Center for
Quantitative
Economics

Long Memory Conditional Heteroscedasticity in Count Data

Mawuli Segnon [†] and Manuel Stapper [†]

82/2019

[†] Department of Economics, University of Münster, Germany

wissen.leben
WWU Münster

Long Memory Conditional Heteroscedasticity in Count Data

Mawuli Segnon
Westfälische Wilhelms-Universität Münster,
Department of Economics (CQE), Germany,

Manuel Stapper
Westfälische Wilhelms-Universität Münster,
Department of Economics (CQE), Germany

Abstract: This paper introduces a new class of integer-valued long memory processes that are adaptations of the well-known FIGARCH(p, d, q) process of Baillie (1996) and HYGARCH(p, d, q) process of Davidson (2004) to a count data setting. We derive the statistical properties of the models and show that reasonable parameter estimates are easily obtained via conditional maximum likelihood estimation. An empirical application with financial transaction data illustrates the practical importance of the models.

Keywords: Count Data, Poisson Autoregression, Fractionally Integrated, INGARCH

1 Introduction

Modeling long-range dependence[1] in time series has a long tradition in many fields dating back to at least Smith (1938), Cox and Townsend (1947), Hurst (1951), among others. Although this phenomenon has been observed and reported since the sixties by Mandelbrot (1963) for speculative markets, its consideration in the modeling of financial or macro time series only started with papers by Ding et al. (1993) and Ding and Granger (1996) that documented in detail the presence of long memory in absolute or squared observations. In addition to the empirical observations, the development of fractionally integrated processes by Granger (1980), Granger (1981), Granger and Joyeux (1980) and Hosking (1981) has unleashed the introduction of competing models of long memory in the literature, see Haldrup and Valdes (2017) for an excellent review of the development in modeling long memory.

Recent empirical observations have revealed the existence of long memory in count data, see Quoreshi (2014), Braccini (2015), Hainaut and Boucher (2014). These observations confirm the theoretical results of Daley et al. (2000) and Hurvich et al. (2009) which imply that, under certain conditions, long memory behavior is preserved when duration processes are converted to count data processes. This in turn implies that

[1] Until now the origin of the long range dependence is subject to debate. For Cox (2014), the origin of long memory remains unclear. However, there exist different theories in the literature to explain the presence of long range dependence.

those count data processes are not only characterized by overdispersion, but also by high persistence of observed autocorrelations that take far longer to decay than the geometrical decay associated with the integer-valued ARMA (INARMA) or GARCH (INGARCH) models.

The presence of long memory in financial count data led Quoreshi (2014) to introduce the integer-valued ARFIMA process. Braccini (2015) and Hainaut and Boucher (2014) independently used the Poisson multifractional process[2] for modeling count data. We contribute to this growing literature by introducing integer-valued FIGARCH (INFIGARCH) and HYGARCH (INHYGARCH) processes. Our primary objective is to develop a more flexible class of processes for the conditional mean that are able to reproduce the long range dependence observed in count data. Introduced by Baillie (1996) and Davidson (2004) respectively, FIGARCH and HYGARCH processes are widely used in empirical applications to model volatility in financial markets. As measures of financial market realized volatility, it seems very important to find stochastic processes that can properly reproduce the stylized facts of count data and avoid the mispricing of volatility.

The rest of the paper is organized as follows. Section 2 introduces the models. Statistical properties of the processes and the inequality conditions that guarantee positivity of the conditional mean are provided in Section 3. We implement a conditional maximum likelihood estimation (CMLE) and present the results of a Monte Carlo simulation in Section 4. An empirical application to financial markets is performed in Section 5, and Section 6 concludes.

2 Model Framework

In the INGARCH framework developed independently by Heinen (2003) and Ferland et al. (2006) a count data time series $\{Y_t\}_{t\in\mathbb{Z}}$ with $Y_t \in \mathbb{N}_0$ is called INGARCH(p, q) process if $Y_t|\mathcal{F}_{t-1}$ follows a Poisson distribution with parameter

$$\lambda_t = \beta_0 + \beta(L)\lambda_t + \alpha(L)Y_t \ , \qquad (1)$$

where $\alpha_i, \beta_j \geq 0$ and $\beta_0 > 0$. $\mathcal{F}_t = \sigma\{Y_t, Y_{t-1}, ...\}$ is called information set at time t, $\beta(L) = \beta_1 L + .. + \beta_q L^q$ and $\alpha(L) = \alpha_1 L + ... + \alpha_p L^p$. If $\alpha(1) + \beta(1) < 1$, the process is weakly and strictly stationary. Instead of the Poisson distribution other distributions on $\mathbb{N}_0$ may be considered.
Subtracting $\beta(L)\lambda_t$ in (1) and rearranging under the assumption that all roots of $[1 - \beta(L)]$ are outside the unit circle gives

$$\lambda_t = \omega + \psi^{\text{INGA}}(L)Y_t = \omega + \sum_{i=1}^{\infty} \psi_i^{\text{INGA}} Y_{t-i} \ , \qquad (2)$$

[2]The Markov-switching multifractional process used by both authors is the discretized version of the Poisson multifractional process originally developed by Calvet and Fisher (2001) for modeling volatility.

where $\omega = \beta_0[1-\beta(1)]^{-1}$ and $\psi^{\mathrm{INGA}}(L) = [1-\beta(L)]^{-1}\alpha(L)$ with $\psi_i^{\mathrm{INGA}} \geq 0$ for all i. In analogy to their continuous counterparts, this form is called the INARCH(∞) representation of an INGARCH(p, q) process. Its marginal expectation is

$$\lambda = \mathrm{E}(Y_t) = \beta_0[1-\alpha(1)-\beta(1)]^{-1} = \omega[1-\psi^{\mathrm{INGA}}(1)]^{-1}$$

Alternatively, the INGARCH(p, q) process can be formalized as ARMA($\max\{p,q\}$, q) process

$$\underbrace{(1-\alpha(L)-\beta(L))}_{=:\Phi(L)}(Y_t-\lambda) = \underbrace{(1-\beta(L))}_{=:B(L)}u_t \tag{3}$$

where $\{u_t\} = \{Y_t - \lambda_t\}$ is a white noise sequence with variance λ. It holds that

$$\begin{aligned}
\mathrm{E}(u_t) &= \mathrm{E}\left[\mathrm{E}(Y_t-\lambda_t|\mathcal{F}_{t-1})\right] = 0 \\
\mathrm{Var}(u_t) &= \mathrm{E}[\mathrm{Var}(Y_t-\lambda_t|\mathcal{F}_{t-1})] + \mathrm{Var}[\mathrm{E}(Y_t-\lambda_t|\mathcal{F}_{t-1})] = \mathrm{E}(\lambda_t) = \lambda \\
\mathrm{Cov}(u_t,u_s) &= \mathrm{E}(u_t u_s) = \mathrm{E}\left[\mathrm{E}(u_t u_s|\mathcal{F}_{t-1})\right] = 0 \qquad \forall t,s \in \mathbb{Z}, s<t
\end{aligned}$$

Further one can easily see that

$$1-\frac{\Phi(L)}{B(L)} = 1-\frac{1-\alpha(L)-\beta(L)}{1-\beta(L)} = \frac{\alpha(L)}{1-\beta(L)} = \psi^{\mathrm{INGA}}(L)$$

allows the form

$$\lambda_t = \beta_0(1-\beta(1))^{-1} + \left[1-\frac{\Phi(L)}{B(L)}\right]Y_t\ . \tag{4}$$

Thus, the autocorrelation is only driven by the coefficients in $\psi^{\mathrm{INGA}}(L)$. The INARCH(∞) representation allows to investigate the effect of past observations on the conditional mean and variance as well as the memory of the process. A shock to the conditional mean of an INGARCH process is known to decay exponentially. Empirical observations, for example in high frequency financial count data, suggest the presence of high persistence, i.e. a slow hyperbolic decay of the ACF.

To take the slow decay into account, we replace the difference in the definition of coefficients by **?** for an integrated GARCH process

$$\psi^{\mathrm{IGARCH}}(L) = \left[1-\frac{\Phi(L)(1-L)}{B(L)}\right]\ . \tag{5}$$

by a fractional difference

$$(1-L)^d = \sum_{k=0}^{\infty}\frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)}L^k\ . \tag{6}$$

The long memory parameter d is thereby restricted to the interval $(0,1)$. The fractionally integrated INGARCH (INFIGARCH) process is then defined by the autoregression coefficients

$$\psi^{\text{INFI}}(L) = 1 - \frac{\Phi(L)(1-L)^d}{B(L)} \ . \tag{7}$$

Remark: *As proved by Conrad and Haag (2006) for $d \in (0,1)$, the slow hyperbolic decay of ψ_i^{INFI} coefficients is responsible for the persistence observed in impulse weights. We note that the* INFIGARCH(p, d, q) *process defined by coefficients in (7) can not be weakly stationary, since we have $\psi^{INFI}(1) = 1$. If $d = 0$, the* INFIGARCH(p, d, q) *process reduces to* INGARCH(p, q) *and for $d = 1$* to an integrated GARCH *process.*

Following Davidson (2004) we further propose a more general class of INGARCH-type models that includes the INFIGARCH(p, d, q) model as a limiting case. The basic idea is to add an amplitude parameter $\eta \in (0,1)$ in equation (7). The lag polynomial $\psi^{\text{INFI}}(L)$ then becomes

$$\begin{aligned} \psi^{\text{INHY}}(L) &= \left[1 - \frac{\Phi(L)\left(1 + \eta((1-L)^d - 1)\right)}{B(L)}\right] \\ &= \eta\psi^{\text{INFI}}(L) + (1-\eta)\psi^{\text{INGA}}(L) \end{aligned} \tag{8}$$

In the following we call this linear combination of INGARCH and INFIGARCH an INHYGARCH(p, d, q) model. Obviously, the process reduces to INFIGARCH when $\eta \to 1$ and to INGARCH when $\eta \to 0$.

3 Statistical Properties and Inequality Constraints

In this section we investigate the statistical properties of INFIGARCH and INHYGARCH processes. To ensure that the Poisson parameters λ_t are (almost surely) positive, the model coefficients have to satisfy certain inequality restrictions. We proceed to discuss these restrictions.

3.1 Inequality constraints

Recall that Poisson parameters for an INGARCH process are strictly positive $\alpha_i, \beta_i \geq 0$ and $\beta_0 > 0$ or equivalently $\psi_i \geq 0$ and $\omega > 0$. The weak stationarity of the process is proven by Ferland et al. (2006) and holds if and only if additionally $\alpha(1) + \beta(1) < 1$. Given that, the process is also strict stationary. If every coefficient ψ_i in the INARCH(∞) representation is strictly positive, $\omega > 0$ can be replaced by $\omega \geq 0$ to ensure almost sure positivity.

The positivity of INFIGARCH coefficients ψ_i^{INFI} is discussed in detail by Conrad and Haag (2006) or Bollerslev and Mikkelsen (1996). The latter give conditions for the special case $p = q = 1$. Sufficient conditions are $\beta_1 - d \leq \phi_1 \leq \frac{2-d}{3}$ and $d\left(\phi_1 - \frac{1-d}{2}\right) \leq \beta_1(d - \beta_1 + \phi_1)$, where $\phi_1 = \alpha_1 + \beta_1$. It is further assumed that $\beta_1 \in (-1,1)\backslash\{0\}$ and $\phi_1 \in (-1,1)\backslash\{\beta_1\}$.

These conditions do not cover every set of parameters yielding positive INARCH(∞) coefficients. Alternatively, Conrad and Haag (2006) prove the necessary and sufficient conditions for coefficients of an INFIGARCH(1, d, 1) process to be positive with $f_k = \frac{k-1-d}{k}$.

The coefficients ψ_i^{INFI} are non-negative if for $\beta_1 \in (0,1)$ either $\psi_1^{\text{INFI}} \geq 0$ and $\phi_1 \leq f_2$ or if for $k > 2$ with $f_{k-1} < \phi_1 \leq f_k$ it holds that $\psi_{k-1}^{\text{INFI}} \geq 0$.
If $\beta_1 \in (-1,0)$, the coefficients are non-negative if either $\psi_1^{\text{INFI}}, \psi_2^{\text{INFI}} \geq 0$ and
$\phi_1 \leq f_2(\beta_1 + f_3)/(\beta_1 + f_2)$ or for $k > 3$ with
$f_{k-2}(\beta_1 + f_{k-1})/(\beta_1 + f_{k-2}) < \phi_1 \leq f_{k-1}(\beta_1 + f_k)/(\beta_1 + f_{k-1})$ it holds that
$\psi_{k-1}^{\text{INFI}}, \psi_{k-2}^{\text{INFI}} \geq 0$.

The difference between the constraints is displayed in Figure 1. The shaded area in red represents the sufficient conditions of Bollerslev and Mikkelsen (1996), while the solid lines frame the constraints by Conrad and Haag (2006).

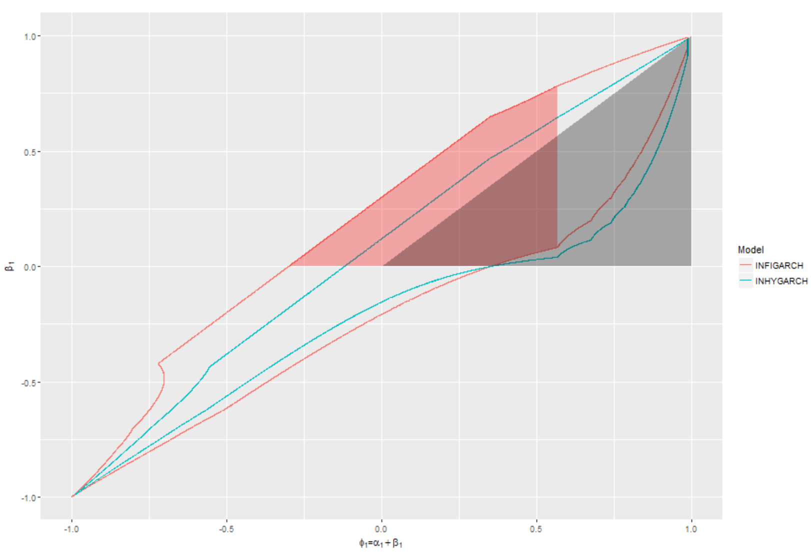

Figure 1: Admissible parameter space for $d = 0.3$ and $\eta = 0.4$.

Equation (8) shows the connection between INHYGARCH, INGARCH and INFIGARCH models. It follows directly that $\psi_i^{\text{INHY}} \geq 0$ if the positivity conditions of INGARCH and INFIGARCH are both fulfilled.

Conrad (2010) argues that this sufficient condition is too restrictive. Alternatively, necessary and sufficient conditions for the almost sure positivity if $p = q = 1$ are given.

If $\beta_1 \in (0,1)$, the coefficients are non-negative if either $\psi_1^{\text{INHY}} \geq 0$ and $\phi_1 \leq f_2$ or for $k > 2$ with $f_{k-1} < \phi_1 \leq f_k$ it holds that $\psi_{k-1}^{\text{INHY}} \geq 0$.

For $\beta_1 \in (-1, 0)$, the coefficients are non-negative if either $\psi_1^{\text{INHY}}, \psi_2^{\text{INHY}} \geq 0$ and $|\beta_1| \leq f_2$ or for $k > 3$ with $f_{k-2} < \phi_1 \leq f_{k-1}$ it holds that $\psi_1^{\text{INHY}}, ..., \psi_{k-1}^{\text{INHY}} \geq 0$.

These restrictions are also shown in Figure 1 together with the positivity constraints of its INFIGARCH and INGARCH components.

3.2 Stationarity

As mentioned before, INGARCH(p, q) processes are proven to be covariance stationary if and only if $\alpha(1) + \beta(1) < 1$. To discuss the stationarity of an INFIGARCH(p, d, q) process we use the INARCH(∞) representation.

From equation (7) it can be seen that $\psi^{\text{INFI}}(1) = 1$ regardless of the parameters, which implies that INFIGARCH processes with a positive intercept do not possess first moments and, hence, are not weakly stationary. As usual in the literature, we use a restricted version of the model where only a finite number of past observations is used in the calculation of the Poisson parameters λ_t. Let this number be R in the following, so that the restricted lag polynomial can be defined as $\psi_R^{\text{INFI}}(L) = \sum_{i=1}^{R} \psi_i^{\text{INFI}} L^i$. We set $R = 1000$ if not stated otherwise which is in line with the relevant literature on FIGARCH processes.

Since $\psi_R^{\text{INFI}}(1) < 1$, the restricted model including a positive intercept $\frac{\beta_0}{1-\beta(1)}$ is weakly stationary with finite marginal expectation. Stationarity implies summable covariances, which contradicts the long memory property. However, its highly persistent positive correlation can be reproduced by a restricted model.

To clarify the relationship between R and the estimation, we conduct a small simulation. For both models, INFIGARCH and INHYGARCH, 1000 time series are simulated following a restricted process with $R = 1000$. Then, the parameters are estimated considering different values of R, ranging from 100 to 2000. Mean point estimates are displayed in figure 2. One can see that only one of the four INFIGARCH parameters is considerably affected by the choice of R, namely the intercept parameter β_0. This effect is barely visible for the INHYGARCH processes.

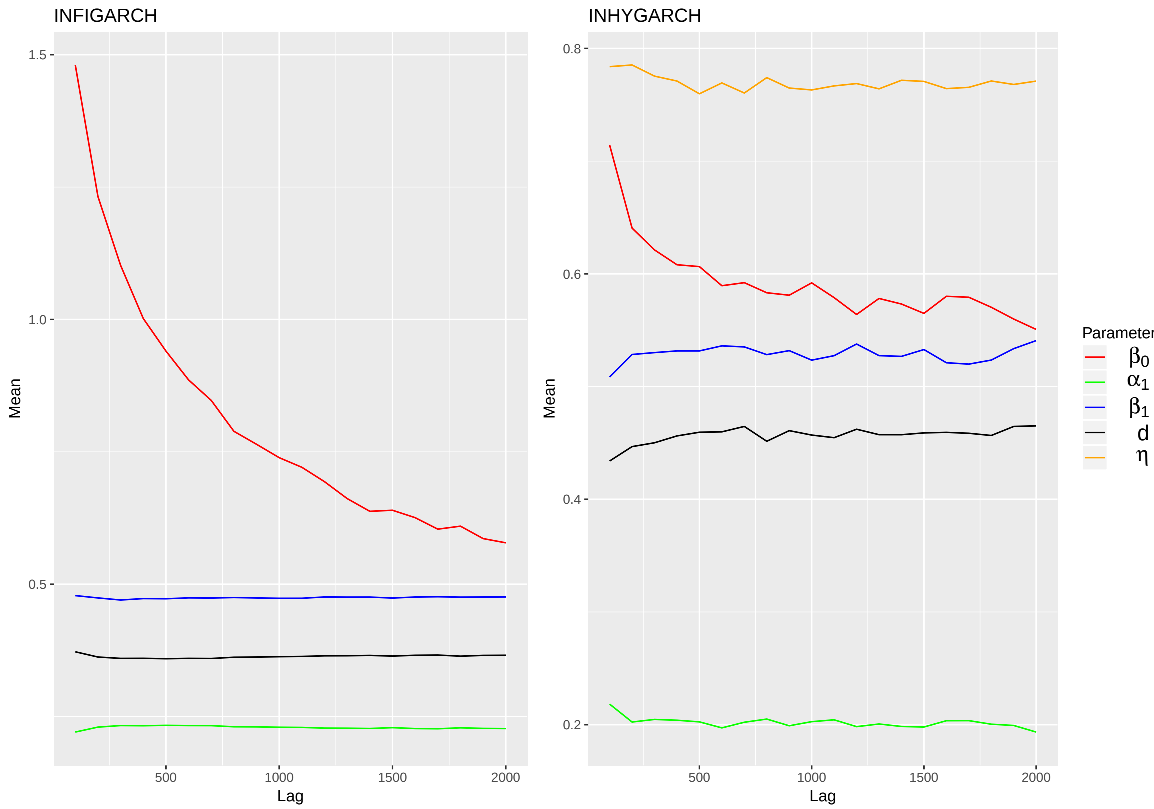

Figure 2: Impact of lag order R on the expectation of the estimators.

By increasing R in the simulation, the mean estimate of the intercept decreased. This simulation supports that the underlying does not have an intercept, and the restricted model can be used to estimate all parameters $(d, \alpha_1, .., \alpha_p, \beta_1, ..., \beta_q)$.

For an INHYGARCH process with $d, \eta \in (0, 1)$ it holds that $\psi^{\text{INHY}}(1) < 1$ and $\psi_R^{\text{INHY}}(1) < 1$, so that the restricted and unrestricted model is stationary according to Ferland et al. (2006).

3.3 Long Memory of INHYGARCH

The INHYGARCH process is covariance stationary for $d, \eta \in (0, 1)$. This contradicts the assumption that the ACF of such process also shows a long memory behavior, because the autocovariances would not be are absolute summable in that case. Still for η close to one, the slow rate of the covariance function comes close to the long memory structure.

Proposition 1: If $d, \eta \in (0, 1)$ and large h, the autocovariance function of the INHYGARCH(p, d, q) is given by

$$\text{Cov}(Y_t, Y_{t-h}) = Ch^{-d-1} \ ,$$

where C is an appropriately defined constant.

Proof: We start with the AR(∞) representation of an INHYGARCH process from (3), which can be rearranged as MA(∞), due to the invertibility of $(1-\psi^{\text{INHY}})$, so that

$$(Y_t - \lambda) = \underbrace{(1-\psi^{\text{INHY}}(L))^{-1}}_{=:\varphi(L)} u_t \ .$$

The covariance is then found by computing

$$\text{Cov}(Y_t - \lambda, Y_{t-h} - \lambda) = \text{E}(u_t^2)\chi(h) \ ,$$

where $\chi(h) = \sum_{i=0}^{\infty} \varphi_i \varphi_{i+|h|}$ and $\text{E}(u_t) = \omega\varphi(1)$.
Using the fundamental theorem for polynomials we deduce from

$$\left(1 + \sum_{i=1}^{\infty} \varphi_i L^i\right)\left(1 - \sum_{i=1}^{\infty} \psi_i^{\text{INHY}} L^i\right) = 1$$

that

$$0 \leq \varphi_j = \psi_j^{\text{INHY}} + \mathcal{O}\left[\left(\sum_{i=1}^{\infty} \psi_i^{\text{INHY}}\right)^{\rho_j}\right] \ , \text{ as } j \to \infty$$

for some constant $0 < \rho < \infty$.
When the ψ_i^{INHY} decay towards zero more slowly than exponentially, i.e. $\psi^{\text{INHY}}/\rho^i \overset{i\to\infty}{\longrightarrow} \infty$ for any $0 < \rho < 1$, the condition $\sum_{i=1}^{\infty} \varphi_i < \infty$ implies that as $h \to \infty$, $\sum_{i=0}^{\infty} \varphi_i \varphi_{i-|h|} \sim C\psi_h^{\text{INHY}}$ for some $0 < C < \infty$ with $f(h) \sim g(h)$ meaning that $f(h)/g(h) \to 1$. These results are similar to those of Theorem 2 in Zaffaroni (2004).

In the following we show that ψ_i^{INHY} decays hyperbolically as i approaches infinity.
By replacing the fractional difference $(1-L)^d$ by its standard binomial expansion

$$(1-L)^d = 1 + \sum_{i=1}^{\infty} \frac{\Gamma(i-d)}{\Gamma(1-d)\Gamma(i+1)} L^i \ ,$$

the HYGARCH lag polynomial in (8) becomes

$$\begin{aligned}
\psi^{\text{INHY}}(L) &= \frac{\Phi(L)}{B(L)}\Big(1 - \eta\big[(1-L)^d - 1\big]\Big) \\
&= \frac{\Phi(L)}{B(L)}\left(1 - \eta\left[\sum_{i=1}^{\infty} \frac{\Gamma(i-d)}{\Gamma(1-d)\Gamma(i+1)} L^i\right]\right) \\
&\sim \frac{\Phi(L)}{B(L)}\left(1 - G(d,\eta)\left[\sum_{i=1}^{\infty} i^{-1-d} L^i\right]\right) \ ,
\end{aligned}$$

where $G(d,\eta) = \eta\Gamma(1-d)^{-1}$ and using Stirling's approximation $\Gamma(i-d)\Gamma(i+1)^{-1} \sim i^{-d-1}$.
From this equation it becomes obvious that ϕ_i^{INHY} can be approximated by $\phi_i^{\text{INHY}} \approx$

ci^{-1-d} for large i, where c is an appropriately defined constant (see Granger and Joyeux (1980) for more details). It follows from Theorem 2 in Zaffaroni (2004) that $\chi(h)$ can be approximated by Ah^{-1-d} so that finally for large h

$$\mathrm{Cov}(Y_t, Y_{t-h}) = \mathrm{E}(u_t^2) A h^{-1-d} \qquad , h = 0, \pm 1, ...$$

■

4 Model Estimation

4.1 Estimation

For the estimation we suppose that $y = (y_{-R+1}, ..., y_0, y_1, ..., y_T)$ is a realization of the restricted INFIGARCH or INHYGARCH process, where R is the number of lags. The complete likelihood of $\theta = (\beta_0, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q, d) \in \Theta^{\mathrm{INFI}}$ or
$\theta = (\beta_0, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q, d, \eta) \in \Theta^{\mathrm{INHY}}$ given y is approximated by conditioning on the first R values, so that

$$L(\theta|y) := \prod_{t=1}^{T} \frac{\lambda_t^{y_t}}{y_t!} \exp(-\lambda_t) \ . \tag{9}$$

The conditional means λ_t depend on the model and the chosen lag order R as described in (7) and (8).
$\lambda_t = \beta_0[1 - \beta(1)]^{-1} - \sum_{i=1}^{R} \psi_i^{\mathrm{INFI}} y_{t-i}$ for an INFIGARCH process and correspondingly using ψ_i^{INHY} for an INHYGARCH process. The admissible parameter space Θ is determined by the model and the constraints discussed in section 3.1.

The conditional maximum likelihood estimates are found by maximizing the log likelihood $\ell(\theta|y)$ with respect to θ

$$\hat{\theta}_{\mathrm{CML}} = \underset{\theta \in \Theta}{\mathrm{argmax}}\ \ell(\theta|y) = \underset{\theta \in \Theta}{\mathrm{argmax}} \sum_{t=1}^{T} [y_t \ln(\lambda_t) - \lambda_t] \ . \tag{10}$$

The estimation procedure described above is implemented in R[3] using functions of the package `tscount`, see Liboschik et al. (2017) and Liboschik et al. (2017). In our implementation, we use a two stage estimation procedure. First, a method of moments estimation yields initial estimates by comparing the mean and sample ACF with theoretical expectation and ACF given a set of parameters. Although these estimates perform rather poorly, we use them as initial values in the likelihood maximization to gain stability in the estimation, and to reduce the computation time in the time consuming maximization.

Not only because of the long time series and long dependence, but also because of a possible application to high frequency financial data, a fast computation is desirable.

[3]see R Core Team (2018) version 3.4.4.

To keep down the computation time as much as possible, every calculation step is either vectorized or — when dealing with loops — written in `C++`.

Under mild regularity assumptions, the CML estimator converges to $\mathcal{N}(\theta_0, \Sigma)$ in distribution, where θ_0 is the true parameter vector and

$$\Sigma = -\left[\mathrm{E}_0\left(\frac{\partial^2 \ell(\theta|y)}{\partial\theta\partial\theta'}\right)\right]^{-1} \tag{11}$$

Σ can simply be estimated by a moment estimator or a sandwich estimator, the latter being more robust against model misspecification. To test for a parameter vector $H_0 : \theta_0 = \theta^*$, the test statistic

$$T(\hat{\theta}_{\mathrm{CML}} - \theta^*)'\widehat{\Sigma}^{-1}(\hat{\theta}_{\mathrm{CML}} - \theta^*) \ .$$

is compared with the corresponding quantile of a χ^2 distribution with degrees of freedom equal to the number of parameters. Alternatively, restrictions can be tested using a likelihood ratio test. This makes sense if the whole parameter vector is tested, so that a restricted maximization of the likelihood is not needed.

4.2 Monte Carlo Simulation

To investigate the performance of the CML estimation for INFIGARCH and INHYGARCH models, we perform a simulation study. The most interesting parameter is d, since it controls the long memory behavior. Parameters less interesting for interpretation are chosen to be the same for all models considered, namely $\alpha_1 = 0.2$, $\beta_1 = 0.5$ and $\eta = 0.85$. Since both models aim to be used for high frequency financial data, there is no problem of collecting enough data. Therefore, we consider sample sizes $T \in \{10000, 20000, 50000\}$, simulated with a burn-in of length 10000.
The intercept is chosen in such way that the marginal mean is between 15 and 20 for better comparability among the settings. For every situation, 1000 time series are simulated and the parameters estimated. In addition, the estimates are tested for their true values. We chose the likelihood ratio test for this purpose, since it only uses the likelihood function and is computed fast. A comprehensive summary of the simulation is displayed in Table 2 in the appendix.
As expected, the estimates for a larger number of observations, have less bias and variation. In contrast to INFIGARCH processes, the LR tests seems to be conservative for INHYGARCH processes. For INFIGARCH, the proportion of rejected null hypotheses is closer to the nominal level of 5% and the test seems to be slightly liberal.
Generally, it seems that the additional parameter of an INHYGARCH process causes estimates and asymptotic tests to be less accurate compared to INFIGARCH processes.

5 Empirical Study

To illustrate the applicability of our models, we use high-frequency financial data. It has often been observed that high-frequency intraday return data of a stock displays a long memory.

5.1 Dataset

High-frequency data is suitable for fitting an INFIGARCH or INHYGARCH model because of the presumed long memory. Further, there is almost unlimited supply of tick data.

We collect trading data for Commerzbank AG via Bloomberg[4] and investigate the number of price changes in a given (short) time interval. A larger number of price changes indicates higher volatility and thus a riskier investment. It is plausible that times with a larger number of price changes tend to be followed by times with many price changes. The INFIGARCH model helps to assess the impact of more turbulent times on the future behavior of a stock. Over the period from March, 26 to October, 26 we create 1-minute intervals starting 5 minutes after the opening auction at 09:00 until 5 minutes before closing auction at 17:30 of a trading day. Further, the five intervals between 13:00 and 13:05 are discarded due to an intraday auction, where the number of trades with a price change is assumed not to follow the usual mechanisms. This results in 495 observations per day and 74250 in total.

[4]Bloomberg (2019)

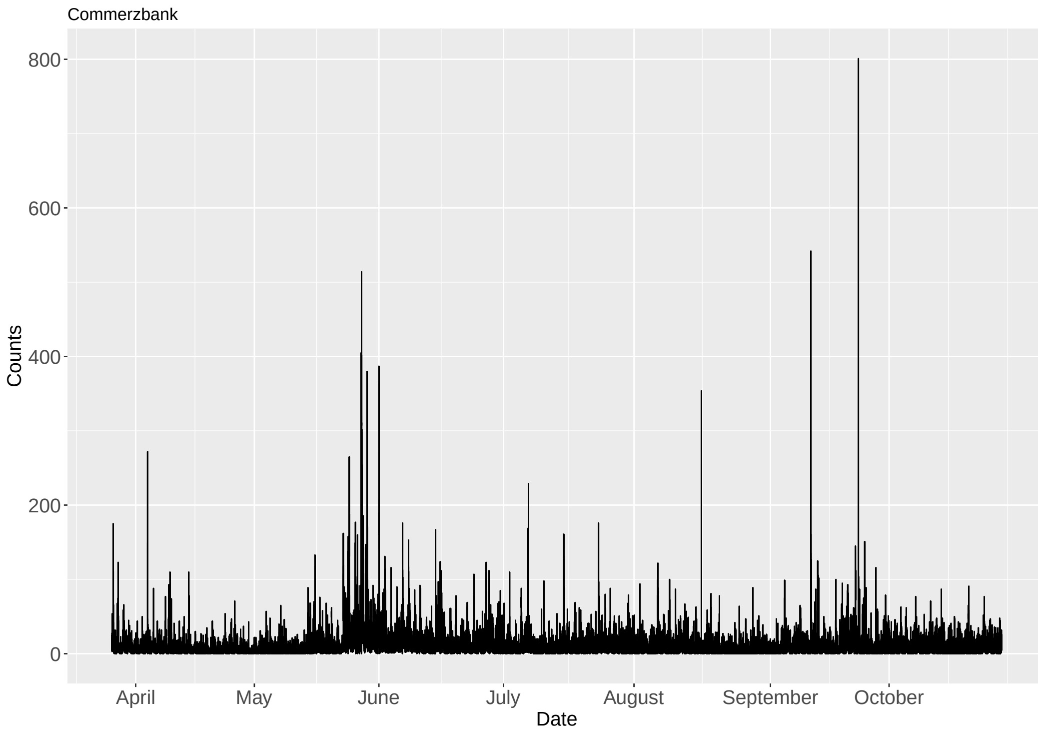

Figure 3: Time series of counts - Number of price changes for Commerzbank stock

The time series of counts is displayed in Figure 3. It catches the eye that there are outliers in the data that need to be handled. There is no systematic structure at which time of the day outliers occur. There is also no manifest link between outliers and news that may have affected the Commerzbank stock. Instead of adapting a robust estimation method, following Brownlees and Gallo (2006) we consider the extreme values to be caused by technical errors and follow the algorithm proposed in section 3.1. Each observation is compared with its neighborhood and discarded if it lies more that 3 (robustly estimated) standard deviations away from the median within the neighborhood. The standard deviation is estimated by the OGK estimator described by Maronna and Zamar (2002). We define the neighborhood as the 100 observations before and after an observation resulting in a window width of 201. By doing so, around 4.3% of the data is discarded. To affect the estimation as little as possible, we do not merge the remaining time series to preserve the dependence structure.
To check for long memory behavior, the ACF is displayed in figure 4. Red lines mark the peaks after one trading day. We find a positive correlation over a long range. Besides, there is a very clear seasonality. The number of price changes tends to be larger in the beginning and at the end of a trading day.

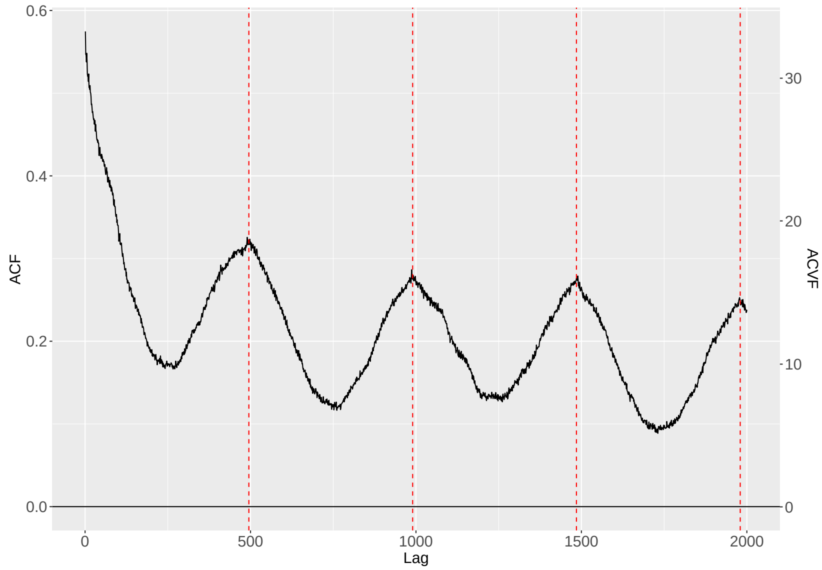

Figure 4: Autocorrelation function - Number of price changes in 1 minute intervals for Commerzbank stock

5.2 Seasonality

The marked seasonality needs to be taken into account for the estimation. We use the entire dataset and compute the sample means for each 1-minute interval of the trading day throughout the 150 trading days. In the following we denote the corresponding interval mean of observation y_t by μ_t. The sequence $\mu_1, ..., \mu_{495}$ displayed in Figure 5 gives an impression of the seasonality.

If one divides the time series $\{y_t\}$ by $\{\mu_t\}$, the resulting time series fluctuates around one and is free of seasonality. The ACF of $\{y_t/\mu_t\}$ is shown in Figure 6. While seasonality is removed from $\{y_t\}$, long memory is preserved. In our framework, we deal with count data, so that a transformation of the time series before estimation is not an option. Alternatively we incorporate the transformation in the estimation.
We do not only include the transformation in the definition of conditional means but also the property that $\mathrm{E}(Y_t/\mu_t|\mu_t) = 1$ so that an intercept parameter is not needed anymore. The model we estimate is then

$$\begin{aligned} Y_t|\mathcal{F}_{t-1} &\sim \mathrm{Pois}(\lambda_t) \\ \lambda_t &= \left[1 + \sum_{i=1}^{1000} \psi_i \left(\frac{Y_{t-i}}{\mu_{t-i}} - 1\right)\right] \mu_t \ . \end{aligned} \tag{12}$$

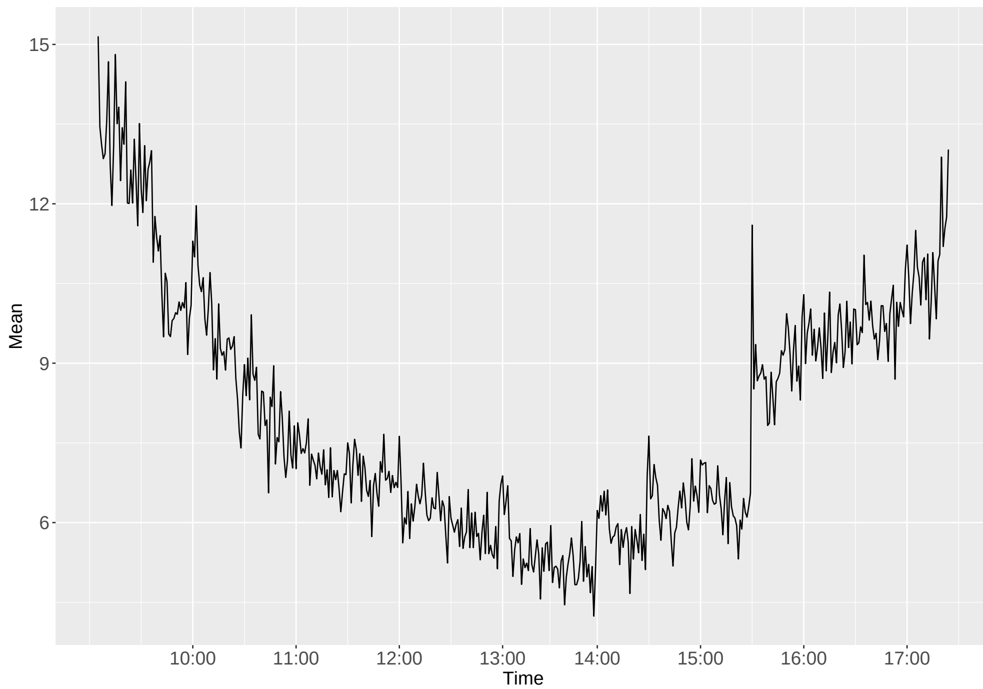

Figure 5: Seasonality structure

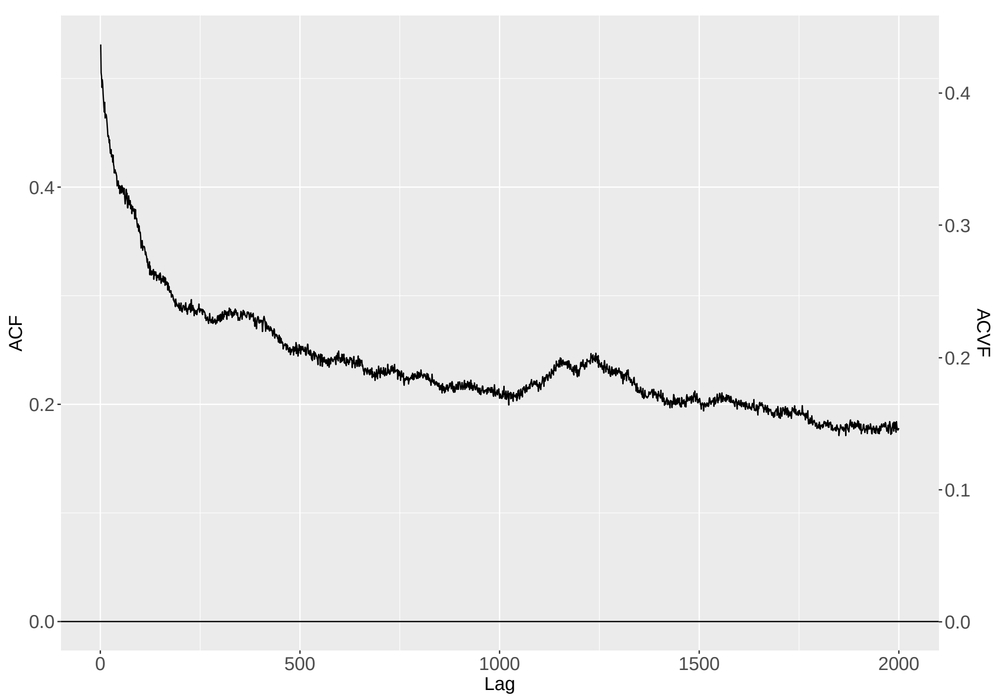

Figure 6: ACF of $\{y_t/\mu_t\}$

In case Y_t is flagged as outlier in the algorithm, the term $\frac{Y_t}{\mu_t}$ is replaced by 1 in the estimation so that the summand simply disappears.

5.3 Results

In definition (12), the coefficients ψ_i can be either from an INHYGARCH or INFIGARCH model. We first fit an INHYGARCH(1, d, 1) model, which yields the estimates

$$\left(\hat{\alpha}_1 \quad \hat{\beta}_1 \quad \hat{d} \quad \hat{\eta} \right)' = \left(-0.4505 \quad 0.7172 \quad 0.5905 \quad 1.0000 \right)' .$$

This parameter vector is on the boundary of the parameter space indicating that an INFIGARCH is a better model.
The results of fitting an INFIGARCH(1, d, 1) model are summarized in Table 1 .

Parameter	α_1	β_1	d
Estimate	-0.4403	0.7083	0.5809
Standard Error	$7.64 \cdot 10^{-3}$	$6.27 \cdot 10^{-3}$	$7.45 \cdot 10^{-3}$
Confidence interval	$[-0.455; -0.425]$	$[0.696; 0.720]$	$[0.566; 0.596]$

Table 1: Estimation results INFIGARCH(1, d, 1)

The long memory parameter d is significantly different from zero, which is not surprising given the ACF in Figure 6. The INFIGARCH model seemed to be an better choice compared to an INHYGARCH model. To check if a conditional Poisson distribution is appropriate, we use a non-randomized PIT histogram, described by Czano et al. (2009) and implemented in the `tscount` package, see Liboschik et al. (2017).
The PIT histogram in Figure 7 exhibits a U-shape which is a clear indicator that the equi-dispersed Poisson distribution does not capture the overdispersion present in the data. As a way out, we use the Negative Binomial distribution, which allows to incorporate an overdispersion through a parameter ϕ. It is estimated in accordance to Christou and Fokianos (2014) by solving

$$\sum_{t=R+1}^{T} \frac{(y_t - \lambda_t)^2}{\lambda_t + \lambda_t^2/\phi} = T - R \ .$$

We get $\hat{\phi} = 3.9140$, giving the PIT histogram displayed in Figure 8.
It is visible that the assumption of a Negative Binomial distribution is justified and the model suits the data well. As visible in Figure 11 in the appendix, the long memory property we had a focus on is also captured by the fitted model. The shape of the ACF from our scaled time series is similar to the theoretical ACF derived from parameter estimates.

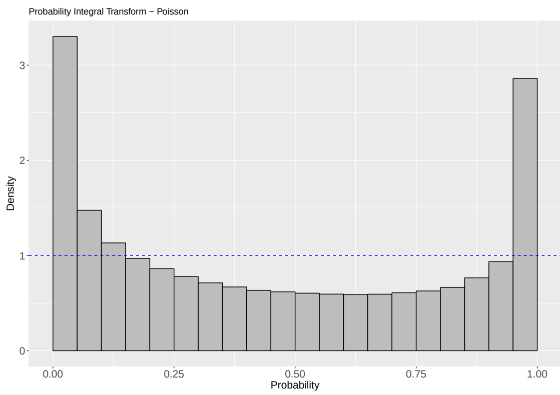

Figure 7: Non-randomized PIT histogram for a conditional Poisson distribution

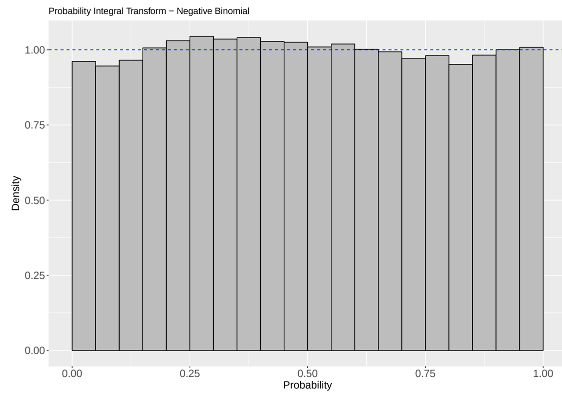

Figure 8: Non-randomized PIT histogram for a conditional Negative Binomial distribution

5.4 Forecasting

As a last step of the application, the out of sample forecasting properties of above model are investigated in the following. The steps that were taken before for the whole time series are therefore applied to a rolling window of 50 trading days. This way, 24750 observations minus detected outliers are used for each estimation of parameters. Afterwards, one subsequent trading day is foretasted by the recursion

$$\hat{Y}_{t+h|t} = \begin{cases} \left[1 + \sum_{i=1}^{1000} \psi_i \left(\frac{Y_{t+1-i}}{\mu_{t+1-i}} - 1\right)\right] \mu_{t+1} & \text{if } h = 1 \\ \left[1 + \sum_{i=1}^{h-1} \psi_i \left(\frac{\hat{Y}_{t+h-i|t}}{\mu_{t+h-i}} - 1\right)\right] \mu_{t+h} + \sum_{i=h}^{1000} \psi_i \left(\frac{Y_{t+h-i}}{\mu_{t+h-i}} - 1\right) & \text{if } h > 1 \end{cases},$$

where ψ_i are the coefficients resulting from parameter estimates, μ_t is the series of interval means calculated from the sample window and $h \in \{1, ..., 495\}$.
This procedure is computationally time intensive but has two objectives. Not only is the goodness of forecasting assessed, but also it is checked whether or not the parameters change during the sample period. A structural break can lead to a false impression of the long memory behavior as for example described by Sibbertsen (2004).

Figure 9 displays the goodness of forecast of our model. It shows the empirical distribution functions of absolute forecast errors $|Y_t - \hat{Y}_{t|t-h}|$ for $h = 1$ and $h = 495$ together with the distribution function of the benchmark. It can be seen that the 495-step ahead forecast dominates the benchmark in the sense of the absolute area, but the one-step ahead forecast has more larger errors compared with the benchmark.

As a byproduct of the forecasting, Figure 10 displays the parameter estimates of the rolling window together with the 95% pointwise confidence intervals. There is of course variation in estimates along time, but the range they fluctuate in is rather narrow looking at the confidence bands. For ϕ, one can see that the change in estimates is more noticeable compared with the other parameters. However, the densities of a Negative Binomial distribution with $\phi = 3$ and $\phi = 5$ do not differ very much. All this raises confidence that the slowly decaying ACF is not caused by a structural break and that the processes characteristics do not change much along time. Nevertheless, the assumption of a time constant overdispersion should be viewed critically.

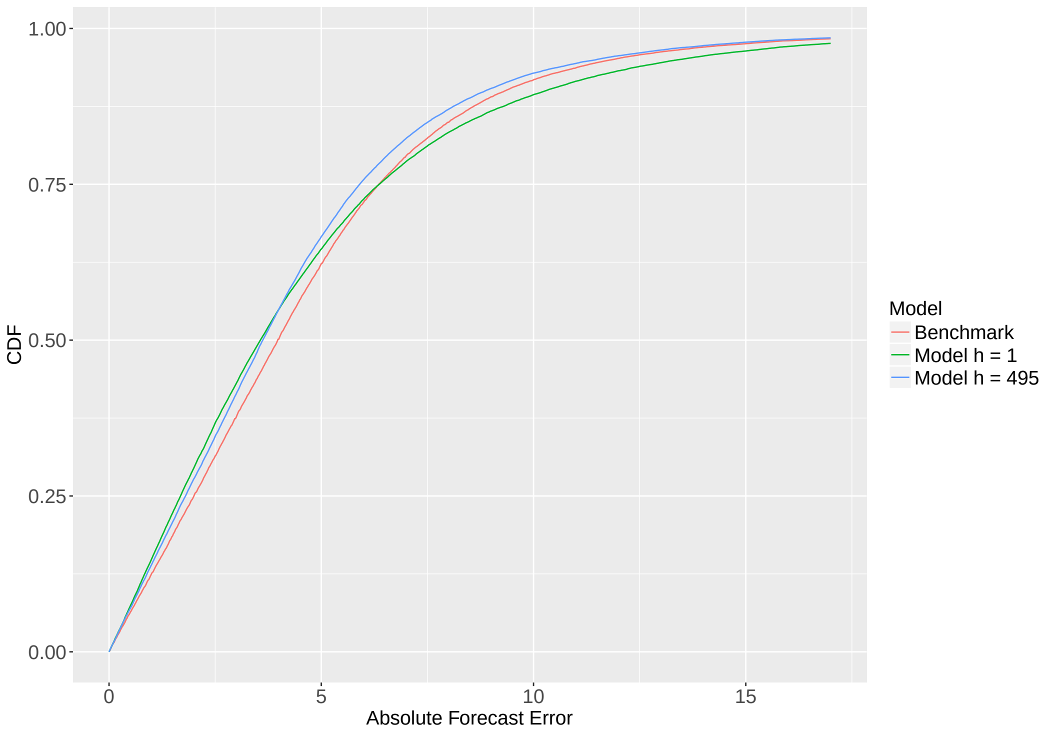

Figure 9: Forecast errors for $h = 1$ and $h = 495$ compared with benchmark

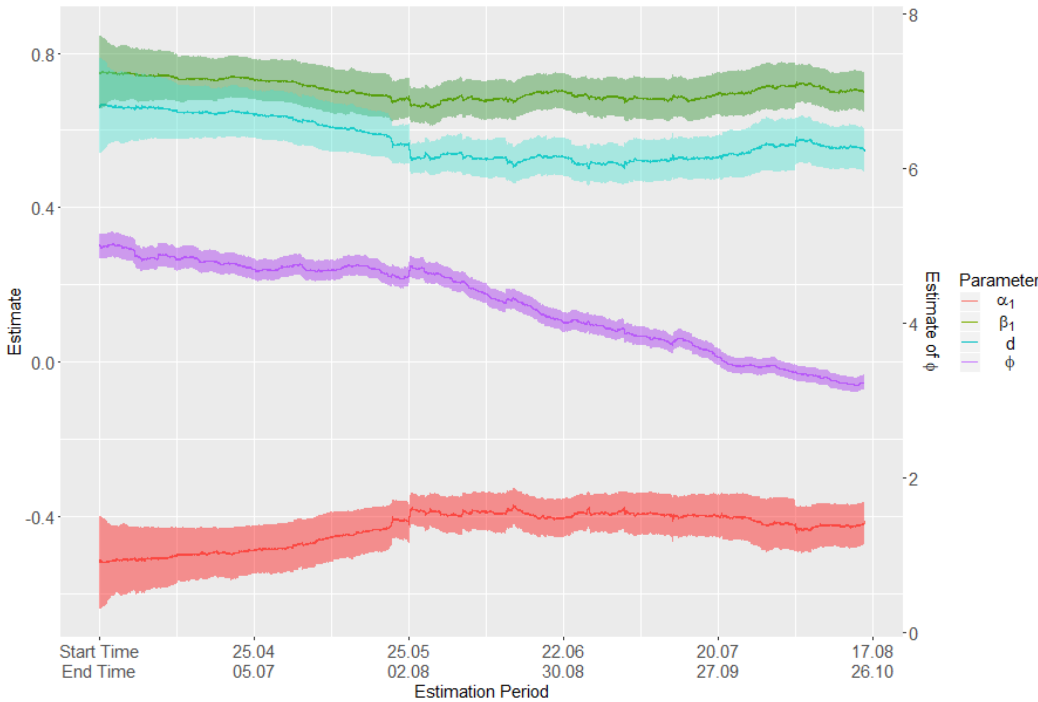

Figure 10: Parameter estimates for rolling window including confidence bands

6 Conclusion

In this article, we presented two models that incorporate a slow decay of the autocorrelation in the INGARCH framework as an alternative to INARFIMA models. Statistical properties were derived and discussed. Via simulation it was shown that parameter estimates are easily obtained by conditional maximum likelihood estimation and restrictions can be tested with different tests. An empirical application underlined the relevance of those models and a way to deal with seasonality in count data models was presented. The forecast properties were compared with a naive approach, revealing advantages and disadvantages of the model as well as potential future research directions.

By dropping the assumption of a time constant overdispersion would be more suitable in the application of this paper. To our knowledge there are no models including a time varying overdispersion so far. Such a model might also be a way to explain the occurrence of extreme values by a very large overdispersion so that an outlier detection is not required. Alternatively, the estimation method can be discussed critically. A robust estimation procedure may also replace an outlier detection step. Further, an estimation technique based on the spectral density seems promising.

Appendix

Model	d	$T+R$	β_0	$\hat{\beta}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{d}$	$\hat{\eta}$	Rejected H_0
INFIGARCH	0.2	10000	1.00	1.0718	0.2080	0.4954	0.1915	-	0.055
				(0.2208)	(0.0314)	(0.0333)	(0.0308)		
	0.2	20000	1.00	1.0331	0.2034	0.4979	0.1962	-	0.057
				(0.1521)	(0.0226)	(0.0230)	(0.0227)		
	0.2	50000	1.00	1.0051	0.2003	0.5000	0.1996	-	0.057
				(0.0940)	(0.0142)	(0.0148)	(0.0141)		
	0.4	10000	0.25	0.2660	0.2048	0.4958	0.3945	-	0.050
				(0.0568)	(0.0310)	(0.0311)	(0.0313)		
	0.4	20000	0.25	0.2556	0.2018	0.4997	0.3975	-	0.057
				(0.0367)	(0.0213)	(0.0224)	(0.0213)		
	0.4	50000	0.25	0.2515	0.2003	0.5001	0.3994	-	0.049
				(0.0224)	(0.0131)	(0.0138)	(0.0131)		
	0.6	10000	0.04	0.0459	0.2009	0.4974	0.5998	-	0.060
				(0.0143)	(0.0296)	(0.0316)	(0.0309)		
	0.6	20000	0.04	0.0427	0.2002	0.4995	0.6001	-	0.059
				(0.0080)	(0.0221)	(0.0223)	(0.0227)		
	0.6	50000	0.04	0.0409	0.2000	0.5000	0.6001	-	0.053
				(0.0041)	(0.0133)	(0.0142)	(0.0135)		
INHYGARCH	0.2	10000	1.75	2.4720	0.1343	0.4335	0.3574	0.7464	0.029
				(0.9992)	(0.1629)	(0.1741)	(0.2562)	(0.2042)	
	0.2	20000	1.75	2.3022	0.1491	0.4453	0.3127	0.7750	0.025
				(0.9132)	(0.1336)	(0.1578)	(0.2129)	(0.1664)	
	0.2	50000	1.75	2.0804	0.1642	0.4730	0.2730	0.7907	0.025
				(0.6984)	(0.1037)	(0.1158)	(0.1551)	(0.1216)	
	0.4	10000	1.00	1.0440	0.2015	0.5143	0.4174	0.8266	0.037
				(0.2107)	(0.0660)	(0.0795)	(0.1030)	(0.0947)	
	0.4	20000	1.00	1.0209	0.2021	0.5064	0.4041	0.8431	0.035
				(0.1525)	(0.0516)	(0.0585)	(0.0738)	(0.0598)	
	0.4	50000	1.00	1.0123	0.2025	0.5037	0.4012	0.8451	0.034
				(0.1053)	(0.0388)	(0.0480)	(0.0572)	(0.0440)	
	0.6	10000	0.80	0.8066	0.2251	0.4854	0.5732	0.8442	0.034
				(0.1137)	(0.0860)	(0.838)	(0.1006)	(0.0330)	
	0.6	20000	0.80	0.8071	0.2203	0.4867	0.5783	0.8463	0.042
				(0.0753)	(0.0667)	(0.0629)	(0.0774)	(0.0220)	
	0.6	50000	0.80	0.8086	0.2187	0.4849	0.5790	0.8477	0.049
				(0.0503)	(0.0494)	(0.0437)	(0.0565)	(0.0151)	

Table 2: Simulation Results - Mean point estimates and standard deviations (in parentheses)

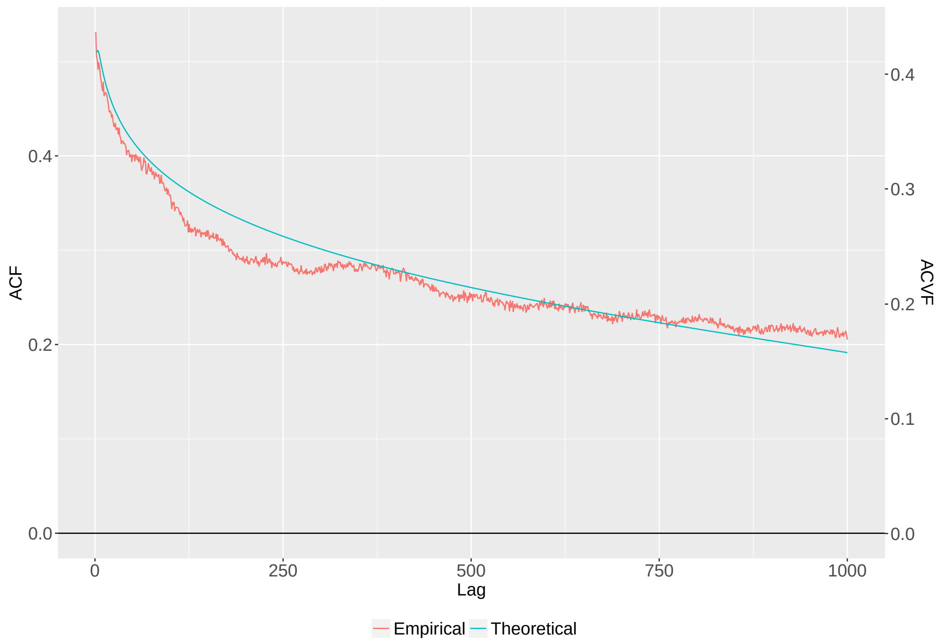

Figure 11: ACF of scaled sample y_t/μ_t and theoretical ACF from parameter estimates in red

References

[Baillie 1996] Baillie, R. T.: Long Memory Processes and Fractional Integration in Economics. In: *Journal of Econometrics* 73 (1996), Juli, Nr. 1, 5 - 58. `http://dx.doi.org/10.1016/j.jeconom.2008.09.034`.. – DOI 10.1016/j.jeconom.2008.09.034.

[Bloomberg 2019] Bloomberg L.P., Retrieved from Bloomberg database via R API

[Bollerslev and Mikkelsen 1996] Bollerslev, T. ; Mikkelsen, H.O.: Modeling and Pricing Long Memory in Stock Market Volatility. In: *Journal of Econometrics* 73 (1996), Nr. 1, 151 - 184. `http://dx.doi.org/10.1016/0304-4076(95)01736-4`. – DOI 10.1016/0304–4076(95)01736–4. – ISSN 0304–4076

[Braccini 2015] Braccini, L.: *Essays in Dynamic Duration and Count Modelling*, University of Pennsylvania, Diss., 2015

[Brownlees and Gallo 2006] Brownlees, C.T. ; Gallo, G.M.: Financial econometric analysis of ultra-high frequency: Data handling concerns. In: *Computational Statistics & Data Analysis* 51 (2006), 2232 - 2245. `http://dx.doi.org/10.1016/j.csda.2006.09.030`. – DOI 10.1016/j.csda.2006.09.030

[Calvet and Fisher 2001] Calvet, L. ; Fisher, A.: Forecasting Multifractal Volatility. In: *Journal of Econometrics* 105 (2001), Nr. 1, 27 - 58. http://dx.doi.org/10.1016/S0304-4076(01)00069-0. – DOI 10.1016/S0304–4076(01)00069–0

[Christou and Fokianos 2014] Christou, V. ; Fokianos, K.: Quasi-Likelihood Inference for Negative Binomial Time Series Models. In: *Journal of Time Series Analysis* 35 (2014), Nr. 1, 55-78. http://dx.doi.org/10.1111/jtsa.12050. – DOI 10.1111/jtsa.12050

[Conrad 2010] Conrad, C.: Non-negativity conditions for the hyperbolic GARCH model. In: *Journal of Econometrics* 157 (2010), Nr. 2, 441-457. https://EconPapers.repec.org/RePEc:eee:econom:v:157:y:2010:i:2:p:441-457

[Conrad and Haag 2006] Conrad, C. ; Haag, B.R.: Inequality Constraints in the Fractionally Integrated GARCH Model. In: *Journal of Financial Econometrics* 4 (2006), Juli, Nr. 3, 413 - 449. http://dx.doi.org/10.1093/jjfinec/nbj015. – DOI 10.1093/jjfinec/nbj015

[Cox 2014] Cox, D.: *Scaling*. https://vimeo.com/102835403. Version: 2014. – Symposium Presentation and Video, Nufflield College Oxford

[Cox and Townsend 1947] Cox, D.R. ; Townsend, M.W.H.: The Use of the Correlogram in Measuring Yarm Irregularity. In: *Proc. Int. Wool Textile Organization, Technical Committee 2* (1947), S. 28–34

[Czano et al. 2009] Czano, C. ; Gneiting, T. ; Held, L.: Predictive Model Assessment for Count Data. In: *Biometrics* 65 (2009), 1254-1261. http://dx.doi.org/10.1111/j.1541-0420.2009.01191.x.. – DOI 10.1111/j.1541–0420.2009.01191.x.

[Daley et al. 2000] Daley, D.J. ; Rolski, T. ; Vesilo, R.: Long-Range Dependent Point Processes and their Palm-Khinchin Distributions. In: *Advances in Applied Probability* 32 (2000), Dezember, Nr. 4, 1051 - 1063. http://dx.doi.org/10.1239/aap/1013540347. – DOI 10.1239/aap/1013540347. – ISSN 0001–8678

[Davidson 2004] Davidson, J.: Moment and Memory Properties of Linear Conditional Heteroscedasticity Models, and a New Model. In: *Journal of Business & Economic Statistics* 22 (2004), Nr. 1, 16-29. http://dx.doi.org/10.1198/073500103288619359. – DOI 10.1198/073500103288619359

[Ding and Granger 1996] Ding, Z. ; Granger, C.W.J.: Modeling Volatility Persistence of Speculative Returns: A New Approach. In: *Journal of Econometrics* 73 (1996), 185-215. http://dx.doi.org/10.1016/0304-4076(95)01737-2. – DOI 10.1016/0304–4076(95)01737–2

[Ding et al. 1993] Ding, Z. ; Granger, C.W.J. ; Engle, R.F.: A Long Memory Property of Stock market Returns and a new Model. In: *Journal of Empirical Finance* 1 (1993), S. 83 – 106

[Ferland et al. 2006] Ferland, R. ; Latour, A. ; Oraichi, D.: Integer-Valued GARCH Process. In: *Journal of Time Series Analysis* 27 (2006), nov, Nr. 6, S. 923–942. http://dx.doi.org/10.1111/j.1467-9892.2006.00496.x. – DOI 10.1111/j.1467–9892.2006.00496.x

[Granger 1980] Granger, C.W.J.: Long Memory Relationships and the Aggregation of Dynamix Models. In: *Journal of Econometrics* 14 (1980), 227 - 238. http://dx.doi.org/10.1016/0304-4076(80)90092-5. – DOI 10.1016/0304–4076(80)90092–5

[Granger 1981] Granger, C.W.J.: Some properties of time series data and their use in econometric model specification. In: *Journal of Econometrics* 16 (1981), Nr. 1, 121 - 130. http://dx.doi.org/10.1016/0304-4076(81)90079-8. – DOI 10.1016/0304–4076(81)90079–8

[Granger and Joyeux 1980] Granger, C.W.J. ; Joyeux, R.: An Introduction to Long-Memory Time Series Models and Fractional Differencing. In: *Journal of Time Series Analysis* 1 (1980), Januar, Nr. 1, 15 - 29. http://dx.doi.org/10.1111/j.1467-9892.1980.tb00297.x. – DOI 10.1111/j.1467–9892.1980.tb00297.x

[Hainaut and Boucher 2014] Hainaut, D. ; Boucher, J.-P.: Frequency and Severity Modelling Using Multifractal Processes: An Application to Tornado Occurrence in the USA and CAT Bonds. In: *Environmental Modeling & Assessment* 19 (2014), Jun, Nr. 3, 207–220. http://dx.doi.org/10.1007/s10666-013-9388-9. – DOI 10.1007/s10666–013–9388–9. – ISSN 1573–2967

[Haldrup and Valdes 2017] Haldrup, N. ; Valdes, J. E. V.: Long memory, fractional integration, and cross-sectional aggregation. In: *Journal of Econometrics* 199 (2017), Nr. 1, 1 - 11. http://dx.doi.org/https://doi.org/10.1016/j.jeconom.2017.03.001. – DOI https://doi.org/10.1016/j.jeconom.2017.03.001. – ISSN 0304–4076

[Heinen 2003] Heinen, A.: Modelling Time Series Count Data: An Autoregressive Conditional Poisson Model / University Library of Munich, Germany. Version: 2003. https://EconPapers.repec.org/RePEc:pra:mprapa:8113. 2003. – MPRA Paper

[Hosking 1981] Hosking, J.R.M.: Fractional Differencing. In: *Biometrics* 68 (1981), April, Nr. 1, 165 - 176. http://dx.doi.org/10.1093/biomet/68.1.165. – DOI 10.1093/biomet/68.1.165

[Hurst 1951] Hurst, H.E.: *Long-term Storage Capacity of Reservoirs*. American Society of Civil Engineers, 1951 https://books.google.de/books?id=aKj4HAAACAAJ

[Hurvich et al. 2009] Hurvich, C. M. ; Deo, R. ; Soulier, P. ; Wang, Y.: Conditions for the Propagation of Memory Parameter from Durations to Counts and Realized Volatility. 25 (2009), 06, S. 764–792

[Liboschik et al. 2017] Liboschik, T. ; Fried, R. ; Fokianos, K. ; Probst, P.: *tscount: Analysis of Count Time Series*, 2017. https://CRAN.R-project.org/package=tscount. – R package version 1.4.1

[Mandelbrot 1963] Mandelbrot, B.: The Variation of Certain Speculative Prices. In: *The Journal of Business* 36 (1963), Nr. 4, 394 - 419. http://www.jstor.org/stable/2350970

[Maronna and Zamar 2002] Maronna, R. ; Zamar, R.: Robust Estimates of Location and Dispersion for High-Dimensional Datasets. In: *Technometrics* 44 (2002), Nr. 4, 307-317. http://dx.doi.org/10.1198/004017002188618509. – DOI 10.1198/004017002188618509

[Quoreshi 2014] Quoreshi, A.M.M.S.: A Long-Memory Integer-Valued Time Series Model, INARFIMA, for Financial Application. In: *Quantitative Finance* 14 (2014), Nr. 12, 2225 - 2235. http://dx.doi.org/10.1080/14697688.2012.711911. – DOI 10.1080/14697688.2012.711911. – ISSN 1469–7688

[R Core Team 2018] R Core Team: *R: A Language and Environment for Statistical Computing*. 3.4.4. Vienna, Austria: R Foundation for Statistical Computing, März 2018. https://www.R-project.org

[Sibbertsen 2004] Sibbertsen, Philipp: Long memory versus structural breaks: An overview. In: *Statistical Papers* 45 (2004), Oct, Nr. 4, 465–515. http://dx.doi.org/10.1007/BF02760564. – DOI 10.1007/BF02760564. – ISSN 1613–9798

[Smith 1938] Smith, H.: An empirical law describing heterogeneity in the yields of agricultural crops. In: *The Journal of Agricultural Science* 28 (1938), 1-23. http://dx.doi.org/10.1017/S0021859600050516. – DOI 10.1017/S0021859600050516

[Zaffaroni 2004] Zaffaroni, P.: Stationarity and Memory of ARMA(∞) Models. In: *Econometric Theory* 20 (2004), Februar, Nr. 1, 147 - 160. http://dx.doi.org/10.1017/S0266466604201062. – DOI 10.1017/S0266466604201062