

Forecasting Extreme Returns in Financial Markets: A Discrete Duration Framework

K. BIEN-BARKOWSKA

*Institute of Econometrics, Warsaw School of Economics,
Madalińskiego 6/8, 02-513 Warsaw, Poland*

Doi: [10.12693/APhysPolA.138.48](https://doi.org/10.12693/APhysPolA.138.48)

*e-mail: katarzyna.bien@sgh.waw.pl

We introduce a new dynamic peaks-over-threshold (POT) model for predicting both the timing and the size of extreme losses in financial markets. The novelty of our approach lies in treating the times at which the magnitude of loss exceeds a sufficiently large threshold as a realization of a discrete random variable. The conditional hazard function with respect to the time in-between consecutive extreme losses — and hence, the risk of an extreme loss over the next time unit — is described using two lifetime distributions: the discrete Weibull and the discrete Burr. To consider the clustering of extreme losses, the scale parameters of these discrete distributions vary with time and have the functional form of autoregressive conditional duration (ACD) models. Accordingly, the probability of an extreme loss over the next unit of time depends on times of extreme losses in the past and the period that has elapsed since the last such event. We demonstrate how to predict the value at risk (VaR) from the discrete-duration POT model and empirically confirm that this new approach provides a good alternative to the ACD-POT models outlined in the literature.

topics: extreme losses, peaks-over-threshold (POT) model, ACD models, discrete durations

1. Introduction

The growing importance of risk management is a well-known fact when considering the globalisation of financial markets, the plethora of innovations in derivative securities, and the cycle of equity/bond market collapses. The protection of individual institutions or the entire financial system against unexpected and often catastrophic movements in security prices is essential, which was clearly demonstrated by the financial crisis of 2008–2009. A primary concept used to quantify market risk is the value at risk (VaR), defined as the maximum loss expected on a portfolio over a certain holding period at a given probability (confidence) level. The VaR has become the industry benchmark for risk calculation and has inspired a growing number of statistical models aimed at measuring risk exposures (see [1] for a survey). One of the relatively new approaches to measuring VaR brings together two different strands of the literature. The first strand is the marked point process (MPP) theory that was primarily used in studies on market microstructure and ultra-high-frequency finance [2]. More recently, this approach is recognized as a well-tailored to predict the arrival times of extreme returns. The second strand is the extreme value theory (EVT), especially the peaks-over-threshold (POT) method for describing the magnitude of extreme loss that surpasses a sufficiently large threshold. The resulting dynamic POT model of [3], being inspired by recent advances in the econometrics of

point processes, allows the user to both capture serial dependence in extreme losses and to maintain the merits of the EVT approach. An alternative approach of modeling extreme losses was also proposed in [4], based on threshold interevent time superstatistics.

Chavez-Demoulin et al. [3] uses a parametric specification of conditional intensity function (CIF) that describes the evolution of the point process comprised of times when extreme losses occur. Over the past years, a variety of studies on this subject have emerged with the objective of modeling the self-exciting property of CIF (characterizing the arrival of extreme events). This was intended to form clustering patterns of large losses, in line with the characteristic features of real-world data. To this end, very often the Hawkes process [5–9], or/and the autoregressive conditional duration (ACD) process [10–12] were applied.

We contribute to this body of literature by introducing a dynamic ACD-POT model that not only describes evident clustering patterns in both the timing and the magnitude of extreme losses, but it also takes into account the discreteness of time intervals between subsequent extreme events. Thus, we do not pursue the continuous-time approach of previously mentioned studies. Instead, we build our model using the discrete-time framework where the time variable is treated as a sequence of indivisible time units upon which extreme events can be observed.

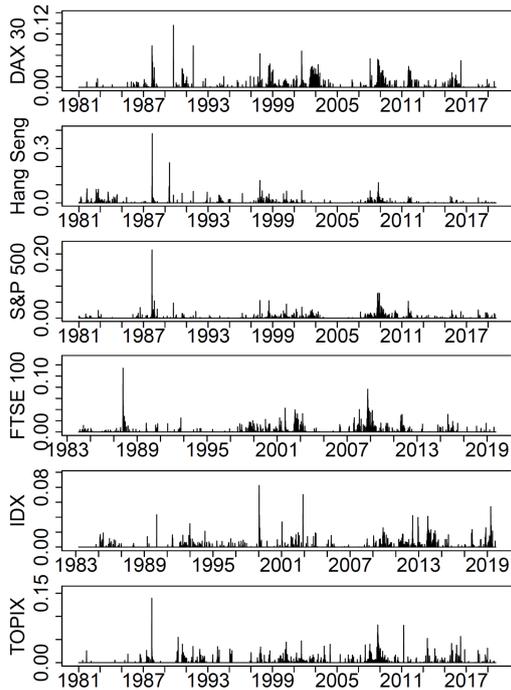


Fig. 1. The times and magnitudes of ground-up threshold exceedances by daily negated log returns on six equity indexes between Jan. 1981 and Aug. 2019. For each stock index, the selected threshold cuts off and exposes 5% of the largest negated log returns (in absolute terms).

The objective of this article is motivated by empirical observation of financial data. Extreme losses are those that exceed (in absolute terms) a sufficiently large threshold u . Figure 1 presents time series of time locations and absolute magnitudes of daily negated log returns that exceed such a large threshold u , for six major international stock indexes (DAX 30, FTSE 100, Hang Seng, IDX, S&P 500, TOPIX) over the period of January 1981–August, 2019 (the daily closing prices for the stock indexes were downloaded from the Datastream database available at the Warsaw School of Economics and provided by the Refinitiv company).

For each of the six stock indexes, the corresponding threshold u has been set to expose only 5% of the largest negated log-returns. To implement the POT method we used the negated log returns, i.e., the log returns pre-multiplied by -1 , and thus the losses are defined as positive variables. In the sequel of the paper we will refer to these extreme losses as “extreme events”. First, we observe that both the magnitude of these losses and the days that they occur are subject to clustering. The latter property can be also deduced from Fig. 2, presenting a quantile-quantile (QQ) plot of inter-threshold-exceedance time spans against an exponential distribution. For all stock indexes, the QQ plots provide evidence that the time location of extreme losses is not

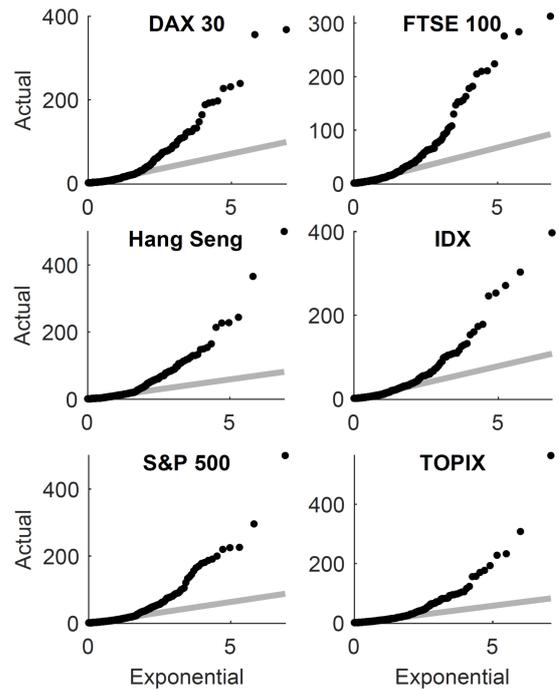


Fig. 2. Exponential quantile-quantile plot of the time intervals (in number of days) between successive extreme losses for six equity indexes between Jan. 1981 and Aug. 2019.

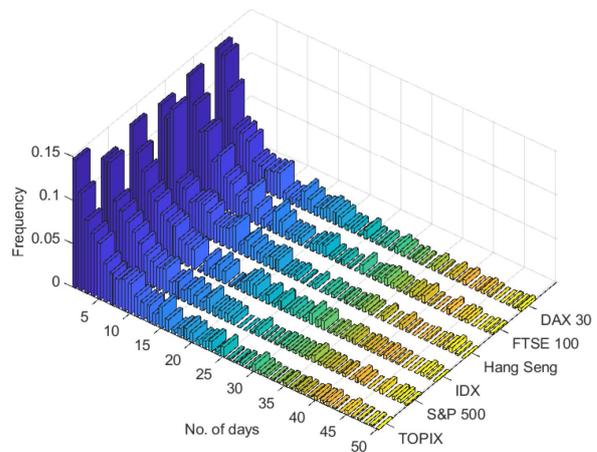


Fig. 3. Frequency histogram for the time intervals (in number of days) between successive extreme losses for six equity indexes between Jan. 1981 and Aug. 2019.

distributed according to the homogeneous Poisson point process. Also interesting, Fig. 3 demonstrates that the daily sampling of data and the clustering phenomenon concentrates the frequency mass of the inter-threshold-exceedance durations (i.e., the time intervals between threshold exceedances) at the lowest values possible (equal to 1–5 days), which emphasizes the discreteness of this data. Indeed, on average about 14% of time spans between any two successive extreme events are only one day long, and about 35% are less or equal to three days.

Intending to allow the data to speak for itself, we aim to model the time interval between successive extreme losses as an integer-valued random variable using discrete lifetime distributions. To this end, we implement the discrete Weibull and the discrete Burr distributions with observation-driven scale parameters as the conditional distributions for inter-exceedance durations, which is in line with the functional form of ACD models. Drawing from the approach of [3], we derive the discrete counterpart to CIF that characterizes the arrival probability of extreme losses and we calculate the one-day-ahead forecasts of VaR. We aim to show that treating inter-exceedance durations not as continuous but as integer-valued variables may result in superior predictions of risk in terms of average VaR levels and obtained model validation results.

Section 2 presents an outline of the point process approach to POT models [3, 5, 7]. Section 3 focuses on the ACD-POT approach [10–12] as an example of the dynamic POT model in continuous time. Section 4 introduces the new ACD-POT model, which was augmented using discretization of conditional distributions for inter-exceedance durations. Finally, Sect. 5 compares different specifications of POT models using empirical data.

2. Point process approach to POT modeling

Let $t \in \mathbb{R}_+$ denote calendar time. According to the point process approach to POT modeling [3], the realization of the stochastic process, $\{t_i, y_i\}_{i \in \{1, 2, \dots, n\}}$, is considered, being a double sequence of: (i) times (i.e., $0 < t_i < t_{i+1}$), when extreme negative returns are observed, and (ii) the corresponding sizes (in absolute terms) y_i of such extreme losses. This can be treated as the sample path of the MPP, reflecting extreme returns occurring at random times t_i . The observed marks of the point process are the realizations y_i of a random variable Y .

The time evolution of the MPP can be characterized by the CIF, i.e., $\lambda(t|\mathcal{F}_t)$:

$$\lambda(t|\mathcal{F}_t) = \lim_{\Delta \downarrow 0} \frac{\Pr[(N(t+\Delta) - N(t)) > 0 | \mathcal{F}_t]}{\Delta}. \quad (1)$$

Note that the CIF is conditional upon \mathcal{F}_t being an information set available at t , consisting of the complete history of event locations and their marks, (i.e., $\mathcal{F}_t \equiv \sigma\{(t_i, y_i), \forall i : t_i \leq t\}$), $N(t)$ denotes a right-continuous counting function (i.e., $N(t) \equiv \max\{n \geq 0 : t_n \leq t\}$). In this framework $\lambda(t|\mathcal{F}_t)$ can be treated as an arrival rate of an extreme loss in the upcoming instant of time given the observed history of the process. If $\lambda(t|\mathcal{F}_t)$ is constant (i.e., $\lambda(t|\mathcal{F}_t) = \lambda$), then the MPP becomes a Poisson MPP with a constant arrival rate λ . However, to capture clustering phenomenon of extreme losses, $\lambda(t|\mathcal{F}_t)$ shall be made dependent on event locations, t_i 's, and y_i 's, where $t_i \leq t$.

The MPP concept provides the grounds for derivation of the conditional VaR measure. From a statistical perspective, the VaR for a confidence level $1 - \tau$, $\tau \in (0, 1)$, constitutes a τ -th quantile of the profit-and-loss distribution. Note, however, that the study relies on the negated log returns — hence, after adopting the notation introduced so far, the conditional VaR at a confidence level $1 - \tau$, forecasted at time t for the next time unit (i.e., $\text{VaR}_{\tau, t+1} \equiv y_{\tau, t+1}$), can be defined as follows:

$$\tau = \Pr(Y_{t+1} > y_{\tau, t+1} | \mathcal{F}_t) = \Pr(Y_{t+1} > u | \mathcal{F}_t)$$

$$\times \Pr(Y_{t+1} - u > y_{\tau, t+1} - u | Y_{t+1} > u; \mathcal{F}_t). \quad (2)$$

Thus, the conditional probability that a return exceeds $y_{\tau, t+1}$, can be decomposed into: (i) the conditional probability that extreme loss occurs in the next unit of time, and (ii) the conditional probability that the magnitude of this loss is larger than $y_{\tau, t+1}$, given that an extreme loss occurred (i.e., a negated return that exceeds u). As proposed in [5], the first component of (2) (i.e., the probability, that an extreme event will be observed over the upcoming time unit — for example, over the next day), can be approximated by the CIF of the point process for extreme losses

$$\Pr(Y_{t+1} > u | \mathcal{F}_t) =$$

$$1 - \Pr[N(t+1) - N(t) = 0 | \mathcal{F}_t] =$$

$$1 - \exp\left(-\int_t^{t+1} \lambda(s|\mathcal{F}_s) ds\right) \approx \lambda(t|\mathcal{F}_t). \quad (3)$$

Accordingly, the point process approach to POT models takes advantage of parametric specifications for both components in (2): (i) $\lambda(t|\mathcal{F}_t)$, that describes timing pattern of an extreme loss; and (ii) the conditional probability that the magnitude of the extreme loss is larger than $y_{\tau, t+1}$ (given that the loss exceeded u). In line with the Pickands–Balkema–de Haan theorem, the latter component can be approximated using the generalized Pareto (GP) distribution [3, 5, 7]. Hence, the conditional probability that a loss magnitude exceeds $y_{\tau, t+1}$ (given that it exceeded u) can be rewritten as follows:

$$\Pr(Y_{t+1} - u > y_{\tau, t+1} - u | Y_{t+1} > u; \mathcal{F}_t) \approx$$

$$1 - F_{\text{GP}}(y_{\tau, t+1} - u | Y_{t+1} > u; \mathcal{F}_t) =$$

$$\left(1 + \xi \frac{y_{\tau, t+1} - u}{\sigma}\right)_+^{-1/\xi}, \quad (4)$$

where $F_{\text{GP}}(\cdot)$ denotes the CDF of the GP distribution with the scale parameter $\sigma \in \mathbb{R}_{>0}$ (the value of the scale parameter can be made time-varying and modeled in an observation-driven fashion) and the shape parameter $\xi \in \mathbb{R}_{\neq 0}$. If $\xi \rightarrow 0$, $F_{\text{GP}}(\cdot)$ tends to the CDF of an exponential distribution. Accordingly, based on (2)–(4), the VaR

at a confidence level $1 - \tau$ predicted at time t for the next unit of time can be derived from the following equality:

$$\tau = \lambda(t|\mathcal{F}_t) \left(1 + \xi \frac{y_{\tau,t+1} - u}{\sigma} \right)_+^{-1/\xi}, \quad (5)$$

as

$$y_{\tau,t+1} = \left[\left(\frac{\tau}{\lambda(t|\mathcal{F}_t)} \right)^{-\xi} - 1 \right] \frac{\sigma}{\xi} + u. \quad (6)$$

Therefore, any upward movement in $\lambda(t|\mathcal{F}_t)$ will result in an instantaneous increase in the forecasted VaR. At the same time, VaR is dependent on the scale σ and shape parameter ξ of the GP distribution for the magnitude of an threshold exceedance.

3. ACD model

A complete description of a point process can be provided by a sequence of event time locations t_i , or a sequence of intervals between them $x_i = t_i - t_{i-1}$, which are also known as inter-event durations or waiting times. Capturing the stylized features that are inherent to financial durations, especially their positive autocorrelation, poses a challenge in the area of empirical finance. One of the seminal solutions to this problem is the ACD model introduced in [13] (for possible applications, see [14–17]), which decomposes a financial duration into two parts: (i) its conditional expectation $\Psi_i = E(x_i|\mathcal{F}_{t_{i-1}})$, and (ii) an independent and identically distributed positive random variable ε_i :

$$x_i = \Psi_i \varepsilon_i, \quad (7)$$

where the density function of ε_i satisfies $f(\varepsilon_i|\mathcal{F}_{t_{i-1}}) = f(\varepsilon_i)$. Hence, the temporal dependence in the duration process is described by the time evolution of conditional expected durations Ψ_i . The literature on high-frequency financial data suggests various functional forms of Ψ_i (for a survey, see [18]). One of the most successful is the logarithmic ACD (LACD) model proposed in [19], which we will use for inter-threshold-exceedance durations in our empirical analysis. However, we additionally enrich its original specification with the explanatory variable $\ln(y_{i-1} - u)$, thus allowing for the potentially negative impact of an extreme loss magnitude on the expected waiting time to the upcoming extreme-loss event as follows:

$$\begin{aligned} \ln(\Psi_i) &= \omega + \beta \ln(\Psi_{i-1}) + \alpha \ln(x_{i-1}) \\ &+ \zeta \ln(y_{i-1} - u). \end{aligned} \quad (8)$$

The observation-driven specification of the LACD model allows us to capture temporal dependence in the duration process, and due to the logarithmic transformation of the conditional expected duration, (8) does not require imposing nonnegativity constraints on the parameter vector.

For parametric ACD models, one must decide on the conditional distribution for inter-event durations. The original ACD model takes advantage of

the Weibull(μ, γ) distribution, where $\mu \in \mathbb{R}_{>0}$ denotes the scale parameter and $\gamma \in \mathbb{R}_{>0}$ is the shape parameter (see [13]). To make the expectation of ε_i equal to 1, the scaling factor for financial durations is defined as $\Phi_{W,i} = \Psi_i [\Gamma(1 + \gamma^{-1})]^{-1}$. Hence, for the Weibull-ACD models, x_i is described by the conditional survival function

$$S_W(x_i|\mathcal{F}_{t_{i-1}}) = \exp\left(-\left(\frac{x_i}{\Phi_{W,i}}\right)^\gamma\right), \quad (9)$$

and its conditional density is given as follows:

$$f_W(x_i|\mathcal{F}_{t_{i-1}}) = \frac{\gamma}{x_i} \left(\frac{x_i}{\Phi_{W,i}}\right)^\gamma \exp\left(-\left(\frac{x_i}{\Phi_{W,i}}\right)^\gamma\right). \quad (10)$$

Another widely used alternative for modeling financial durations is the Burr distribution, i.e., Burr(μ, κ, η) [20]. For the Burr-ACD class of models, the conditional survival function of x_i is given as

$$S_B(x_i|\mathcal{F}_{t_{i-1}}) = \left(1 + \eta \Phi_{B,i}^{-\kappa} x_i^\kappa\right)^{-\eta^{-1}}, \quad (11)$$

where $0 < \eta < \kappa$, and

$$\Phi_{B,i} = \Psi_i \frac{\eta^{1+\kappa^{-1}} \Gamma(\eta^{-1} + 1)}{\Gamma(1 + \kappa^{-1}) \Gamma(\eta^{-1} - \kappa^{-1})}.$$

For $\eta \rightarrow 0$ the Burr distribution converges to the Weibull distribution with the shape parameter κ . In turn, for $\eta \rightarrow 0$ and $\kappa = 1$ its limiting case is an exponential distribution. If $\eta = 1$, the log-logistic distribution is obtained. Accordingly, the conditional density function of x_i in the Burr-ACD model is as follows:

$$f_B(x_i|\mathcal{F}_{t_{i-1}}) = \frac{\kappa \Phi_{B,i}^{-\kappa} x_i^{\kappa-1}}{\left(1 + \eta \Phi_{B,i}^{-\kappa} x_i^\kappa\right)^{\eta^{-1}+1}}. \quad (12)$$

After replacing waiting times $x_i = t_i - t_{i-1}$ with the backward recurrence times (i.e., $x(t) = t - t_{\check{N}(t)}$, where $\check{N}(t) \equiv \max\{n \geq 0 : t_n < t\}$), one can derive the CIF for the ACD models as $\lambda(t|\mathcal{F}_t) = f(x(t)|\mathcal{F}_{t_{\check{N}(t)}})/S(x(t)|\mathcal{F}_{t_{\check{N}(t)}})$. It is important to note here, that the concept of CIF is analogous to the concept of the conditional hazard function, much more widely used term in the ACD literature. In this way, there exists a relationship: $h(x(t)|\mathcal{F}_{t_{\check{N}(t)}}) = \lambda(t|x(t), \mathcal{F}_{t_{\check{N}(t)}})$ [2]. Thus, the CIF according to the Weibull-ACD or the Burr-ACD processes is correspondingly given as follows:

$$\lambda_W(t|\mathcal{F}_t) = \gamma \Phi_{W,\check{N}(t)+1}^{-\gamma} (t - t_{\check{N}(t)})^{\gamma-1}, \quad (13)$$

$$\lambda_B(t|\mathcal{F}_t) = \frac{\kappa \Phi_{B,\check{N}(t)+1}^{-\kappa} (t - t_{\check{N}(t)})^{\kappa-1}}{1 + \eta \Phi_{B,\check{N}(t)+1}^{-\kappa} (t - t_{\check{N}(t)})^\kappa}. \quad (14)$$

Based on (13), or — alternatively — on (14), one obtains the ground intensity needed for the VaR predictions (6). The use of (13)–(14) implies that a clustering of event times might arise due to a co-acting of both factors: (i) instantaneous jumps in the conditional expected duration directly after the arrival of an extreme loss, and (ii) the shape

of the baseline intensity function determined by parameters γ (for the Weibull distribution) or κ and η (for the Burr distribution). More specifically, the CIF corresponding to the Weibull-ACD models decreases monotonically between any two subsequent event times if $\gamma < 1$, is constant if $\gamma = 1$ (the case of the exponential distribution), and increases monotonically if $\gamma > 1$. For the Burr-ACD models, if $\kappa \leq 1$, the CIF between any two subsequent event times is monotonically decreasing and, if $\kappa > 1$, it is non-monotone and has an inverted U shape. Any abrupt changes in the intensity level can only take place instantaneously after extreme losses are observed. The magnitude of these shifts is determined based on the time-varying conditional expected durations, see (8).

4. A discrete duration POT model

We propose two feasible extensions of the ACD models to the discrete time framework. Let us now define the calendar time as a discrete positive variable, i.e., $t \in \mathbf{Z}_{>0}$. Similarly to the case of continuous time, we consider the sample path of the process given by the double sequence $\{t_i, y_i\}_{i=1,2,\dots,n}$ ($0 < t_i < t_{i+1}$). However, once the time is now a discrete value, t_i is a positive integer and, basically, denotes the number of a time unit (i.e., a day) over which the i -th extreme loss was observed. As before, $y_i \in \mathbb{R}_+$ denotes the observed magnitude of an i -th extreme loss. The inter-threshold-exceedance duration X_i , taking the values $x_i = t_i - t_{i-1}$ ($x_i \in \mathbf{Z}_{>0}$), denotes the time interval (in the number of days) between successive losses that surpass the threshold u . It is characterized by the conditional probability $f(x_i|\mathcal{F}_{t_{i-1}}) = \Pr(X_i = x_i|\mathcal{F}_{t_{i-1}})$ and the conditional survival function which — if evaluated at x_i — delivers the probability that X_i is at least equal to x_i , i.e., that the period between extreme losses lasts at least x_i (given $\mathcal{F}_{t_{i-1}}$)

$$S(x_i|\mathcal{F}_{t_{i-1}}) = \Pr(X_i \geq x_i|\mathcal{F}_{t_{i-1}}) = \sum_{k=i}^{\infty} f(x_k|\mathcal{F}_{t_{i-1}}). \quad (15)$$

Let us also introduce integer-valued backward recurrence times $X(t)$ in a close analogue to the case of continuous time (i.e., $x(t) = t - t_{\check{N}(t)}$). The discrete backward recurrence times count the days that have passed since the last extreme loss. In this discrete framework, the conditional hazard function is given as follows:

$$h(x(t)|\mathcal{F}_{t_{\check{N}(t)}}) = \frac{f(x(t)|\mathcal{F}_{t_{\check{N}(t)}})}{S(x(t)|\mathcal{F}_{t_{\check{N}(t)}})}. \quad (16)$$

Accordingly, in a discrete-duration framework, the conditional hazard rate is the conditional probability that the extreme event takes place over the day t , given \mathcal{F}_{t-1} (i.e., given the information set up to (and including) day No. $t-1$). For example, the value of the hazard function corresponding to an extreme loss just one

day after the previous one yields the following: $\Pr[X(t) = 1|\mathcal{F}_{t_{\check{N}(t)}}]/\Pr[X(t) \geq 1|\mathcal{F}_{t_{\check{N}(t)}}] = \Pr[X(t) = 1|\mathcal{F}_{t_{\check{N}(t)}}]$, the conditional probability of experiencing an extreme loss exactly two days after the preceding one (given that there was no intervening extreme loss in-between) yields: $\Pr[X(t) = 2|\mathcal{F}_{t_{\check{N}(t)}}]/\Pr[X(t) \geq 2|\mathcal{F}_{t_{\check{N}(t)}}]$, etc.

There is a considerable strand of literature on discrete lifetime distributions associated with well-known continuous distributions (for a survey, see [21, 22]). According to one approach, the discrete distribution based on the continuous distribution can be obtained by an appropriate grouping on the time axis (for more about this idea, see [23]). In this setup, the continuous lifetime distribution characterized by the survival function, $S(\cdot)$, may serve as an underlying distribution for the discrete random variable Z . If so, the probability function of Z is given as follows:

$$P(Z = z) = S(z) - S(z + 1), \quad z = 0, 1, 2, \dots \quad (17)$$

Such a discretization of a continuous lifetime distribution maintains the functional form of the survival function. As stated in [24], many characteristics and reliability properties of the discrete distribution based on the underlying continuous distribution, shall, hence, remain unchanged.

Because time intervals between extreme losses can take only positive values (i.e., $x_i = 1, 2, 3, \dots$), while using the discretization technique given in (17), we implement the right-shifted discrete analogues of the Weibull and Burr distributions. Accordingly, for the LACD model, where the durations are (conditionally) discrete Weibull (DW) distributed (i.e., for the discrete-Weibull-LACD model), the conditional probability function for X_i (i.e., $f(x_i|\mathcal{F}_{t_{i-1}}) = \Pr(X_i = x_i|\mathcal{F}_{t_{i-1}})$), can be derived as follows:

$$f_{\text{DW}}(x_i|\mathcal{F}_{t_{i-1}}) = S_W(x_i - 1|\mathcal{F}_{t_{i-1}}) - S_W(x_i|\mathcal{F}_{t_{i-1}}), \quad (18)$$

and for the LACD model, where the durations are (conditionally) discrete Burr (DB) distributed (i.e., for the discrete-Burr-LACD model), the probability function of X_i can be derived as

$$f_{\text{DB}}(x_i|\mathcal{F}_{t_{i-1}}) = S_B(x_i - 1|\mathcal{F}_{t_{i-1}}) - S_B(x_i|\mathcal{F}_{t_{i-1}}). \quad (19)$$

The conditional survival functions, $S(x_i|\mathcal{F}_{t_{i-1}}) = \Pr(X_i \geq x_i|\mathcal{F}_{t_{i-1}})$, are correspondingly given as

$$S_{\text{DW}}(x_i|\mathcal{F}_{t_{i-1}}) = S_W(x_i - 1|\mathcal{F}_{t_{i-1}}), \quad (20)$$

$$S_{\text{DB}}(x_i|\mathcal{F}_{t_{i-1}}) = S_B(x_i - 1|\mathcal{F}_{t_{i-1}}). \quad (21)$$

A discrete analogue of the Weibull distribution was first introduced in [25]. As with the case of the Weibull distribution, it is characterized by one shape and one scale parameter, that in our case has an observation-driven functional form based on (8). Some examples of conditional probability and conditional hazard functions are presented in Fig. 4.

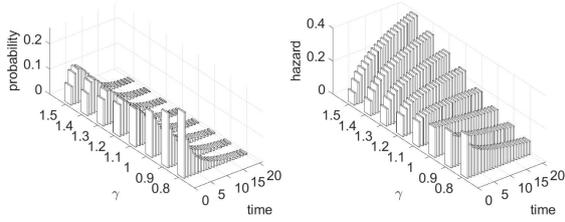


Fig. 4. Discrete-Weibull-LACD model: feasible conditional probability and conditional hazard functions evaluated at $\Psi_i = 5$.

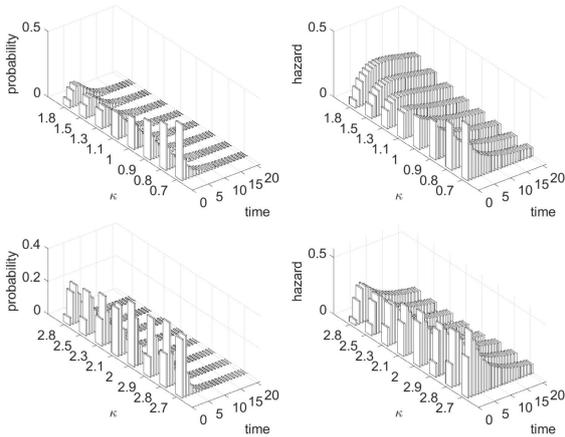


Fig. 5. Discrete-Burr-LACD model: feasible conditional probability and conditional hazard functions evaluated at $\Psi_i = 5$ for $\eta = 0.3$ (upper panel) and $\eta = 1.5$ (lower panel).

Feasible shapes of these functions were plotted for $\Psi_i = 5$. Similar to the case of the Weibull distribution, the hazard function for the discrete Weibull distribution can be monotonically increasing for $\gamma > 1$ and decreasing for $\gamma < 1$. For $\gamma = 1$ the hazard function is flat, and one yields the geometric distribution as a discrete analogue to the exponential distribution.

A discrete counterpart to the Burr distribution was examined in [24]. When used as the conditional distribution of time intervals between extreme losses, it has a time-varying scale parameter and two constant shape parameters κ and η . Figure 5 illustrates several feasible shapes of conditional probability and hazard functions. The discrete Burr distribution allows us to capture a wider plethora of clustering patterns than the discrete Weibull distribution. More specifically, the hazard can be monotonically decreasing for $\kappa \leq 1$ and of inverted U-shape for $\kappa > 1$.

Our study aims to account for the temporal dependence in both x_i , and $y_i - u$. To fulfill the latter objective, we allow for the time-varying scale parameter of the GP distribution describing the magnitudes of extreme losses over the threshold u . The scale parameter σ (4), will be thus specified as a linear function of the hazard function $h(x_i|\mathcal{F}_{t_{i-1}})$ as follows:

$$\sigma_i = \omega_s + \zeta_s h(x_i|\mathcal{F}_{t_{i-1}}), \quad (22)$$

where $\omega_s \in \mathbb{R}_{>0}$, $\zeta_s \in \mathbb{R}_{\geq 0}$, so that $\sigma_i \in \mathbb{R}_{>0}$. Thus, the awaited size of a threshold exceedance shall be proportional to the conditional probability of experiencing such an extreme loss over day t_i . For positive value of ζ_s , the observed magnitude of the extreme losses shall, hence, increase in periods where the frequency of extreme losses rises. The conditional density for the magnitude of loss over the threshold u (i.e., the threshold exceedance $y_i - u$) is given by the GP density:

$$g(y_i - u|y_i > u, x_i, \mathcal{F}_{t_{i-1}}) = \frac{1}{\sigma_i} \left(1 + \xi \frac{y_i - u}{\sigma_i} \right)_+^{-1/\xi - 1}. \quad (23)$$

The unknown parameters of the model can be estimated using the Maximum Likelihood (ML) method. Note that the parametric specifications of $f(x_i|\mathcal{F}_{t_{i-1}})$ and $g(y_i|x_i, \mathcal{F}_{t_{i-1}})$ have already been introduced – see (10) for the Weibull-LACD model, (12) for the Burr-LACD model, (18) for the discrete-Weibull-LACD model, and (19) for the discrete-Burr-LACD model when specifying $f(x_i|\mathcal{F}_{t_{i-1}})$. Based on the realized sample path $\{x_i, y_i\}_{i=1, \dots, n}$, the log likelihood function of the discrete-duration LACD-POT model can be derived as

$$\mathcal{L} = \sum_{i=1}^n \left[\ln \left(f(x_i|\mathcal{F}_{t_{i-1}}) \right) + \ln \left(g(y_i - u|y_i > u, x_i, \mathcal{F}_{t_{i-1}}) \right) \right] \quad (24)$$

Note that (18)–(21) provide the grounds for deriving the parametric form of a conditional hazard function in line with (16). The conditional hazard function corresponds to the conditional probability of an extreme event over the t -th day, given information set up to $t - 1$, namely $\Pr(Y_t > u|\mathcal{F}_{t-1}) = h(x(t)|\mathcal{F}_{t_{\bar{N}(t)}})$. Hence, at $(t-1)$, the one-day-ahead forecast of VaR can be derived as follows:

$$y_{\tau, t} = \left[\left(\frac{\tau}{h(x(t)|\mathcal{F}_{t_{\bar{N}(t)}})} \right)^{-\xi} - 1 \right] \frac{\sigma_{N(t)}}{\xi} + u.$$

5. Estimation results and diagnostic checks

For each of the stock indexes and the threshold \hat{u} exposing the 5% largest negated returns we estimated four specifications of the LACD-POT models. These are, namely the Weibull-LACD-POT and the Burr-LACD-POT model, as well as their two discrete counterparts: the discrete-Weibull-LACD-POT and the discrete-Burr-LACD-POT model. The threshold u was set equal to the 95%-quantile of negated returns in order to achieve a reasonable compromise between the sufficient amount of observations for the robustness of estimates, and the reliable assumption of the GP distribution for excessive extreme losses. The latter has also been validated with the D-test [22].

TABLE I

ML parameter estimates of the discrete-Burr-LACD POT model. T-ratios given in brackets.

	DAX 30	FTSE 100	Hang Seng	IDX	S&P 500	TOPIX
Model for the probability of threshold exceedances						
ω	9.186 (3.750)	0.694 (4.089)	0.940 (5.011)	0.532 (2.839)	0.368 (3.779)	0.985 (2.299)
α	0.318 (7.281)	0.230 (5.081)	0.297 (5.066)	0.221 (5.221)	0.260 (6.571)	0.236 (3.663)
β	0.609 (11.643)	0.520 (6.520)	0.414 (2.980)	0.597 (7.133)	0.621 (11.075)	0.432 (2.222)
ζ	-0.081 (-2.542)	-0.196 (-4.027)	-0.178 (-3.543)	-0.116 (-2.457)	-0.126 (-4.430)	-0.100 (-2.466)
κ	0.869 (16.172)	0.900 (14.773)	0.863 (13.144)	0.794 (12.733)	0.930 (15.821)	0.864 (15.728)
η	0.264 (2.794)	0.334 (3.232)	0.313 (2.532)	0.253 (1.501)	0.255 (2.686)	0.349 (3.489)
Model for the sizes of threshold exceedances						
ω_s	0.515 (8.599)	0.298 (6.383)	0.480 (5.529)	0.253 (5.588)	0.282 (5.609)	0.346 (5.429)
ζ_s	4.736 (6.329)	5.532 (7.443)	6.867 (6.890)	3.636 (5.790)	5.189 (7.255)	4.845 (6.139)
ξ	0.103 (1.915)	0.089 (1.448)	0.245 (3.846)	0.186 (2.926)	0.127 (2.505)	0.140 (2.784)

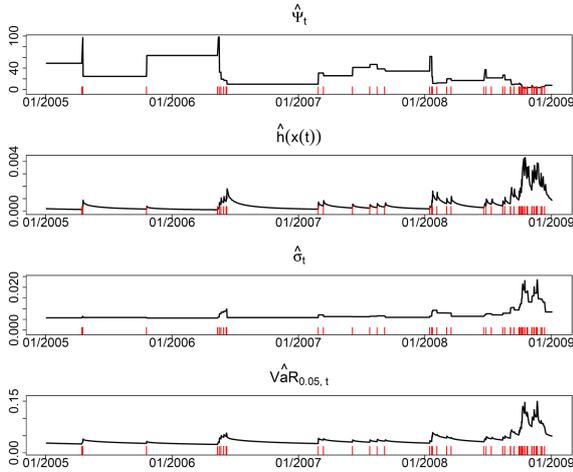


Fig. 6. Discrete-Burr-LACD-POT estimates for the DAX index in 2005-2008. Vertical lines denote days of extreme losses.

For both scale factors in all of considered specifications we relied on the same functional form: (i) Ψ_i for the conditional distribution of inter-threshold-exceedance durations (8), and (ii) σ_i for the conditional distribution of threshold exceedance magnitudes (22), where the functional form of $h(x_i|\mathcal{F}_{t_{i-1}})$ was derived in line with the choice of the conditional distribution for the durations. The data set was split into an in-sample period (Jan. 1981–Dec. 2015) used for model estimation and an out-of-sample period (Jan. 2016–Sep. 2019) serving only evaluative purposes. Table I presents

the estimation results corresponding to the discrete-Burr-LACD-POT model. The log-likelihood function was pre-programmed and maximised using the BFGS algorithm of the constrained maximum likelihood (CML) application of the Gauss mathematical and statistical system. With some rare exceptions, all the parameter estimates are statistically different from zero. Apart from positive and statistically significant coefficients $\hat{\alpha}$ and $\hat{\beta}$ that reflect positive autocorrelation in durations, we document a negative impact of the magnitude of previously observed extreme loss on the waiting time to the next extreme event ($\hat{\zeta} < 0$) for all stock indexes. Hence, large magnitudes of losses tend to increase the awaited arrival rate of subsequent extreme losses. We obtained decreasing hazard functions as $\hat{\kappa} < 1$, which remains in line with a strong clustering of extreme losses. Moreover, we can see that for all equity indexes under investigation, the awaited magnitude of extreme losses is boosted by the probability of experiencing a huge loss over that day ($\hat{\zeta}_s > 0$).

Figure 6 illustrates the functioning of the discrete-Burr-LACD-POT model based on its estimated time-varying components. For the sake of lucidity, the time evolution of $\hat{\Psi}_t$, $\hat{h}(x(t))$, and $\hat{\sigma}_t$ was plotted only for the DAX index over a short five-year subperiod between Jan. 2005 and Dec. 2008. The obtained plot of $\hat{\Psi}_t$ is a step function that features jumps only at days directly following extreme losses. In turbulent periods, when the distribution of extreme-loss days is dense, the scale parameter of the discrete Burr

BIC for the LACD-POT models.

TABLE II

Conditional distribution	DAX 30	FTSE 100	Hang Seng	IDX	S&P 500	TOPIX
Weibull	9.431	9.066	10.021	8.700	8.974	9.054
Burr	9.336	8.954	9.910	8.634	8.897	8.903
discrete Weibull	9.196	8.850	9.789	8.490	8.792	8.801
discrete Burr	9.186	8.832	9.779	8.497	8.785	8.777

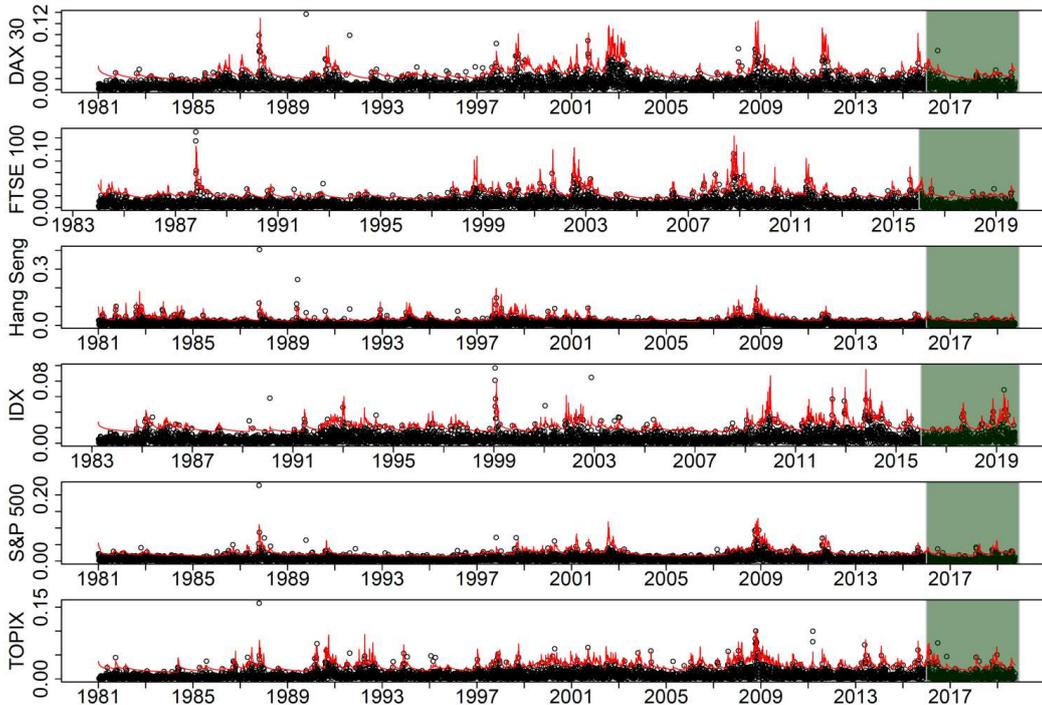


Fig. 7. VaR forecasts for $\tau = 0.01$ from the discrete-Burr-LACD-POT model and daily negated log returns from six equity indexes between Jan. 1981 and Sep. 2019. Shaded areas denote the out-of-sample periods.

distribution wanders down in stepwise motion whereas very long inter-exceedance durations induce a jump in the upward direction. The plot of the hazard function \hat{h}_t combines the evolution of $\hat{\Psi}_t$ with the downward sloping baseline hazard in-between any two subsequent extreme-loss days. Hence, the conditional probability of observing an extreme loss over the next day climbs if driven by preceding extreme losses — both in terms of their times and magnitudes. It also monotonically decays instantaneously after an extreme loss is experienced, which causes a clustering of the days of extreme loss. The scale parameter of the GP distribution keeps changing its level in reaction to upward and downward jumps of the hazard function (i.e., only at days following extreme losses). All of these factors contribute to the time evolution of the VaR, that was depicted in the lowest panel of Fig. 6.

The goodness-of-fit tests confirmed the validity of the conditional GP distribution for the magnitudes of threshold exceedances. We relied on two diagnostic tools: the D-test proposed in [26] and

the χ^2 test for uniformity of probability integral transforms (PIT). To conserve space, we do not report the detailed results here. However, both of these tests were not able to reject the null hypothesis that the conditional distribution of the threshold exceedances corresponds to the GP distribution.

In order to determine the superiority of the discrete-Burr-LACD-POT model, Table II compares the Bayesian information criteria (BIC) for all models under investigation. For five out of six stock indexes, the BIC achieved the smallest value for the discrete-Burr-LACD model, pinpointing the best goodness of fit. The only exception is the IDX, where the discrete-Weibull-LACD-POT model rendered the best fit among the four specifications. Apart from the case of the IDX, the discrete-Weibull-LACD-POT model turned out to be the second best in our study. There is also a considerable difference in the BIC value among the models based on the continuous distributions for inter-exceedance durations and those based on discrete distributions, whereas the latter solution allows for a better goodness of fit.

TABLE III

In-sample results from VaR backtesting procedures. The columns entitled “UC” and “CC” include the p-values of the UC and CC tests, respectively. The average VaR levels correspond to the negated log returns pre-multiplied by 100.

	τ	Weibull			Burr			Discrete Weibull			Discrete Burr		
		LACD-POT			LACD-POT			LACD-POT			LACD-POT		
		UC	CC	$\overline{\text{VaR}}_\tau$	UC	CC	$\overline{\text{VaR}}_\tau$	UC	CC	$\overline{\text{VaR}}_\tau$	UC	CC	$\overline{\text{VaR}}_\tau$
DAX 30	0.05	0.979	0.518	2.02	0.721	0.053	2.00	0.945	0.790	1.99	0.906	0.906	1.98
	0.025	0.723	0.532	2.55	0.673	0.079	2.50	0.752	0.864	2.50	0.961	0.423	2.48
	0.01	0.035	0.035	3.32	0.151	0.036	3.21	0.035	0.108	3.23	0.077	0.076	3.19
	0.005	0.048	0.142	3.95	0.691	0.051	3.80	0.099	0.257	3.84	0.242	0.504	3.78
	0.0025	0.293	0.576	4.63	0.293	0.049	4.44	0.293	0.576	4.48	0.080	0.216	4.40
	0.001	0.462	0.763	5.61	0.702	0.929	5.35	0.462	0.763	5.42	0.702	0.929	5.31
FTSE 100	0.05	0.704	0.370	1.63	0.275	0.033	1.62	0.527	0.191	1.61	0.195	0.406	1.61
	0.025	0.233	0.245	2.02	0.206	0.120	1.98	0.446	0.625	1.99	0.182	0.300	1.98
	0.01	0.540	0.104	2.57	0.868	0.156	2.50	0.468	0.728	2.52	0.402	0.078	2.49
	0.005	0.055	0.003	3.03	0.666	0.366	2.93	0.114	0.082	2.95	0.452	0.277	2.91
	0.0025	0.265	0.040	3.52	0.846	0.114	3.39	0.175	0.026	3.41	0.265	0.040	3.35
	0.001	0.630	0.011	4.23	0.093	0.001	4.04	0.630	0.011	4.06	0.630	0.011	3.98
Hang Seng	0.05	0.026	0.007	2.40	0.015	0.001	2.39	0.018	0.060	2.40	0.009	0.028	2.40
	0.025	0.129	0.004	3.01	0.369	0.086	2.96	0.045	0.132	2.97	0.084	0.178	2.95
	0.01	0.185	0.175	3.98	0.700	0.232	3.86	0.122	0.278	3.89	0.151	0.331	3.85
	0.005	0.184	0.141	4.89	0.391	0.270	4.69	0.242	0.180	4.75	0.184	0.141	4.67
	0.0025	0.698	0.109	5.96	0.545	0.833	5.67	0.545	0.088	5.76	0.409	0.068	5.65
	0.001	0.777	0.961	7.69	0.777	0.961	7.23	0.549	0.836	7.40	0.549	0.836	7.22
IDX	0.05	0.003	0.000	1.49	0.054	0.000	1.45	0.000	0.002	1.48	0.001	0.002	1.47
	0.025	0.070	0.001	1.79	0.416	0.000	1.73	0.095	0.232	1.77	0.082	0.094	1.75
	0.01	0.083	0.015	2.27	0.625	0.436	2.17	0.295	0.539	2.21	0.295	0.539	2.18
	0.005	0.085	0.063	2.69	0.289	0.196	2.56	0.221	0.155	2.59	0.221	0.155	2.55
	0.0025	0.089	0.013	3.18	0.327	0.051	3.00	0.327	0.051	3.03	0.145	0.022	2.98
	0.001	0.850	0.016	3.94	0.850	0.016	3.69	0.850	0.016	3.71	0.850	0.016	3.64
S&P 500	0.05	0.010	0.008	1.64	0.013	0.000	1.63	0.008	0.028	1.63	0.012	0.010	1.63
	0.025	0.084	0.120	2.02	0.932	0.051	1.99	0.270	0.483	1.99	0.531	0.420	1.98
	0.01	0.436	0.107	2.58	0.807	0.183	2.52	0.726	0.928	2.52	0.269	0.245	2.50
	0.005	0.582	0.860	3.04	0.242	0.504	2.97	0.691	0.924	2.96	0.691	0.924	2.93
	0.0025	0.293	0.576	3.55	0.545	0.833	3.46	0.130	0.319	3.43	0.201	0.441	3.40
	0.001	0.777	0.961	4.29	0.365	0.663	4.19	0.549	0.836	4.12	0.777	0.961	4.09
TOPIX	0.05	0.114	0.002	1.84	0.166	0.000	1.80	0.198	0.298	1.82	0.138	0.059	1.81
	0.025	0.168	0.034	2.29	0.855	0.004	2.23	0.334	0.568	2.26	0.369	0.086	2.24
	0.01	0.502	0.122	2.97	0.859	0.225	2.87	0.573	0.448	2.91	0.375	0.324	2.87
	0.005	0.032	0.026	3.54	0.805	0.443	3.42	0.070	0.057	3.46	0.184	0.141	3.40
	0.0025	0.080	0.013	4.18	0.515	0.144	4.05	0.545	0.088	4.06	0.861	0.126	3.99
	0.001	0.549	0.026	5.12	0.365	0.025	5.00	0.549	0.026	4.93	0.777	0.024	4.86

The superiority of the discrete-distribution-LACD-POT models has also been evidenced using VaR backtesting methods. Two widely recognized VaR verification procedures were performed: the unconditional coverage (UC) test [37] and the conditional coverage (CC) test [1]. Note that $\text{VaR}_{\tau,t+1}$ is intended to reflect a ‘promise’ that the actual return at $t+1$ will be worse than $\text{VaR}_{\tau,t+1}$ for an average $\tau \times 100\%$ of the time. Both backtesting methods rely on the comparison of ex ante (i.e.,

one-day-ahead) VaR predictions and actual returns over a given period of time. A VaR exceedance occurs when an actual loss is larger than the VaR forecasted for that day. The null hypothesis of the UC test states that the fraction of VaR exceedances (i.e., actual losses that exceeded the one-step-ahead forecasts of VaR) according to a risk model, matches the assumed VaR probability level τ (i.e., the coverage probability for VaR). The CC test additionally points toward the issue of independence by focusing

TABLE IV

Out-of-sample results from VaR backtesting procedures (Symbol “-” means that the test could not be performed). The columns entitled “UC” and “CC” include the p-values of the UC and CC tests, respectively. The average VaR levels correspond to the negated log returns pre-multiplied by 100.

	τ	Weibull			Burr			Discrete Weibull			Discrete Burr		
		LACD-POT			LACD-POT			LACD-POT			LACD-POT		
		UC	CC	$\overline{\text{VaR}}_\tau$	UC	CC	$\overline{\text{VaR}}_\tau$	UC	CC	$\overline{\text{VaR}}_\tau$	UC	CC	$\overline{\text{VaR}}_\tau$
DAX 30	0.05	0.070	0.167	1.82	0.567	0.348	1.71	0.237	0.415	1.78	0.180	0.231	1.73
	0.025	0.066	0.184	2.29	0.613	0.191	2.14	0.108	0.274	2.23	0.349	0.449	2.16
	0.01	0.091	0.239	2.97	0.802	0.969	2.76	0.091	0.239	2.88	0.192	0.427	2.78
	0.005	0.679	0.918	3.52	0.679	0.918	3.27	0.357	0.654	3.42	0.679	0.918	3.29
	0.0025	0.770	0.958	4.13	0.294	0.577	3.82	0.770	0.958	3.99	0.294	0.577	3.84
	0.001	0.981	1.000	5.00	0.981	1.000	4.62	0.981	1.000	4.82	0.981	1.000	4.63
FTSE 100	0.05	0.005	0.001	1.52	0.021	0.000	1.43	0.008	0.002	1.48	0.021	0.006	1.44
	0.025	0.108	0.150	1.87	0.168	0.039	1.73	0.066	0.091	1.82	0.168	0.039	1.74
	0.01	0.348	0.068	2.38	0.557	0.118	2.15	0.557	0.841	2.30	0.557	0.841	2.17
	0.005	0.357	0.654	2.80	0.679	0.918	2.51	0.357	0.654	2.69	0.679	0.918	2.51
	0.0025	0.770	0.958	3.25	0.730	0.942	2.88	0.730	0.942	3.10	0.730	0.942	2.88
	0.001	0.981	1.000	3.89	0.365	0.663	3.42	0.981	1.000	3.69	0.981	1.000	3.41
Hang Seng	0.05	0.003	0.012	2.08	0.003	0.012	2.04	0.005	0.020	2.07	0.001	0.004	2.06
	0.025	0.169	0.232	2.55	0.066	0.092	2.48	0.250	0.334	2.51	0.169	0.232	2.49
	0.01	0.036	0.110	3.30	0.193	0.428	3.18	0.036	0.110	3.22	0.036	0.110	3.17
	0.005	0.032	0.100	3.99	0.032	0.100	3.82	0.032	0.100	3.87	0.032	0.100	3.80
	0.0025	0.295	0.577	4.81	0.295	0.577	4.57	0.295	0.577	4.65	0.295	0.577	4.55
	0.001	0.981	1.000	6.14	0.981	1.000	5.78	0.981	1.000	5.91	0.981	1.000	5.75
IDX	0.05	0.572	0.142	1.53	0.307	0.014	1.51	0.239	0.136	1.52	0.307	0.059	1.52
	0.025	0.745	0.084	1.86	0.471	0.017	1.81	0.772	0.959	1.83	0.772	0.959	1.82
	0.01	0.091	0.240	2.36	0.487	0.785	2.28	0.559	0.843	2.30	0.804	0.970	2.28
	0.005	0.358	0.656	2.81	0.957	0.999	2.69	0.138	0.333	2.71	0.358	0.656	2.69
	0.0025	-	-	3.33	0.295	0.577	3.17	0.295	0.577	3.19	-	-	3.15
	0.001	-	-	4.14	-	-	3.90	-	-	3.91	-	-	3.86
S&P 500	0.05	0.000	0.000	1.76	0.000	0.000	1.60	0.000	0.000	1.71	0.000	0.000	1.63
	0.025	0.066	0.096	2.19	0.248	0.071	1.92	0.108	0.157	2.11	0.108	0.157	1.97
	0.01	0.557	0.118	2.83	0.802	0.173	2.40	0.348	0.068	2.68	0.802	0.173	2.47
	0.005	0.679	0.918	3.37	0.959	0.049	2.81	0.959	0.049	3.17	0.959	0.049	2.89
	0.0025	0.294	0.577	3.95	0.152	0.018	3.26	0.770	0.958	3.69	0.361	0.659	3.35
	0.001	0.981	1.000	4.80	0.101	0.261	3.92	0.981	1.000	4.45	0.101	0.261	4.02
TOPIX	0.05	0.180	0.353	1.82	0.135	0.292	1.83	0.469	0.565	1.80	0.135	0.292	1.82
	0.025	0.907	0.889	2.28	0.472	0.573	2.26	0.604	0.828	2.25	0.768	0.791	2.26
	0.01	0.348	0.068	2.95	0.489	0.228	2.91	0.557	0.118	2.91	0.941	0.220	2.90
	0.005	0.959	0.999	3.53	0.625	0.888	3.47	0.959	0.999	3.46	0.959	0.999	3.44
	0.0025	0.361	0.659	4.17	0.361	0.659	4.10	0.730	0.942	4.06	0.730	0.942	4.03
	0.001	0.981	1.000	5.11	0.981	1.000	5.06	0.365	0.663	4.94	0.981	1.000	4.91

on the Markov property of VaR exceedances. Thus, according to the null of the CC test, the proportion of losses that are larger than the VaR according to a risk model is equal to the coverage probability for VaR (τ) and the VaR exceedances are independent over time.

We forecasted one-day-ahead VaR at different levels of τ (i.e., $\tau \in \{0.05, 0.025, 0.01, 0.005, 0.0025, 0.001\}$), for each of four LACD-POT models, both in-sample and out-of-sample. Figure 7

illustrates the plots of obtained VaR forecasts at $\tau = 0.01$. Table III presents the p-values from implementations of the UC and CC tests as well as the average VaR levels for the negated log returns pre-multiplied by 100, for the in-sample period ($\overline{\text{VaR}}_\tau$), and Table IV covers similar information for the out-of-sample period. The in-sample backtesting results do not stay in favour of the Weibull-LACD-POT model. Using 0.05 significance level for the test statistics, the VaR numbers

based on Weibull-LACD-POT models are misspecified in six instances of UC test implementations (out of 36 in general), because the null of the correct coverage had to be rejected. In the case of the Burr-LACD-POT model the null had to be rejected only two times. Generally, based on the results of the CC test, the discrete-Weibull-LACD-POT models render VaR numbers with a much better goodness of fit than the standard Weibull-LACD-POT models. Similarly, the discrete-Burr-LACD-POT models turn out to be better than the Burr-LACD-POT models. In the case of the Burr-LACD-POT (Weibull-LACD-POT) models, in sample, in 12 (17) implementations of CC test, p-values were less than 0.05. However for its augmented version that accounts for the discreteness of inter-exceedance durations after implementing the discrete Burr (discrete Weibull) distribution, the null was rejected in only 8 (6) cases. For the out-of-sample period, which is much shorter (it covers only about 3.5 years), the results are not so unambiguous. The CC null had to be rejected in 8 (3) implementations of the test for the Burr-ACD-POT VaR model (Weibull-LACD-POT) and in 5 (4) instances for the discrete Burr-LACD-POT (discrete-Weibull-LACD-POT) models.

A final major observation from our analysis is that in the LACD-POT models augmented with the discrete distributions for inter-exceedance durations the average levels of VaR are evidently lower than the corresponding average VaR levels derived from the standard LACD-POT models. In the case of the discrete Weibull distribution this observation holds true for all stock indexes and for all coverage probability levels τ , both in- and out-of-sample. Accordingly, the better fit of VaR forecasts derived from the augmented LACD-POT model does not come at the cost of higher VaR numbers. This observation also means that the discrete-Weibull-LACD-POT model does not imply higher capital reserves for covering market risk when being used by financial institutions. In the case of the discrete-Burr-LACD-POT models these results are not so clear — the average VaR levels are lower than the corresponding values for DAX 30, FTSE 100 and S&P 500 (in-sample). For the rest of indexes, this property holds true only at low coverage levels (i.e., for $\tau = 0.005$ or lower).

6. Conclusions

Forecasting extreme returns stays at the forefront of both econometrics and econophysics because it allows for effective management of market risk. Adequate econometric tools for modeling extreme losses should account for their clustering in time. This study has proposed a new discrete-duration POT model that is suitably tailored for the integer-valued times between extreme losses. Our model is inspired by the ACD-POT approach used in [10–12]. However, we describe the conditional distribution of

times between extreme losses in a different fashion — we use the discrete Weibull and the discrete Burr distribution. Similarly to the original ACD-POT approach, the conditional distribution of time intervals between extreme losses in our model is time-varying, characterized by an observation-driven scale parameter that accommodates any changes in the arrival rate of extreme losses. The magnitudes of extreme losses are depicted by the conditional GP distribution with the scale parameter that depends on the conditional probability of experiencing an extreme loss over that day. Based on empirical data for six international stock indexes, we compared the goodness of fit for the discrete-Weibull-LACD-POT and discrete-Burr-LACD-POT models with their continuous-duration counterparts. Based on BIC, we provided an evidence of the superiority of the LACD-POT models augmented with the discrete conditional distributions for inter-exceedance durations. The backtesting methods for VaR obtained with discrete-duration LACD-POT model also stayed in its favour. The performance of this novel model sheds new light on econometric models of extreme risk but requires more investigation. Further extensions can focus on different discrete distribution families for inter-exceedance time spans or different specifications of scale parameters. A more thorough exploration is also needed in terms of other types of financial data and a wider range of VaR backtesting methods.

References

- [1] P.F. Christoffersen, *Elements of Financial Risk Management*, Elsevier, 2012.
- [2] N. Hautsch, *Econometrics of Financial High-Frequency Data*, Springer, 2013.
- [3] V. Chavez-Demoulin, A.C. Davidson, J.A. McGill, *Quant. Finance* **5**, 227 (2005).
- [4] M. Denys, M. Jagielski, T. Gubiec, R. Kutner, H.E. Stanley, *Acta Phys. Pol. A* **129**, 913 (2015).
- [5] V. Chavez-Demoulin, J.A. McGill, *J. Bank. Finance* **36**, 3415 (2012).
- [6] O. Grothe, V. Korniichuk, H. Manner, *J. Economet.* **182**, 269 (2014).
- [7] R. Herrera, A.E. Clements, *J. Bank. Finance* **88**, 161 (2018).
- [8] M. Jagielski, R. Kutner, D. Sornette, *Physica A* **483**, 68 (2017).
- [9] T. Stindl, F. Chen, *Extremes* **22**, 705 (2019).
- [10] R. Herrera, *Energy Econ.* **38**, 64 (2013).
- [11] R. Herrera, B. Schipp, *J. Empir. Finance* **23**, 33 (2013).
- [12] R. Herrera, N. González, *Int. J. Forecast* **30**, 477 (2014).
- [13] R.F. Engle, J.R. Russell, *Econometrica* **66**, 1127 (1998).

- [14] K. Bień-Barkowska, *Acta Phys. Pol. A* **127**, 7 (2015).
- [15] K. Bień-Barkowska, *Int. J. Comput. Math.* **94**, 2223 (2017).
- [16] L. Ponta, M. Trinh, M. Roberto, E. Scalas, S. Cincotti, *Physica A* **521**, 173 (2019).
- [17] V. Pyrlík, *Physica A* **392**, 6041 (2013).
- [18] M. Pacurar, *J. Econ. Surv.* **22**, 711 (2008).
- [19] L. Bauwens, P. Giot, *Annales d'Économie et de Statistique* **60**, 117 (2000).
- [20] J. Grammig, K.-O. Maurer, *Economet. J.* **3**, 16 (2000).
- [21] C. Bracquemond, O. Gaudoin, *Int. J. Reliabil. Qual. Saf. Eng.* **10**, 69 (2003).
- [22] S. Chakroborty, *J. Stat. Distrib. Appl.* **2**, 2 (2015).
- [23] R.L. Prentice J.D. Kalbfleisch, *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, 2002.
- [24] H. Krishna, P. S. Pundir, *Stat. Methodol.* **6**, 177 (2009).
- [25] T. Nakagawa, S. Osaki, *IEEE T. Reliab.* **R-24**, 300 (1975).
- [26] M. Fernandes, J. Grammig, *J. Economet.* **127**, 35 (2005).
- [27] P.H. Kupiec, *J. Deriv.* **3**, 73 (1995).