{s}+t)}} = \lambda_0 + (\alpha k + \lambda_{\underline{s}} - \lambda_0)e^{-\beta(d+t)}$, and we obtain

$$P\{T_{i+k} < \bar{s} + \underline{t}_\epsilon | \mathcal{G}\} = \int_0^{\underline{t}_\epsilon} \Big(\lambda_0 + (\alpha k + \lambda_{\underline{s}} - \lambda_0)e^{-\beta(d+t)}\Big) e^{-\int_0^t \lambda_0 + (\alpha k + \lambda_{\underline{s}} - \lambda_0)e^{-\beta(d+\tau)} d\tau} dt$$

$$= \int_0^{\underline{t}_\epsilon} (\lambda_0 + (\alpha k + \lambda_{\underline{s}} - \lambda_0)e^{-\beta(d+t)}) \exp\left(-\frac{1}{\beta}(e^{-\beta d}(\alpha k + \lambda_{\underline{s}} - \lambda_0) - e^{-\beta(d+t)}(\alpha k + \lambda_{\underline{s}} - \lambda_0) + \lambda_0 t\beta)\right)$$

$$= e^{-\frac{e^{-\beta d}}{\beta}(\alpha k+\lambda_{\underline{s}})}\left[-\left((\alpha k+\lambda_{\underline{s}}-\lambda_0)\frac{e^{-\beta d}}{\lambda_0\epsilon}\right)^{-\frac{\lambda_0}{\beta}} e^{\frac{\lambda_0}{\beta}(e^{-\beta d}+\epsilon)} + e^{\frac{e^{-\beta d}}{\beta}(\alpha k+\lambda_{\underline{s}})}\right],$$

which leads to (8). In the opposite limit case where all the k jumps occurred at the end of the interval $[\underline{s},\bar{s})$, then $T_\ell = \bar{s}$ for all ℓ and $t_\epsilon(\bar{s})$ coincides with $\bar{t}_\epsilon$, so $\lambda_{\bar{s}+t|NJ_{[\bar{s},\bar{s}+t)}} = \lambda_0 + (\lambda_{\underline{s}}-\lambda_0)e^{-\beta(d+t)} + k\alpha e^{-\beta t}$, and we obtain that $P\{T_{i+k} < \bar{s}+\bar{t}_\epsilon|\mathcal{G}\}$ coincides with

$$\int_0^{\bar{t}_\epsilon} (\lambda_0 + (\alpha k + (\lambda_{\underline{s}}-\lambda_0)e^{-\beta d})e^{-\beta t})\exp\left(-\int_0^t \lambda_0 + (\alpha k + (\lambda_{\underline{s}}-\lambda_0)e^{-\beta d})e^{-\beta\tau}d\tau\right)dt$$

$$= \int_0^{\bar{t}_\epsilon} \left(\lambda_0 + (\alpha k + \lambda_{\underline{s}} - \lambda_0)e^{-\beta t}\right) e^{\frac{1}{\beta}\left((\lambda_{\underline{s}}-\lambda_0)e^{-\beta(d+t)}+e^{-\beta t}\alpha k-\lambda_0 t\beta-e^{-\beta d}(\lambda_{\underline{s}}-\lambda_0)-k\alpha\right)}$$

$$= 1-\exp\left(-\frac{1}{\beta}\Big(-(\lambda_{\underline{s}}-\lambda_0+\alpha k)e^{-\beta(d+\bar{t}_\epsilon)} + \lambda_0\bar{t}_\epsilon\beta + e^{-\beta d}(\lambda_{\underline{s}}-\lambda_0)+k\alpha\Big)\right)$$

$$= e^{-\frac{e^{-\beta d}\lambda_{\underline{s}}+k\alpha}{\beta}}\left(-\left((e^{\beta d}\alpha k+\lambda_{\underline{s}}-\lambda_0)\frac{e^{-\beta d}}{\lambda_0\epsilon}\right)^{-\frac{\lambda_0}{\beta}} e^{\frac{\lambda_0}{\beta}(e^{-\beta d}+\epsilon)} + e^{\frac{e^{-\beta d}\lambda_{\underline{s}}+k\alpha}{\beta}}\right),$$

which leads to (9).

As for the monotonicity:

the first derivative in d of $LB = P\{T_{i+k} < \bar{s}+\underline{t}_\epsilon|\mathcal{G}\} = P\{T_{i+k} < \underline{s}+d+\underline{t}_\epsilon|\mathcal{G}\}$ is

$$-e^{-e^{-\beta d}\cdot\frac{\alpha k+\lambda_{\underline{s}}}{\beta}}\left(\frac{\alpha k+\lambda_{\underline{s}}-\lambda_0}{\lambda_0\epsilon}e^{-\beta d}\right)^{-\frac{\lambda_0}{\beta}} e^{\frac{\lambda_0}{\beta}(e^{-\beta d}+\epsilon)}\left(e^{-\beta d}(\alpha k+\lambda_{\underline{s}}-\lambda_0)+\lambda_0\right),$$

which is strictly negative, as stated. Furthermore, $\frac{d}{dk}P\{T_{i+k} < \underline{s}+d+\underline{t}_\epsilon|\mathcal{G}\}$ is given by

$$\alpha e^{-e^{-\beta d}\frac{\alpha k+\lambda_{\underline{s}}}{\beta}}\left(\frac{\alpha k+\lambda_{\underline{s}}-\lambda_0}{\lambda_0\epsilon}e^{-\beta d}\right)^{-\frac{\lambda_0}{\beta}} e^{\frac{\lambda_0}{\beta}(e^{-\beta d}+\epsilon)}\cdot\frac{e^{-\beta d}(\alpha k+\lambda_{\underline{s}}-\lambda_0)+\lambda_0}{\beta(\alpha k+\lambda_{\underline{s}}-\lambda_0)}$$

and is strictly positive as soon as $\alpha > 0$.

The first derivative in d of $UB = P\{T_{i+k} < \underline{s}+d+\bar{t}_\epsilon|\mathcal{G}\}$ is

$$-e^{-\frac{e^{-\beta d}\lambda_{\underline{s}}+k\alpha}{\beta}}\left((e^{\beta d}\alpha k+\lambda_{\underline{s}}-\lambda_0)\frac{e^{-\beta d}}{\lambda_0\epsilon}\right)^{-\frac{\lambda_0}{\beta}} e^{\frac{\lambda_0}{\beta}(e^{-\beta d}+\epsilon)}(\lambda_{\underline{s}}-\lambda_0)\cdot\frac{\alpha k+e^{-\beta d}(\lambda_{\underline{s}}-\lambda_0)+\lambda_0}{e^{-\beta d}\alpha k+\lambda_{\underline{s}}-\lambda_0}$$

and is strictly negative. Finally $\frac{d}{dk}P\{T_{i+k} < \underline{s}+d+\bar{t}_\epsilon|\mathcal{G}\}$ is

$$\alpha e^{-\frac{1}{\beta}(e^{-\beta d}\lambda_{\underline{s}}+k\alpha)}\left((e^{\beta d}\alpha k+\lambda_{\underline{s}}-\lambda_0)\frac{e^{-\beta d}}{\lambda_0\epsilon}\right)^{-\frac{\lambda_0}{\beta}} e^{\lambda_0\frac{e^{-\beta d}+\epsilon}{\beta}}\cdot\frac{e^{\beta d}\alpha k+\lambda_0 e^{\beta d}+\lambda_{\underline{s}}-\lambda_0}{(e^{\beta d}\alpha k+\lambda_{\underline{s}}-\lambda_0)\beta} > 0.$$

A.3 For Section 3.2: Duration between consecutive jump times

Proof of Proposition 2. First, we prove (10). Note that conditioning on $\lambda_{T_n} = \ell$ does not lead us to know the value of the r.v. T_n. However, if we also condition on the event 'no jumps occurred within $(T_n, T_n + t)$', by the recursive formula, we obtain

$$\lambda_{T_n+t|\{\lambda_{T_n}=\ell\},NJ_{(T_n,T_n+t)}} = \lambda_0 + (\ell - \lambda_0 + \alpha)e^{-\beta t}.$$

Thus we can write

$$P\{T_{n+1} - T_n = t|\lambda_{T_n} = \ell\} \tag{33}$$

$$= P\{\text{jump at } T_n + t|NJ_{(T_n,T_n+t)}, \{\lambda_{T_n} = \ell\}\}P\{NJ_{(T_n,T_n+t)}|\lambda_{T_n} = \ell\}$$

$$= \lambda_{T_n+t|NJ_{(T_n,T_n+t)},\{\lambda_{T_n}=\ell\}} \exp\left(-\int_{T_n}^{T_n+t} \lambda_{s|\{\lambda_{T_n}=\ell\},NJ_{(T_n,s)}} ds\right)$$

$$=\left(\lambda_0 + (\ell - \lambda_0 + \alpha)e^{-\beta t}\right) e^{-\int_0^t \left(\lambda_0+(\ell-\lambda_0+\alpha)e^{-\beta\sigma}\right)d\sigma} =\left(\lambda_0 + (\ell - \lambda_0 + \alpha)e^{-\beta t}\right) e^{\frac{1}{\beta}(e^{-\beta t}-1)(\ell-\lambda_0+\alpha)-\lambda_0 t},$$

as we claimed. Finally, the equality

$$P\{T_{n+1} - T_n > \tau|\lambda_{T_n} = \ell\} = \exp\left(\frac{1}{\beta}(e^{-\tau\beta}(\ell + \alpha - \lambda_0) - \beta\lambda_0\tau - \ell - \alpha + \lambda_0)\right)$$

is obtained by integrating in t from τ to $+\infty$ the expression (33). □

Proof of Theorem 2. For brevity, we write j for indicating 'jump', and, along the whole proof, we condition on $\lambda_{T_{n-1}}$. We first compute $P\{T_{n+1} - T_n = t| \ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\}$ because, after that, to obtain $P\{T_{n+1} - T_n > \tau| \ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\}$ is straightforward through the relation

$$P\{T_{n+1} - T_n > \tau| \ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\} = \int_\tau^{+\infty} P\{T_{n+1} - T_n = t| \ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\}dt.$$

Note that

$$P\{T_{n+1} - T_n = t| \ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\} =$$

$$P\{\text{j at } T_{n-1} + s + t, NJ_{(T_{n-1}+s,T_{n-1}+s+t)}| \text{ j at } T_{n-1} + s, NJ_{(T_{n-1},T_{n-1}+s)}, \lambda_{T_{n-1}}\}$$

$$= P\Big(\text{j at } T_{n-1}+s+t|\ NJ_{(T_{n-1}+s,T_{n-1}+s+t)\cup(T_{n-1},T_{n-1}+s)},\ \text{j at } T_{n-1}+s, \lambda_{T_{n-1}}\Big)\cdot$$

$$P\Big(NJ_{(T_{n-1}+s,T_{n-1}+s+t)}\Big|\ \text{j at } T_{n-1}+s, NJ_{(T_{n-1},T_{n-1}+s)}, \lambda_{T_{n-1}}\Big).$$

$$= \lambda_{T_{n-1}+s+t|\ j\ at\ T_{n-1}+s,\ NJ_{(T_{n-1},T_{n-1}+s)},\ \lambda_{T_{n-1}}} e^{-\int_{T_{n-1}+s}^{T_{n-1}+s+t} \lambda_{\tau|\ j\ at\ T_{n-1}+s,\ NJ_{(T_{n-1},T_{n-1}+s)},\ \lambda_{T_{n-1}}} d\tau} \quad (34)$$

By recursive formula (27) we evaluate the state of λ at the needed times:

- $\lambda_{T_{n-1}+s|\ NJ_{(T_{n-1},T_{n-1}+s)},\ \lambda_{T_{n-1}}} = \lambda_0 + (\lambda_{T_{n-1}} - \lambda_0 + \alpha)e^{-\beta s}$;

- $\lambda_{T_{n-1}+s+t|\ j\ at\ T_{n-1}+s,\ NJ_{(T_{n-1},T_{n-1}+s)},\ \lambda_{T_{n-1}}}$

  $= \lambda_0 + [\lambda_{T_{n-1}+s|\ NJ_{(T_{n-1},T_{n-1}+s)},\ \lambda_{T_{n-1}}} - \lambda_0 + \alpha]e^{-\beta t}$: using the previous formula this becomes

  $= \lambda_0 + [(\lambda_{T_{n-1}} - \lambda_0 + \alpha)e^{-\beta s} + \alpha]e^{-\beta t}$;

- for $\tau \in (T_{n-1}+s, T_{n-1}+s+t)$, we have $\lambda_{\tau|\ j\ at\ T_{n-1}+s,\ NJ_{(T_{n-1},T_{n-1}+s)},\ \lambda_{T_{n-1}}}$

  $= \lambda_0 + [\lambda_{T_{n-1}+s|\ NJ_{(T_{n-1},T_{n-1}+s)},\ \lambda_{T_{n-1}}} - \lambda_0 + \alpha]e^{-\beta(\tau - T_{n-1}-s)}$

  $= \lambda_0 + [(\lambda_{T_{n-1}} - \lambda_0 + \alpha)e^{-\beta s} + \alpha]e^{-\beta(\tau - T_{n-1}-s)}$.

Using these expressions, from eq. (34) we obtain $P\{T_{n+1}-T_n = t|\ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\}$

$$\Big\{\lambda_0 + \Big[(\lambda_{T_{n-1}} - \lambda_0 + \alpha)e^{-\beta s} + \alpha\Big]e^{-\beta t}\Big\}\, e^{-\lambda_0 t} e^{-\int_{T_{n-1}+s}^{T_{n-1}+s+t} [(\lambda_{T_{n-1}}-\lambda_0+\alpha)e^{-\beta s}+\alpha]e^{-\beta(\tau-T_{n-1}-s)} d\tau} \quad (35)$$

To compute the integral in (35), we change variable by $\sigma = \tau - T_{n-1} - s$, so σ varies from 0 to t, and we obtain

$$P\{T_{n+1}-T_n = t|\ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\}$$

$$= \Big\{\lambda_0 + \Big[(\lambda_{T_{n-1}} - \lambda_0 + \alpha)e^{-\beta s} + \alpha\Big]e^{-\beta t}\Big\}\, e^{-\lambda_0 t} e^{\frac{1}{\beta}\Big[(\lambda_{T_{n-1}}-\lambda_0+\alpha)e^{-\beta s}+\alpha\Big](e^{-\beta t}-1)} \quad (36)$$

We compute now

$$P\{T_{n+1}-T_n > \tau|\ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\} = \int_{\tau}^{+\infty} P\{T_{n+1}-T_n = t|\ T_n - T_{n-1} = s, \lambda_{T_{n-1}}\}dt$$

$$e^{-\beta(s+\tau)}\alpha - e^{-\beta(s+\tau)}\lambda_0 - \lambda_0\tau\beta - e^{-\beta s}\lambda_{T_{n-1}} + e^{-\beta\tau}\alpha - e^{-\beta s}\alpha + e^{-\beta s}\lambda_0 - \alpha$$

$$- \lim_{x \to +\infty} e^{-\frac{1}{\beta}(-\lambda_{T_{n-1}} e^{-\beta(s+x)} - \alpha e^{-\beta(s+x)} + e^{-\beta(s+x)} \lambda_0 + \lambda_0 x \beta + e^{-\beta s} \lambda_{T_{n-1}} + e^{-\beta s} \alpha - e^{-\beta s} \lambda_0 - \alpha e^{-\beta x} + \alpha)}$$

$$= e^{-\lambda_0 \tau} \exp\left(\frac{e^{-\beta\tau} - 1}{\beta} \Big(e^{-\beta s} (\lambda_{T_{n-1}} - \lambda_0 + \alpha) + \alpha \Big) \right). \tag{37}$$

In order to study the monotonicity wrt s, we compute

$$\frac{d}{ds} P\Big\{ T_{n+1} - T_n > \tau | \; T_n - T_{n-1} = s, \lambda_{T_{n-1}} \Big\} =$$

$$\Big(e^{-\beta s} - e^{-\beta(s+\tau)} \Big) \Big(\lambda_{T_{n-1}} - \lambda_0 + \alpha \Big) e^{-\lambda_0 \tau} \exp\left\{ \frac{e^{-\beta\tau} - 1}{\beta} \Big(e^{-\beta s} (\lambda_{T_{n-1}} - \lambda_0 + \alpha) + \alpha \Big) \right\},$$

which is greater than 0 for any $s, \tau > 0$. □

A.4 For Section 4.1: Probability of a jump in the next time interval

Proof of Theorem 3.

As mentioned in Section 4.1, through relation (12) we express $\lambda_{n\delta}$ in terms of $\tilde{\lambda}_n$, then we relate $\lambda_{n\delta}$ to $\lambda_{(n-1)\delta}$ using (13), and then use again (12) to connect $\lambda_{(n-1)\delta}$ and $\tilde{\lambda}_{n-1}$. Namely, using (12) it turns out that

$$\lambda_{n\delta} = \lambda_0 + \frac{\beta}{e^{-\beta\delta} - 1} (\delta\lambda_0 + \log(1 - \tilde{\lambda}_n)). \tag{38}$$

If $\omega_{n-1} = 0$, we know by (28) that $\lambda_{n\delta} = \lambda_0 + (\lambda_{(n-1)\delta} - \lambda_0) e^{-\beta\delta}$, which, compared with eq. (38), gives

$$(\lambda_{(n-1)\delta} - \lambda_0) e^{-\beta\delta} = \frac{\beta}{e^{-\beta\delta} - 1} (\delta\lambda_0 + \log(1 - \tilde{\lambda}_n)). \tag{39}$$

In addition, as in eq. (38),

$$\lambda_{(n-1)\delta} = \lambda_0 + \frac{\beta}{e^{-\beta\delta} - 1} (\delta\lambda_0 + \log(1 - \tilde{\lambda}_{n-1})),$$

which, replaced in eq. (39), leads to

$$e^{-\beta\delta} (\delta\lambda_0 + \log(1 - \tilde{\lambda}_{n-1})) = \delta\lambda_0 + \log(1 - \tilde{\lambda}_n),$$

hence the thesis.

If $\omega_{n-1} = 1$, by the recursive formula (13) we only obtain upper and lower bounds for $\lambda_{n\delta}$ in terms of $\lambda_{(n-1)\delta}$: the lower bound is obtained assuming that the jump within $[(n-1)\delta, n\delta)$ occurred at $(n-1)\delta$, the upper bound would be achieved in the limit case in which the jump occurred at $n\delta$. Following the same substitution strategy as above, for each of the two limit cases, the result is obtained. □

A.5 For Section 4.2: Probability of a cluster of consecutive jumps and of a jump immediately after

Proof of Proposition 3. The proof is straightforward, using the recursive formula (3) and the bounds in (2). In fact, by (3) with $t = (n+j)\delta$ and $s = n\delta$, we have

$$\lambda_{(n+j)\delta} = \lambda_0 + (\lambda_{n\delta} - \lambda_0)e^{-\beta j\delta} + \alpha \sum_{T_\ell \in [n\delta,(n+j)\delta)} e^{-\beta\left((n+j)\delta - T_\ell\right)}.$$

Since, by assumption, on $[n\delta,(n+j)\delta)$ we have j jumps in consecutive j intervals, then for any $T_\ell \in [n\delta,(n+j)\delta)$ there is an interval $[i\delta,(i+1)\delta)$ containing $T_\ell =: T_{\ell_i}$, $i = n, ..., n+j-1$, and we have $-\beta\Big((n+j)\delta - T_{\ell_i}\Big) \in \Big[-\beta\Big((n+j)\delta - i\delta\Big), -\beta\Big((n+j)\delta - i\delta - \delta\Big)\Big)$, thus

$$\sum_{i=n}^{n+j-1} e^{-\beta\left((n+j)\delta - i\delta\right)} \leq \sum_{T_\ell \in [n\delta,(n+j)\delta)} e^{-\beta\left((n+j)\delta - T_\ell\right)} \leq \sum_{i=n}^{n+j-1} e^{-\beta\left((n+j)\delta - i\delta - \delta\right)}$$

or, with $p = n+j-i$,

$$\sum_{p=1}^{j} e^{-\beta p\delta} \leq \sum_{T_\ell \in [n\delta,(n+j)\delta)} e^{-\beta\left((n+j)\delta - T_\ell\right)} \leq e^{\beta\delta} \sum_{p=1}^{j} e^{-\beta p\delta}.$$

Since $\sum_{p=1}^{j} e^{-\beta\delta p} = e^{-\beta\delta}\frac{1-e^{-\beta\delta j}}{1-e^{-\beta\delta}}$, we obtain

$$\lambda_{(n+j)\delta} \in \left[\lambda_0 + e^{-\beta\delta j}(\lambda_{n\delta} - \lambda_0) + \alpha e^{-\beta\delta}\frac{1-e^{-\beta\delta j}}{1-e^{-\beta\delta}},\ \lambda_0 + e^{-\beta\delta j}(\lambda_{n\delta} - \lambda_0) + \alpha\frac{1-e^{-\beta\delta j}}{1-e^{-\beta\delta}}\right),$$

which proves the thesis. □

Proof of Theorem 4. By subsequently conditioning, we see that

$$\tilde{\lambda}_n^{(k)} = P\{\omega_{n+k-1} = 1 | \omega_i = 1,\ i = n, \ldots, n+k-2, \lambda_{n\delta}\}$$

$$\cdot P\{\omega_{n+k-2} = 1 | \omega_i = 1,\ i = n, \ldots, n+k-3, \lambda_{n\delta}\} \cdots P\{\omega_{n+1} = 1 | \omega_n = 1, \lambda_{n\delta}\} \cdot \tilde{\lambda}_n$$

and note that, for each $j \in \{1, 2, ..., k-1\}$,

$$P\{\omega_{n+j} = 1 | \omega_i = 1,\ i = n, \ldots, n+j-1, \lambda_{n\delta}\}$$

$$= E\left[E\left[\mathbf{1}_{\{\omega_{n+j}=1\}} | \mathcal{F}_{(n+j)\delta-}\right] | \omega_i = 1,\ i = n, \ldots, n+j-1, \lambda_{n\delta}\right]$$

$$= E\left[\tilde{\lambda}_{n+j} \mid \omega_i = 1,\ i = n, \ldots, n+j-1, \lambda_{n\delta}\right].$$

Using (12) we obtain that the above display equals

$$E\left[1 - \exp\left(-\frac{1}{\beta}(\lambda_{(n+j)\delta} - \lambda_0)(1 - e^{-\beta\delta}) - \delta\lambda_0\right) \mid \omega_i = 1,\ i = n, \ldots, n+j-1, \lambda_{n\delta}\right]. \tag{40}$$

From Proposition 3 it follows that the expression in (40) belongs to

$$\left[1 - \exp\left(-\frac{1}{\beta}(A_{j,n} - \lambda_0)(1 - e^{-\beta\delta}) - \delta\lambda_0\right), 1 - \exp\left(-\frac{1}{\beta}(B_{j,n} - \lambda_0)(1 - e^{-\beta\delta}) - \delta\lambda_0\right)\right)$$

and thus also $P\{\omega_{n+j} = 1 | \omega_i = 1,\ i = n, \ldots, n+j-1, \lambda_{n\delta}\}$ does. Therefore the sought approximation for the probability of occurrence of k consecutive jumps is as in the statement of this Proposition. □

Figure legends

Figure 1. Light blue line with circles, left y-axis: empirical probability P_{emp} of at *least* k consecutive jumps after the observation time 10/3/2022 (as in Table 2 for BHP and GE) for all selected assets. Pink dashed line with asterisks, right y-axis: $\hat{\alpha}/\hat{\beta}$, which is directly related to $E[N_1] = \frac{\lambda_0}{1-\frac{\alpha}{\beta}}$.

References

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

Bacry, E., S. Delattre, M. Hoffmann, and J.-F. Muzy (2013). 'Modelling microstructure noise with mutually exciting point processes'. *Quantitative finance Vol. 13*, pp. 65–77.

Bacry, E., I. Mastromatteo, and J. Muzy (2015). 'Hawkes processes in finance'. *Market Microstructure and Liquidity Vol. 1*, p. 59.

Bowsher, C. G. (2007). 'Modelling security market events in continuous time: Intensity based, multivariate point process models'. *Journal of Econometrics Vol. 141*, pp. 876–912.

Busch, T., B. J. Christensen, and M. Ø. Nielsen (2011). The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Econometrics 160*(1), 48–57.

Christensen, K., R. Oomen, and R. Renò (2022). The drift burst hypothesis. *Journal of Econometrics 227*(2), 461–497.

Clinet, S. and N. Yoshida (2017). 'Statistical inference for ergodic point processes and application to Limit Order Book'. *Stochastic Processes and their Applications Vol. 127*, pp. 1800–1839.

Cont, R. and C. Mancini (2011). 'Nonparametric tests for pathwise properties of semimartingales'. *Bernoulli Vol. 17*, pp. 781–813.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics 7*(2), 174–196.

Corsi, F. and R. Renò (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics 30*(3), 368–380.

Cui, L., A. Hawkes, and H. Yi (2020). 'An elementary derivation of moments of Hawkes processes'. *Advances in Applied Probability Vol. 52*, pp. 102–137.

Da Fonseca, J. and R. Zaatour (2014). 'Hawkes process: Fast calibration, application to trade clustering, and diffusive limit'. *Journal of Futures Markets Vol. 34*, pp. 548–579.

Ferriani, F. and P. Zoi (2020). 'The dynamics of price jumps in the stock market: an empirical study on Europe and US'. *The European Journal of Finance*, pp. 1–25.

Figueroa-López, J. E. and C. Li (2020). 'Optimal kernel estimation of spot volatility of stochastic differential equations'. *Stochastic Processes and their Applications Vol. 130*, pp. 4693–4720.

Hawkes, A. G. (2018). 'Hawkes processes and their applications to finance: a review'. *Quantitative Finance Vol. 18*, pp. 193–198.

Huang, X. and G. Tauchen (2005). 'The relative contribution of jumps to total price variance'. *Journal of Financial Econometrics 3*, pp. 456–499.

Jacod, J., L. Huidi, and V. Todorov (2023). 'Systematic jump risk'. *Working paper*.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts.* New York: Chapman and Hall/CRCr.

Li, L., H. Deng, A. Dong, Y. Chang, and H. Zha (2014). 'Identifying and labeling search tasks via query-based hawkes processes'. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 731–740.

Maheu, J. M. and T. H. McCurdy (2004). 'Jump dynamics, and volatility components for individual stock returns'. *The Journal of Finance Vol. 59 (2)*, pp. 755–793.

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping brownian motion. *Giornale dell'Istituto Italiano degli Attuari, Vol.64, pp.19-47*.

Mancini, C. (2009). 'Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps'. *Scand. J. Stat. Vol. 36*, pp. 270–296.

Mancini, C., V. Mattiussi, and R. Renó (2015). 'Spot volatility estimation using delta sequences'. *Finance Stoch. Vol. 19*, pp. 261–293.

Purcaru, O. and M. Denuit (25 (4), 781-793, 2002). On the stochastic increasingness of future claims in the bühlmann linear credibility premium. *Blätter der DGVFM*.

Rambaldi, M., V. Filimonov, and F. Lillo (2018). 'Detection of intensity bursts using Hawkes processes: An application to high-frequency financial data'. *Physical Review E Vol. 97*, p. 032318.

Rangan, A. and R. Grace (1988). 'A non-Markov model for the optimum replacement of self-repairing systems subject to shocks'. *J. Appl. Probab. Vol. 25*, pp. 375–382.

Reynaud-Bouret, P. and E. Roy (2007). 'Some non asymptotic tail estimates for Hawkes processes'. *Bulletin of the Belgian Mathematical Society-Simon Stevin Vol. 13*, pp. 883–896.

Swishchuk, A., R. Zagst, and G. Zeller (2021). 'Hawkes processes in insurance: Risk model, application to empirical data and optimal investment'. *Insurance: Mathematics and Economics*.

Zhuang, J., Y. Ogata, and D. Vere-Jones (2002). 'Stochastic declustering of space-time earthquake occurrences'. *J. Am. Stat. Assoc. Vol. 97*, pp. 369–380.

Notes

[1]Similar considerations hold for Da Fonseca and Zaatour (2014), Bacry et al. (2013) and Cui et al. (2020), who give formulas, or procedures, for the moments and/or the autocorrelation function of the number of jumps on a given interval for either an Exponential Hawkes process or bivariate generalizations.

[2] For each asset we consider a univariate model for the times of all the jumps. It is possible to consider the times of negative jumps and positive jumps separately or identify the jump times according to a rule of interest such as a specific range of jump size. Of course, in order to apply our formulas the times of the considered subset of jumps must be consistent with an EHM and the observations need to be sufficiently numerous to deliver reliable quantifications. By contrast, if we want to investigate the interaction between the two subsets of negative and positive jumps we need a bivariate model, but this is beyond the scope of our paper.

[3]A point process is *simple* if, a.s., at every t it makes at most one jump.

[4]The symbol $\Phi \not\equiv 0$ means that there is at least one value of $x \geq 0$ such that $\Phi(x) \neq 0$.

[5]For a fixed t, the probability of the set of paths that have a jump exactly at t is 0: there are no *fixed times* of discontinuity. This is justified a few lines before Section 4.1.

[6]A reformulation of (14) suitable for the implementation of the recursion is given in Remark 2 in the Web Appendix.

[7]We exclude indices from our analysis, as they function similarly to portfolios by diversifying risk across multiple assets. Such diversification can obscure the presence of jumps: a jump in one asset may be offset by an opposite movement in another, effectively smoothing out aggregate behavior.

[8]BAX: Baxter International; BHP: BHP Billiton Ltd.; JNPR: Juniper Networks; MRO: Marathon Oil Corporation; C: Citigroup; GE: General Electric Company; PNC: PNC Financial Services Group; WFC: Wells Fargo & Company; GLW: Corning; JNJ: Johnson & Johnson; WMB: Williams Companies; MRK: Merck & Co.; LMT: Lockheed Martin Corporation; ACN: Accenture PLC; T: AT&T; UNH: UnitedHealth Group; BAC: Bank of America Corporation; IBM: International Business Machines Corp.; TXN: Texas Instruments; LLY: Eli Lilly & Co.

[9]To save space, we do not set out the coefficients estimated for models HAR-J-CJ and HAR-J-CJP. Briefly, at 5% significance level, we find that $\gamma^{(d)}$ coefficient is significant for eight assets, $\gamma^{(w)}$ for seven assets and $\gamma^{(m)}$ for thirteen assets out of the twenty tickers considered. All the estimation results are available on request.

[⋆]*Acknowledgements*: we thank Giacomo Bormetti, Lorenzo Braccini, Francesco Calvori, Fabrizio Cipollini, Giampiero M. Gallo, Jean Jacod, Fabrizio Lillo, Davide Pirino, Roberto Renó, Simone Scotti, Peter Tankov, Viktor Todorov and Nakahiro Yoshida for their precious comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily reflect those of their institutions. *Funding*: this project benefited from funds provided by the INdAM-GNAMPA Association of Italian Mathematicians located in Rome; from the Europlace Institute of Finance of the Institut Louis Bachelier, in Paris; and from the University of Pisa project 2022 “Fostering the role of public services evaluation for innovation, inclusion and sustainability’.

Tables

Table 1: Selected assets: Name, Ticker, Sector, Sample period, Average trading volume over 5 minutes, maximum likelihood estimated parameters of the best fitting Exponential Hawkes model, with standard errors in brackets

Name	Ticker	Sector	Sample period	Volume	$\check{N}_T$	λ_0	α	β
ACCENTURE	ACN	Technology	2/1/03 - 8/3/22	8003	1522	23.53	9.64	13.66
						(-0.32)	(0.57)	(1.81)
BANK OF AMERICA	BAC	Financial	2/1/03 - 8/3/22	252337	2543	106.29	787.09	3908.33
						(0.87)	(94.38)	(1106.59)
BAXTER INTERNATIONAL	BAX	Healthcare	2/1/03 - 8/3/22	14820	1095	49.55	521.28	3850.24
						(0.91)	(27.3)1	(880.7)6
BHP BILLITON	BHP	Basic	2/1/03 - 8/3/22	9594	2539	15.96	23.16	26.29
						(0.19)	(0.05)	(4.80)
CITIGROUP	C	Financial	2/1/03 - 8/3/22	55513	3030	125.04	647.46	3061.30
						(0.85)	(81.96)	(827.93)
GENERAL ELECTRIC	GE	Industrial	2/1/03 - 10/3/22	19375	21786	782.31	560.33	1786.70
						(0.72)	(103.94)	(328.88)
CORNING	GLW	Technology	2/1/03 - 9/3/22	32870	885	40.55	471.92	3794.28
						(0.92)	(20.55)	(762.21)
INTERNATIONAL BUSINESS MACHINES	IBM	Technology	2/1/03 - 9/3/22	18017	817	36.79	681.15	4881.82
						(0.92)	(39.75)	(1185.31)
JOHNSON JOHNSON	JNJ	Healthcare	2/1/03 - 10/3/22	26509	990	43.78	669.48	4326.83
						(0.91)	(48.61)	(1052.02)
JUNIPER NETWORKS	JNPR	Technology	2/1/03 - 10/3/22	21102	732	33.03	752.38	5473.83
						(0.92)	(43.15)	(1448.12)
ELI LILLY & CO.	LLY	Healthcare	2/1/03 - 10/3/22	14961	781	35.77	680.53	5468.86
						(0.93)	(30.59)	(1154.36)
LOCKHEED MARTIN CORPORATION	LMT	Industrial	2/1/03 - 10/3/22	6006	1024	46.28	405.74	2981.18
						(0.91)	(21.12)	(632.45)
MERCK & CO.	MRK	Healthcare	2/1/03 - 10/3/22	39400	914	39.43	964.86	5498.55
						(0.90)	(94.35)	(1587.84)
MARATHON OIL	MRO	Basic	2/1/03 - 10/3/22	34065	669	30.61	442.74	3528.25
						(0.92)	(19.98)	(851.72)
PNC FINANCIAL SERVICES GROUP	PNC	Financial	2/1/03 - 10/3/22	8149	2274	97.62	571.47	3185.40
						(0.88)	(52.74)	(747.36)
AT&T, INC.	T	Technology	2/1/03 - 10/3/22	94055	991	44.06	689.54	4591.53
						(0.91)	(47.10)	(1202.25)
TEXAS INSTRUMENTS	TXN	Technology	2/1/03 - 10/3/22	30277	521	23.48	683.22	4927.01
						(0.92)	(40.12)	(1099.25)
UNITEDHEALTH GROUP	UNH	Healthcare	2/1/03 - 10/3/22	16149	9461	361.57	676.32	2509.62
						(0.78)	(114.2)8	(562.43)
WELLS FARGO & COMPANY	WFC	Financial	2/1/03 - 10/3/22	79111	2236	91.12	696.32	3150.32
						(0.85)	(99.82)	(876.73)
WILLIAMS COMPANIES	WMB	Basic	2/1/03 - 10/3/22	28068	1338	59.58	616.46	4141.56
						(0.91)	(39.99)	(912.83)

Table 2: Bounds (18) for the probability of at *least k* consecutive jumps after the observation time 10/3/2022: BHP and GE stock prices.

asset	k	Lower bound	Upper bound	P_{emp}	$dummyP_{emp}$	P_{sim}	$dummyP_{sim}$
BHP	1	8.22×10^{-04}	9.01×10^{-02}	6.85×10^{-03}	1	1.04×10^{-02}	1
BHP	2	1.65×10^{-05}	8.20×10^{-03}	5.18×10^{-04}	1	1.67×10^{-04}	1
BHP	3	5.29×10^{-09}	7.55×10^{-04}	7.82×10^{-05}	1	3.83×10^{-06}	1
BHP	4	2.32×10^{-11}	7.02×10^{-05}	1.89×10^{-05}	1	6.00×10^{-08}	1
BHP	5	1.29×10^{-13}	6.60×10^{-05}	$5.40\ \times10^{-05}$	1	0	0
BHP	6	8.68×10^{-16}	6.26×10^{-07}	0	0	0	0
BHP	7	6.87×10^{-18}	6.00×10^{-08}	0	0	0	0
BHP	8	6.24×10^{-20}	5.81×10^{-09}	0	0	0	0
BHP	9	6.39×10^{-22}	5.68×10^{-10}	0	0	0	0
BHP	10	7.30×10^{-24}	5.61×10^{-11}	0	0	0	0

asset	k	Lower bound	Upper bound	P_{emp}	$dummyP_{emp}$	P_{sim}	$dummyP_{sim}$
GE	1	3.95×10^{-02}	1.90×10^{-01}	6.00×10^{-02}	1	5.36×10^{-02}	1
GE	2	2.50×10^{-03}	3.82×10^{-02}	6.90×10^{-03}	1	3.54×10^{-03}	1
GE	3	2.12×10^{-04}	8.00×10^{-03}	$1.18\times10^{-}03$	1	3.45×10^{-04}	1
GE	4	2.20×10^{-05}	1.74×10^{-03}	2.67×10^{-04}	1	4.23×10^{-05}	1
GE	5	2.65×10^{-05}	3.93×10^{-04}	8.90×10^{-05}	1	5.88×10^{-05}	1
GE	6	3.59×10^{-07}	9.11×10^{-05}	3.24×10^{-05}	1	8.80×10^{-07}	1
GE	7	5.36×10^{-08}	2.17×10^{-05}	1.62×10^{-05}	1	1.60×10^{-07}	1
GE	8	8.66×10^{-09}	5.28×10^{-06}	8.09×10^{-06}	0	1.00×10^{-08}	1
GE	9	1.50×10^{-09}	1.31×10^{-06}	2.70×10^{-06}	0	0	0
GE	10	2.73×10^{-10}	$3.32\text{-}\times10^{-07}$	0	0	0	0

Table 3: Probability that a cluster of k jumps observed up to time $\bar{s}$= 4/6/2021 has not yet finished: application of formulas (8) and (9).

Asset	Sector	Volume	$\hat{N}_T$	half-life	$E[N_1]$	$\underline{s}$	$\lambda_{\bar{s}}$	$\lambda_{\underline{s}}$	d/δ	k	LB	UB
BHP	Basic	9594	2539	0.026368193	134.00	02/05/2018	56.2597	20.7154	50227	188	0	1
WMB	Basic	28068	1338	0.000167364	70.00	03/12/2020	322.0753	59.5782	5	1	0.12678	0.22008
PNC	Financial	8149	2274	0.000217601	118.96	03/12/2020	396.5198	103.4024	5	1	0.22111	0.31293
C	Financial	55513	3030	0.000226423	158.58	03/12/2020	469.5299	125.0974	5	1	0.27294	0.37272
WFC	Financial	79111	2236	0.000220025	116.97	03/12/2020	454.8404	91.1188	5	1	0.23387	0.33818
BAC	Financial	252337	2543	0.000177351	133.09	03/12/2020	457.9521	106.2909	5	1	0.20196	0.31669
UNH	Healthcare	16149	9461	0.000276196	494.95	03/12/2020	764.7506	361.619	5	1	0.55079	0.64004
JNJ	Healthcare	26509	990	0.000160197	51.79	03/12/2020	318.1685	43.7777	5	1	0.10734	0.20452
MRK	Healthcare	39400	914	0.00012606	47.82	03/12/2020	350.0235	39.4252	5	1	0.084481	0.20654
GE	Industrial	19375	21786	0.000387949	1139.75	03/12/2020	1596.7547	898.2484	5	2	0.9135	0.94303
ACN	Technology	8003	1522	0.050746579	79.91	02/01/2003	53.3665	23.5263	346370	1472	0	1
GLW	Technology	32870	885	0.000182682	46.31	03/12/2020	256.4118	40.5496	5	1	0.10567	0.18103
T	Technology	94055	991	0.000150962	51.84	03/12/2020	311.6633	44.0591	5	1	0.099797	0.19801

Table 4: Probability that a cluster of k jumps observed up to time $\bar{s}$= 16/12/2021 has not yet finished: application of formula (5).

Ticker	Sector	Volume	$\hat{N}_T$	half-life	$E[N_T]$	$\lambda_{\bar{s}}$	P
MRO	Basic	34065	669	0.00020	35.01	37.37	0.02827
PNC	Financial	8149	2274	0.00022	118.96	214.80	0.16741
WFC	Financial	79111	2236	0.00022	116.98	127.61	0.1113
BAC	Financial	252337	2543	0.00018	133.09	123.43	0.07663
JNJ	Healthcare	26509	990	0.00017	51.79	48.74	0.02527
MRK	Healthcare	39400	914	0.00013	47.82	41.32	0.01145
TXN	Technology	30277	521	0.000143	27.26	25.46	0.01048
GLW	Technology	32870	885	0.00018	46.31	66.46	0.04987

Table 5: The table shows the R^2 from HAR-J-CJ (J-CJ-R^2) and HAR-J-CJP (J-CJP-R^2) models; the RMSE losses computed as the ratio between the RMSE of the HAR-J-CJP model and that of the benchmark model HAR-J-CJ. DM test represents the t-statistics from Diebold-Mariano tests of equal predictive accuracy between HAR-J-CJ and HAR-J-CJP. A t-statistic greater than 1.64 in absolute value indicates a rejection of the null of equal predictive accuracy at the 10% significance level. These statistics are marked with an asterisk. The sign of the t-statistics indicates which forecast performed better for each loss function: a negative t-statistic indicates that the HAR-J-CJ produced larger average loss than the HAR-J-CJP forecast, while a positive sign indicates the opposite. HAR-J-CJ is as in equation (22) while HAR-J-CJP is presented in equation (24).

	$h=1$				$h=5$				$h=22$			
	J-CJ-R^2	J-CJP-R^2	RMSE	DM test	J-CJ-R^2	J-CJP-R^2	RMSE	DM test	J-CJ-R^2	J-CJP-R^2	RMSE	DM test
ACN	0.0404	0.0444	0.9979	$-1.863^{\star}$	0.1706	0.1738	0.9981	-1.2503	0.3263	0.3324	0.9955	$-1.6479^{\star}$
BAC	0.1356	0.1441	0.995	$-1.8456^{\star}$	0.393	0.3942	0.9991	-0.4788	0.5401	0.5407	0.9993	-0.4697
BAX	0.0202	0.0205	0.9999	-0.9504	0.0383	0.1719	0.9279	$-1.8612^{\star}$	0.0264	0.0571	0.9841	-1.0776
BHP	0.0995	0.1095	0.9944	$-2.2741^{\star}$	0.3086	0.3228	0.9897	$-1.7438^{\star}$	0.5111	0.5349	0.9754	-1.3519
C	0.2474	0.2483	0.9994	-0.1271	0.3702	0.3757	0.9957	-0.6412	0.5381	0.5401	0.9978	-0.5112
GE	0.0059	0.0097	0.9981	-0.5067	0.0036	0.0037	0.9999	-0.3139	0.0036	0.0036	1.0025	-0.1797
GLW	0.0024	0.0044	0.999	$-2.2406^{\star}$	0.0101	0.0108	0.9997	-0.8709	0.0361	0.0364	0.9999	-0.7299
IBM	0.0171	0.072	0.9716	-1.3419	0.0677	0.0735	0.9969	-1.272	0.1026	0.1134	0.994	-1.4954
JNJ	0.0061	0.0135	0.9963	$-1.9788^{\star}$	0.0261	0.0272	0.9994	-1.1229	0.0438	0.0442	0.9998	-0.7144
JNPR	0.0008	0.0028	0.999	$-1.9607^{\star}$	0.0039	0.0041	0.9999	-1.4852	0.0028	0.0028	1.0001	-0.576
LLY	0.0146	0.0375	0.9883	$-1.9673^{\star}$	0.0547	0.0571	0.9987	-0.6791	0.0913	0.0925	0.9994	-0.8131
LMT	0.0253	0.2329	0.8871	-0.9892	0.0695	0.1133	0.9762	-1.0003	0.122	0.1285	0.9963	-0.8182
MRK	0.0176	0.1392	0.9361	-1.0323	0.0274	0.047	0.9899	$-1.6367^{\star}$	0.0277	0.0315	0.9981	-1.0135
MRO	0.1765	0.1785	0.9988	-1.3789	0.2032	0.2044	0.9992	-1.067	0.0621	0.0621	1.0065	-0.3611
PNC	0.0132	0.0352	0.9888	-0.5806	0.0557	0.0845	0.9846	$-1.7946^{\star}$	0.1676	0.1854	0.9893	-0.735
T	0.0905	0.4569	0.7727	$-1.9418^{\star}$	0.1907	0.2602	0.9561	$-1.834^{\star}$	0.2246	0.2408	0.9895	-0.8894
TXN	0.0334	0.0567	0.9879	$-1.7485^{\star}$	0.1	0.103	0.9983	-1.0068	0.1327	0.1354	0.9984	-1.2799
UNH	0.2583	0.5123	0.8109	$-1.8337^{\star}$	0.0287	0.057	0.9853	$-1.7354^{\star}$	0.0059	0.0119	0.997	-1.0408
WFC	0.1116	0.1631	0.9706	$-1.8219^{\star}$	0.2819	0.2939	0.9916	-1.3296	0.2621	0.2639	0.9987	-0.9964
WMB	0.0361	0.0464	0.9946	-0.9483	0.1304	0.1317	0.9993	-0.7991	0.2875	0.2906	0.9978	-1.0253

Table 6: The table shows the R^2 from HAR-RV-CJ (RV-CJ-R^2) and HAR-RV-CJP (RV-CJP-R^2) models; the RMSE losses computed as the ratio between the RMSE of the HAR-RV-CJP model and that of the benchmark model HAR-RV-CJ. DM test represents the t-statistics from Diebold–Mariano–West tests of equal predictive accuracy between HAR-RV-CJ and HAR-RV-CJP. A t-statistic greater than 1.64 in absolute value indicates a rejection of the null hypothesis of equal predictive accuracy at the 10% significance level. These statistics are marked with an asterisk. The sign of the t-statistics indicates which forecast performed better for each loss function: a negative t-statistic indicates that the HAR-RV-CP produced larger average loss than the HAR-RV-CJP forecast, while a positive sign indicates the opposite. HAR-RV-CJ is as in equation (25) while HAR-RV-CJP is presented in equation (26).

	$h=1$				$h=5$				$h=22$			
	RV-CJ R^2	RV-CJP R^2	RMSE	DM test	RV-CJ R^2	RV-CJP R^2	RMSE	DM test	RV-CJ R^2	RV-CJP R^2	RMSE	DM test
ACN	0.3316	0.3316	1.0781	-0.1482	0.4936	0.4962	0.9974	-1.0399	0.3963	0.4018	0.9955	-1.1287
BAC	0.5336	0.5344	0.9991	-0.4402	0.6694	0.6699	0.9992	-0.4979	0.6872	0.6894	0.9965	-0.9332
BAX	0.2008	0.201	0.9999	-0.7259	0.3108	0.3593	0.9642	-0.87	0.2018	0.2224	0.987	-0.97
BHP	0.6322	0.6329	0.9991	-0.8134	0.7247	0.7257	0.9982	-0.7741	0.6056	0.6083	0.9966	-1.0683
C	0.5035	0.506	0.9975	-0.5346	0.5029	0.5074	0.9955	-0.7319	0.6104	0.613	0.9966	-0.921
GE	0.007	0.0107	0.9981	-0.5053	0.0056	0.0058	0.9999	-0.3889	0.0072	0.0072	1.097	-0.2458
GLW	0.4016	0.4019	0.9998	-0.7233	0.5629	0.5647	0.9979	-0.7181	0.5082	0.5125	0.9956	-0.9619
IBM	0.546	0.5506	0.9949	-1.0797	0.6609	0.6643	0.995	-1.532	0.4601	0.4676	0.993	-1.4372
JNJ	0.292	0.2922	0.9998	-0.1909	0.4215	0.4217	0.9998	-0.3297	0.2697	0.2706	0.9993	-0.6453
JNPR	0.4065	0.4066	1.021	-0.4443	0.5406	0.5412	0.9993	-1.3214	0.4292	0.431	0.9984	$-1.6992^{\star}$
LLY	0.4587	0.459	0.9998	-0.3283	0.5659	0.567	0.9987	-0.9487	0.3525	0.3542	0.9987	-0.8183
LMT	0.5741	0.6191	0.9457	$-1.7707^{\star}$	0.6556	0.6645	0.987	$-1.6815^{\star}$	0.4513	0.4543	0.9973	-1.1036
MRK	0.3054	0.4979	0.8502	$-1.8592^{\star}$	0.4551	0.4817	0.9752	$-1.9397^{\star}$	0.3824	0.3886	0.9949	-0.9557
MRO	0.5307	0.5312	0.9995	-0.6719	0.5999	0.6017	0.9978	-0.792	0.4449	0.4457	0.9993	-0.8505
PNC	0.1677	0.1792	0.993	-0.4936	0.4081	0.4223	0.9879	-0.9244	0.5654	0.5684	0.9966	-1.2151
T	0.4947	0.6906	0.7824	$-1.7581^{\star}$	0.6473	0.6833	0.9476	$-1.8096^{\star}$	0.5414	0.5515	0.9889	-1.0348
TXN	0.544	0.5463	0.9975	-0.9561	0.6515	0.6518	0.9996	-0.6114	0.5458	0.5494	0.996	$-1.8541^{\star}$
UNH	0.2591	0.5124	0.8112	-1.0336	0.0314	0.0595	0.9854	-1.0351	0.0139	0.0197	0.997	-1.0405
WFC	0.64	0.6412	0.9984	-0.516	0.699	0.6991	0.9997	-0.3435	0.6196	0.6223	0.9965	-1.2304
WMB	0.3259	0.3267	0.9994	-0.2541	0.4745	0.4746	0.9999	-0.3183	0.4348	0.4358	0.9991	-0.6442

Figures

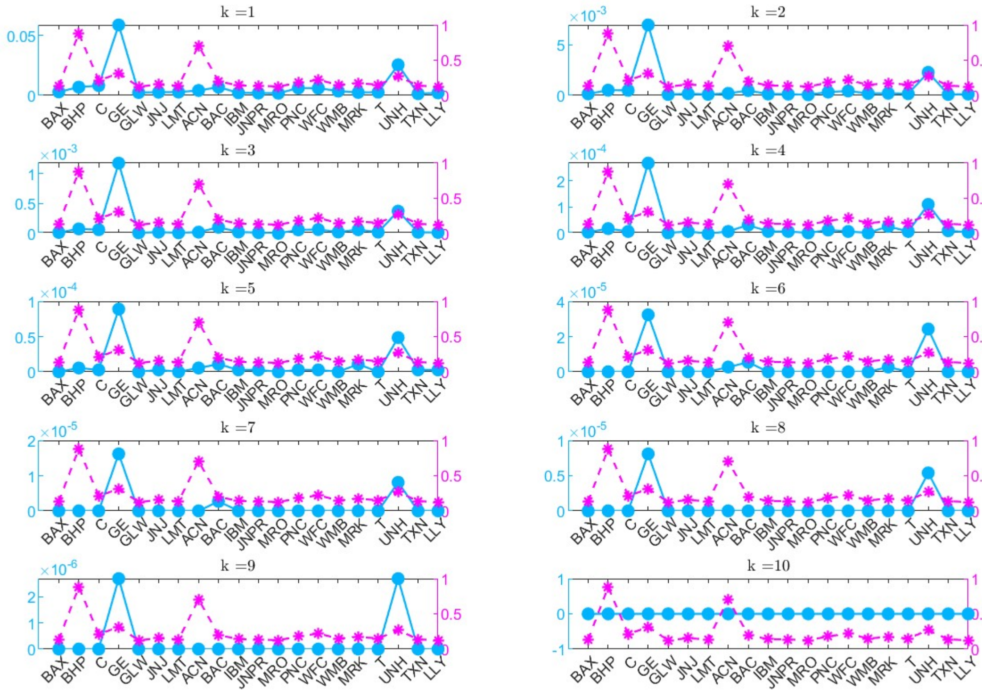

Figure 1