

Herrera, Rodrigo; Schipp, Bernhard

Working Paper

Extreme value models in a conditional duration intensity framework

SFB 649 Discussion Paper, No. 2011-022

Provided in Cooperation with:

Collaborative Research Center 649: Economic Risk, Humboldt University Berlin

Suggested Citation: Herrera, Rodrigo; Schipp, Bernhard (2011) : Extreme value models in a conditional duration intensity framework, SFB 649 Discussion Paper, No. 2011-022, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin

This Version is available at:

<https://hdl.handle.net/10419/56698>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Extreme value models in a conditional duration intensity framework

Rodrigo Herrera*
Bernhard Schipp**



* Universidad de Talca, Chile

** Technische Universität Dresden, Germany

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



Extreme value models in a conditional duration intensity framework

R. Herrera^{a,1,*}, B. Schipp^b

^a*Departamento de Modelación y Gestión Industrial, Facultad de Ingeniería, Universidad de Talca, Camino Los Nihes Km.1- Curicó. Chile*

^b*Faculty of Business and Economics, Technische Universität Dresden, D-01062 Dresden. Germany*

Abstract

The analysis of return series from financial markets is often based on the Peaks-over-threshold (POT) model. This model assumes independent and identically distributed observations and therefore a Poisson process is used to characterize the occurrence of extreme events. However, stylized facts such as clustered extremes and serial dependence typically violate the assumption of independence. In this paper we concentrate on an alternative approach to overcome these difficulties. We consider the stochastic intensity of the point process of exceedances over a threshold in the framework of irregularly spaced data. The main idea is to model the time between exceedances through an Autoregressive Conditional Duration (ACD) model, while the marks are still being modelled by generalized Pareto distributions. The main advantage of this approach is its capability to capture the short-term behaviour of extremes without involving an arbitrary stochastic volatility model or a prefiltration of the data, which certainly impacts the estimation. We make use of the proposed model to obtain an improved estimate for the Value at Risk. The model is then applied and illustrated to transactions data from Bayer AG, a blue chip stock from the German stock market index DAX.

JEL classification: C22; C58 ; F30;

Keywords: Extreme value theory, autoregressive conditional duration, value at risk, self-exciting point process, conditional intensity.

*Corresponding author

Email addresses: rodriherrera@utalca.cl (R. Herrera), Bernhard_Schipp@tu-dresden.de (B. Schipp)

¹The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Ökonomisches Risiko", Humboldt-Universität zu Berlin is gratefully acknowledged. Moreover, Rodrigo Herrera thanks the Fritz Thyssen Stiftung for the financial support via a postdoctoral Fellowship.

1. Introduction

In recent years there has been a noticeable increase in the frequency and impact of extreme events and financial crises. These events range from currency crashes (Asian East 1997, Russia in 1998, Argentina in 2001), to liquidity crises (LTCM in 1998), to stock market crashes (Black Monday in 1987, Dot.com in 2000), to the US subprime market spillovers from 2007 through to 2009. An important aspect of these extreme events is that their impact is exacerbated by simultaneous occurrence in a multiple class of assets.

Critical questions are being asked concerning some of the quantitative methods used in risk management under the Basel II proposals. Why do extreme events occur? What measures are being taken to meet the current and certainly extreme crisis? Can researchers study former extreme events and learn how to avoid them in the future? Both theoretical and more practical oriented questions are on the actual agenda of academics, practitioners and regulators understanding of the dynamics of asset markets under stress (see Embrechts, 2009 and references therein).

The estimation of the Value at Risk (VaR) and related risk measures is a current topic of interest in finance, for which many approaches of varying sophistication have been derived. According to Chavez-Demoulin et al. (2005) two main approaches can be distinguished: the time series and the extreme value approach. The first emphasizes modelling the temporal features (e.g., volatility clustering and fat tails) with ARCH-type and stochastic volatility models. However, the study of extreme dependence may reveal contrasts which are obscured when we only concentrate on examining the conditional second moment of a time series. Interestingly, and unlike the situation for GARCH processes (see Davis and Mikosch, 2009), there is no extremal clustering for stochastic volatility processes in either the light- or heavy-tailed cases. That is, large values of the processes do not come in clusters, which mean that the large sample behaviour of maxima is the same as that of the maxima of the associated iid sequence. On the other hand, Mikosch (2003) showed in a simulation study that for the GARCH case the expected cluster size in a set of various log-return series is smaller than for the fitted GARCH model, i.e., there is less dependence in the tails for the returns and volatilities than for the prescribed GARCH model. While these models imply some information about extreme events, still little is known about the extremes per se.

The extreme value approach makes inference on the VaR using results from Extreme Value Theory (EVT), which only focuses on the tail of the distribution (see Embrechts et al. 1997 for an introduction). The majority of the approaches on EVT for VaR estimation concern the estimation of unconditional quantiles (see for example Danielsson and De Vries, 2000, Coles, 2001 and Cotter and Dowd, 2006). An exception is the work of McNeil and Frey (2000), which addressed the conditional quantile problem and proposed a method for applying EVT to the conditional return

distribution by using a two-stage method, combining GARCH models for forecasting volatility and EVT techniques applied to the residuals from the GARCH analysis. Although this methodology works quite well in practice it has the major drawback addressed by Mikosch (2003). Thus, one should be cautious with the interpretation of the results of this method since there is no theory in the extremal clustering behaviour based on the residuals of a GARCH model.

A novel form to deal with the cluster on extremes is to use a cluster point process version of Peaks Over Threshold (POT) model introduced preliminarily in McNeil et al. (2005); Chavez-Demoulin et al. (2005) and Herrera and Schipp (2009), where the cluster of extreme data are modelled as self-exciting point processes without involving a prefiltration of data. A point process is described as self-exciting when the past evolution impacts the probability of future events. The marks contain the information that is associated with these events. The main characteristic of these models is that the intensity of occurrence of extreme events can depend on past extreme events and the size of the exceedances, allowing more realistic models.

In this paper we concentrate on a different alternative. We model the stochastic intensity of the point process of exceedances within the framework of irregularly spaced data. Contrary to the classical POT methodology, where the time of occurrence of the extreme events is modelled, the methodology proposed models the inter-exceedance times between extreme events. To this end, we use a methodology similar to an Autoregressive Conditional Duration (ACD) model (see Engle and Russell, 1998 for more reference), while the marks still being modelled by generalized Pareto distributions. Like the GARCH models, the ACD models and their alternatives (see Hautsch, 2004; Bauwens and Hautsch, 2009 and Bauwens and Hautsch, 2006) have proven to be very useful in capturing the clustering effects. For this reason, it seems natural to model the cluster behaviour of extreme observations by means of this class of processes.

In the following we will explain our motivation in investigating extreme events in a stock market as a marked point process of exceedances. The classical POT model for iid data assumes that if the threshold u has been chosen highly enough then the exceedances over this threshold, the extreme events, occur in time according to a homogeneous Poisson process. In addition, the size of the excess returns over the threshold, the mark sizes, are independently and identically distributed according to the generalised Pareto distribution (GPD). Figure 1.1 shows in the top panel the negative daily percentage log returns of Bayer shares between 2 January 1990 and 18 January 2008, and the times and sizes of the negative daily percentage log returns exceeding a threshold $u = 1,5$. Observe that this contradicts the classical model assumption of no cluster at the extremes. Indeed, under a homogeneous Poisson process the inter-exceedance times should be independent exponential random variables. The lower left picture shows an exponential probability plot for the

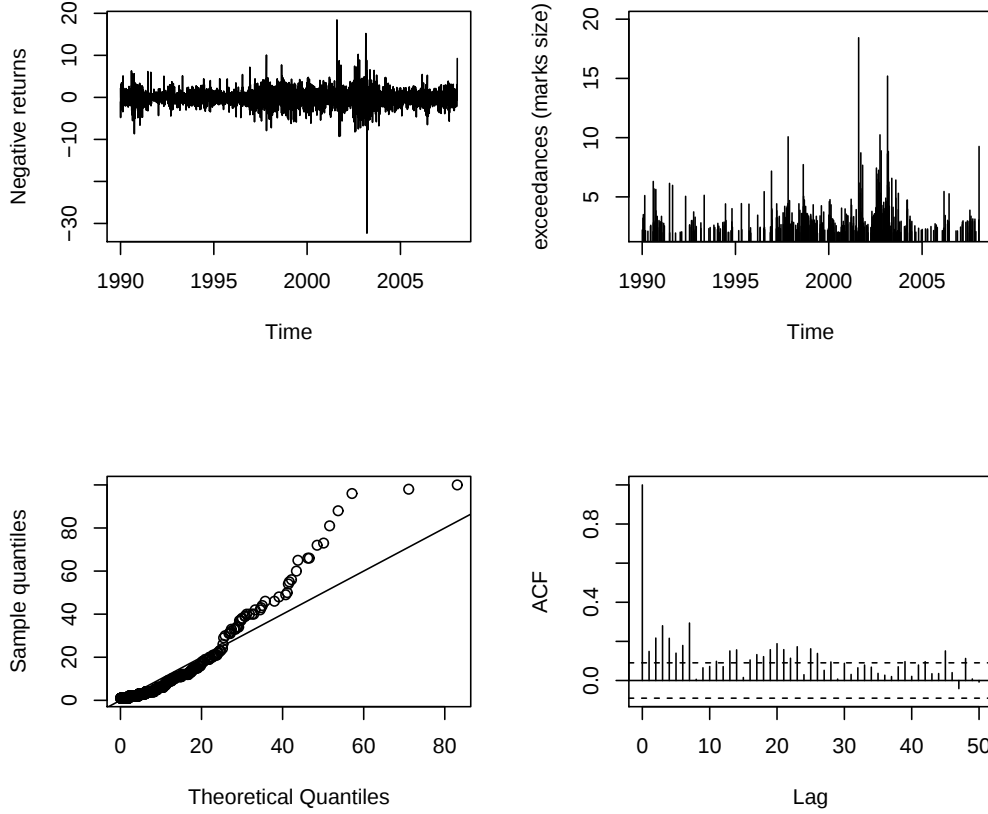


Figure 1.1: Upper left panel shows Bayer daily percentage loss data from 02.01.1990 to 18.01.2008. Upper right panel displays the 732 largest losses (the exceedances). Lower left panel shows a QQ-plot of inter-exceedance times against an exponential reference distribution. Finally, the lower right panel displays the autocorrelogram for the inter-exceedance times for the exceedances.

inter-event times, these are clearly far from exponential, giving evidence against a Poisson process of exceedances. Furthermore, the autocorrelogram plot suggests clustering of the inter-exceedance times. This hypothesis is moreover reaffirmed by the Ljung-Box statistic using 10 lags. The null hypothesis of white noise is easily rejected with the Ljung-Box statistic of 217.63 well above the critical value of 18.307 at the 5% level, rejecting the Null hypothesis.

Since extreme events are inherently irregularly spaced in time and their inter-exceedance times provide strong evidence of correlation, it seems natural to study the timing of transactions as an autoregressive conditional duration model.

The main contribution of this paper from the point of view of extreme value theory is that we can capture the short-term behaviour of extremes without involving an arbitrary stochastic volatility model or the prefiltration of the data, which certainly impacts the measures of risk. Further-

more, contrary to the models proposed in McNeil et al. (2005); Chavez-Demoulin et al. (2005) and Herrera and Schipp (2009), whose self-exciting functions are restricted to monotone decreasing functions, the models proposed in this paper allow hazard functions that are both monotonically decreasing and increasing. This has a logical interpretation in periods of financial turmoil, for instance, in an initial phase the VaR of a stock market initially should increase, then become close to constant and after this period decreases. To the best of our knowledge, this is the first research to take into account the incidence of the inter-exceedance times and models for irregularly spaced data in extreme value models.

The results of the application to the Bayer stock index indicate that the estimation of such models can be straightforward, derived through conditionals intensities. Different models were proposed, having in mind the simplicity of the structure of the conditional intensities. The empirical results show that characteristics associated with previous extreme losses as well the time between these extreme events have a significant impact on the dynamic aspects and size of future extreme events. In a VaR context the results of our backtesting procedure, which dynamically adjusts quantiles incorporating the new information daily, allows us to statistically conclude that the models proposed are suitable for the estimation of different risk measures as the VaR, according to the restriction imposed by Basel Committee on Banking Supervision (1996, 2006).

This paper is organized as follows. In section 2 we outline relevant aspects of the classical POT model, and then in section 3 we describe the ACD-POT model theory that is central to the paper and discuss a conditional generalized Pareto distribution based approach for the exceedances. In addition, we make use of the models proposed to obtain an expression and its estimate for the VaR one day ahead predictive distribution of the returns, conditionally on the past and current data. In section 4 the models are applied to transactions data from the Bayer index. Conclusions and future works are resumed in section 5.

2. The Peaks over threshold model

In the following we will explain our motivation in investigating the consequences of modelling the inter-exceedance times in the POT framework in relation to the classical approach where only the time in which the extreme events happen are taken account.

Suppose $Y_1, \dots, Y_n$ are random variables with distribution function F which belongs to the maximum domain of attraction of $H_{\xi, \mu, \sigma}$. Then $H_{\xi, \mu, \sigma}$ is a generalized extreme value distribution

$$H_{\xi, \mu, \sigma}(y) = \begin{cases} \exp \left\{ - \left(1 + \xi \frac{y - \mu}{\sigma} \right)^{-1/\xi} \right\} & \xi \neq 0, \\ \exp \left\{ - \exp \left(- \frac{y - \mu}{\sigma} \right) \right\} & \xi = 0, \end{cases}$$

where $1 + \xi y > 0$, $\xi, \mu \in \mathbb{R}$ and $\sigma > 0$ are the shape location and scale parameter respectively.

A point process N can be viewed as a random distribution of indistinguishable points in a defined state space. For instance, the basic model for threshold exceedances in extreme value theory is based on constructing a two-dimensional point process $\{(t_i, y_i)\}$ with state space $\mathcal{T} \times \mathcal{Y} = [0, 1) \times (u, \infty)$. The time events t_i are the time of the i -th peak exceedance and we shall refer to this process as the ground process, while $y_i - u$ is the value of the exceedances for a sufficiently high threshold u and we will call this process the process of the marks. This two dimensional point process will look like as a non-homogeneous Poisson process with intensity defined for all subsets of the form $A = [t_1, t_2) \times (y, \infty)$ where t_1 and t_2 are times of occurrence of extreme events. This representation is as follows

$$\lambda(t, y) = \lambda(y) = \frac{1}{\sigma} \left(1 + \xi \frac{y - \mu}{\sigma} \right)_+^{-1/\xi - 1}, \quad (2.1)$$

where $y_+ = \max(y, 0)$ and μ, σ, ξ are precisely the parameters of the generalized extreme value distribution. Furthermore, the intensity measure of the subset A for any $y \geq u$ may be expressed in the form of an one-dimensional Poisson process with intensity

$$\tau(y) = \int_y^\infty \lambda(s) ds = -\ln H_{\xi, \mu, \sigma}(y).$$

If we accept that the point process of exceedances is an one-dimensional Poisson with intensity $\tau > 0$, then the process has independent increments, i.e., the number of events t_i that occur in disjoint time intervals are mutually independent, which implies lack of memory in the evolution of the process. In addition, the number of extreme events t_i in any interval of length $(t_2 - t_1) \geq 0$ is Poisson distributed with mean

$$\int_{t_1}^{t_2} \int_y^\infty \lambda(s) ds dt = \tau(y) (t_2 - t_1).$$

Another important result of extreme value theory is the limiting conditional probability that $Y > u + y$ given $Y > u$

$$\frac{\tau(u + y)}{\tau(u)} = \left(1 + \frac{\xi y}{\sigma + \xi(u - \mu)} \right)^{-1/\xi} = \bar{G}_{\xi, \beta}(x),$$

which is just the survival función of the GPD, i.e., $\bar{G} = 1 - G$, with scaling parameter $\beta = \sigma + \xi(u - \mu)$ for $0 \leq y < y_F$. Here y_F is the right endpoint with values $y_F = \infty$ if $\xi > 0$ and $y_F = -\beta/\xi$ if $\xi < 0$.

Regardless of this approach, in this paper the statistical approach is based on viewing the

high level of exceedances as a *marked point processes* (MPP). In many stochastic process models, a point process arises as the component that carries the information about the events t in time or space of objects that may themselves have a stochastic structure and stochastic dependency relations. In this paper we define a MPP N as a set of observations, occurrence times and marks $\{(t_i, y_i)\}$ on the space $\mathcal{T} \times \mathcal{Y}$, whose history $\mathcal{H}_t = (\{t_1, y_1\}, \dots, \{t_{t-1}, y_{t-1}\})$ consists only of the occurrence times and marks $\{t_1, y_1\}, \dots, \{t_{t-1}, y_{t-1}\}$ up to time t but not including t . Moreover, we define a point process N_g “the ground process” which refers to the stochastic process of the inter-exceedance times. This point process has a conditional density function $p(t | \mathcal{H}_t)$ and its corresponding survival distribution function $S(t | \mathcal{H}_t)$. The conditional (finite) intensity function (or hazard function) for the ground process N_g is given by

$$\lambda_g(t | \mathcal{H}_t) = \frac{p(t | \mathcal{H}_t)}{S(t | \mathcal{H}_t)} \quad (2.2)$$

while the conditional intensity function for the MPP N is given by

$$\lambda(t, y | \mathcal{H}_t) = \lambda_g(t | \mathcal{H}_t) f(y | \mathcal{H}_t, t) \quad (2.3)$$

where $f(y | \mathcal{H}_t, t)$ is the density function of the marks conditional on t and $\mathcal{H}_t$.

Thus, the conditional intensity function with respect to the internal history $\mathcal{H}_t$ determines the probability structure of N uniquely. Furthermore, we say that a MPP $N = \{(t_i, y_i)\}$ on $\mathcal{T} \times \mathcal{Y}$ has independent marks, if given the ground process N_g the marks y_i are mutually independent random variables such that its distribution depends only on the corresponding location t_i . In addition, we define a MPP as having unpredictable marks for $\mathcal{T}$, if the distribution of the marks at t_i is independent of the locations and marks $\{(t_j, y_j)\}$ for which $t_j < t_i$. For a more formal introduction to marked point processes we refer to Daley and Vere-Jones (2003, p. 246).

According to our definition of MPP, the marks are conditionally independent of the associated ground process. Therefore, the product of mark densities simply has to be multiplied with the likelihood of the ground process. Letting N be a MPP on $[t_0, T) \times \mathcal{Y}$ for some finite positive T and let $(t_1, y_1), \dots, (t_{N(T)}, y_{N(T)})$ be a realization of N , we can obtain the log-likelihood L of such a realization in terms of the conditional densities or intensities as

$$\begin{aligned} L &= \sum_{i=1}^{N(T)} \log p_i(t_i | \mathcal{H}_t) + \sum_{i=1}^{N(T)} \log f_i(y_i | \mathcal{H}_t, t) \\ &= \sum_{i=1}^{N(T)} \log \lambda_g(t_i | \mathcal{H}_t) - \int_{t_0}^T \lambda_g(s | \mathcal{H}_t) ds + \sum_{i=1}^{N(T)} \log f_i(y_i | \mathcal{H}_t, t) \end{aligned} \quad (2.4)$$

Observe that an alternative description of the non-homogeneous Poisson process (2.1) is rewritten as a special case of a MPP in terms of a ground process N_g with rate of the one-dimensional Poisson process of exceedances of the level u , i.e., $\tau = \lambda_g(t \mid \mathcal{H}_t) = -\ln H_{\xi, \mu, \sigma}^z(u)$, and a GPD function for the marks $f(y \mid \mathcal{H}_t, t) = \frac{1}{\beta} \left(1 + \xi \frac{y-u}{\beta}\right)^{-1/\xi-1}$

$$\lambda(t, y) = \tau \frac{1}{\beta} \left(1 + \xi \frac{y-u}{\beta}\right)^{-1/\xi-1}. \quad (2.5)$$

This is exactly the idea that we want to explore in the next section. We will concentrate on models where the conditional intensity for the ground process will be parameterized in terms of interval between two consecutive extreme events $x_i = t_i - t_{i-1}$ so that the impact of a duration between successive events depends upon the number of intervening extreme events. The main area of application of these models has traditionally been in the modelling of high frequency financial data, however their structure would also seem appropriate for modelling extreme events and the tremors that follow these.

3. The Autoregressive conditional duration peaks over threshold model (ACD-POT)

As we saw in the introduction, in contrast to iid data, exceedances of a high threshold for daily financial returns do not necessarily occur according to a homogeneous Poisson process. Thus, the classical POT model is not directly applicable to financial return data.

We move away from homogeneous Poisson models for the occurrence times of exceedance of high thresholds and consider autoregressive conditional duration models for the conditional intensity of the ground process $\lambda_g(t \mid \mathcal{H}_t)$. In particular, we propose a set of models, which show autocorrelation between inter-exceedance times, clustered extremes and non iid exceedances or marks size.

Following Engle and Russell (1998) we define a model for the conditional intensity of the ground point process of exceedances depending only on a fixed number of the most recent inter-exceedance times $x_i = t_i - t_{i-1}$. Let ψ_i be the expectation of the i -th inter-exceedance time given by

$$\mathbb{E}(x \mid x_{i-1}, \dots, x_1) = \psi_i(x_{i-1}, \dots, x_1; \theta) \equiv \psi_i, \quad (3.1)$$

where θ is a parameter vector. We assume that ψ_i correspond to the ACD class of models. In general the assumption is based on that for a strictly positive function with positive support $\varphi(\cdot) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ the standardized durations

$$\varepsilon_i = \frac{x_i}{\varphi(\psi_i)} \quad (3.2)$$

are iid random variables. To derive a general expression for the conditional intensity let p be the density function of (3.2)

$$p\left(\frac{x_i}{\varphi(\psi_i)} \mid \mathcal{H}_t; \theta\right) = p\left(\frac{x_i}{\varphi(\psi_i)} \mid \theta\right), \quad (3.3)$$

where θ is a parameter vector. This implies that the time dependence of the duration process is summarized by the conditional expected duration sequence. If we define again a MPP on $[t_0, T) \times \mathcal{Y}$ for some finite positive time T and let $(t_1, y_1), \dots, (t_{N(T)}, y_{N(T)})$ be a realization of N over the interval $[0, T)$, one can easily show that the conditional expected intensity of the interexceedances times between extreme events, the ground process, can be expressed as a multiplicative effect between the baseline hazard function and a shift given by the expected duration

$$\lambda_g(t \mid \mathcal{H}_t; \theta) = \lambda_0\left(\frac{t - t_{N(T)}}{\varphi(\psi_{N(T)})}\right) \frac{1}{\varphi(\psi_{N(T)})} \quad (3.4)$$

Replacing further the form of the ground process defined in (2.2) by one of this type and plugging the conditional expected intensity (3.4) in (2.5) we obtain the ACD-POT model in its more general form.

$$\lambda(t, y \mid \mathcal{H}_t; \theta) = \frac{\lambda_0\left(\frac{t - t_{N(T)}}{\varphi(\psi_{N(T)})}\right)}{\varphi(\psi_{N(T)}) \beta} \left(1 + \xi \frac{y - u}{\beta}\right)_+^{-1/\xi - 1}. \quad (3.5)$$

Effectively we have combined the one-dimensional intensity in (3.4) with a generalized Pareto density. Under this model the conditional rate of crossing the threshold $x \geq u$ at time t , given the history $\mathcal{H}_t$ up to that time, is

$$\tau(t, y \mid \mathcal{H}_t; \theta) = \int_y^\infty \lambda(t, s \mid \mathcal{H}_t; \theta) ds = \frac{\lambda_0\left(\frac{t - t_{N(T)}}{\varphi(\psi_{N(T)})}\right)}{\varphi(\psi_{N(T)})} \left(1 + \xi \frac{y - u}{\beta}\right)_+^{-1/\xi}.$$

Observe that the distribution of the marks are assumed to be independent of the behavior of inter-exceedance times. Indeed, the implied distribution of the marks when an extreme event takes place is given by

$$\frac{\tau(t, u + y \mid \mathcal{H}_t; \theta)}{\tau(t, u \mid \mathcal{H}_t; \theta)} = \left(1 + \frac{\xi y}{\beta}\right)_+^{-1/\xi} = \bar{G}_{\xi, \beta}(y).$$

The estimation of the parameters of this model can be performed separately. In the first instance we can estimate the likelihood for the ground process, the conditional intensity of inter-exceedance

times, and in the second instance, the likelihood of the generalized Pareto distribution of the marks. The model proposed in (3.5) is, therefore, a model with unpredictable marks because of the fact that the distribution of the last mark at t is independent of the history $\mathcal{H}_t$.

Nevertheless, we also consider the case of predictable marks, i.e., the marks are conditionally generalized Pareto, given the history $\mathcal{H}_t$ up to the time of the mark. To this end, we parameterized the scaling parameter $\beta(t, y | \mathcal{H}_t)$ such that it depends on the history². In this way, we assume that in a period of turmoil the temporal intensity of the inter-exceedance times and the magnitude of the marks increase.

$$\lambda(t, y | \mathcal{H}_t; \theta) = \frac{\lambda_0 \left(\frac{t - t_{N(T)}}{\varphi(\psi_{N(T)})} \right)}{\varphi(\psi_{N(T)}) \beta(t, y | \mathcal{H}_t)} \left(1 + \xi \frac{y - u}{\beta(t, y | \mathcal{H}_t)} \right)_+^{-1/\xi - 1}. \quad (3.6)$$

The features of this model immediately follow those of the first model proposed. The conditional rate of crossing the threshold $x \geq u$ at time t , given the history $\mathcal{H}_t$ up to that time, is in this case

$$\tau(t, y | \mathcal{H}_t; \theta) = \int_y^\infty \lambda(t, s | \mathcal{H}_t; \theta) ds = \frac{\lambda_0 \left(\frac{t - t_{N(T)}}{\varphi(\psi_{N(T)})} \right)}{\varphi(\psi_{N(T)})} \left(1 + \xi \frac{y - u}{\beta(t, y | \mathcal{H}_t)} \right)_+^{-1/\xi},$$

while the implied distribution of the marks when an extreme observation occurs is given by

$$\frac{\tau(t, u + y | \mathcal{H}_t; \theta)}{\tau(t, u | \mathcal{H}_t; \theta)} = \left(1 + \xi \frac{y - u}{\beta(t, y | \mathcal{H}_t)} \right)_+^{-1/\xi} = \bar{G}_{\xi, \beta(t, y | \mathcal{H}_t)}(y).$$

Note that the marginal distribution of the marks will now be a conditional GPD. In the practical application, using a likelihood ratio test, we will formally test the hypothesis that the marks are unpredictable.

One main purpose of this paper is to develop a methodology to obtain an expression and its estimate for the quantile of the one day ahead predictive distribution of the returns, conditionally on the past and current data. In particular, the Value-at-risk (VaR) and Expected shortfall (ES) measures of risk. These measures have become standard measures in financial risk management due to their conceptual simplicity, computational facility and ready applicability. Following we derive these measures for the ACD-POT models.

The VaR is defined as the q -th quantile of a distribution F given by

$$\text{VaR}_\alpha^t = y_\alpha^t = \inf \{ y \in \mathbb{R} : F_{y_{t+1} | \mathcal{H}_t}(y) = \alpha \},$$

²We can also parameterized the shape parameter ξ . However, the behaviour of the estimation is severely affected. For this reason it is reasonable to take the shape parameter to be constant.

which is solution to $\mathbb{P}(y_{t+1} > y_\alpha^t \mid \mathcal{H}_t) = 1 - \alpha$. Observe that

$$\begin{aligned}\mathbb{P}(y_{t+1} > y_\alpha^t \mid \mathcal{H}_t) &= \mathbb{P}(y_{t+1} - u > y_\alpha^t - u \mid \mathcal{H}_t) \\ &= \mathbb{P}(y_{t+1} - u > y_\alpha^t - u \mid y_{t+1} > u, \mathcal{H}_t) \mathbb{P}(y_{t+1} > u \mid \mathcal{H}_t).\end{aligned}\quad (3.7)$$

The first term in the right hand side of equation (3.7) can be approximated via

$$\mathbb{P}(y_{t+1} - u > y_\alpha^t - u \mid y_{t+1} > u, \mathcal{H}_t) = \left(1 + \xi \frac{y_\alpha^t - u}{\beta(t, y \mid \mathcal{H}_t)}\right)_+^{-1/\xi},$$

while

$$\begin{aligned}\mathbb{P}(y_{t+1} > u \mid \mathcal{H}_t) &= \mathbb{P}(N(t, t+1) = 1 \mid \mathcal{H}_t) \\ &= 1 - \exp(-\lambda(t, s \mid \mathcal{H}_t; \theta)).\end{aligned}$$

Thus the VaR is defined by

$$\text{VaR}_\alpha^t = u + \frac{\beta(t, y \mid \mathcal{H}_t)}{\xi} \left(\left(\frac{1 - \alpha}{1 - \exp(-\lambda(t, s \mid \mathcal{H}_t; \theta))} \right)^{-\xi} - 1 \right). \quad (3.8)$$

The last equation implies that the VaR is only defined for our models if $1 - \exp(-\lambda(t, s \mid \mathcal{H}_t; \theta)) > 1 - \alpha$. In the case of expected shortfall (ES), it is defined as the average of all losses which are greater or equal to VaR, i.e. the average loss in the worst $(1 - \alpha)\%$ cases $ES_\alpha^t = E[Y \mid Y > \text{VaR}_\alpha^t]$. In the models proposed the ES is given by

$$ES_\alpha^t = \frac{\text{VaR}_\alpha^t}{1 - \xi} + \frac{\beta(t, y \mid \mathcal{H}_t) - \xi u}{1 - \xi}. \quad (3.9)$$

In the following sections we introduce the models that we want to utilize to parameterize the expected conditional duration function ψ_i , the distribution of probability of the standardized durations ε_i and the models for the scale parameter $\beta(t, y \mid \mathcal{H}_t)$.

3.1. ACD models for the expected conditional duration

In this section, we consider models that allow for additive as well as multiplicative components in the conditional duration function ψ . In addition, we introduce parameterizations that allow not only for linear but also for more flexible innovations impact curves.

The ACD model

The most popular autoregressive conditional duration model, introduced by Engle and Russell (1998), is based on a linear parameterization of the conditional mean function

$$\psi_i = w + \sum_{j=1}^p a_j x_{i-j} + \sum_{j=1}^q b_j \psi_{i-j},$$

where $w > 0$, $a, b \geq 0$. In order to ensure the stationarity and existence of the unconditional expected duration we need $\sum_{j=1}^p a_j + \sum_{j=1}^q b_j < 1$.

The logarithmic ACD (Log-ACD) model

In order to prevent ψ_i becoming negative, Bauwens and Giot (2000) introduced the Logarithmic ACD model in which the autoregression bears on the logarithm of the conditional expected duration.

$$\psi_i = \exp \left\{ w + \sum_{j=1}^p a_j \log x_{i-j} + \sum_{j=1}^q b_j \log \psi_{i-j} \right\}.$$

Note that the functional form of ψ_i implies a multiplicative relationship between past durations which is quite different from a linear form.

The Box-Cox ACD (BCACD) model

Dufour and Engle (2000) propose the Box-Cox-ACD model as a more flexible alternative based on a Box-Cox transformation of the past innovations due to the Log-ACD model implies a relatively rigid adjustment process of the conditional mean to recent durations and thus, in general, an over adjustment of the conditional mean after very short durations.

$$\psi_i = w + \sum_{j=1}^p \frac{a_j}{\delta} \left(\varepsilon_{i-j}^\delta - 1 \right) + \sum_{j=1}^q b_j \psi_{i-j}.$$

This specification includes the Log-ACD model for the Box-Cox parameter $\delta \rightarrow 0$ and a linear specification for $\delta = 1$.

The EXponential ACD (EXACD) model

Dufour and Engle (2000) also introduce the so called EXponential ACD Model since this model is similar to the one Nelson (1991) devised for the GARCH model. In this model the news effects are modelled with a piece-wise linear specification.

$$\psi_i = w + \sum_{j=1}^p \left\{ a_j \varepsilon_{i-j} + \delta_j |\varepsilon_{i-j} - 1| \right\} + \sum_{j=1}^q b_j \psi_{i-j}.$$

Thus, for durations shorter than the conditional mean ($\varepsilon_{i-j} < 1$), the news impact curve has a slope $a_j - \delta_j$ and an intercept $w + \delta_j$. Durations longer than the conditional mean ($\varepsilon_{i-j} > 1$), also have a linear effect, but with a slope $a_j + \delta_j$ and intercept $w - \delta_j$.

Strict stationarity of the conditional mean for the models Log-ACD, BCACD and EXACD is guaranteed by constraints on coefficients b 's, i.e., all roots of $1 - \sum_{j=1}^q b_j v^j$ must lie outside the unit circle.

3.2. Distributional assumptions for the standardized durations

The second important ingredient in the parameterization of our ACD-POT model is the distributional assumption for the innovation process. In this section we explore two alternatives the Burr and the generalized gamma distribution. The major advantage of these distributions over the exponential and Weibull distribution, the most common distributions utilized in ACD models, is that these have non-monotonic hazard functions taking bathtub shaped or inverted U-shaped forms. In a bathtub shaped form the hazard rate initially decreases, during the middle phase the hazard rate is essentially constant, and in the final phase the hazard increases. Inverted U-shaped forms are the counterparts, the hazard rate initially increases, then becomes close to constant and ultimately decreases. This feature is of particular importance if we are interested in modelling risk measures such as the VaR or the expected shortfall.

Generalized Gamma distribution

Lunde (1999) as well Zhang et al. (2001) propose the use of a generalized gamma distribution to characterize the standardized durations because one can then obtain a non-monotonic hazard function and a time-varying conditional mean duration. A three parameter generalized gamma density is given by

$$f(x | \gamma, k) = \frac{\gamma x^{k\gamma-1}}{\lambda^{k\gamma} \Gamma(k)} \exp \left\{ - \left(\frac{x}{\lambda} \right)^\gamma \right\}, x > 0.$$

It includes the exponential distribution ($\gamma = k = 1$), the Weibull distribution ($k = 1$), the half-normal ($\gamma = 1/2, k = 1$) and the ordinary gamma distribution ($k = 1$). Under the restriction that $\lambda = 1$ we chose $\varphi(\psi_i) = \phi_i = \psi_i \frac{\Gamma(k)}{\Gamma(k + \frac{1}{\gamma})}$ which implies a conditional density of the standardized duration given by

$$p \left(\frac{x_i}{\phi_i} | \mathcal{H}_i; \theta \right) = \frac{\gamma \psi_i}{x_i \Gamma(k)} \left(\frac{x_i}{\phi_i} \right)^{k\gamma} \exp \left\{ - \left(\frac{x_i}{\phi_i} \right)^\gamma \right\}$$

and a conditional density of the durations

$$p(x_i | \mathcal{H}_t; \theta) = \frac{\gamma \Gamma(k + \frac{1}{\gamma})}{x_i} \left(\frac{x_i}{\phi_i}\right)^{k\gamma} \exp\left\{-\left(\frac{x_i}{\phi_i}\right)^\gamma\right\},$$

where θ is once more a parameter vector. Note that if $k = 1$, then we get the Weibull-ACD model, while for $k = \gamma = 1$ the model reduces to an Exponential-ACD model. The hazard function implied by the generalized gamma model may now be written as

$$\lambda_g(x_i | \mathcal{H}_t; \theta) = \frac{\frac{\gamma x_i^{k\gamma-1}}{\phi_i^{k\gamma} \Gamma(k)} \exp\left\{-\left(\frac{x_i}{\phi_i}\right)^\gamma\right\}}{I\left(k, \left(\frac{x_i}{\phi_i}\right)^\gamma\right)},$$

where is the upper incomplete gamma integral $I\left(k, \left(\frac{x_i}{\phi_i}\right)^\gamma\right) = \int_{\left(\frac{x_i}{\phi_i}\right)^\gamma}^{\infty} u^{k-1} \exp(-u) du$.

In addition, the shape properties of the conditional hazard function can be derived from its parameters values. If $k\gamma < 1$, the hazard rate is decreasing for $\gamma \leq 1$ and U-shaped for $\gamma > 1$. Conversely, if $k\gamma > 1$, the hazard rate is increasing for $\gamma \geq 1$, and inverted U-shaped for $\gamma < 1$. Finally, if $k\gamma = 1$, the hazard is decreasing for $\gamma < 1$, constant for $\gamma = 1$, and increasing for $\gamma > 1$.

The conditional intensity in this case takes the form

$$\lambda(t, y | \mathcal{H}_t; \theta) = \frac{\frac{\gamma x_i^{k\gamma-1}}{\phi_i^{k\gamma} \Gamma(k)} \exp\left\{-\left(\frac{x_i}{\phi_i}\right)^\gamma\right\}}{I\left(k, \left(\frac{x_i}{\phi_i}\right)^\gamma\right)} \frac{1}{\beta(t, y | \mathcal{H}_t)} \left(1 + \xi \frac{y - u}{\beta(t, y | \mathcal{H}_t)}\right)_+^{-1/\xi-1}.$$

The conditional log-likelihood function of this model on a set of observed inter-exceedance times and of marks or size of the exceedances can easily be derived of (2.4)

$$\begin{aligned} L = & \sum_{i=1}^{N(T)} \left\{ \log \gamma + (k\gamma - 1) \log \left(\frac{x_i}{\phi_i}\right) - \log(\Gamma(k) \phi_i) - \left(\frac{x_i}{\phi_i}\right)^\gamma \right\} \\ & - (1 + 1/\xi) \sum_{i=1}^{N(T)} \log \left(1 + \xi \frac{y_i - u}{\beta(t, y | \mathcal{H}_t)}\right)_+ \end{aligned}$$

Burr distribution

Another flexible alternative is the Burr distribution introduced in the context of ACD models by Grammig and Maurer (2000). The density function is defined by

$$f(x | \lambda, k, \gamma) = \frac{\lambda k t^{k-1}}{(1 + \gamma^2 \lambda t^k)^{\gamma^{-2}+1}}$$

In this case we define $\phi(\psi_i) = \phi_i = \psi_i^{\frac{\gamma^2(1+\frac{1}{k})}{\Gamma(1+\frac{1}{k})\Gamma(\gamma^{-2}-\frac{1}{k})}}$, where $0 < \gamma^{-2} < k$. We choose the density (3.3) to be Burr under the restriction that $\lambda = 1$,

$$p\left(\frac{x_i}{\phi_i} | \mathcal{H}_i; \theta\right) = \frac{k \phi_i^{1-k} x_i^{k-1}}{(1 + \gamma^2 \phi_i^{-k} x_i^k)^{\gamma^{-2}+1}}.$$

The conditional density of x_i is then

$$p(x_i | \mathcal{H}_i; \theta) = \frac{k \phi_i^{-k} x_i^{k-1}}{(1 + \gamma^2 \phi_i^{-k} x_i^k)^{\gamma^{-2}+1}},$$

which is a Burr density with parameter $\lambda = \phi_i^{-k}$. The implied conditional hazard function is

$$\lambda_g(x_i | \mathcal{H}_i; \theta) = \frac{k \phi_i^{-k} x_i^{k-1}}{1 + \gamma^2 \phi_i^{-k} x_i^k}, \quad (3.10)$$

which is non-monotonic function with respect to duration. From (3.10) can be obtained: the Weibull-ACD model if $\gamma^2 \rightarrow 0$, Exponential-ACD model if $\gamma^2 \rightarrow 0$ and $k = 1$ and Log-Logistic ACD for $\gamma^2 = 1$. The conditional intensity in this case takes the form

$$\lambda(t, y | \mathcal{H}_i; \theta) = \frac{k \phi_i^{-k} x_i^{k-1}}{1 + \gamma^2 \phi_i^{-k} x_i^k} \frac{1}{\beta(t, y | \mathcal{H}_i)} \left(1 + \xi \frac{y - u}{\beta(t, y | \mathcal{H}_i)}\right)_+^{-1/\xi - 1}.$$

The conditional log-likelihood function of this model on a set of observed inter-exceedance times and of marks or size of the exceedances can be easily obtained from (2.4), i.e.,

$$\begin{aligned} L = & \sum_{i=1}^{N(T)} \left\{ \log k - k \log \phi_i + (k-1) \log x_i - (1 + \gamma^{-2}) \log \left(1 + \gamma^2 \phi_i^{-k} x_i^k\right) \right\} \\ & - (1 + 1/\xi) \sum_{i=1}^{N(T)} \log \left(1 + \xi \frac{y_i - u}{\beta(t, y | \mathcal{H}_i)}\right)_+ \end{aligned}$$

3.3. Models for the time varying scale parameter

In this section we consider different models to parameterize the scaling parameter $\beta(t, y | \mathcal{H}_t)$ such that it depends on the history. This feature implies that the marks are conditionally generalized Pareto, given the history $\mathcal{H}_t$ up to the time of the mark. Under these models we assume that in a period of turmoil the temporal intensity of the inter-exceedance times and the magnitude of the marks increase. We will specify and estimate two alternatives forms for the scaling parameter $\beta(t, y | \mathcal{H}_t)$, the lineal

$$\beta(t, y | \mathcal{H}_t) = \omega + \beta_1 y_{i-1} + \beta_2 \varphi(\psi_i)$$

and the exponential

$$\beta(t, y | \mathcal{H}_t) = \omega + \beta_1 y_{i-1} + \beta_2^{\varphi(\psi_i)},$$

where $\omega, \beta_1, \beta_2 \in \mathbb{R}_+$.

The two models have in common a natural interpretation; the scaling parameter $\beta(t, y | \mathcal{H}_t)$ depends on the last mark and the conditional mean function of the inter-exceedance times, given the information up to time t , but not including t^3 .

4. Empirical application

In this section we present the main empirical results obtained by testing the models introduced in the previous sections. For the empirical test we chose the transaction data from Bayer AG as announced already in the introduction. In this study we only concentrate on the left tail, so that the daily returns are obtained as $r_t = -100 \ln(p_t / p_{t-1})$, where p_t denotes the (closing) stock price at day t . The sample period spans from 2 January 1990 to January 18, 2008, two days before January 20, when the Global stock markets suffered their biggest falls since September 11, 2001. A second sample is used for backtesting the estimation of the VaR in the Bayer AG from 20 January 2008 to January 16, 2009. In the backtest we daily update the new information that becomes available for the parameter estimates previously obtained. Thus, we dynamically adjust quantiles, which allows us to improve as accurately as possible the estimation of the risk measures.

In order to summarize adequately the large quantity of empirical results obtained, we use a classification scheme for the ACD-POT models. The first lower-case letter describes the type of

³In the first instance we take different models into account to find the best approach. Our analyses with financial time series have suggested that by model comparisons based on the likelihood ratio statistic, these models keep the formulation easy to understand.

distributional assumption with respect to the ACD model (generalized gamma or Burr). The following capital letters denote the type of ACD model: ACD, Log-ACD, BCACD or EXACD⁴. The last small letter denotes the models for the time varying scale parameter: l (lineal), e (exponential) or u (for unpredictable marks as in (3.5) or equivalently with scale parameter β constant). For example, a model gLog-ACDu means that we are working with a Log-ACD model for the expected conditional duration with generalized gamma distribution and unpredictable marks. In total we have 24 models.

All parametric models are estimated using quasi maximum likelihood. We adopt different models according to our scheme classification to test the different ACD models with different distributional assumptions regardless the possibility of a model with unpredictable marks (marks with iid GPD) or predictable marks (marks with GPD whose scale parameter is time-dependent). Observe that the former is equivalent to fitting separately a ACD model to the inter-exceedance times and a GPD to the excess losses over the threshold.

An important point is the choice of the threshold, which implies a balance between bias and variance. The threshold must be set high enough so that the exceedances are distributed generalized Pareto. However, the choice of the optimal threshold is still considered an open problem and different approaches have been proposed to overcome this difficulty. In this paper we choose to work with the 10% of the maxima of the sample⁵.

In relation to the measures of goodness of fit we utilize the W-statistics to assess our success in modelling the temporal behaviour of the exceedances of the threshold u . This statistic states that if the GPD parameter model is correct, then the residuals are approximately independent unit exponential variables. In addition, to check that there is no further time series structure the autocorrelation function (ACF) for the residuals is also included. Similarly, we provide empirical evidence on the accuracy of actual VaR measures derived from the models. The first of them is an unconditional coverage test proposed by McNeil and Frey (2000). The idea is to test if the fraction of violations obtained for a particular risk measure, is significantly different from the theoretical one. A violation of the VaR is defined as occurring when the ex-post return is lower than the VaR. The second approach proposed by Berkowitz et al. (2009) tests for uncorrelatedness of the violations. In particular, we suggest the well-known Ljung-Box test of the violation sequence's autocorrelation function. All these methods are resumed briefly in the Appendix.

It is quite evident that the performance of the models considered varies depending on the kind

⁴For a meaningful comparison of alternatives and for simplicity, we limit the dynamic structure of the ACD-POT models to the first lag order only.

⁵The choice of the threshold is done with help of the the mean excess (ME) function, which is a popular tool used to determine the adequacy of the GPD model in practice.

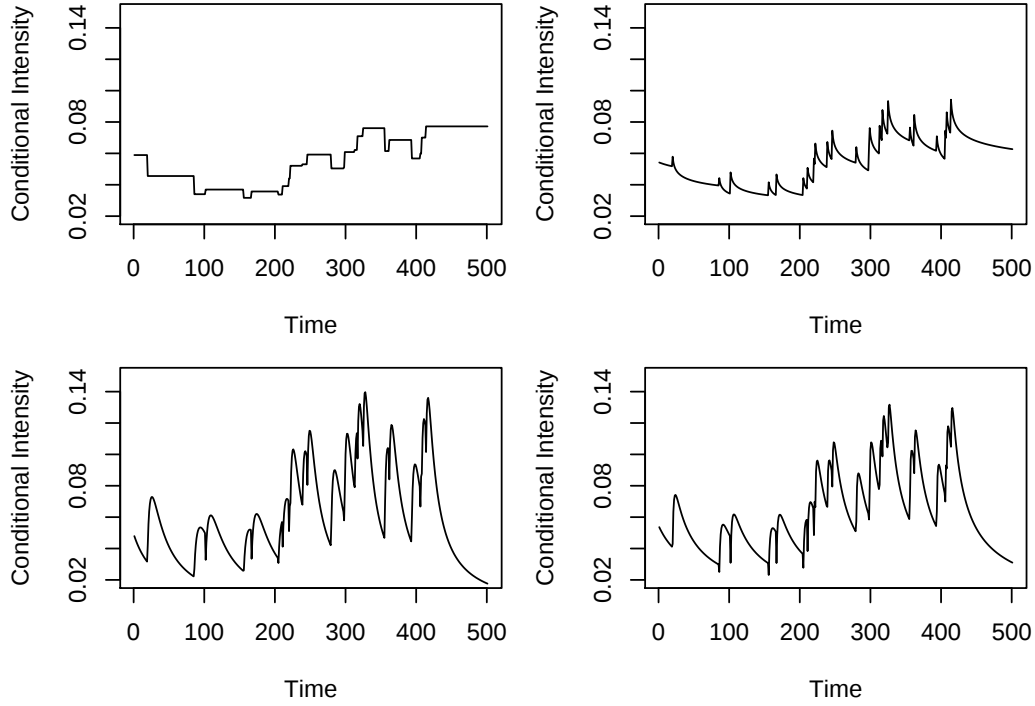


Figure 4.1: The path of the conditional intensities of an ACD-POT model under four types of distributional assumptions: in the top panel the exponential (left) and the Weibull (right) distributions. In the bottom panel the Burr (left) and the generalized gamma (right) distributions.

of distributional assumption or the model for the time varying scale parameter, that we fitted. To gain understanding about how each class of models and distribution assumptions affect the estimations we visually compare the behaviour of the estimate conditional intensity of a small sample for the Bayer returns.

In the first instance we only concentrate on the kind of distributional assumption. In Figure 4.1 we display the path of a ACD-POT model with four types of distributional assumptions: in the top panel the exponential (left) and the Weibull (right), as especial cases of the Burr or generalized gamma distribution, and in the bottom panel the Burr (left) and the generalized gamma (right) distribution. Observe that in the case of the exponential and Weibull distributions we have a flat or monotone conditional intensity, respectively. On the other hand, both the Burr and the generalized gamma distribution show a non-monotone conditional intensity. This important feature allows to the last two distributions rapidly adapting the conditional intensity to periods of high volatility which are associated with clustering of short interexceedance times.

In relation to the type of ACD model for conditional mean duration, Figure 4.2 shows conditional intensities for the four models proposed under the assumption of a generalized gamma

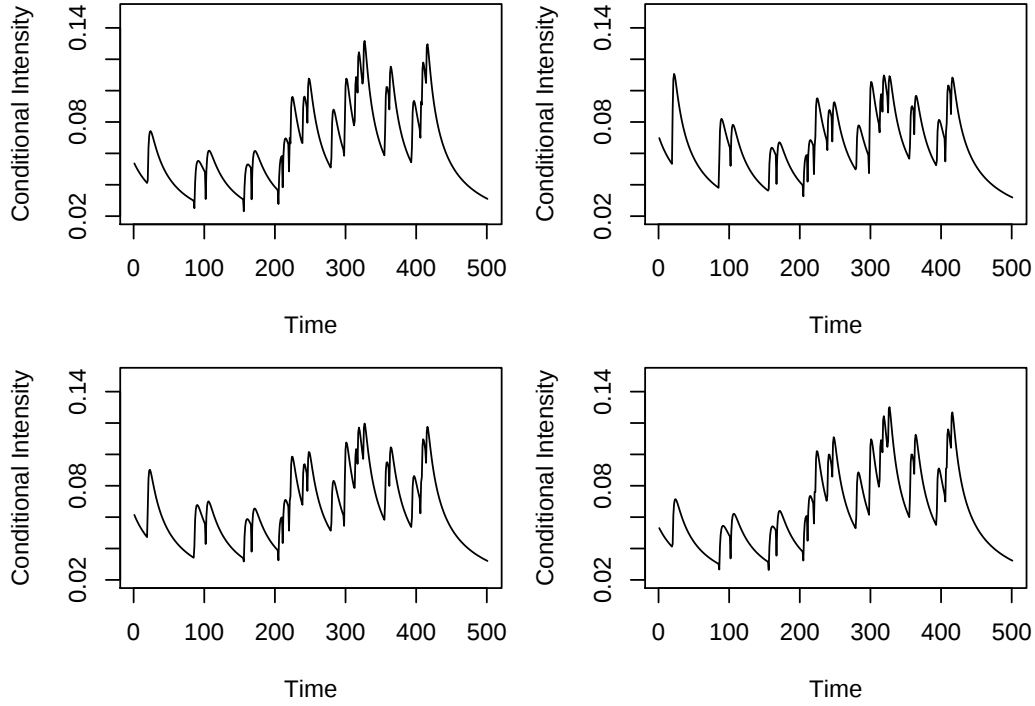


Figure 4.2: The path of the conditional intensities for the four models proposed under the assumption of a generalized gamma distribution for the innovations: the gACDu (top left), the gLog-ACDu (top right), the gEXACDu (bottom left) and the gBCACDu model (bottom right).

distribution for the innovations: the ACD (top left), the Log-ACD (top right), the EXACD (bottom left) and the BCACD model (bottom right). At this stage, it is not yet possible to reach a conclusion on the appropriateness of each model. However, there are some important features of the models to keep in mind, before we make a choice. On the one hand, the Log-ACD allows for nonlinear effects of short and long durations in the conditional mean, without requiring the estimation of additional parameters in comparison to the standard ACD model. While on the other hand, the BCACD and the EXACD models offer a captivating compromise between the need of greater flexibility and the burden of higher complexity.

Empirical results

The maximum log-likelihood estimates of the ACD-POT models proposed in section 3 for the returns are displayed in Tables A.1 and A.2 in the Appendix. For the inter-exceedance times, the generalized gamma seems to be the best distribution between the two choices. The results on ACD models for the expected conditional duration lead to markedly favour the Log-ACD specifications, followed by the ACD one. Finally, the models with time varying scale parameters lead to a better fit. Indeed, the results suggest that the models with predictable marks react more quickly to in-

creasing and decreasing cluster of extremes, which means that the size of the exceedances has an effect on the probability of further exceedances in the near future.

According to the AIC of the models proposed, the best fitted model for the Bayer index is a ACD model with generalized gamma distribution and lineal form for the scale parameter (gACDI) with AIC of 4067.59. We observe further that $k = 53.767$ (36.539), $\gamma = 0.121$ (0.041), which implies that $k\gamma > 1$ and $\gamma < 1$ so that the hazard rate is inverted U-shaped. This should not come as a surprise if one is aware of the intimate relationship between durations and cluster of extremes. Furthermore, this is the sort of hazard function that earlier authors have found to be realistic in modeling the dynamics of "price durations" in stock markets (see for instance Zhang et al., 2001 and Grammig and Maurer, 2000). In relation to the results of estimation of the conditional GPD model to the exceedances we obtained $\xi=0.503$ (0.100), $\omega= 0.094$ (0.044), $\beta_1=5.7e-07$ (0.001) and $\beta_2= 4.492$ (0.826). This result indicates that the lineal form to parameterize the scaling parameter $\beta(t, y | \mathcal{H}_t)$, such that it depends on the history, was a good choice. Interestingly, the size of the last exceedance is not as important as the expectation of the i -th inter-exceedance time.

Although the model of choice identified by the AIC may be seen as the best among the existing models because it shows the best global fit, this does not mean that no better model is possible for backtesting. So, we usually check whether the major features of the given data can be reproduced by the estimated models, for instance, the cluster of extreme events. If this important feature is not reproduced, we must consider further models whose AIC values can be compared with those of the previous best model. To this end, we include two other models to have a comparison of different alternatives in the backtest. The second best alternative is a Log-ACD and the third is a BCACD, both with generalized gamma distribution and lineal form for the scale parameter. We concentrate on these three alternatives and test the reliability of these models by investigating the conditional GPD assumption of the marks in the models fitted, the quality of the times component of our model and the performance in-sample of the estimated VaRs.

The results on the goodness of fit in sample are displayed in Figure A.1. More general test on miss specification in an ACD model were proposed by Meitz and Teräsvirta (2006). Here, we first assess the conditional GPD assumption of the marks in the models fitted. To this end, we provide the W-statistic explained in details in the Appendix. This statistic forms an iid sequence of exponential random variables with mean one if the marks are GPD. According to the QQ-plots displayed in Figure A.1, we do not observe a substantial deviation from an exponential distribution. In addition, to check that there is no further time series structure, the autocorrelation function (ACF) for the residuals (middle panel) is also included. The autocorrelations is negligible at nearly all lags. Finally, to appraise the quality of the times component of our model, we employ

the residual analysis method for point processes resumed briefly in the Appendix. This is based on the change of time scale using the estimated conditional intensity. We investigated whether the transformed time-scale version of the data constitutes a homogeneous Poisson process. The residual analysis for the three models indicates that the ACD-POT models in their three alternatives in the changed time scale seems to be acceptable. Hence, for the returns the ACD-POT specification seems to be appropriate.

In relation to the performance in-sample of the estimates VaRs, Tables A.3 displays the results for the unconditional coverage test and the Ljung-Box test, for all the models for three different VaR levels (0.05, 0.01, and 0.001). For each model proposed in the last section we give the number of violations or failures in the VaR, the unconditional coverage test and the Ljung-Box test (the last two in brackets). In the case of these three models fitted the conditional VaR is correctly estimated for all the confidence levels⁶.

Backtesting the models

Backtesting provides invaluable feedback about the accuracy of the models proposed to risk managers. The performance of VaR w.r.t backtesting has been carried out with the daily returns for one year, i.e., from January 20, 2008 to January 16, 2009. The backtest method consists on comparing the estimated conditional VaR for one day time horizon t , given knowledge of returns up to and including t for three different confidence levels (0.95, 0.99, and 0.999). For each day in the back test we reestimate the models, something that immediately reveals possible stability problems of a model. Then, we reestimated the risk measures for each return series according to the formula (3.8).

Table A.4 reports the results on the VaR backtesting exercise for all confidence levels, while the VaR violations for the 0.99 confidence level under the gLog-ACDl, gACDl and gBCACDl models are shown in Figure A.2. The performances for the models are similar for the results on VaR forecasting, although we observe some differences. For instance, the gLog-ACDl model tends to lightly underestimate the $VaR_{0.95}$, while the gBCACDl do the same for the $VaR_{0.99}$. However, the unconditional coverage test and the Ljung-Box test for all the confidence levels indicate that no severe clustering of exceedances is present and that the VaR violations can be considered as independent, respectively. In addition, according to the “traffic light” approach the three models are all classified in the green zone (see for more reference Basel Committee on Banking Supervision (2006)). Finally, due to the shortness of the time horizon we do not find a VaR violation for the 0.999 quantile, and therefore the Box-Ljung test p-values are not reported.

⁶The null hypothesis is rejected whenever the p-value of the binomial test and the Ljung_Box test are less than 5 percent.

To summarize, the results of our backtesting procedure with a dynamic adjustment of quantiles incorporating the new information daily allows us to statistically conclude that the models proposed are suitable for the estimation of different risk measures, as for example, the VaR according to the restriction imposed by Basel Committee on Banking Supervision (1996, 2006). Moreover, these models allow us to take the heavy-tailness or the stochastic nature of the cluster of extreme events into consideration.

5. Conclusions and future works

This paper proposed a new technique for modelling extreme events of stationary sequences as is the case of the most financial returns. We make use of a new class of self-exciting point process models that seem particularly well suited. The idea was to create a model being able to incorporate stylized facts such as clustering of extreme events and autocorrelation of the inter-exceedance times of extreme events, i.e., properties that are observed in practice.

The model can be interpreted as a combination between the classical Peaks over Threshold (POT) model from Extreme Value Theory and the class of Autoregressive Conditional Duration (ACD) models that are popular in finance. For this reason we call it ACD-POT models.

We observe that under this methodology the estimation of such models can be straightforwardly derived through conditionals intensities. Different models were proposed having in mind the simplicity of the structure of the conditional intensities. However, other more complicated structures could also be adopted.

With regard to the empirical application the models and their estimations with returns from Bayer AG were more than satisfactory. Our empirical results show that characteristics associated with previous extreme losses as well the time between these extreme events have a significant impact on the dynamic aspects and size of future extreme events.

On average, the best three models fit well in-sample for the VaR for different levels of risk, i.e., in terms of capital requirement; the models keep necessary capital to guarantee the desired confidence level. For these models the VaR is backtested through a comparison with the actual losses over an out-of-the-sample period of one year. The backtesting results indicate that the proposed methodology performs well in forecasting the risk dynamically and provides therefore certainly more precise estimate as the information in the data sample is exploited more efficiently. This refers particularly to clustering of extreme events and the inter-exceedance times.

In summary the ACD-POT models do a very good job of modelling the inter-exceedance times associated with waiting times between extreme events and the size of exceedances. These models are presented as a useful starting point along other extensions. Other possible directions for future

research emerge from the results. For instance, being interested in long term behaviour rather than in short term forecasting, the simulation of ACD-POT models is possible to calculate the measures of risk with other time horizons. Other options would be different distributional assumptions for the standardized durations or other flexible forms for the self-exciting models, which could be used incorporating other characteristics of the series such as trend of increasingly exceedances or different regimes as aftershocks. Another idea is to combine ACD with another class of self-exciting models, such as Hawkes- or ETAS- (Epidemic Type Aftershock Sequence) models. This could help to characterize other important features such as slow decay of autocorrelation or a power-law decay between jumps. Finally, the application of these models are not only limited to daily returns. A natural extension is to use this methodology to high frequency data to estimate intraday measures of risk.

References

- Basel Committee on Banking Supervision, 1996. Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements. Basel Committee on Banking Supervision.
- Basel Committee on Banking Supervision, 2006. Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version. Basel Committee on Banking Supervision.
- Bauwens, L., Giot, P., 2000. The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks. *Annales d'Economie et de Statistique* 60, 117–149.
- Bauwens, L., Hautsch, N., 2006. Stochastic conditional intensity processes. *Journal of Financial Econometrics* 4, 450–493.
- Bauwens, L., Hautsch, N., 2009. Modelling financial high frequency data using point processes. In: T. G. Andersen, R. A. Davis, J.-P. K., Mikosch, T. (Eds.), *Handbook of Financial Time Series*. Vol. 6. Springer, pp. 953–979.
- Berkowitz, J., Christoffersen, P., Pelletier, D., 2009. Evaluating value-at-risk models with desk-level data. *Forthcoming in Management Science*.
- Chavez-Demoulin, V., Davison, A., McNeil, A., 2005. A point process approach to value-at-risk estimation. *Quantitative Finance* 5, 227–234.
- Coles, S., 2001. *An introduction to Statistical Modelling of Extreme Values*. Springer.
- Cotter, J., Dowd, K., 2006. Extreme spectral risk measures: An application to futures clearinghouse margin requirements. *Journal of Banking and Finance* 30, 3469–3485.
- Daley, D., Vere-Jones, D., 2003. *An Introduction to the Theory of Point Processes*. Springer.
- Danielsson, J., De Vries, C., 2000. Value-at-risk and extreme returns. *Annales d'Economie et de Statistique* 60, 239–270.
- Davis, R., Mikosch, T., 2009. Extreme value theory for GARCH processes. In: T. G. Andersen, R. A. Davis, J.-P. K., Mikosch, T. (Eds.), *Handbook of Financial Time Series*. Vol. 6. Springer, pp. 187–200.
- Dufour, A., Engle, R., 2000. The ACD model: predictability of the time between consecutive trades. Tech. rep., University of Reading, ISMA Centre.
- Embrechts, P., 2009. Linear Correlation and EVT: Properties and Caveats. *Journal of Financial Econometrics* 7, 30–39.
- Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. *Modelling Extremal Events*. Springer, Berlin.
- Engle, R., Russell, J., 1998. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica* 66, 1127–1162.
- Grammig, J., Maurer, K., 2000. Non-monotonic hazard functions and the autoregressive conditional duration model. *Econometrics Journal* 3, 16–38.

- Hautsch, N., 2004. Modelling irregularly spaced financial data: theory and practice of dynamic duration models. Springer.
- Herrera, R., Schipp, B., 2009. Self-exciting extreme value models on stock market crashes. In: Schipp, B., Krämer, W. (Eds.), Statistical Inference, Econometric Analysis and Matrix Algebra. Physica, pp. 209–231.
- Lunde, A., 1999. A generalized gamma autoregressive conditional duration model. Tech. rep., Aalborg University.
- McNeil, A., Frey, R., 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance* 7, 271–300.
- McNeil, A. J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press.
- Meitz, M., Teräsvirta, T., 2006. Evaluating models of autoregressive conditional duration. *Journal of Business & Economic Statistics* 24, 104 – 124.
- Mikosch, T., 2003. Modeling dependence and tails of financial time series. In: Finkenstaedt, B., Rootzen, H. (Eds.), Extreme values in finance, telecommunications and the environment. Chapman & Hall, pp. 185 – 286.
- Nelson, D., 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59, 347–370.
- Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association* 83, 9–27.
- Zhang, M., Russell, J., Tsay, R., 2001. A nonlinear conditional autoregressive duration model with applications to financial transactions data. *Journal of Econometrics* 104, 179–207.

Appendix A. Goodness of fit

Point process residual analysis involves rescaling or thinning the original point process to obtain a new point process that is homogeneous Poisson. The common element of residual analysis techniques is the construction of an approximate homogeneous Poisson process from the data points and an estimated conditional intensity function $\bar{\lambda}_g(t | \mathcal{H}_t; \theta)$. Suppose we observe a one-dimensional point process $\{t_1, \dots, t_n\}$ on $[0, T)$ with conditional intensity $\lambda_g(t | \mathcal{H}_t; \theta)$. It is well known that the points

$$\tau_i = \int_0^{t_i} \lambda_g(s | \mathcal{H}_t; \theta) ds, \quad (\text{A.1})$$

for $i = 1, \dots, N(T)$ constitute a homogeneous Poisson process of rate 1 on an interval $[0, N(T)]$ hence is part of a transformed time axis. This new point process is called the residual process. If the estimated model $\bar{\lambda}_g(t | \mathcal{H}_t; \theta)$ is close to the true conditional intensity, then the residual process resulting from replacing $\lambda_g(t | \mathcal{H}_t; \theta)$ with $\bar{\lambda}_g(t | \mathcal{H}_t; \theta)$ in (A.1) should closely resemble a homogeneous Poisson process of rate 1. Ogata (1988) used this random rescaling method of residual analysis to assess the fit of one dimensional point process models for earthquake occurrences. The resulting property of exponentially distributed durations enables us to test for the presence of a homogeneous Poisson process via a Kolmogorov-Smirnov test.

In the case of the marks, we provide the W-statistics to assess our success in modelling the temporal behavior of the exceedances of the threshold u . The W-statistic is defined by

$$W = \xi^{-1} \ln \left(1 + \xi \frac{y - u}{\beta(t, y | \mathcal{H}_t)} \right).$$

This statistic states that if the GPD parameter model is correct, then the residuals are approximately independent unit exponential variables. In practice, the independence assumption can be checked via an ACF plot of the residuals.

In relation to the accuracy of VaR estimates, we use a Bernoulli test with a null hypothesis of estimating correctly the VaR_{α, t_i} at time t_i against the alternative that the method systematically underestimates the returns $r_{t_{i+1}}$. Thus, the indicator for a violation at time t_i is Bernoulli $I_t := 1_{\{r_{t_{i+1}} > \{VaR_{\alpha, t_i}\}\}} \sim Be(1 - \alpha)$. How it is described by McNeil and Frey (2000), I_t and I_s are independent for $t, s \in T$, then

$$\sum_{t_i \in T} I_{t_i} \sim B(n, 1 - \alpha). \quad (\text{A.2})$$

Expression (A.2) is a two-tailed test that is asymptotically distributed as binomial. We perform

the null hypothesis that it is a method that correctly estimates the risk measures against the alternative that the method has a systematic estimation error and gives too few or too many violations. A p -value less than 0.05 is interpreted as evidence against the null hypothesis. Moreover, we implement a test statistics proposed by Berkowitz et al. (2009) for the autocorrelations of de-meanned violations $\{I_t - \alpha\}$, which form a martingale difference sequence. This is a Ljung-Box statistic, which is a joint test of whether the first m autocorrelations of $\{I_t - \alpha\}$ are zero by calculating

$$LB(m) = T(T+2) \sum_{k=1}^m \frac{\gamma_k^2}{T-k}$$

where T is the sample size, γ_k is the sample autocorrelation at lag k and $LB(m)$ is asymptotically chi-square with m degrees of freedom.

Models	Parameters							Loglik	AIC	
	ACD Model				POT model					
	w	a ₁	b ₁	δ	γ	k	ξ			ω
bACDu	1.342 (0.691)	0.187 (0.057)	0.731 (0.073)		0.993 (0.135)	1.532 (0.157)	1.091 (0.072)	0.144 (0.048)		-2053.55 4121.11
bACDI	1.030 (0.497)	0.216 (0.059)	0.740 (0.060)		0.963 (0.126)	1.468 (0.143)	0.557 (0.095)	0.095 (0.044)	2.9e-08 (0.01)	4.11 (0.77)
bACDe	1.169 (0.623)	0.206 (0.062)	0.738 (0.068)		1.004 (0.135)	1.536 (0.159)	0.663 (0.164)	0.123 (0.046)	0.05 (0.04)	0.51 (0.21)
bLog-ACDu	0.303 (0.141)	0.142 (0.035)	0.778 (0.068)		0.997 (0.125)	1.545 (0.152)	1.091 (0.072)	0.144 (0.048)		-2052.87 4119.73
bLog-ACDI	0.270 (0.109)	0.158 (0.034)	0.780 (0.056)		0.960 (0.119)	1.473 (0.139)	0.554 (0.097)	0.100 (0.044)	2.6e-07 (0.01)	4.08 (0.79)
bLog-ACDe	0.280 (0.128)	0.152 (0.035)	0.781 (0.063)		1.002 (0.124)	1.544 (0.152)	0.675 (0.162)	0.124 (0.046)	0.05 (0.04)	0.53 (0.22)
bEXACDu	0.133	0.158	0.889	0.000	0.980	1.519	1.091	0.144		-2054.38 4124.75
	0.100	0.037	0.041	0.008	0.128	0.150	0.071	0.048		
bEXACDI	0.065 (0.073)	0.175 (0.038)	0.913 (0.030)	0.000 (0.007)	0.959 (0.121)	1.464 (0.137)	0.562 (0.094)	0.091 (0.044)	1.1e-07 (0.010)	4.079 (0.763)
bEXACDe	0.089 (0.087)	0.169 (0.039)	0.905 (0.035)	0.000 (0.008)	0.992 (0.127)	1.525 (0.150)	0.672 (0.160)	0.122 (0.046)	0.051 (0.044)	0.525 (0.213)
bBCACDu	0.318 (0.147)	0.174 (0.053)	0.885 (0.045)	0.824 (0.460)	0.981 (0.128)	1.521 (0.150)	1.091 (0.072)	0.144 (0.048)		-2054.31 4124.62
bBCACDI	0.271 (0.116)	0.195 (0.049)	0.909 (0.033)	0.956 (0.120)	0.796 (0.359)	1.463 (0.137)	0.558 (0.095)	0.093 (0.044)	8.9e-07 (0.010)	4.098 (0.770)
bBCACDe	0.285 (0.151)	0.202 (0.159)	0.925 (0.024)	1.230 (0.086)	1.960 (0.482)	1.766 (0.152)	0.661 (0.162)	0.120 (0.046)	0.052 (0.044)	0.500 (0.206)

Table A.1: Results of the estimation of all ACD-POT models with distributional assumption Burr for the standardized durations of the inter-exceedance times for the Bayer returns. Standard deviations are given in parentheses. Loglike are the results of the maximization of the log-likelihood estimation and AIC is the Akaike Information Criterion.

Models	Parameters										Loglik	AIC
	ACD model					POT Model						
	w	a_1	b_1	δ	γ	k	ξ	ω	β_1	β_2		
gACDu	0.833 (0.326)	0.135 (0.035)	0.775 (0.057)		0.157 (0.046)	33.093 (19.491)	1.091 (0.072)	0.144 (0.048)			-2042.72	4099.45
gACDI	0.754 (0.280)	0.157 (0.036)	0.767 (0.051)		0.121 (0.041)	53.767 (36.539)	0.503 (0.100)	0.094 (0.044)	5.7e-07 (0.010)	4.492 (0.826)	-2024.8	4067.59
gACDe	0.769 (0.190)	0.146 (0.036)	0.773 (0.052)		0.097 (0.021)	86.674 (41.234)	0.657 (0.166)	0.123 (0.046)	0.055 (0.045)	0.513 (0.212)	-2032.27	4082.55
gLog-ACDu	0.175 (0.078)	0.122 (0.031)	0.831 (0.053)		0.162 (0.044)	31.198 (17.101)	1.091 (0.072)	0.144 (0.048)			-2042.61	4099.23
gLog-ACDI	0.184 (0.072)	0.139 (0.030)	0.814 (0.049)		0.179 (0.046)	24.668 (12.629)	0.474 (0.105)	0.097 (0.043)	3.2e-07 (0.009)	4.731 (0.887)	-2025.03	4068.07
gLog-ACDe	0.161 (0.069)	0.128 (0.031)	0.833 (0.049)		0.108 (0.029)	69.870 (37.581)	0.653 (0.165)	0.124 (0.046)	0.055 (0.045)	0.510 (0.208)	-2031.9	4081.8
gEXACDu	0.643 (0.241)	0.343 (0.064)	0.599 (0.112)	0.095 (0.128)	0.974 (0.000)	0.949 (0.000)	0.952 (0.071)	0.346 (0.083)			-2133.07	4282.15
gEXACDI	0.675 (0.102)	0.308 (0.075)	0.582 (0.040)	0.156 (0.083)	0.989 (0.000)	0.980 (0.000)	0.638 (0.063)	0.229 (0.053)	0.182 (0.055)	0.504 (0.027)	-2138.13	4296.27
gEXACDe	0.053 (0.057)	0.129 (0.023)	0.916 (0.028)	0.001 (0.010)	0.132 (0.023)	47.045 (21.023)	0.653 (0.164)	0.120 (0.046)	0.055 (0.045)	0.504 (0.205)	-2033.52	4087.05
gBCACDu	0.214 (0.081)	0.133 (0.044)	0.905 (0.033)	0.889 (0.394)	0.160 (0.046)	31.695 (18.085)	1.090 (0.072)	0.145 (0.048)			-2043.57	4103.14
gBCACDI	0.205 (0.074)	0.157 (0.039)	0.916 (0.029)	0.776 (0.314)	0.139 (0.043)	41.114 (25.380)	0.502 (0.100)	0.093 (0.043)	2.6e-08 (0.008)	4.501 (0.830)	-2025.24	4070.48
gBCACDe	0.197 (0.075)	0.145 (0.043)	0.915 (0.030)	0.838 (0.360)	0.120 (0.041)	56.519 (38.521)	0.664 (0.161)	0.123 (0.046)	0.053 (0.045)	0.520 (0.208)	-2033.16	4086.31

Table A.2: Results of the estimation of all ACD-POT models with distributional assumption generalize gamma for the standardized durations of the inter-exceedance times for the Bayer returns. Standard deviations are given in parentheses. Loglike are the results of the maximization of the log-likelihood estimation and AIC is the Akaike Information Criterion.

Tables and Figures

Models	Number of violations			Models	Number of violations		
	$VaR_{0.95}$	$VaR_{0.99}$	$VaR_{0.999}$		$VaR_{0.95}$	$VaR_{0.99}$	$VaR_{0.999}$
bACDu	389 (0.00, 0.00)	103 (0.00, 0.00)	2 (0.35, 1)	gACDu	386 (0.00, 0.00)	95 (0.00) (0.00, 0.00)	0 (0.02, -)
bACDI	248 (0.40, 0.15)	63 (0.02, 0.06)	7 (0.28, 0.24)	gACDI	245 (0.52, 0.22)	54 (0.30, 0.48)	6 (0.48, 0.51)
bACDe	294 (0.00, 0.00)	55 (0.24, 0.26)	3 (0.64, 0.86)	gACDe	291 (0.00, 0.00)	46 (0.94, 0.88)	0 (0.02, -)
bLog-ACDu	389 (0.00, 0.00)	102 (0.00, 0.00)	1 (0.10, 1)	gLog-ACDu	390 (0.00, 0.00)	90 (0.00, 0.00)	0 (0.02, -)
bLog-ACDI	251 (0.30, 0.11)	56 (0.18, 0.25)	7 (0.28, 0.24)	gLog-ACDI	243 (0.60, 0.36)	43 (0.61, 0.92)	5 (0.82, 0.61)
bLog-ACDe	298 (0.00, 0.00)	51 (0.56, 0.38)	2 (0.35, 1)	gLog-ACDe	289 (0.00, 0.00)	41 (0.42, 0.93)	0 (0.02, -)
bEXACDu	388 (0.00, 0.00)	102 (0.00, 0.00)	1 (0.10, 1)	gEXACDu	387 (0.00, 0.00)	61 (0.05, 0.09)	3 (0.64, 0.86)
bEXACDI	248 (0.40, 0.15)	64 (0.01, 0.02)	7 (0.28, 0.24)	gEXACDI	400 (0.00, 0.00)	144 (0.00, 0.00)	10 (0.03, 0.00)
bEXACDe	293 (0.00, 0.00)	53 (0.17, 0.62)	2 (0.35, 1)	gEXACDe	288 (0.00, 0.00)	43 (0.61, 0.92)	0 (0.02, -)
bBCACDu	386 (0.00, 0.00)	102 (0.00, 0.00)	1 (0.10, 1)	gBCACDu	389 (0.00, 0.00)	93 (0.00, 0.00)	0 (0.02, -)
bBCACDI	250 (0.33, 0.11)	64 (0.02, 0.02)	7 (0.28, 0.24)	gBCACDI	244 (0.57, 0.31)	53 (0.17, 0.29)	6 (0.48, 0.51)
bBCACDe	289 (0.00, 0.00)	70 (0.00, 0.00)	6 (0.48, 0.51)	gBCACDe	293 (0.00, 0.00)	43 (0.69, 0.92)	0 (0.02, -)

Table A.3: Some VaR in-sample results all models. The number of violations for all VaR confidence levels is displayed for each model. Values in parentheses are p-values for the unconditional coverage test and the Ljung-Box statistic with 5 lags. The number of observations in the sample is 4607.

Model	Number of violations		
	$VaR_{0.95}$	$VaR_{0.99}$	$VaR_{0.999}$
gACDl	18 (0.34, 0.38)	4 (0.54, 0.81)	0 (1, -)
gLog-ACDl	20 (0.14, 0.72)	3 (0.77, 0.86)	0 (1, -)
gBCACDl	19 (0.22, 0.48)	5 (0.22, 0.76)	0 (1, -)

Table A.4: Some VaR backtesting results for three models. Values in parentheses are p-values for the unconditional coverage test and the Ljung-Box test iwth 5 lags. Values smaller than a p-value 0.05 indicate failure. The number of observations in the sample is 288.

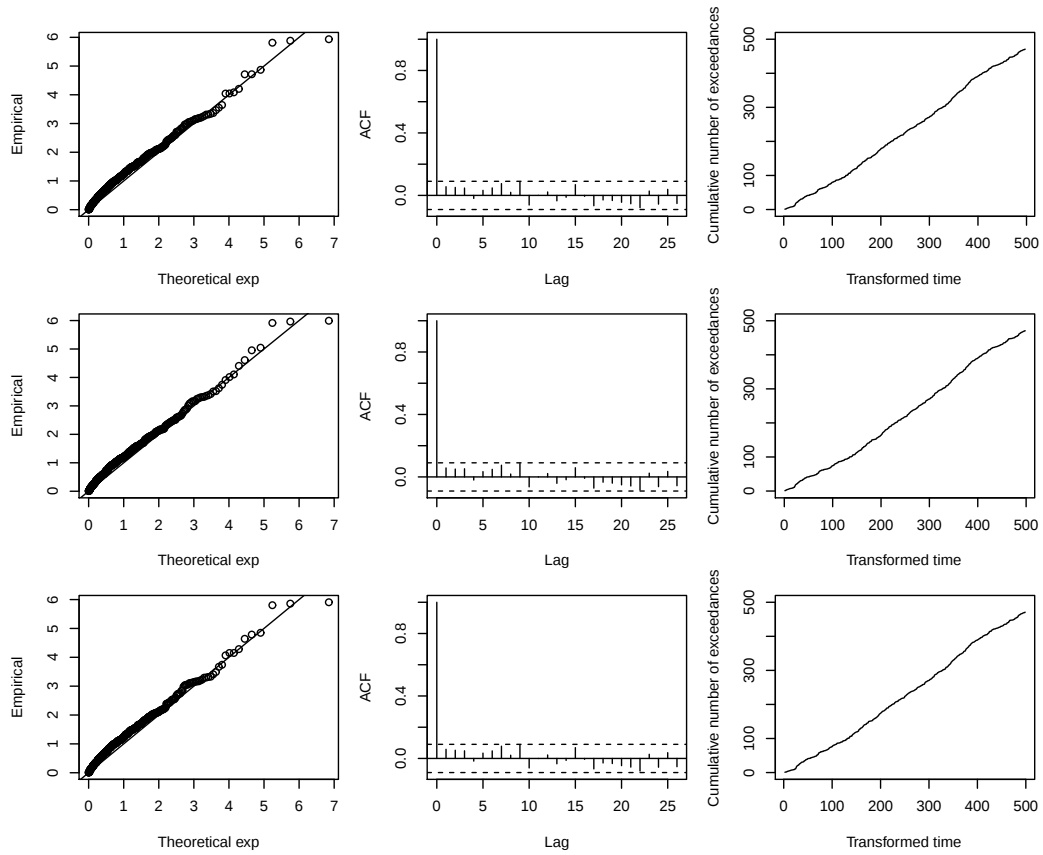


Figure A.1: QQ-plots of the residuals (left), autocorrelation function of the residuals (middle) and cumulative numbers of the residual process versus the transformed time $\{\tau_i\}$ (right), for the returns of the Bayer index for the gACDl (upper panel), gLog-ACDl (middle panel) and gBCACDl (lower panel) models.

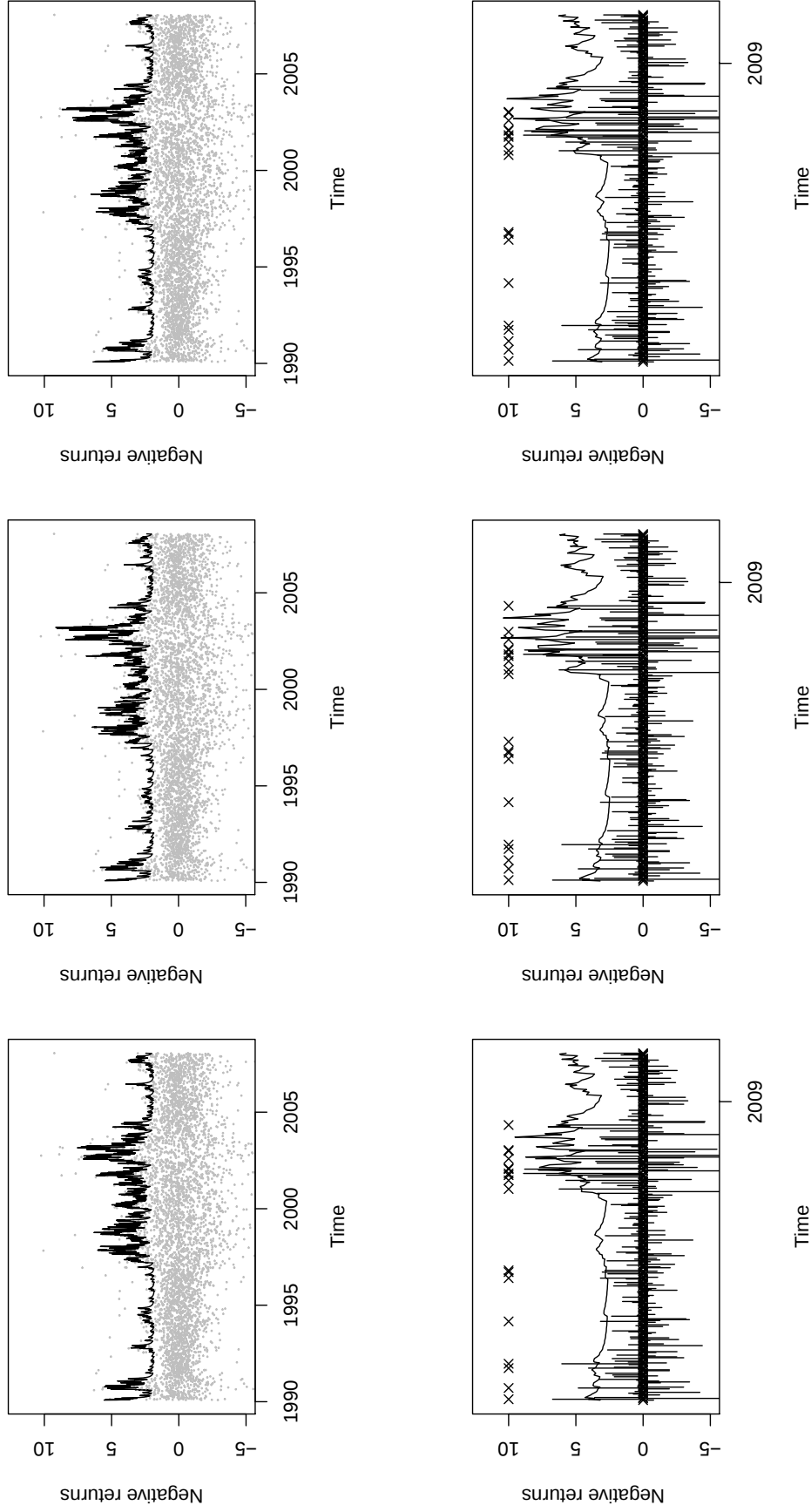


Figure A.2: In-sample and Backtest estimation of the VaR for the 0.99 confidence level for the models fitted to the Bayer returns. The black line is the VaR estimation. In the top panel from left to right we have the VaR estimates in-sample for the gACD, gLog-ACD and gBCACD models. In the bottom panel from left to right we have the backtesting results for the VaR estimates for the gACD, gLog-ACD and gBCACD models. The x marks at the top of the figures indicate the violations of the VaR at the 0.99 confidence level.

SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Localising temperature risk" by Wolfgang Karl Härdle, Brenda López Cabrera, Ostap Okhrin and Weining Wang, January 2011.
- 002 "A Confidence Corridor for Sparse Longitudinal Data Curves" by Shuzhuan Zheng, Lijian Yang and Wolfgang Karl Härdle, January 2011.
- 003 "Mean Volatility Regressions" by Lu Lin, Feng Li, Lixing Zhu and Wolfgang Karl Härdle, January 2011.
- 004 "A Confidence Corridor for Expectile Functions" by Esra Akdeniz Duran, Mengmeng Guo and Wolfgang Karl Härdle, January 2011.
- 005 "Local Quantile Regression" by Wolfgang Karl Härdle, Vladimir Spokoiny and Weining Wang, January 2011.
- 006 "Sticky Information and Determinacy" by Alexander Meyer-Gohde, January 2011.
- 007 "Mean-Variance Cointegration and the Expectations Hypothesis" by Till Strohsal and Enzo Weber, February 2011.
- 008 "Monetary Policy, Trend Inflation and Inflation Persistence" by Fang Yao, February 2011.
- 009 "Exclusion in the All-Pay Auction: An Experimental Investigation" by Dietmar Fehr and Julia Schmid, February 2011.
- 010 "Unwillingness to Pay for Privacy: A Field Experiment" by Alastair R. Beresford, Dorothea Kübler and Sören Preibusch, February 2011.
- 011 "Human Capital Formation on Skill-Specific Labor Markets" by Runli Xie, February 2011.
- 012 "A strategic mediator who is biased into the same direction as the expert can improve information transmission" by Lydia Mechtenberg and Johannes Münster, March 2011.
- 013 "Spatial Risk Premium on Weather Derivatives and Hedging Weather Exposure in Electricity" by Wolfgang Karl Härdle and Maria Osipenko, March 2011.
- 014 "Difference based Ridge and Liu type Estimators in Semiparametric Regression Models" by Esra Akdeniz Duran, Wolfgang Karl Härdle and Maria Osipenko, March 2011.
- 015 "Short-Term Herding of Institutional Traders: New Evidence from the German Stock Market" by Stephanie Kremer and Dieter Nautz, March 2011.
- 016 "Oracally Efficient Two-Step Estimation of Generalized Additive Model" by Rong Liu, Lijian Yang and Wolfgang Karl Härdle, March 2011.
- 017 "The Law of Attraction: Bilateral Search and Horizontal Heterogeneity" by Dirk Hofmann and Salmai Qari, March 2011.
- 018 "Can crop yield risk be globally diversified?" by Xiaoliang Liu, Wei Xu and Martin Odening, March 2011.
- 019 "What Drives the Relationship Between Inflation and Price Dispersion? Market Power vs. Price Rigidity" by Sascha Becker, March 2011.
- 020 "How Computational Statistics Became the Backbone of Modern Data Science" by James E. Gentle, Wolfgang Härdle and Yuichi Mori, May 2011.
- 021 "Customer Reactions in Out-of-Stock Situations – Do promotion-induced phantom positions alleviate the similarity substitution hypothesis?" by Jana Luisa Diels and Nicole Wiebach, May 2011.

SFB 649, Ziegelstraße 13a, D-10117 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 022 "Extreme value models in a conditional duration intensity framework"
by Rodrigo Herrera and Bernhard Schipp, May 2011.

SFB 649, Ziegelstraße 13a, D-10117 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

