An ACD-ECOGARCH(1,1) model

CLAUDIA CZADO * and STEPHAN HAUG*

AbstractIn this paper we introduce an ACD-ECOGARCH$(1,1)$ model. An exponential autoregressive conditional duration model is used to describe the dependence structure in durations of ultra-high-frequency financial data. The innovation process of the ACD model then defines the interarrival times of a compound Poisson process. We use this compound Poisson process as the background driving Lévy process of an exponential continuous time GARCH$(1,1)$ process. The dynamics of the random time transformed log-price process are then described by the latter process. To estimate its parameters we construct a quasi maximum likelihood estimator under the assumption that all jumps of the log-price process are observable. Finally the model is fitted for illustrative purpose to General Motors tick-by-tick data of the New York Stock Exchange.

Keywords and phrases: ultra-high-frequency data, ECOGARCH, ACD, QMLE, leverage effect.

This is a pre-copy-editing, author-produced PDF of an article accepted for publication in *Journal of Financial Econometrics* following peer review. The definitive publisher-authenticated version
An ACD-ECOGARCH(1,1) Model.
Journal of Financial Econometrics 2009; doi: 10.1093/jjfinec/nbp023 is available online at:
http://jfec.oxfordjournals.org/cgi/content/abstract/nbp023?ijkey=ZC6HpnJyVZwzHSi&keytype=ref

1. Introduction

The fundamental characteristic of tick-by-tick, or also called ultra-high-frequency, data is the irregular spacing of the observation times. This feature prevents the application of standard discrete time econometric models to analyse such kind of data. The reason is of course that in such models the durations between two observations are assumed to be constant. Thus new econometric methods have to be developed for the analysis of ultra-high-frequency data. In doing so one has to deal with several problems. One such problem is that the random durations seem not to be independent but show an autoregressive dependence structure given the past observations. Therefore Engle and Russell (1998) introduce the autoregressive conditional duration model (ACD) to describe such a behaviour. Based on the ACD model there were several extensions of the GARCH process developed to model irregularly sampled financial time series. Here we have to mention the ACD-GARCH model of Ghysels and Jasiak (1998) and the work of Engle (2000). Both approaches are summarised and compared in Meddahi, Renault, and Werker (2006). The authors also propose a further specification of a GARCH model for irregularly spaced data, which incorporates the advantages of the previous two models. Grammig and Wellner (2002) extend the UHF-GARCH model of Engle (2000) by modelling the interdependence of intraday volatility and trading intensity. All of these models are based on the discrete time weak or strong GARCH process. A different way to model tick-by-tick data is to assume the existence of an underlying continuous time model. Such

*Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, D-85748 Garching, Germany. *Email:* {cczado,haug}@ma.tum.de

an approach was developed in Maller, Müller, and Szimayer (2008). They specify a discrete time approximation of the continuous time GARCH(1, 1) process (COGARCH) defined in Klüppelberg, Lindner, and Maller (2004), which is suitable for irregularly spaced observation times. This idea can be extended to other continuous time GARCH or stochastic volatility models as long as the approximation has a tractable form. We refer to Lindner (2009) and the references therein for an overview of continuous time approximations of GARCH and stochastic volatility models.

Bollerslev, Litvinova, and Tauchen (2006) report that a model, which is applied to high-frequency financial data, has to be able to describe the so called leverage effect. A continuous time model with a tractable discretisation, which further incorporates a leverage effect, is the exponential COGARCH process recently introduced by Haug and Czado (2007). However the exponential COGARCH as well as the other approaches based on continuous time models can not directly deal with a dependence structure in the durations between observations. Therefore we will combine in the present paper the ACD model and the exponential COGARCH(1, 1) process to address both, the dependence structure in the durations and the leverage effect. Before we explain the structure of the current paper, we will recall from Haug and Czado (2007) the definition of an exponential COGARCH(1, 1) process, abbreviated to ECOGARCH(1, 1). For a detailed discussion of the ACD model we refer to Engle and Russell (1998). A survey of the several extensions and enhancements of the ACD model is provided by Pacurar (2008).

Haug and Czado (2007) consider a zero mean Lévy processes $L := (L_t)_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, P)$. In this paper we only consider the case of L being the compound Poisson process (CPP)

$$L_t = \sum_{k=1}^{J_t} Z_k, \quad t > 0, \qquad L_0 = 0\,, \tag{1}$$

where $(J_t)_{t\geq 0}$ is an independent Poisson process with intensity $\lambda > 0$ and $(Z_k)_{k\in\mathbb{N}}$ is an i.i.d. sequence of random variables independet of J. The jump sizes Z_k are assumed to have a symmetric distribution function $F_{0,1/\lambda}$ with mean 0 and variance $1/\lambda$. In this case the ECOGARCH(1, 1) process is defined as follows:

Let L be the CPP (1). Then the ECOGARCH(1, 1) process G is defined as the stochastic process satisfying,

$$G_t = \int_0^t \sigma_{s-} dL_s = \sum_{k=1}^{J_t} \sigma_{t_k -} Z_k\,, \quad t > 0, \quad G_0 = 0\,,$$

where $(t_k)_{k\in\mathbb{N}}$ are the jump times of L. The log-volatility process $(\log(\sigma_t^2))_{t\geq 0}$ is an Ornstein-Uhlenbeck process with state space representation

$$\begin{aligned} \log(\sigma_t^2) &= \mu + b_1 X_t \\ X_t &= e^{-a_1 t} X_0 + \int_0^t e^{-a_1(t-s)} dM_s\,, \quad t > 0\,, \end{aligned}$$

and parameters $\mu, a_1, b_1 \in \mathbb{R}$. Here $X_0 \in \mathbb{R}$ is independent of the driving CPP L and $M_t = \sum_{k=1}^{J_t} [\theta Z_k + \gamma |Z_k|] - \gamma\lambda K t$, with $K = \int_{\mathbb{R}} |x| F_{0,1/\lambda}(dx)$, is a zero mean CPP with parameters $(\theta, \gamma) \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$.

The paper is now organised as follows. In Section 2 we show how to extend the ECOGARCH model by incorporating an ACD model. Since estimation of the ACD model is rather straightforward and very well explained in Engle and Russell (1998), we concentrate in Section 3 on introducing a quasi maximum likelihood estimator (QMLE) for the parameters of the ECOGARCH$(1,1)$ process under the assumption of full observation of the sample path. An illustrative data example is presented in Section 4. Concluding remarks are made in Section 5.

2. ACD-ECOGARCH(1,1) model

We assume to have ultra-high-frequency observations $P_{T_1}, \dots, P_{T_n}$ of an asset log-price at transaction times T_i, $i = 1, \dots, n$. The observed durations ΔT_i will be modeled by an exponential ACD(p,q) model, $p, q \in \mathbb{N}$, as introduced in Engle and Russell (1998), i.e.

$$\Delta T_i = \psi_i \Delta t_i \,,$$

where

$$\psi_i = \mathbb{E}(\Delta T_i | \Delta T_{i-1}, \dots \Delta T_1) = \omega + \sum_{j=1}^{p} \alpha_j \Delta T_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} \,.$$

We have to assume an exponential distribution for the innovations Δt_i since they will be the interarrival times of the driving compound Poisson process L of the ECOGARCH$(1,1)$ process. This implies that the observations G_{t_i} of the ECOGARCH$(1,1)$ process are given by $G_{t_i} := P_{T_i}\,,\ i = 1, \dots, n\,.$ The ACD(p,q) model can of course be replaced by one of the many extension introduced in the recent years as long as the innovations are exponentially distributed. We refer to Pacurar (2008) for an extensive overview of ACD models.
Here we want to mention that this approach is different to the one proposed by Ghysels and Jasiak (1998). They use the expected conditional duration ψ_i and not the innovation Δt_i as an input variable for the volatility equation at time T_i. In contrast Engle (2000) assumes in his simplest volatility model that the variance per unit time follows a GARCH$(1,1)$ equation. Thus he considers the model

$$\tilde\sigma_i^2 = \mathbb{V}\mathrm{ar}(r_i/\sqrt{\Delta T_i} \mid r_j, \Delta T_j, j < i;\ \Delta T_i) = c + a(r_i/\sqrt{\Delta T_i})^2 + b\tilde\sigma_{i-1}^2 \,,$$

where $r_i = P_{T_i} - P_{T_{i-1}}$. He further proposes specifications of $\tilde\sigma_i^2$ including the duration ΔT_i as well as the expected duration ψ_i.

3. Estimation in the ACD-ECOGARCH(1,1) model

To estimate the parameters in our model we will follow a two step estimation strategy. In a first step the MLE $\widehat{\boldsymbol\vartheta}_n^d$ of the parameter $\boldsymbol\vartheta^d \in \mathbb{R}^{p+q+1}$ of the ACD(p,q) model will be computed as described in Engle and Russell (1998). In our example in Section 4 we consider the cases $p = q = 1$ and $p = q = 2$. Given the observed durations ΔT_i, $i = 1, \dots, n$, and the MLE $\widehat{\boldsymbol\vartheta}_n^d$ we can compute the fitted innovations

$$\widehat{\Delta t_i} = \frac{\Delta T_i}{\widehat{\psi}_i}\,, \qquad i = 1, \dots, n\,,$$

where $\widehat{\psi}_i = \widehat{\omega}_n + \sum_{j=1}^{p} \widehat{\alpha_j}_n \Delta T_{i-j} + \sum_{j=1}^{q} \widehat{\beta_j}_n \widehat{\psi}_{i-j}\,.$ In the following we will denote them by Δt_i for ease of notation. Hence after the first estimation step the data is given by the pairs

$\{(G_{t_i}^{\Delta t_i}, \Delta t_i)\,,\ i = 1, \dots, n\}$, where $G_{t_i}^{\Delta t_i} := G_{t_i} - G_{t_{i-1}}$.
Since we assume to observe G at n consecutive jump times $0 = t_0 < t_1 < \cdots < t_n$, the state process X of the log-volatility process has the following autoregressive representation

$$b_1 X_{t_i} \;=\; b_1 e^{-a_1 \Delta t_i} X_{t_{i-1}} + b_1 \theta Z_i + b_1 \gamma \left(|Z_i| - \frac{\lambda K}{a_1}(1 - e^{-a_1 \Delta t_i}) \right) \tag{2}$$

(cf. Example 3.1 in Haug and Czado (2007)). Here we used the fact $J_{t_{i-1}} + 1 = J_{t_i} = i$. This implies that the left-hand limit $\log(\sigma_{t_i-}^2)$ of the log-volatility process at the jump times $0 < t_1 < \cdots < t_n$ is given by

$$\log(\sigma_{t_i-}^2) = \mu + b_1 e^{-a_1 \Delta t_i} X_{t_{i-1}} - b_1 \gamma \frac{\lambda K}{a_1}(1 - e^{-a_1 \Delta t_i})\,. \tag{3}$$

In Haug and Czado (2007), Proposition 3.1, it is shown that the leverage effect depends on the sign of θb_1. To identify θ as the leverage parameter we will set b_1 equal to one in the following. The observations of the ECOGARCH process are then $G_{t_i} = \sum_{k=1}^{J_{t_i}} \sigma_{t_k-} Z_k = G_{t_{i-1}} + \sigma_{t_i-} Z_i$, which implies that the return at time t_i is equal to $G_{t_i}^{\Delta t_i} := \sigma_{t_i-} Z_i$.
Given the data $\{(G_{t_i}^{\Delta t_i}, \Delta t_i)\,,\ i = 1, \dots, n\}$, we now aim at estimating the remaining unknown parameters $\boldsymbol{\vartheta}^g := (a_1, \theta, \gamma, \mu, \lambda, K) =: (\boldsymbol{\vartheta}, \lambda, K)$ in our model. But equation (3) contains an identifiability problem. The constant term in (3) is given by $\mu^* := \mu - \gamma \frac{\lambda K}{a_1}$. In the QML approach, which we will take, only the constant term μ^* is identifiable and not μ, K and λ. Because of that we will estimate the rate λ given only the jump times $t_1, \dots, t_n$ of the compound Poisson process through the MLE $\widehat{\lambda}_n := \frac{n}{\sum_{i=1}^n \Delta t_i}$. Different estimators of K and thus μ^* are analysed in Czado and Haug (2009) with regard to their finite sample properties. In the following we work with the approximation $\widehat{K}_n := (\frac{\pi}{2}\widehat{\lambda}_n)^{-1/2}$, which is motivated by the fact that $K = (\frac{\pi}{2}\lambda)^{-1/2}$ in case $F_{0,1/\lambda}$ is a normal distribution.
To derive a contrast function, which we can maximise with respect to the unknown parameters, we followed Engle (2000) by assuming the log-likelihood is of the following form

$$\log \rho_{(\lambda,\boldsymbol{\vartheta})}(\mathbf{G}_n^{\Delta}, \boldsymbol{\Delta}\mathbf{t}_n | X_0) \;=\; \sum_{i=1}^{n} \left(\log \rho^g_{(\lambda,\boldsymbol{\vartheta})}(G_{t_i}^{\Delta t_i} | \mathbf{G}_{i-1}^{\Delta}, \boldsymbol{\Delta}\mathbf{t}_i) + \log \rho^d_{\lambda}(\Delta t_i | \boldsymbol{\Delta}\mathbf{t}_{i-1}) \right)\,,$$

where $\mathbf{G}_k^{\Delta} = (G_{t_1}^{\Delta t_1}, \dots, G_{t_k}^{\Delta t_k})$, $1 \le k \le n$, and $\boldsymbol{\Delta}\mathbf{t}_k$ is defined analogously. One should remember that the distribution of the durations Δt_i is independent of the current and past returns. Since the conditional distribution of the returns is unknown we will follow a QML approach in the second estimation step by choosing a Gaussian quasi log-likelihood as contrast function. This is analogously to QML estimation in discrete time conditionally heteroscedastic time series models, see e.g. Berkes and Horváth (2003), Berkes, Horváth, and Kokoszka (2003), Hall and Yao (2003), Jeantheau (1998) and the monograph Straumann (2005).

Since the Gaussian quasi log-likelihood depends on the volatilities $\sigma_{t_i}^2$, which are unobservable, it can not be evaluated numerically. Thus we need an approximation of the volatility, which is given by

$$\widehat{\sigma}_{t_i-}^2(\boldsymbol{\vartheta}, \widehat{\lambda}_n) := \exp\left(\mu + e^{-a_1 \Delta t_i} \widehat{X}_{t_{i-1}}(\boldsymbol{\vartheta}, \widehat{\lambda}_n) - \gamma \frac{\widehat{\lambda}_n \widehat{K}_n (1 - e^{-a_1 \Delta t_i})}{a_1} \right)\,,\ i = 1, \dots, n,$$

with $\widehat{X}_{t_i}(\boldsymbol{\vartheta},\widehat{\lambda}_n) = e^{-a_1\Delta t_i}\widehat{X}_{t_{i-1}}(\boldsymbol{\vartheta},\widehat{\lambda}_n) + \theta\frac{G_{t_i}^{\Delta t_i}}{\widehat{\sigma}_{t_i-}(\boldsymbol{\vartheta},\widehat{\lambda}_n)} + \gamma\left(\frac{|G_{t_i}^{\Delta t_i}|}{\widehat{\sigma}_{t_i-}(\boldsymbol{\vartheta},\widehat{\lambda}_n)} - \frac{\widehat{\lambda}_n\widehat{K}_n(1-e^{-a_1\Delta t_i})}{a_1}\right)$.
QML estimates $\widehat{\boldsymbol{\vartheta}}_n$ are then obtained by maximising

$$L(\boldsymbol{\vartheta}|\mathbf{G}_n^{\Delta},\boldsymbol{\Delta}_n,\widehat{\lambda}_n) := -\frac{1}{2}\sum_{i=1}^{n}\left(\log(\widehat{\sigma}^2_{t_i-}(\boldsymbol{\vartheta},\widehat{\lambda}_n)) + \frac{(G_{t_i}^{\Delta t_i})^2}{\widehat{\sigma}^2_{t_i-}(\boldsymbol{\vartheta},\widehat{\lambda}_n)/\widehat{\lambda}_n}\right)$$

with respect to $\boldsymbol{\vartheta}$. An extensive simulation study of the small sample properties of $\widehat{\boldsymbol{\vartheta}}_n$ is given in Czado and Haug (2009), showing satisfactory results.

4. An illustrative example

As an illustration of the potential usefulness of the ACD-ECOGARCH model, it is now applied to tick-by-tick data traded on NYSE: GM (General Motors). The data was extracted from the Trade and Quote database released by the NYSE. The time period under consideration spans four weeks starting form 6th of May 2002 until the 31st of May. Due to the Memorial Day there was no trading on the 27th of May at the NYSE. Only transaction between 9:30am and 4:00pm are considered. If equal transaction times T_i occurred, the corresponding trades are combined to a single trade at an average price. We omitted consecutive zero returns if they occurred at the beginning or end of the trading day. On average we have about 1960 observations per day.
The data will be analysed on a daily basis to get insight about varying parameter values over the observation period. Since durations in ultra-high-frequency data are characterised by an intraday seasonality, as e.g. reported in Bauwens, Giot, Grammig, and Veredas (2004), Engle and Russell (1998) or Tay, Ting, Tse, and Warachka (2009), we diurnally adjusted them at first. For that purpose we fitted a cubic smoothing spline to the durations of each day of the week. The diurnally adjusted durations are then computed by dividing each durations with the corresponding smoothing spline value. If we would aim at estimating one model for the whole data set, then the overnight durations have to be adjusted as e.g. explained in Bauwens and Giot (2000). Typically the volatility also shows a deterministic time-of-day effect, see e.g. Engle (2000). We therefore computed diurnally adjusted returns by dividing each returns with the corresponding value of a cubic smoothing spline fitted to the absolute returns. The resulting smoothing splines are shown in Figure 1. The volatility smoothing splines for Wednesday and Thursday show a rather atypical behaviour of slightly increasing during the first half of the trading day and decreasing afterwards. The shapes of the remaining splines are conform with results reported in the literature. Further we have to take into account a market microstructure noise on this fine level. To address this problem we will follow Engle (2000), by considering mid quotes, which are the average of the last bid and ask quote just before the trade, as our price data. In particular this means, if we have observation points $T_1,\dots,T_n$, then the log-price P_{T_i} is given by

$$P_{T_i} = \frac{1}{2}\left(\log(b_{T_i-}) + \log(a_{T_i-})\right), \qquad i = 1,\dots,n,$$

where b_{T_i-} (a_{T_i-}) denotes the last bid (ask) quote just before or at time T_i.

However one has to be aware that this choice of price measure reduces the econometric issues of bid ask bounce and price discreteness but it does not eliminate these problems

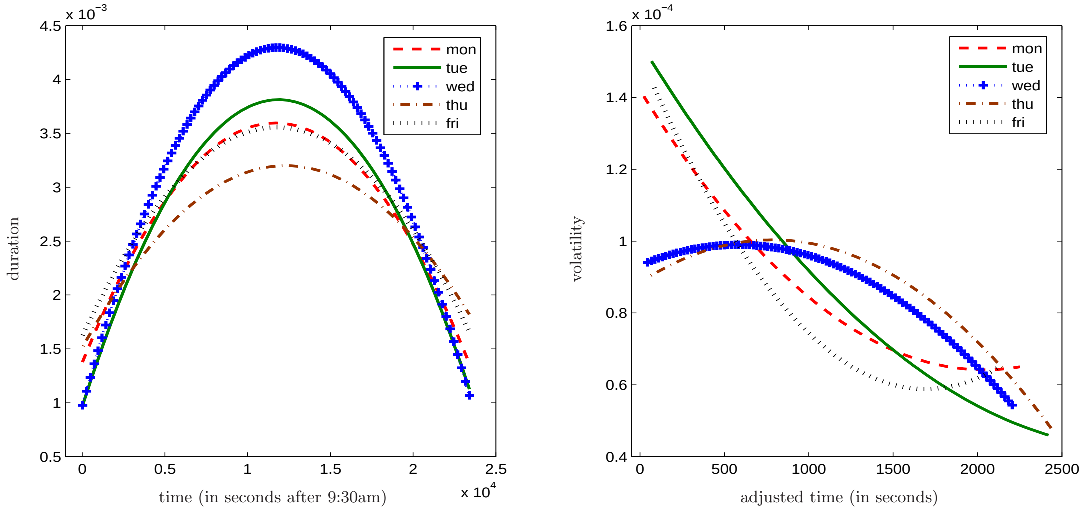

Figure 1. *Cubic smoothing splines of durations (*left*) and absolute returns (*right*).*

as mentioned by Engle (2000). More sophisticated models or approaches in dealing with market microstructure noise can be found e.g. in Aït-Sahalia, Mykland, and Zhang (2005) or Engle and Sun (2007) and references therein.

Autocorrelation in the diurnally adjusted durations was tested through a Ljung-Box test with 15 degrees of freedom. The hypothesis of no correlation is rejected for all 19 days. An ACD model was estimated for each day. We just considered the cases $p, q \in \{1, 2\}$ and chose p and q such that the test statistic of the Ljung-Box test with 15 degrees of freedom applied to the fitted innovations $\widehat{\Delta} t_i = \frac{\Delta T_i}{\widehat{\psi}_i}$, $i = 1, \ldots, n,$ was minimal. The hypothesis of no correlation in the fitted innovations could not be rejected at the 0.05 significance level on each of the 19 days. Except for two days, where we fitted an ACD$(1, 1)$ model, an ACD$(2, 2)$ model was utilised. The sums over the estimated coefficients vary mainly between 0.6 and 0.9, but significantly smaller values such as 0.18 are also obtained for two of the days. Given the fitted innovations we define through $t_0 = 0, t_i = t_{i-1} + \widehat{\Delta} t_i$, $i = 1, \ldots, n$, the observations of the ECOGARCH process as $G_{t_i} = P_{T_i}$, $i = 1, \ldots, n$.

The parameter $\boldsymbol{\vartheta}^g$ of the ECOGARCH$(1, 1)$ process is then estimated as explained in Section 3[1]. The estimated parameter values suggest that we have a leverage effect, which is the case if $\widehat{\theta}_n < 0$, on 9 of the 19 days. On these days we observe different types of leverage effects. We have the case that a positive jump to the log-price increases the log-volatility less than a negative one ($-\widehat{\gamma}_n < \widehat{\theta}_n < 0$), the case that a negative jump in the price process decreases the log-volatility less than a positive one ($\widehat{\gamma}_n < \widehat{\theta}_n < 0$) and also the case that a negative jump in the log-price processes increases while a positive one decreases the log-volatility ($\widehat{\theta}_n < -|\widehat{\gamma}_n|$). From equation (3) we see that mostly long durations will decrease the volatility as long as $\widehat{\gamma}_n$ is positive, which is the case for 17 of the 19 days. The parameter a_1 reflects strong dependence in the log-volatility process for most of the days by taking on values between 0.0074 and 0.1577. However we also have two days with almost no correlation since $\widehat{a_1}_n$ is equal to 3.9915 and 4.9336 on those days. The estimated parameters $(\widehat{a_1}_n, \widehat{\theta}_n, \widehat{\gamma}_n)$

[1]All calculations are done using MATLAB 7.6

along with bootstrapped standard errors[2] on days with strong persistence in the log-volatility are shown in Figure 2. Due to the assumptions in the exponential ACD model the jump rate λ should be one, which is confirmed by estimates $\widehat{\lambda}_n$ being close to one. Estimation results for μ along with the full set of estimated parameters can be found in Czado and Haug (2009).

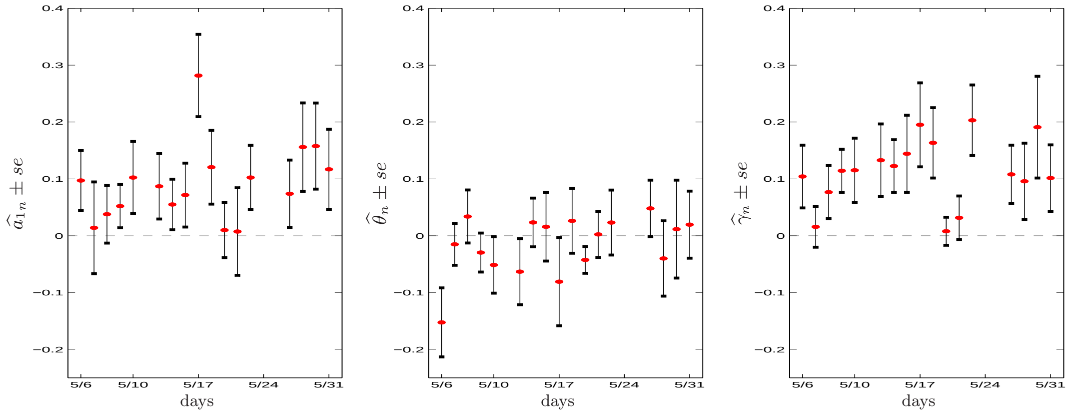

FIGURE 2. *Estimated parameters $(\widehat{a}_{1n}, \widehat{\theta}_n, \widehat{\gamma}_n)$ together with bootstrapped standard errors on those days with strong persistence in the log-volatility.*

Given the parameter estimate $\widehat{\boldsymbol{\vartheta}}_n^g$ we are able to estimate the volatility, which allows us to compute the fitted innovations $\widehat{Z}_i = G_{t_i}^{\Delta t_i}/\widehat{\sigma}_{t_i-}$. Due to our assumptions there should be no correlation in the squared fitted innovations. Therefore we performed a Ljung-Box test for the squared fitted innovations $\widehat{Z}_i^2$ on each day. The hypothesis of no correlation is rejected at the 0.05 significance level only on May 15th. Over the remaining days the average p value is equal to 0.76. The degrees of freedom df were chosen such that $df \approx \sqrt{n}$.
Mentionable is now that we obtained a suitable fit for most of the days although there is no direct dependence between the volatility and the observed durations in our model. The log-volatility process (2) depends only on the i.i.d. sequence of innovations $(\Delta t_i)_{i=1,\ldots,n}$ and not on the observed or conditional durations. Therefore the dependence structure in the durations does not influence the volatility process, which is in contrast to the results in Engle (2000), Ghysels and Jasiak (1998) and Meddahi, Renault, and Werker (2006).

5. Conclusions

In this paper we introduced the ACD-ECOGARCH(1, 1) model to analyse ultra-high-frequency data. The ACD model of Engle and Russell (1998) describes the dependence in the durations whereas the ECOGARCH models the stochastic volatility of the price process. Both models are linked through the innovation sequence of the ACD model. We have shown that the model can be estimated by a QML approach. The limiting properties of the proposed estimator are however not investigated. But this is to the best of our knowledge also an open problem in the discrete time EGARCH model except for some special cases, see e.g. Straumann (2005). The application to General Motors stock prices demonstrates the potential usefulness of this

[2]We computed 999 residual-based bootstraps. For details on residual-based bootstrap see e.g. Corradi and Iglesias (2008) and references therein.

class of models. They are able to identify and quantify a leverage effect present in the data. On a daily basis different behaviours are detected.
We like to note that this model assumes that all jumps are observable and that the driving Lévy process is of compound Poisson type. Although we successfully fitted the model to ultra-high-frequency data these assumptions certainly will not hold for all kinds of data. Relaxing these assumptions would also allow to consider other distributions than the exponential distribution for the innovations in the ACD model. Therefore the development of estimation and prediction methods for the general ECOGARCH model is necessary and subject of current research.

Acknowledgements

We are indebted to two anonymous referees for a number of very helpful comments. Furthermore, we would like to thank George Tauchen and an associate editor for further helpful suggestions. All of them helped to improve the paper considerably. C. Czado acknowledges the support by the DFG through grant CZ 86/1-3.

References

Aït-Sahalia, Y., P.A. Mykland, and L. Zhang (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies* 18: 351–416.

Bauwens, L., and P. Giot (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., P. Giot, J. Grammig, and D. Veredas (2004). A comparison of financial duration models via density forecasts. *International Journal of Forecasting* 20: 589–609.

Berkes, I., and L. Horváth (2003). The rate of consistency of the quasi-maximum likelihood estimator. *Statistic and Probability Letters* 61: 133–143.

Berkes, I., L. Horváth, and P.S. Kokoszka (2003). GARCH processes: Structure and estimation. *Bernoulli* 9: 201–227.

Bollerslev, T., J. Litvinova, and G. Tauchen (2006). Leverage and volatility feedback effects in high-frequency data. *Journal of Financial Econometrics* 4: 353–384.

Corradi, V., and E.M. Iglesias (2008). Bootstrap refinements for QML estimators of the GARCH(1,1) parameters. *Journal of Econometrics* 144: 500–510.

Czado, C., and S. Haug (2009). Finite sample properties of the QMLE in the ACD-ECOGARCH(1, 1) model. available from: http://www-m4.ma.tum.de.

Engle, R.F. (2000). The econometrics of ultra-high-frequency data. *Econometrica* 68: 1–22.

Engle, R.F., and J.R. Russell (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica* 66: 1127–1162.

Engle, R.F., and Z. Sun (2007). When is noise not noise: A microstructure estimate of realized volatility. available from: http://hdl.handle.net/2451/26358.

Ghysels, E., and J. Jasiak (1998). GARCH for irregularly spaced financial data: The ACD-GARCH model. *Studies in Nonlinear Dynamics and Econometrics* 2: 133–149.

Grammig, J., and M. Wellner (2002). Modeling the interdependence of volatility and inter-transaction duration processes. *Journal of Econometrics* 106: 369–400.

Hall, P., and Q. Yao (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* 71: 285–317.

Haug, S., and C. Czado (2007). An exponential continuous time GARCH process. *Journal of Applied Probability* 44: 960–976.

Jeantheau, T. (1998). Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* 14: 70–86.

Klüppelberg, C., A.M. Lindner, and R.A. Maller (2004). A continuous time GARCH process driven by a Lévy process: stationarity and second order behaviour. *Journal of Applied Probability* 41: 601–622.

Lindner, A.M. (2009). Continuous time approximations to GARCH and stochastic volatility models. In T.G. Andersen, R.A. Davis, J.-P. Kreiß and T. Mikosch (eds.), *Handbook of Financial Time Series*, pp. 481–496. Berlin: Springer.

Maller, R., G. Müller, and A. Szimayer (2008). GARCH modelling in continuous time for irregularly spaced time series data. *Bernoulli* 14: 519-542.

Meddahi, N., E. Renault, and B.J.M. Werker (2006). GARCH and irregularly spaced data. *Economic Letters* 90: 200–204.

Pacurar, M. (2008). Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. *Journal of Economic Surveys* 22: 711–751.

Straumann, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models.* Lecture Notes in Statistics 181. Berlin: Springer.

Tay, A., C. Ting, Y.K. Tse, and M. Warachka (2009). Using high-frequency transaction data to estimate the probability of informed trading. *Journal of Financial Econometrics* 7: 288–311.