Finite-Sample Properties of GARCH Models in the Presence of Time-Varying Unconditional Variance. A Simulation Study

Oliver Old

Diskussionsbeitrag Nr. 519

Januar 2020

Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung
FernUniversität in Hagen
Universitätsstraße 41
58084 Hagen
http://www.fernuni-hagen.de/ls_statistik/
Sekretariat.Statistik@FernUni-Hagen.de

Finite-Sample Properties of GARCH Models in the Presence of Time-Varying Unconditional Variance. A Simulation Study

Oliver Old*

Abstract

In this paper, the finite-sample properties of symmetric GARCH and asymmetric GJR-GARCH models in the presence of time-varying long term variance are considered. In particular, the deterministic spline-GARCH model is investigated by Monte-Carlo simulation, where the true parameter values are taken from estimated real equity index data. As a proxy for the behaviour of equity indices of developed countries, the S&P500 Index is estimated with the Quasi-Maximum-Likelihood (QML) method for different conditional heteroscedastic models (GARCH, GJR-GARCH, spline-GARCH and spline-GJR-GARCH). The estimated S&P500 parameter values are used to simulate a broad range of 6 different time-series lengths {100, 500, 1000, 5000, 10000, 25000} and 4 different numbers of spline knots {1,4,9,14}, combining to a total amount of 60 different model setups. To the best of my knowledge, there exist only a few limited simulation studies that focus on the spline-GARCH model. The main contribution of this paper is therefore to highlight the behaviour of the QML estimates when the long-term variance is implemented by the spline-GARCH model. Beside this, the paper provides a least-square approach to get useful starting values for the numerical estimation routine.

Keywords: Finite-sample distribution; spline-GARCH model; time-varying unconditional variance; simulation study.

*Research Assistant at the chair of Applied Statistics & Methods of Empirical Social Research (Univ.-Prof. Dr. Hermann Singer), Department of Economics and Business Administration, FernUniversität Hagen, Universitätsstraße 41, 58097 Hagen, Tel.: +49 2331/9871275. Home page: https://www.fernuni-hagen.de/ls_statistik/, E-mail: oliver.old@fernuni-hagen.de.

1 Introduction

The ability to model important stylised-facts of financial returns series like volatility clustering, first explored by Engle (1982) with the so-called Autoregressive Conditional Heteroscedastic (ARCH) model and later generalised by Bollerslev (1986) (GARCH), made considerable progress in the description and forecasting of volatility, in particular for a short period. So one key assumption of these models is the fluctuation of the short-term conditional variance around a constant long-term unconditional variance in a mean-reverse-process. Along with this assumption, many long empirical time series, especially financial ones, reveal a high persistent volatility in a near unit-root state. This so-called integrated GARCH (IGARCH) effect (Engle and Bollerslev, 1986) is based on the often violated assumption of a constant unconditional variance over varying states of volatility. This specious assumption is a consequence of neglected structural breaks and regime switches in GARCH models as proved by Mikosch and Starica (2004) and Hillebrand (2005) beside others. A remedy is to find change points or regimes in between the unconditional variance as well as the estimated parameters are locally constant, but vary among these segments. Some prominent representatives of this approach are Time-Varying models (Mercurio and Spokoiny, 2004; Medeiros and Veiga, 2009; Čižek and Spokoiny, 2009), Smooth-Transition models (González-Rivera, 1998) and Markov-Regime-Switching models (Hamilton, J.D., Susmel, R., 1994; Cai, 1994). Another approach is to mitigate the assumption of a constant and stationary unconditional variance for the whole sample and each possible segment within. This is done by decomposing the variance in a short-term stationary part and in a multiplicative linked long-term non-stationary part. Thereby smoothing the short-term volatility process for a lower volatility persistence. Within this proposal, the parameters are estimated globally for the whole sample. Beside some others, an early semi-parametric approach by Feng (2004) modeled the unconditional variance as a scale function by kernel estimation and the parameters of the conditional variance by maximum-likelihood estimation. In this context, Engle and Rangel (2008) proposed to model the long term variance as an exponential function with a quadratic truncated power basis function, the so-called spline-GARCH-model. As the knots of the spline basis functions are arranged equidistant over all sampled time points, within the spline-GARCH framework there is no need to identify break points or segments in advance. Beside smoothing the long-term volatility process, another issue of their paper is to analyse the economic source of volatility. They analysed the in-sample properties by means of a comprehensive study of the short-term volatility of various real equity indices and their impact due to various exogenous economic determinants. In the spline-GARCH-model, the high-frequent and the low-frequent volatility have the same time index. So if low-frequent exogenous variables are included, the low-frequent component is averaged and therefore constant over a fixed time span. Following the MIxed DAta Sampling approach (Ghysels et al., 2007), Engle et al. (2013) addresses this issue with a fully parametric approach called GARCH-MIDAS. Within this framework, the long-term volatility component, which included exogenous variables in a different frequency, could vary in the same frequency as the short-term component by estimating a rolling window. The ability of models to capture structural breaks in the volatility process in an immanent way is indicated by the persistence of volatility. As mentioned by Engle et al. (2013), although they are conceived to, the spline-GARCH, as well as the GARCH-MIDAS models do not capture all breaks, in particular not fundamental ones.

The main contribution of this study is the examination of the finite-sample properties of the parameters of univariate GARCH and GJR-GARCH models, when the innovation series is smoothed by a long-term component, in particular by a spline-GARCH-model (Engle

and Rangel, 2008) and to explore under which circumstances, within the applied simulation setup, the estimated parameters are consistent and the asymptotic theory for maximum-likelihood-estimators holds. This paper offers, therefore, a comprehensive simulation study of 10 different Data-Generating-Processes (DGP) explored with 6 different time series lengths, each with $M = 1000$ replications, resulting in 60000 simulated paths. The linkage between the volatility of equity returns and its exogenous sources are not further illuminated. The results of this study provide some evidence for empirical researchers, in particular, whether some of the central assumptions of the spline-GARCH model are applicable in the same way for different time series lengths and different numbers of knots. A desirable and important side effect arises, as the standard GARCH and GJR-models are further illuminated under the broad applied simulation setup within this study.
This paper is organized as follows. Section 2 gives a short recap of classic GARCH models with the problems caused by and introduce the applied model specifications. Within this section, the relationship between long-memory processes and structural breaks and their impact to the assumptions of GARCH models are briefly discussed. Section 3 describes the simulation setup, the origin of the Data Generating Processes and the simulation results. Section 4 concludes.

2 Model Specifications

Let p_t be the observed price, y_t the resulting log-returns and ϵ_t the innovations of a financial asset at time $t \in \mathbb{Z}$. Here the time is measured in days. $\Psi_{t-1} = \{p_{t-1}, p_{t-2}, ...\}$ is the information set the observer have up to $t-1$. The log-returns series

$$y_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \cdot 100 \tag{1}$$

$$y_t = \mu_t + \epsilon_t \tag{2}$$

$$\mu_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} \tag{3}$$

is multiplied by 100 to get a percentage of returns. The conditional mean $\mathrm{E}[y_t|\Psi_{t-1}] = \mu_t$ is a dynamic linear function of lagged values of the dependent variable and possibly exogenous independent variables. Without theoretical reason, the mean process (3) will be assumed to follow an AR(2)-process without a constant term, i.e. $\mathrm{E}[y_t] = 0$ by assumption. ϕ_1 and ϕ_2 are constant autoregressive parameters. ϵ_t is the deviation of a return to the expected return at time t, with respect to Ψ_{t-1}. In time series literature this unexpected difference is often called innovation. The variation around the mean of the return series is measured by the variance respectively the standard deviation. This variation is called volatility, which is indicated by the innovation series. Unlike the innovation series, the volatility is not observable and has to be estimated by data. By the assumptions of an efficient market (Fama, 1970), one key property of the innovation series is the independence of past values $\mathrm{E}[\epsilon_t|\Psi_{t-1}] = 0$, i.e. ϵ_t is a martingale difference. From this orthogonality condition, it could be derived that the innovations are uncorrelated $\mathrm{Cov}[\epsilon_t, \epsilon_s] = 0$ for $t \neq s$. But the marginal distribution of financial time series innovations often appear to be leptokurtic, i.e. $\kappa(\epsilon_t) \geq \kappa(z_t)$, whereby $\kappa(z_t)$ is the kurtosis of the process generating variable and $\kappa(\epsilon_t)$ is the kurtosis of the innovation series. So even if the observed innovations are uncorrelated, they are not necessarily independent[1], which will be illuminated later in this paper.
Figure 1 depicts the observed daily spot-prices p_t from Standard & Poor's 500 composite

[1] $\mathrm{E}[g(\epsilon_t)f(\epsilon_s)] \neq \mathrm{E}[g(\epsilon_t)]\mathrm{E}[f(\epsilon_s)]$ for $t \neq s$ and arbitrary functions g, f

stock market index (S&P 500) and the resulting log-returns y_t for the period from January 2, 1980 to December 31, 2018. Table 1 summarises the drawn sample, which is used throughout this paper for illustrative purposes and later as Data Generating Processes (DGP) for the simulation study. The S&P 500 sample was chosen because it stands as a proxy for developed countries' stock market indices. The estimated innovation series from the AR(2)-process will be used as an independent variable for modeling the different conditional variances. Table 2 presents the descriptive statistics and the AR(2) model.

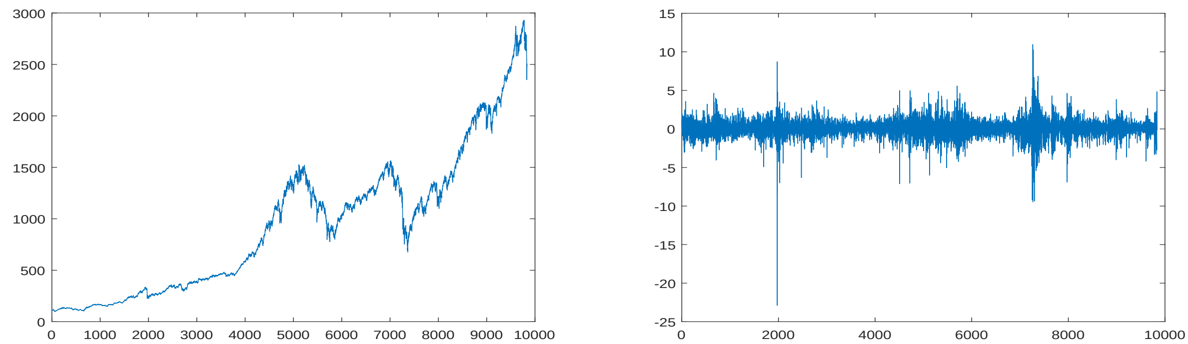

Figure 1: S&P 500 Index. Spot-prices p_t (left), log-returns y_t (right)

Definition	Sample Period	Observations non-trading days adjusted	Frequency	Source
S&P500 equity Index	02.01.1980 - 31.12.2018	9835	daily	Thomson-Reuters-Datastream

Table 1: Sample

	y_t	$\hat{\epsilon}_t$
Mean	0.0322	0.0343
Standard Deviation	1.1055	1.1040
Skewness	-1.1429	-1.2496
Kurtosis	29.5017	30.0626
Minimum	-22.8997	-23.1424
Maximum	10.9572	10.5913
$\hat{\epsilon}_t = y_t - 0.0269 y_{t-1} - 0.0422 y_{t-2}$		AR(2)-model

Table 2: Descriptive Statistics and AR(2)-model

2.1 Short Term Volatility

2.1.1 symmetric GARCH model

Some well-known stylised facts about financial time series like volatility clustering, leptokurtic unconditional distribution and no independent elements in the innovation series are captured by symmetric models of the ARCH-family, like the famous GARCH(P, Q) model

(Bollerslev, 1986). In GARCH models the innovation series

$$\epsilon_t = \sqrt{h_t} z_t \tag{4}$$

is generated by a random variable z_t rescaled by the conditional variance h_t of the ϵ_t series. If for the standardised innovations z_t the assumptions $\mathrm{E}[z_t] = 0$ and $\mathrm{E}[z_t^2] = 1$ hold, then the innovation series is generated by a (semi-) strong GARCH process and if additionally z_t is assumed i.i.d., the innovation series is generated by a strong GARCH process, following the definitions from Drost and Nijman (1993). Hence $\mathrm{E}[z_t] = 0$, $\mathrm{E}[\epsilon_t] = 0$. Hereby, z_t could be assumed to be Gaussian or differently distributed. Holding the assumption of a (semi)-strong GARCH process with $z_t \sim \mathcal{N}(0, 1)$, the innovation series is conditional Gaussian $\epsilon_t|\Psi_{t-1} \sim \mathcal{N}(0, h_t)$. Within this paper all simulated processes are generated by a strong - GARCH process with $z_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$ and with a kurtosis

$$\kappa(z_t) = \mathrm{E}[z_t^4]/(\mathrm{E}[z_t^2])^2 = 3. \tag{5}$$

The following equations (6a) and (6b)

$$\mathrm{E}[\epsilon_t^2|\Psi_{t-1}] = h_t = \alpha_0 + (\sum_{p=1}^{P} \alpha_p z_{t-p}^2 + \sum_{q=1}^{Q} \beta_q) h_{t-q} \tag{6a}$$

$$\mathrm{E}[\epsilon_t^2|\Psi_{t-1}] = h_t = \alpha_0 + (\alpha_1 z_{t-1}^2 + \beta_1) h_{t-1} \tag{6b}$$

reveal the important property of volatility-clustering. As there is empirical evidence for the superiority of GARCH models with order $P = 1$ and $Q = 1$ (Hansen and Lunde, 2001, 2005), hereafter and for the subsequent simulation study, only the GARCH(1,1) will be considered. To have a clear and concise notation of the kurtosis, the autocorrelation and the volatility persistence later, this paper follows the account from He and Teräsvirta (1999), where $\nu_{S,t} = \alpha_1 z_t^2 + \beta_1$ and $\eta_e = \mathrm{E}[\nu_t^e]$. So for the GARCH(1,1) case $\eta_{S,1} = \alpha_1 + \beta_1$ and $\eta_{S,2} = 3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2$ apply to $z_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$, where the subscript S denotes symmetric GARCH models.
Since the conditional variance series h_t cannot be negative by definition, Bollerslev (1986) recommend to constrain the parameters $\alpha_0 > 0$, $\alpha_p \geq 0, p = 1, ..., P$ and $\beta_q \geq 0, q = 1, ..., Q$.

Applying the law of iterated expectation,

$$\mathrm{E}[\epsilon_t^2] = \mathrm{E}[\mathrm{E}[\epsilon_t^2|\Psi_{t-1}]] = \mathrm{E}[h_t]\mathrm{E}[z_t^2] = \frac{\alpha_0}{1 - \eta_{S,1}} = \sigma^2 \tag{7}$$

provides the second moment of the unconditional distribution of the innovation series. So even if the observed innovation series is conditionally heteroscedastic, it is unconditionally homoscedastic. The assumption of a time-invariant unconditional variance, provides a stationary mean-reverse-process. As by the law of large numbers, the sample variance approaches the unconditional variance as $T \to \infty$, the unconditional variance can also be assumed to be the long-term variance. To build a model for a weak stationary short term GARCH (1,1) process,

$$\mathrm{E}[h_t] < \infty \Leftrightarrow \eta_{S,1} < 1 \tag{8}$$

the parameter constraints derive from (7). But a constant unconditional variance is rarely observed, in particular not in long time series. To highlight this problem is one of the key features of the spline-GARCH model. To capture the stylised-fact of a leptokurtic unconditional distribution, the fourth-moment conditions

$$\kappa(\epsilon_t) = \frac{\mathrm{E}[\epsilon_t^4]}{(\mathrm{E}[\epsilon_t^2])^2} = \kappa(z_t)\frac{\mathrm{E}[h_t^2]}{(\mathrm{E}[h_t])^2} \tag{9}$$

of a GARCH(1,1) have to be taken into considerations. $\kappa(\epsilon_t)$ is the kurtosis of the innovation series ϵ_t, with $\kappa(z_t)$ as defined in (5). It, therefore, follows that

$$\mathrm{E}[\epsilon_t^4] = \kappa(z_t)\mathrm{E}[h_t^2] \tag{10a}$$

$$\begin{aligned} \mathrm{E}[h_t^2] &= \alpha_0^2 + \eta_{S,2}\mathrm{E}[h_{t-1}^2] + 2\alpha_0\eta_{S,1}\mathrm{E}[h_{t-1}] \\ &= \frac{\alpha_0^2 + 2\alpha_0\eta_{S,1}\mathrm{E}[h_{t-1}]}{1-\eta_{S,2}} = \frac{\alpha_0^2(1+\eta_{S,1})}{(1-\eta_{S,2})(1-\eta_{S,1})} \end{aligned} \tag{10b}$$

$$\mathrm{E}[h_t^2] < \infty \Leftrightarrow \eta_{S,2} < 1$$

with $\mathrm{E}[h_t]$ like defined in (7). The kurtosis for a specific GARCH(1,1) model in (9) is accordingly defined as

$$\kappa(\epsilon_t) = \kappa(z_t)\frac{1-\eta_1^2}{1-\eta_{S,2}} \geq \kappa(z_t). \tag{11}$$

Here α_1 is the decisive parameter. For $\alpha_1 = 0$, there is no autoregressive conditional heteroscedasticity and the process is distributed as the process generating series z_t. For a large α_1 GARCH(1,1) models a large kurtosis. Closely associated with the kurtosis, is the decaying pattern of the process. The general autocorrelation function of the univariate GARCH family is given by

$$\rho_1 = \frac{\bar{\eta}_1(1-\eta_1^2) - \eta_{S,1}(1-\eta_2)}{3(1-\eta_{S,1}) - (1-\eta_2)} \tag{12a}$$

$$\rho_j = (\eta_{S,1})^{j-1}\rho_1 \quad \text{for} \quad j \geq 1 \tag{12b}$$

where $\bar{\eta}_1 = 3\alpha_1 + \beta_1$ and $\eta_{S,1} = \alpha_1 + \beta_1$ is the exponential decay rate of the autocorrelation function. Financial time series often display heavy tails and a slowly decaying pattern of its autocorrelation function. So the chosen model has to be capable to capture these properties. Within this study, no parameter constraints are imposed. The different conditional variance models will be estimated by maximum-likelihood, with $z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ and $\boldsymbol{\theta} \in \mathbb{R}^v$,

$$\begin{aligned} L_T(\boldsymbol{\theta}) &= \sum_{t=1}^{T} l_t(\boldsymbol{\theta}) \\ l_t(\boldsymbol{\theta}) &= -\frac{1}{2}\ln(2\pi h_t) - \frac{1}{2}\left(\frac{\epsilon_t^2}{h_t}\right) \\ \hat{\boldsymbol{\theta}} &= \arg\max_{\boldsymbol{\theta}} L_T(\boldsymbol{\theta}) \end{aligned} \tag{13}$$

where $\boldsymbol{\theta} = (\alpha_0, \alpha_1, \beta_1)'$ is a $(v \times 1)$ unknown parameters vector, where v is the number of elements of the parameter vector. The initial values $\boldsymbol{\theta}_0$ are chosen by empirical knowledge. Even if the assumption of a normal distribution of the process generating variable z_t is violated, the maximum-likelihood approach yields consistent and approximately normal distributed estimators $\hat{\boldsymbol{\theta}}$. This so-called quasi-maximum-likelihood approach requires the computation of robust-standard errors $\mathbf{se}(\hat{\boldsymbol{\theta}})$, as described by Bollerslev and Wooldridge (1992)

$$\begin{aligned}
\boldsymbol{J}(\hat{\boldsymbol{\theta}}) &= \boldsymbol{G}_T(\hat{\boldsymbol{\theta}})\boldsymbol{G}'_T(\hat{\boldsymbol{\theta}}) \\
\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}) &= \boldsymbol{H}^{-1}(\hat{\boldsymbol{\theta}}) \cdot \boldsymbol{J}(\hat{\boldsymbol{\theta}}) \cdot \boldsymbol{H}^{-1}(\hat{\boldsymbol{\theta}}) \\
\mathbf{se}(\hat{\boldsymbol{\theta}}) &= \left[\text{diag}\left(\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})\right)\right]^{\frac{1}{2}}
\end{aligned} \tag{14}$$

where $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$ is the $(v \times v)$ asymptotic covariance matrix, $\boldsymbol{H}(\hat{\boldsymbol{\theta}})$ is the $(v \times v)$ Hessian, a consistent estimator of the Fisher information matrix, $\boldsymbol{G}_T$ is a $(v \times T)$ gradient matrix and $\boldsymbol{J}(\hat{\boldsymbol{\theta}})$ is the $(v \times v)$ Outer Product of the Gradients (OPG) matrix. Every following example and every DGP is estimated by a quasi-maximum-likelihood approach, with a BFGS optimisation with line searching algorithms. Despite the findings by Fiorentini et al. (1996), gradients and Hessians are computed using finite differencing for demonstrative purpose, as commonly used in most software packages. The optimisation algorithm, gradients and Hessians are adapted from Dennis and Schnabel (1983) pseudo-codes. The models and the likelihood functions are implemented by self-written MATLAB code.

Example 1. S&P500 (cf. tables 1, 2) GARCH(1,1)

$$\begin{aligned}
\hat{h}_t &= \underset{(0.0005)}{0.0154} + \underset{(0.0077)}{0.0847}\epsilon^2_{t-1} + \underset{(0.0060)}{0.9032}\hat{h}_{t-1} \\
\hat{\eta}_{S,1} &= 0.9879 \Rightarrow \text{E}[h_t] < \infty \\
\hat{\eta}_{S,2} &= 0.9903 \Rightarrow \text{E}[h_t^2] < \infty \\
\hat{\sigma}^2 &= 1.273 \quad (\hat{s}^2 = 1.219) \\
\hat{\kappa}(\epsilon_t) &= 7.46 \quad (\hat{k}^s = 30.06) \\
\hat{\rho}_1 &= 0.29 \quad (\hat{r}_1 = 0.12) \quad \hat{\rho}_{50} = 0.16 \quad (\hat{r}_{50} = 0.03) \quad \hat{\rho}_{100} = 0.09 \quad (\hat{r}_{100} = 0.03)
\end{aligned}$$

The robust standard-errors are presented in parentheses under the estimated parameter values. $\hat{s}^2$ is the sample variance, $\hat{k}^s$ is the sample kurtosis and $\hat{r}_j$ are the sample autocorrelation functions.

2.1.2 Asymmetric GJR-GARCH model

Another important stylised fact of financial time series first described by Black (1976): "[...] A negative return will be tied to a rise in volatility, and a positive return will be tied to a fall in volatility". This so-called leverage effect is neglected in symmetric GARCH models (6b). There only the size, but not the sign of each innovation affects the volatility process. A remedy are asymmetric GARCH models like the Exponential GARCH model (Nelson, 1991), the Power-GARCH model (Ding et al., 1993), the Threshold-GARCH model (Zakoian, 1994) and the Glosten-Jagannathan-Runkel (Glosten et al., 1993) (GJR)-GARCH model. As its volatility process is modeled in the same way as the GARCH(P, Q) model and its asymptotic behaviour is well-known, the GJR-GARCH model

$$h_t = \alpha_0 + (\sum_{p=1}^{P} \left(\alpha_p + \delta_p \mathbb{1}_{\epsilon_{t-p}<0}\right) z^2_{t-p} + \sum_{q=1}^{Q} \beta_q) h_{t-q} \tag{15a}$$

$$h_t = \alpha_0 + \left((\alpha_1 + \delta_1 \mathbb{1}_{\epsilon_{t-1}<0}) z^2_{t-1} + \beta_1\right) h_{t-1} \tag{15b}$$

will be applied within this paper. Here the conditional variance h_t depends also on the sign of ϵ_{t-p}. The function $\mathbb{1}_{\epsilon_{t-p}<0}$ indicates the subset $\epsilon_{t-p} < 0$ of the observed innovation series. For values of ϵ_t within this subset, the additional parameter δ_p have to be estimated. If $\delta_p > 0$, then the leverage effect exists. For the sake of consistency, the GJR-GARCH model will also be considered with order $P = 1$ and $Q = 1$ and generated by $z_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. For the GJR-GARCH(1,1) model $\nu_{A,t} = (\alpha_1 + \delta_1 \mathbb{1}_{\epsilon_t<0})z_t^2 + \beta_1$ and consequently $\eta_{A,1} = \alpha_1 + \beta_1 + \frac{1}{2}\delta_1$, $\bar{\eta}_{A,1} = 3(\alpha_1 + \frac{1}{2}\delta_1) + \beta_1$ and $\eta_{A,2} = 3\alpha_1^2 + 2\alpha_1\beta_1 + 3\alpha_1\delta_1 + \beta_1\delta_1 + \beta_1^2 + \frac{3}{2}\delta_1^2$, where the subscript A denotes asymmetric GARCH models.
As in the symmetric case of the GARCH(1,1) model, the unconditional variance σ^2 in the asymmetric GJR-GARCH(1,1) case is also constant

$$\mathrm{E}[\epsilon_t^2] = \mathrm{E}[\mathrm{E}[\epsilon_t^2|\Psi_{t-1}]] = \mathrm{E}[h_t]\mathrm{E}[z_t^2] = \frac{\alpha_0}{1 - \eta_{A,1}} = \sigma^2 \tag{16}$$

and so the process is weakly stationary if

$$1 - \eta_{A,1} < 1. \tag{17}$$

The fourth moment conditions, the kurtosis and the autocorrelation function given by equations (9) - (12b) are similar to the GARCH(1,1) model. As in the GARCH(1,1) case, optimised parameter values $\hat{\boldsymbol{\theta}} = (\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_1, \hat{\delta}_1)'$ are determined by maximum-likelihood estimation, cf. (13).

Example 2. S&P500 (cf. tables 1, 2) GJR-GARCH(1,1)

$$\begin{aligned}
\hat{h}_t &= \underset{(0.0025)}{0.0215} + (\underset{(0.0018)}{0.0216} + \underset{(0.0081)}{0.1259}\mathbb{1}_{\epsilon_{t-1}<0})\epsilon_{t-1}^2 + \underset{(0.0018)}{0.8980}\hat{h}_{t-1} \\
\hat{\eta}_{A,1} &= 0.983 \Rightarrow \mathrm{E}[h_t] < \infty \\
\hat{\eta}_{A,2} &= 0.9915 \Rightarrow \mathrm{E}[h_t^2] < \infty \\
\hat{\sigma}^2 &= 1.233 \quad (\hat{s}^2 = 1.219) \\
\hat{\kappa}_\epsilon &= 12.25 \quad (\hat{k}^s = 30.06) \\
\hat{\rho}_1 &= 0.33 \quad (\hat{r}_1 = 0.12) \quad \hat{\rho}_{50} = 0.14 \quad (\hat{r}_{50} = 0.03) \quad \hat{\rho}_{100} = 0.06 \quad (\hat{r}_{100} = 0.03)
\end{aligned}$$

The robust standard-errors are presented in parentheses under the estimated parameter values. $\hat{s}^2$ is the sample variance, $\hat{k}^s$ is the sample kurtosis and $\hat{r}_j$ are the sample autocorrelation functions.

2.1.3 Long-Range Dependence and Structural Breaks

The innovations ϵ_t are uncorrelated, the absolute and squared innovations are correlated, i.e. the innovations are not independent, which is a well-known fact of financial time series innovations, as noted above. Estimating the Sample Autocorrelation Function (SACF) of ϵ_t^2 for long periods mostly reveals a slowly decaying pattern in the first lags, approximating to a positive constant for larger lags. This is called the long-memory or long-range-dependence (LRD) property of a time series (Beran et al., 2013, p. 19ff). A common measure of LRD is

$$\rho_j \sim c_\rho |j|^{2d-1} \qquad d \in (0, 0.5) \tag{18a}$$

$$\sum_{j=-\infty}^{\infty} \rho_j = \infty \tag{18b}$$

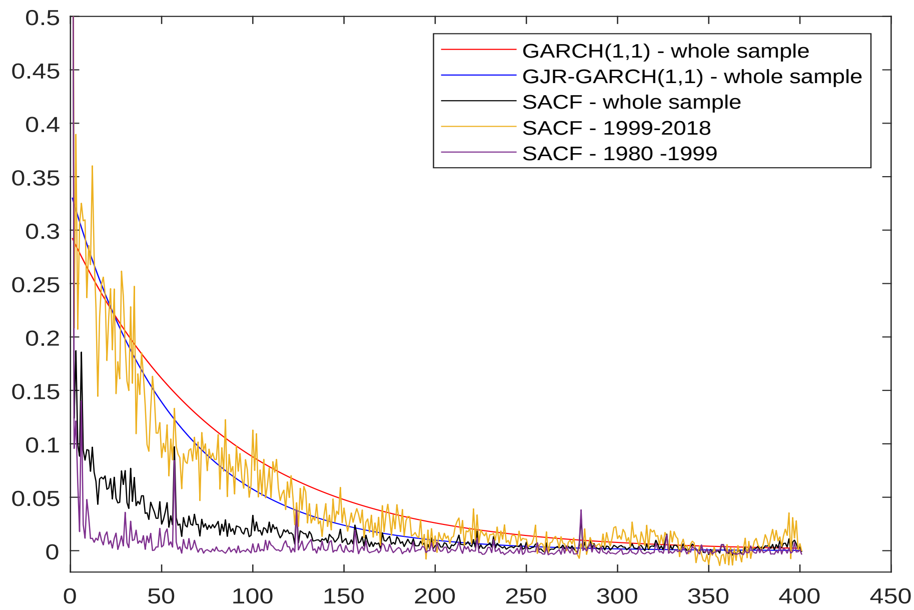

Figure 2: Sample autocorrelation function for the whole sample (black), first half (purple) and the second half (yellow). The red curve is the ACF generated by the estimated GARCH(1,1) parameters, the blue curve is the ACF generated by the estimated GJR-GARCH(1,1) parameters.

where c_ρ is a constant. For large j the autocorrelation converges to zero if $d < 0.5$. The higher d is estimated, the longer the memory of the process is. This measure cannot distinguish between stationary long-memory processes and non-stationary time series (Mikosch and Starica, 2004). The autocorrelation of ARCH type models has an exponential decaying pattern as (12a) and (12b) reveal. Therefore, the persistence of the variance in GARCH(1,1) models is characterised by the sum $\alpha_1 + \beta_1 = \eta_{S,1}$ and in GJR-GARCH(1,1) models by the sum $\alpha_1 + 0.5\delta_1 + \beta_1 = \eta_{A,1}$, the so-called volatility persistence (hereafter VP). The estimation of the VP for long financial series with GARCH models often appear in a nearly unit-root state, i.e. $\hat{\eta}_{S,1} \approx 1$ or $\hat{\eta}_{A,1} \approx 1$. Estimators of Example 1 and 2, where $\hat{\eta}_{S,1} = 0.9879$ respectively $\hat{\eta}_{A,1} = 0.983$, are in line with this assumption. This undisputed stylised-fact of long financial time series motivated Engle and Bollerslev (1986) to the so-called integrated GARCH (IGARCH) model, where $\alpha_1 + \beta_1$ are assumed to sum up to 1. Even though the IGARCH model got some good in-sample and out-of-sample estimation results, there is a lack in the theoretical reasoning of the random walk process of the variance. Diebold (1986) firstly supposed that persistence in volatility is due to the failure in modeling regime switches for the intercept α_0. Later Lamoureux and Lastrapes (1990) proved, that in small samples, the VP is considerably lower than in large samples and that for long periods there are disregarded changes in the structure of the process, which results in the appearance of a very high persistence in a near non-stationary state. As the variance is not observable, it is not possible to assess if a long-memory process generated the data, or if there are neglected structural breaks (Hillebrand and Medeiros, 2008). Hillebrand (2005) stated, that before estimating a GARCH model, a change point detection test is needed. If there are changes in the structure of the process, which were ignored in the estimation of global parameters, the resulting high persistence volatility is spurious, what he called "spurious-almost-integration". Hillebrand (2005) proved, that the reason for a VP almost one, are neglected parameter changes, i.e. different regimes of the unconditional variance σ^2 which are not accounted for. If there are

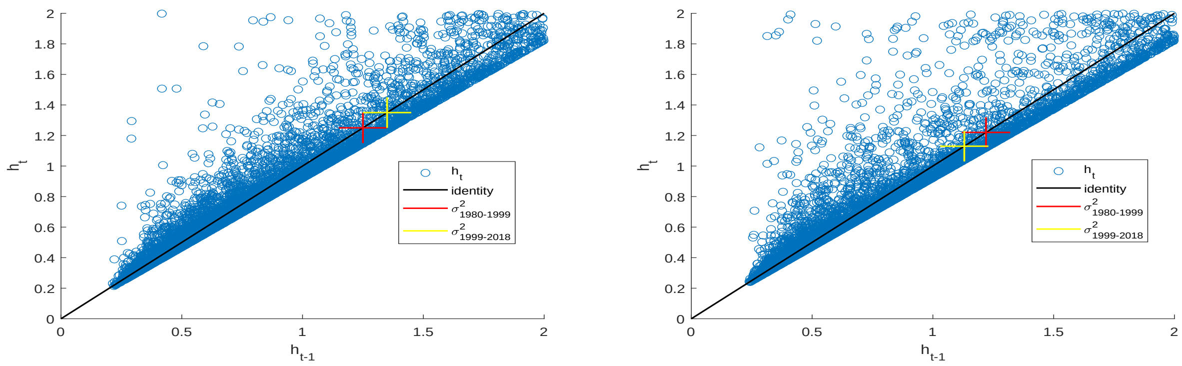

Figure 3: GARCH(1,1) model with $\hat{\beta}_1 = 0.9032$ (left-hand side), GJR-Garch model with $\hat{\beta}_1 = 0.898$ (right-hand side) and the identity curve (black). The red crosses identify $\sigma^2_{1980-1999}$ and the yellow crosses identify $\sigma^2_{1999-2018}$. Both unconditional variances lie on the identity line. h_t values greater than 2 are not displayed for clarity.

one or more breaks in the structure of the time series process with different unconditional means in each segment, global estimated parameters of a GARCH model capture these different means. In particular, $\hat{\beta}_1$ is picking up the slope of the identity line, which crosses the different means of each segment, and $\hat{\alpha}_1$ or $(\hat{\alpha}_1 + \frac{1}{2}\hat{\delta}_1)$ are assumed to be $\hat{\alpha}_1 \approx 1 - \hat{\beta}_1$ or $(\hat{\alpha}_1 + \frac{1}{2}\hat{\delta}_1) \approx 1 - \hat{\beta}_1$, as figure 3 shows (cf. Hillebrand, 2004). The process is getting integrated. Hillebrand (2004, 2005) shows for the GARCH(1,1) case, that in the occurrence of structural breaks $\hat{\beta}_1$ is globally overestimated and $\hat{\alpha}_1$ is globally underestimated. Figure 2 reveals, that there are breaks in the structure of the process and that both GARCH-models are more sensitive to the high-volatility period than to the low-volatility period.

In Example 1 and 2 the unconditional variances are estimated over the whole sample. In figure 4 the estimated unconditional variances are displayed for the whole sample, as well as for the first and the second half of the sample. There is evidence for the drawn sample, that the unconditional variance is not constant, as well in the GARCH(1,1) model as in the GJR-GARCH(1,1) model [2]. These findings are in accordance with those of Hillebrand (2004, 2005) and Mikosch and Starica (2004). So there is evidence, that the estimated high VP in long financial time series is not trustworthy and the so-called IGARCH effect is more due to neglected changes in the process than to a true LRD.

[2]For the GARCH(1,1) model, the unconditional variance is higher in the second segment and in the case of the GJR-GARCH(1,1) model the other way around. This is due to the crash in October 1987, which is illuminated by the spike around time point 2000. There were negative returns from about 23% in one day. The GJR-GARCH framework weights these negative returns due to an extra parameter and therefore the unconditional variance (c.f. (15b) is more sensitive for negative returns.

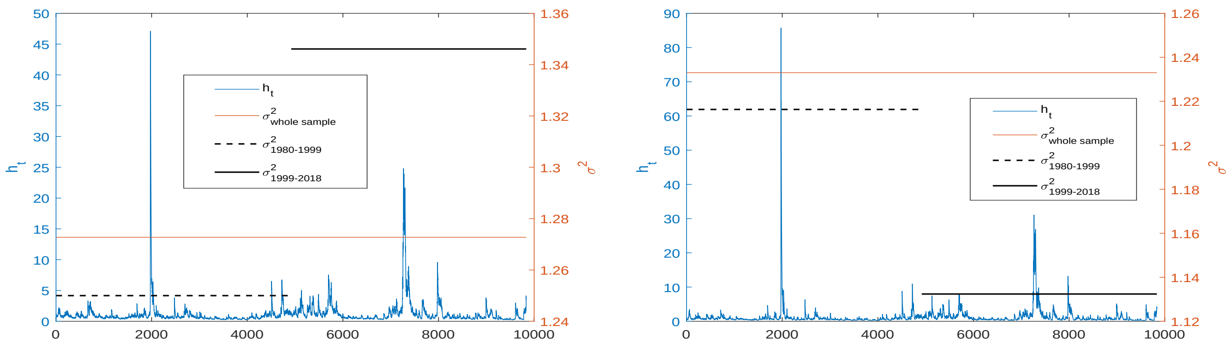

Figure 4: Conditional and unconditional variances. GARCH(1,1) (left-hand side), GJR-GARCH(1,1)(right-hand side)

2.2 Long Term Volatility

As underlined in the previous section, the assumption of a constant unconditional variance could lead to spurious integrated processes. To face these problems, several approaches like Markov-Regime-Switching (Hamilton, J.D., Susmel, R., 1994; Cai, 1994), Time-Varying GARCH (Mercurio and Spokoiny, 2004; Medeiros and Veiga, 2009; Čižek and Spokoiny, 2009) or Smooth-Transition models (González-Rivera, 1998), just to mention the most influential models, have been established. More recent approaches, tend to decompose the conditional variance process in a short term and a long term part. An early exploration in this field has been done by Engle and Lee (1999). They decomposed the conditional variance by the two aforementioned components additively. More recent approaches, decompose the conditional variance by multiplicative long- and short-run components. Two important models are the so-called GARCH-MIDAS (Engle et al., 2013) and the Spline-GARCH model (Engle and Rangel, 2008). A multivariate extension to the spline-GARCH model dealing with a factor framework is the factor-spline-GARCH model by Rangel and Engle (2012). Amado et al. (2018) give a good outline over further models with multiplicative decomposed conditional variances. Conrad and Kleen (2018) examined the statistical properties of multiplicative GARCH models. To the best of my knowledge, there exists a few other studies dealing with estimation conditional variance by splines. So, for example, Audrino and Bühlmann (2009) build stochastic B-spline basis functions to model the logarithm of the general conditional variance, Brownlees and Gallo (2010) modeled the long-term volatility part as penalized B-spline. The use of the term "nonparametric" for spline smoothing or spline interpolation is widespread in literature, but somewhat misleading, as spline basis functions are estimated by parameters. In that regard, Eilers and Marx (1996) recommend to use the terms "overparametric techniques" or "anonymous models". From the statistical point of view, the term "smoothing" will be preferred within this paper.

2.2.1 Spline-GARCH model

The principal reasons for introducing the Spline-GARCH model by Engle and Rangel (2008) [3] were to explain the sources of financial time series volatility by exogenous macro-economic variables. As these variables are typically measured in a different frequency, the spline-GARCH model seems to be the first one capable to embedding those variables. Beside

[3]The original paper from Engle and Rangel (2008) will be referred to as "original framework".

this, it has been shown, that the problems with a VP in a near unit-root state could be mitigated, which is a gratifying side effect that will be further illuminated within this paper. The innovation series (19)

$$\epsilon_t = z_t \sigma_t \tag{19}$$
$$\sigma_t^2 = h_t \tau_t \tag{20}$$

is also generated by a random variable z_t assumed to be standard normal and i.i.d. and the conditional variance (20) is decomposed into a short-term part h_t and a long-term part

$$\tau_t^o = c \cdot \exp\left(w_0 t + \sum_{i=1}^{K} w_i ((t - t_{i-1})_+)^2 + \gamma x_t \right) \tag{21a}$$
$$\tau_t = \exp\left(c' + w_0 \frac{t}{T} + \sum_{i=1}^{K} w_i \left(\frac{(t - t_{i-1})_+}{T} \right)^2 \right). \tag{21b}$$

In the original framework in (21a) τ_t^o is constructed as an exponential spline for modeling exogenous sources of volatility embedded by the variable x_t. To the end of this paper, some modifications to the spline-GARCH model has been made in (21b). As recommended by Laurent (2013), the time is rescaled by T to keep the optimisation numerically stable. The constant is modeled as $c' = \exp(c)$ to ensure $\tau_t > 0$. For purposes of this study, only the deterministic part of the spline function, but no exogenous variables will be considered. The spline bases are truncated power functions

$$(t - t_i)_+^2 = \begin{cases} (t - t_i)^2 & \text{if } t > t_i \\ 0 & \text{otherwise} \end{cases}$$
$$t_0 = 0; t_1, t_2, ..., t_{K-1}$$

with equidistant knots as illustrated in figure 5. Engle and Rangel (2008) recommend to estimate different spline orders by models with a range of different numbers of knots and choose the optimal model by the Bayesian Information Criterion (BIC). In the context of the S&P500 sample, the GARCH(1,1)-spline(9) and the GJR-GARCH(1,1)-spline(9) models (cf. tables 4 and 5) are the optimal choice and will be used for illustrative purposes within this section. Spline smoothing with truncated power series basis functions has some serious drawbacks. de Boor (2001, p. 84ff) shows that truncated spline functions tend to be linear dependent, if the knots are very nonuniform and if the distance between two adjacent knots are too close. Additionally, if the knots are too close to each other the estimators are getting insignificant. Both problems are relevant for the spline-GARCH framework if the knots/observation ratio is very high.

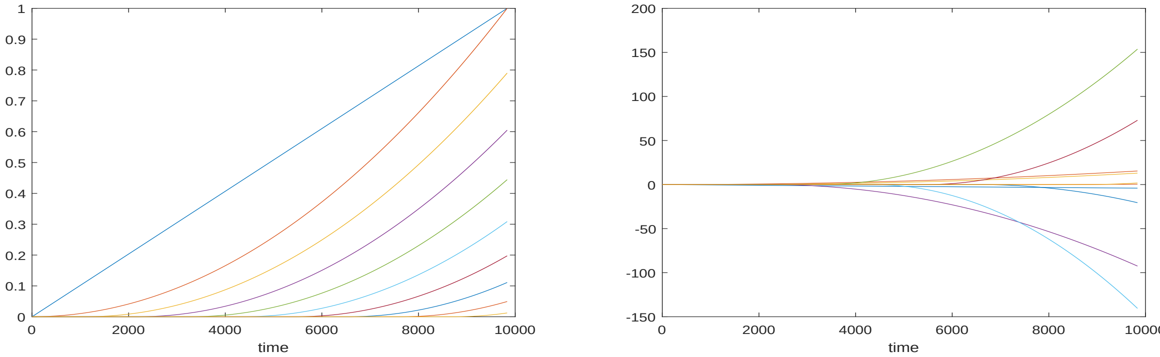

Figure 5: Spline basis function with 9 equidistant knots $\frac{t}{T}$, $\left(\frac{(t-t_0)_+}{T}\right)^2$,..., $\left(\frac{(t-t_8)_+}{T}\right)^2$ (left-hand side) and spline basis function scaled by estimated parameters $\hat{\omega}_0$, $\hat{\omega}_i$,...,$\hat{\omega}_9$ (right-hand side)

As noted above, classic GARCH models, with the assumption of a constant unconditional variance, are only capable to capture some of the most important stylised-facts for short periods. So within the original framework, the short-term volatility h_t is modeled as a GARCH(1,1) model and within this study extended by the GJR-GARCH(1,1) model. Both models are now smoothed by the long-term volatility τ_t. The definition of the conditional variances in the spline-GARCH model

$$h_t = \alpha_0 + \alpha_1 \left(\frac{\epsilon_{t-1}^2}{\tau_{t-1}}\right) + \beta_1 h_{t-1} = \alpha_0 + \left(\alpha_1 z_{t-1}^2 + \beta_1\right) h_{t-1} \tag{22a}$$

$$\begin{aligned} h_t &= \alpha_0 + (\alpha_1 + \delta_1 \mathbb{1}_{\epsilon_{t-1}<0}) \left(\frac{\epsilon_{t-1}^2}{\tau_{t-1}}\right) + \beta_1 h_{t-1} \\ &= \alpha_0 + \left((\alpha_1 + \delta_1 \mathbb{1}_{\epsilon_{t-1}<0}) z_{t-1}^2 + \beta_1\right) h_{t-1} \end{aligned} \tag{22b}$$

shows the similarity to the standard GARCH models in (6a) and (6b), with the difference, that the process of the innovations in (19) is now smoothed by τ_t. Multiplying both sides with τ_t results in the well-known form of (20). As τ_t is deterministic it holds that $\mathrm{E}[\tau_t] = \tau_t$ and $\mathrm{E}[\tau_t z_s] = 0 \quad \forall t, s$. To insulate the time-varying effect of τ_t and to avoid identification problems within the model, the variance is targeted as described in Engle and Mezrich (1996) resulting in a so-called unit-GARCH process. Hereafter α_0 is modeled as

$$\alpha_0 = (1 - \eta_{S,1}) \tag{23a}$$

$$\begin{aligned} \alpha_0 &= (1 - \eta_{A,1}) \\ &\Rightarrow \mathrm{E}[h_t] = \sigma^2 = 1 \\ &\Rightarrow \mathrm{E}[h_t^2] = \frac{(1-\eta_1)(1+\eta_1)}{1-\eta_2} \end{aligned} \tag{23b}$$

This representation reveals the linkage between the unconditional variance σ^2 and the long-term volatility τ_t, as

$$\mathrm{E}[\epsilon_t^2] = \mathrm{E}[h_t \tau_t z_t^2] = \tau_t \mathrm{E}[h_t] = \tau_t. \tag{24}$$

As intended, the spline-GARCH framework is capable to model the unconditional variance time-varying. So even if the exogenous sources of volatility are hidden within this paper, their impact on long-term patterns is to some extent picked up by the deterministic spline basis functions. In Conrad and Kleen (2018), the kurtosis for multiplicative GARCH models is derived. But the assumptions made in their paper are not met by the spline-GARCH model, because here the time-varying unconditional variance is represented by a deterministic function (24).

The likelihood function contains the conditional variance σ_t as defined in (20) and a vector of unknown parameters $\boldsymbol{\theta} = (\alpha_1, \beta_1, \delta_1, c', w_0, w_1, ..., w_K)$

$$
\begin{aligned}
l_t(\boldsymbol{\theta}) &= -\frac{1}{2}\ln(2\pi\sigma_t) - \frac{1}{2}\left(\frac{\epsilon_t^2}{\sigma_t}\right) \\
\hat{\boldsymbol{\theta}} &= \arg\max_{\boldsymbol{\theta}} L_T(\boldsymbol{\theta})
\end{aligned} \tag{25}
$$

The optimisation of a spline-GARCH model is more demanding and more sensible to chosen initial values than in a GARCH(1,1) or in a GJR-GARCH(1,1) model. So the choice of good starting values is required. Therefore a two-step estimation procedure is recommended. In the first step the GARCH(1,1) and the GJR-GARCH(1,1) parameters $\hat{\boldsymbol{\theta}}_{0\mathrm{G}}$ are estimated like in (13). With these parameter values, the $\hat{h}_{0t}$ series is evaluated. In the second step, an ordinary least squares (OLS) approach is applied. Before calculating the OLS estimator, τ_t is approximated by τ_{0t}, which is received by the following transformation

$$
\epsilon_t = \sqrt{\hat{h}_{0t}\tau_{0t}}z_t \Rightarrow \epsilon_t^2 = \hat{h}_{0t}\tau_{0t}z_t^2 \Rightarrow \tau_{0t} = \epsilon_t^2/(\hat{h}_{0t}z_t^2),
$$

where z_t is a standard-normally distributed random number and ϵ_t is the observed innovation series. After taking the logarithm of the computed τ_{0t} values

$$
\ln\tau_{0t} = c_0 + w_{00}\frac{t}{T} + \sum_{i=1}^{K} w_{0i}\left(\frac{(t-t_{i-1})_+}{T}\right)^2 \tag{26}
$$

the initial parameter vector $\hat{\boldsymbol{\theta}}_{0\tau_t}$ is obtained by OLS estimation

$$
\underbrace{\begin{bmatrix} \ln\tau_{01} \\ \ln\tau_{02} \\ . \\ . \\ . \\ \ln\tau_{0T} \end{bmatrix}}_{\boldsymbol{Y}} = \underbrace{\begin{bmatrix} 1 & \frac{1}{T} & \left(\frac{(1-t_0)_+}{T}\right)^2 & . & . & \left(\frac{(1-t_{K-1})_+}{T}\right)^2 \\ 1 & \frac{2}{T} & \left(\frac{(2-t_0)_+}{T}\right)^2 & . & . & \left(\frac{(2-t_{K-1})_+}{T}\right)^2 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & \frac{T}{T} & \left(\frac{(T-t_0)_+}{T}\right)^2 & . & . & \left(\frac{(T-t_{K-1})_+}{T}\right)^2 \end{bmatrix}}_{\boldsymbol{X}} \underbrace{\begin{bmatrix} c_0 \\ w_{00} \\ w_{01} \\ . \\ . \\ w_{0K} \end{bmatrix}}_{\hat{\boldsymbol{\theta}}_{0\tau_t}} + \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_T \end{bmatrix}}_{\boldsymbol{U}} \tag{27a}
$$

$$
\hat{\boldsymbol{\theta}}_{0\tau_t} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}, \tag{27b}
$$

where $\hat{\boldsymbol{\theta}}_{0\tau_t} = (\hat{c}_0', \hat{w}_0, \hat{w}_1, ..., \hat{w}_K)$. To improve the initial values, the second step is replicated and $L_T(\hat{\boldsymbol{\theta}}_{0\tau_t})$ is evaluated $x = 1000$ times. The parameter vector $\hat{\boldsymbol{\theta}}_{0\tau_t}$ with the largest $L_T(\hat{\boldsymbol{\theta}}_{0\tau_t})$ was chosen. The resulting starting values are $\hat{\boldsymbol{\theta}}_0 = (\hat{\boldsymbol{\theta}}_{0\mathrm{G}}, \hat{\boldsymbol{\theta}}_{0\tau_t})$.

Example 3.

3.1. S&P500 (cf. tables 1, 2) Spline(9)-GARCH(1,1)

$$
\begin{aligned}
\hat{h}_t &= 0.0319 + \underset{(0.00085)}{0.0881}\left(\frac{\epsilon_{t-1}^2}{\hat{\tau}_{t-1}}\right) + \underset{(0.0009)}{0.88}\,\hat{h}_{t-1} \\
\hat{\boldsymbol{\theta}}_{\tau_t} &= (\underset{(0.018)}{0.276}, \underset{(0.024)}{-4.08}, \underset{(0.034)}{15.39}, \underset{(0.144)}{16.41}, \underset{(0.061)}{-152.92}, \underset{(0.134)}{345.81}, \underset{(0.42)}{-455.28}, \underset{(0.66)}{369.37}, \underset{(0.219)}{-185.18}, \underset{(0.325)}{7.98}, \underset{(5.96)}{148.69})
\end{aligned}
$$

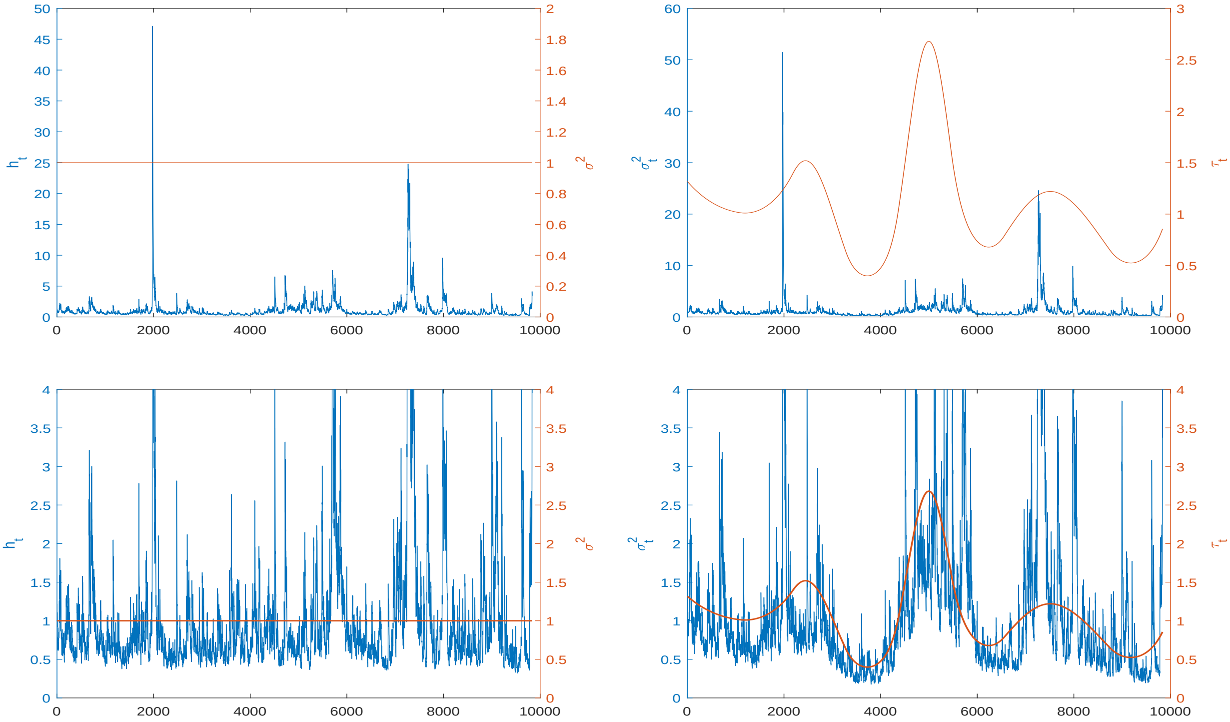

Figure 6: S&P500 (cf. tables 1, 2) GARCH(1,1) and spline(9)-GARCH(1,1) models. GARCH(1,1) model with variance targeting and constant unconditional variance σ^2 (left column) and spline(9)-GARCH(1,1) model with time-varying unconditional variance τ_t (right column). In the top row, all estimated variance values are plotted. To zoom in, in the bottom row only variance values in the range $[0, 4]$ are displayed.

3.2. S&P500 (cf. tables 1, 2) Spline(9)-GJR-GARCH(1,1)

$$\hat{h}_t = 0.03695 + (\underset{(0.004)}{0.0101} + \underset{(0.007)}{0.1491}\mathbb{1}_{\epsilon_{t-1}<0}) \left(\frac{\epsilon^2_{t-1}}{\hat{\tau}_{t-1}}\right) + \underset{(0.001)}{0.8784}\hat{h}_{t-1}$$

$$\hat{\boldsymbol{\theta}}_{\tau_t} = (\underset{(0.0012)}{0.357}, \underset{(0.034)}{-3.26}, \underset{(0.04)}{16.38}, \underset{(0.023)}{-1.95}, \underset{(0.058)}{-108.24}, \underset{(0.072)}{284.21}, \underset{(0.135)}{-390.81}, \underset{(0.233)}{322.22}, \underset{(0.02)}{-161.93}, \underset{(0.19)}{12.06}, \underset{(1.46)}{97.59})$$

Regarding example 3, figures 6 and 7, corroborate some of the theoretical considerations made. One intended consequence is the reduced VP, which declines from 0.988 to 0.968 for the GARCH(1,1) case and from 0.983 to 0.963 for the GJR-GARCH(1,1) case[4]

3 Simulation Study

In this section, the finite-sample properties of the GARCH parameters $(\alpha_1, \beta_1, \delta_1)$ in the presence of a time-varying unconditional variance τ_t will be stressed. To the best of my knowledge there exist only a few limited simulation studies for the spline-GARCH-model so far. Goldman and Wang (2015) compared their so-called spline-threshold-GARCH model with the original spline-GARCH model by the means of a single simulation of $T = 5000$ datapoints with $M = 200$ replications and a fixed number of $K = 9$ knots. Goldman and Shen (2017) conducted a similar simulation setup with $M = 400$ replications and a broader

[4]To keep in mind, that GARCH(1,1) and GJR-GARCH(1,1) models with constant unconditional variance were estimated with intercept. But the estimation with variance targeting yield to nearly the same VP (0.984 or 0.978), cf. tables 4 and 5.

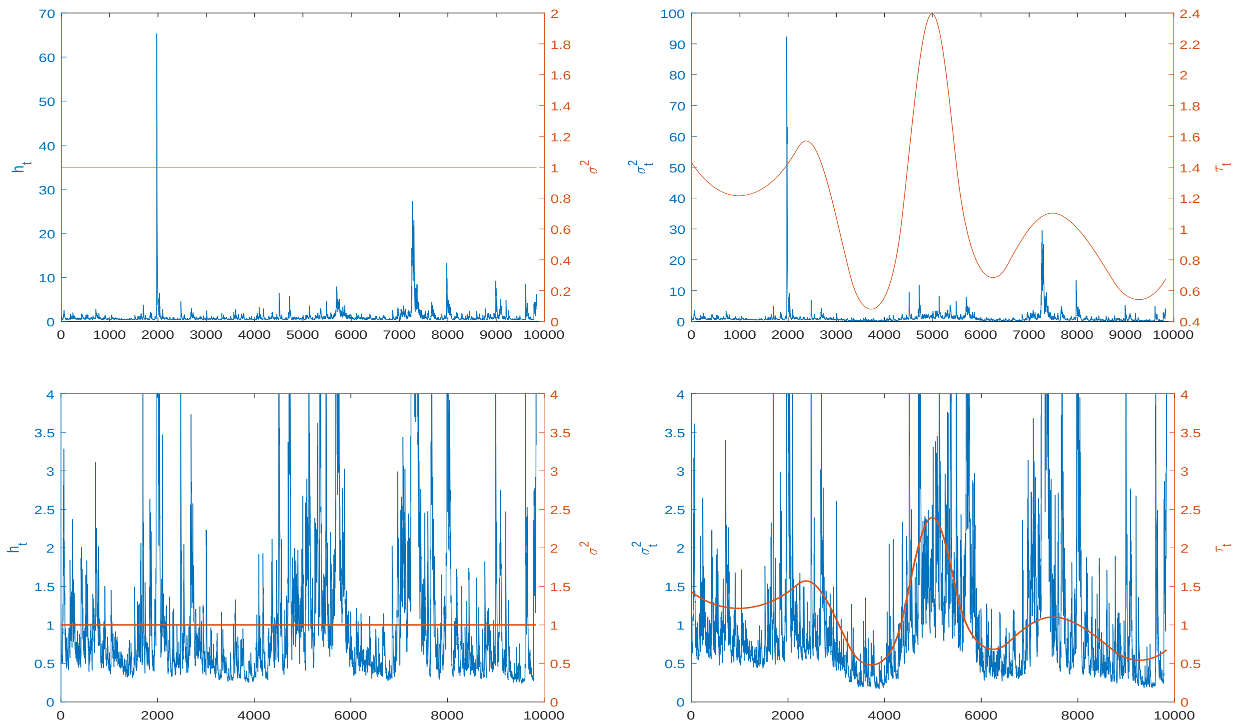

Figure 7: S&P500 (cf. tables 1, 2) GJR-GARCH(1,1) and spline(9)-GJR-GARCH(1,1) models. GJR-GARCH(1,1) model with variance targeting and constant unconditional variance σ^2 (left column) and spline(9)-GJR-GARCH(1,1) model with time-varying unconditional variance τ_t (right column). In the top row, all estimated variance values are plotted. To zoom in, in the bottom row only variance values in the range $[0, 4]$ are displayed.

set of reference models. Within this paper, a more extensive simulation study is conducted, based on the continuously used S&P500 sample.

With the initial sample of 9835 spot prices p_t and the resulting log-returns y_t, the innovation series ϵ_t is obtained by an AR(2) model (cf. table 1). With the innovation series, 10 different conditional variance models are estimated. The obtained estimates from each of the $N = 10$ models are used as DGP for the simulation study (cf. tables 4 and 5). Each DGP is then applied to 6 different time series lengths $T \in \{100, 500, 1000, 5000, 10000, 25000\}$, leading to $N \times 6$ different model setups with $M = 1000$ replications each. So there were 60000 paths simulated and the related parameters estimated. Each simulation is generated by $z_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$. Table 3 summarises the simulation setup.

As aforementioned, to obtain finite second or higher-order moments, some equality, inequality and positivity restrictions have to be imposed. These restrictions may lead to likelihood optimisation problems near the imposed boundaries, in particular, if the actual optimisation algorithm is built to solve unconstrained optimisation problems (Silvennoinen, 2006, p. 155-167). Therefore the positivity constraint by the exponential form of τ_t in (21b) is the only restriction imposed within this paper. For this reason, some of the proposed replications didn't converge, wherefore $m = 1, ..., M$ ist indexed. After simulating $M = 1000$ replications for $N = 10$ different DGP applied to six different time series lengths, each of the $N \times M \times 6$ simulated time series are estimated with the quasi-maximum-likelihood approach described in equations (13) and (25).

As mentioned in section 2.1.1, every DGP is estimated by the presented optimisation routine and implemented in MATLAB by following Dennis and Schnabel (1983) pseudo-codes. For computational purposes, the estimation of the 60000 simulated times series, is done by the

	Notation	
$t = 1, ..., T$	$T \in \{100, 500, 1000, 5000, 10000, 25000\}$	Time Series Lengths
$n = 1, ..., N$	$N \in \{1, ..., 10\}$	DGP
$m = 1, ..., M$	$M \in \{1000\}$	Replications
$i = 1, ..., K$	$K \in \{1, 4, 9, 14\}$	Knots
$N \times M \times 6$ time series generated by $z_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$		
N Data Generating Processes (cf. tables 4 and 5)		
$\tilde{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta}_n} L_T(\boldsymbol{\theta}_n)$		
$\tilde{\boldsymbol{\theta}}_n = (\tilde{\alpha}_{1_n}, \tilde{\beta}_{1_n})$		GARCH(1,1)
$\tilde{\boldsymbol{\theta}}_n = (\tilde{\alpha}_{1_n}, \tilde{\beta}_{1_n}, \tilde{\delta}_{1_n})$		GJR-GARCH(1,1)
$\tilde{\boldsymbol{\theta}}_n = (\tilde{\alpha}_{1_n}, \tilde{\beta}_{1_n}, \tilde{c}'_n, \tilde{w}_{n0}, \tilde{w}_{i_n}, ..., \tilde{w}_{K_n})$		SPLINE-GARCH(1,1)
$\tilde{\boldsymbol{\theta}}_n = (\tilde{\alpha}_{1_n}, \tilde{\beta}_{1_n}, \tilde{\delta}_{1_n}, \tilde{c}'_n, \tilde{w}_{n0}, \tilde{w}_{i_n}, ..., \tilde{w}_{nK})$		spline-GJR-GARCH(1,1)
$N \times 6$ Model Setups		
$\bar{\hat{\boldsymbol{\theta}}}_{nT} = (\bar{\hat{\alpha}}_{1_{nT}}, \bar{\hat{\beta}}_{1_{nT}})$		GARCH(1,1)
$\bar{\hat{\boldsymbol{\theta}}}_{nT} = (\bar{\hat{\alpha}}_{1_{nT}}, \bar{\hat{\beta}}_{1_{nT}}, \bar{\hat{\delta}}_{1_{nT}})$		GJR-GARCH(1,1)
$\bar{\hat{\boldsymbol{\theta}}}_{nT} = (\bar{\hat{\alpha}}_{1_{nT}}, \bar{\hat{\beta}}_{1_{nT}}, \bar{\hat{c}}'_{nT}, \bar{\hat{w}}_{0_{nT}}, \bar{\hat{w}}_{i_{nT}}, ..., \bar{\hat{w}}_{K_{nT}})$		spline-GARCH(1,1)
$\bar{\hat{\boldsymbol{\theta}}}_{nT} = (\bar{\hat{\alpha}}_{1_{nT}}, \bar{\hat{\beta}}_{1_{nT}}, \bar{\hat{\delta}}_{1_{nT}}, \bar{\hat{c}}'_{nT}, \bar{\hat{w}}_{0_{nT}}, \bar{\hat{w}}_{i_{nT}}, ..., \bar{\hat{w}}_{K_{nT}})$		spline-GJR-GARCH(1,1)
$\approx N \times M \times 6$ Replications		
$\hat{\boldsymbol{\theta}}_{mnT} = \arg\max_{\tilde{\boldsymbol{\theta}}_n} L_T(\tilde{\boldsymbol{\theta}}_n)$		
$\hat{\boldsymbol{\theta}}_{mnT} = (\hat{\alpha}_{1_{mnT}}, \hat{\beta}_{1_{mnT}})$		GARCH(1,1)
$\hat{\boldsymbol{\theta}}_{mnT} = (\hat{\alpha}_{1_{mnT}}, \hat{\beta}_{1_{mnT}}, \hat{\delta}_{1_{mnT}})$		GJR-GARCH(1,1)
$\hat{\boldsymbol{\theta}}_{mnT} = (\hat{\alpha}_{1_{mnT}}, \hat{\beta}_{1_{mnT}}, \hat{c}'_{mnT}, \hat{w}_{0_{mnT}}, \hat{w}_{i_{mnT}}, ..., \hat{w}_{K_{mnT}})$		spline-GARCH(1,1)
$\hat{\boldsymbol{\theta}}_{mnT} = (\hat{\alpha}_{1_{mnT}}, \hat{\beta}_{1_{mnT}}, \hat{\delta}_{1_{mnT}}, \hat{c}'_{mnT}, \hat{w}_{0_{mnT}}, \hat{w}_{i_{mnT}}, ..., \hat{w}_{K_{mnT}})$		spline-GJR-GARCH(1,1)

Table 3: Simulation Setup

integrated MATLAB algorithm *fminunc* and the DGP parameters $\tilde{\boldsymbol{\theta}}_n$ are used as starting values for accelerating convergence. Every other calculation is implemented by an own MATLAB code.

3.1 Simulation Setup

The occurrence of a VP in a near-unit root state, may cause the spurious assumption of a long-memory process and ignoring the existence of breaks in the structure of the process (c.f. section 2.1.3). So one intention of the spline-GARCH model is to mitigate the integrated GARCH effect, by allowing the unconditional variance to vary over time and so expose slow-moving regime switches, which are covered by global parameter estimation. In tables 4 and 5 are the estimates of the 10 DGP listed. To be concise, GARCH(1,1) or GJR-GARCH(1,1) models are displayed as zero knot spline-GARCH models. To emphasise that the DGP are themselves estimates of an unknown process and to distinguish these parameters from the estimators of the replications, the DGP parameters are specified by $\tilde{\boldsymbol{\theta}}_{nT}$. Referring to the deterministic character of the spline part of the model, no initial data points were eleminated. The initial values $h_0 = 1$ and $\epsilon_0^2 = 1$ are chosen arbitrarily for the simulations. The initial values for the estimation procedure $\hat{h}_0$ and $\hat{\epsilon}_0^2$ are chosen by $1/T \sum_{t=1}^{T} \epsilon_t^2$, following the suggestion from Bollerslev (1986).

3.1.1 Sample Statistics

The subscripts m describe the m^{th} replication of DGP n and the corresponding time series length T. The vector of estimated parameters $\hat{\boldsymbol{\theta}}_{mnT}$ contains $v(n)$ elements of the m^{th} replication. $\bar{\hat{\boldsymbol{\theta}}}_{nT}$ is the related mean vector over M_{nT} converged replications. $\tilde{\boldsymbol{\theta}}_n$ is the vector of the particular DGP. Every parameter vector is a $[v \times 1]$ column vector. The resulting $[v \times M_{nT}]$ matrix $\hat{\boldsymbol{\Theta}}_{mnT}$ contains all of the converged estimates. $\hat{\boldsymbol{\Theta}}^c_{mnT} = \hat{\boldsymbol{\Theta}}_{mnT} - \bar{\hat{\boldsymbol{\theta}}}_{nT}$ is a centered matrix. The considered statistics

$$\hat{\mathrm{E}}[\hat{\boldsymbol{\theta}}_{nT}] = \frac{\sum_{m=1}^{M_{nT}} \hat{\boldsymbol{\theta}}_{mnT}}{M_{nT}} = \bar{\hat{\boldsymbol{\theta}}}_{nT} \tag{28}$$

$$\widehat{\mathrm{Cov}}[\hat{\boldsymbol{\theta}}_{nT}, \hat{\boldsymbol{\theta}}'_{nT}] = \frac{\hat{\boldsymbol{\Theta}}^c_{mnT} \hat{\boldsymbol{\Theta}}^{c'}_{mnT}}{M_{nT} - 1} = \hat{\boldsymbol{s}}(\hat{\boldsymbol{\theta}}_{nT}, \hat{\boldsymbol{\theta}}'_{nT}) \tag{29}$$

$$\widehat{\mathrm{Var}}[\hat{\boldsymbol{\theta}}_{nT}] = \mathrm{diag}\left(\hat{\boldsymbol{s}}(\hat{\boldsymbol{\theta}}_{nT}, \hat{\boldsymbol{\theta}}'_{nT})\right) = \hat{\boldsymbol{s}}^2(\hat{\boldsymbol{\theta}}_{nT})$$

$$\widehat{\mathrm{Std}}[\hat{\boldsymbol{\theta}}_{nT}] = \sqrt{\hat{\boldsymbol{s}}^2(\hat{\boldsymbol{\theta}}_{nT})} = \hat{\boldsymbol{s}}(\hat{\boldsymbol{\theta}}_{nT})$$

$$\widehat{\mathrm{bias}} = |\bar{\hat{\boldsymbol{\theta}}}_{nT} - \tilde{\boldsymbol{\theta}}_n| \tag{30}$$

$$\hat{\mathrm{E}}[\bar{\hat{\boldsymbol{\theta}}}_{nT} - \tilde{\boldsymbol{\theta}}_n]^2 = \widehat{\mathrm{MSE}} = \hat{\boldsymbol{s}}^2(\hat{\boldsymbol{\theta}}_{nT}) + \widehat{\mathrm{bias}}^2 \tag{31}$$

$$\widehat{\mathrm{RMSE}} = \sqrt{\widehat{\mathrm{MSE}}}$$

in equations (28)-(31) are the sample statistics of the estimated $M_{nT} \times N \times 6$ parameter vectors. (28) is the arithmetic mean over all replications of each model set up. (29) describes the estimations for the second moments of the distribution of the estimated parameters. The estimator for the covariance $\hat{\boldsymbol{s}}(\hat{\boldsymbol{\theta}}_{nT}, \hat{\boldsymbol{\theta}}'_{nT})$ is a Gramian matrix. The bias (30) is expressed in absolute values. The RMSE (31) measures the trade-off between variance and unbiasedness. To evaluate the statistics of the VP, some adjustments have to be made,

$$\mathrm{bias} = \|\hat{\eta}_{1_{nT}} - \tilde{\eta}_{1_n}\|_2 \tag{32}$$

$$\mathrm{Var}[\hat{\eta}_{S,1_{nT}}] = \mathrm{Var}[\hat{\alpha}_{1_{nT}}] + \mathrm{Var}[\hat{\beta}_{1_{nT}}] + 2\mathrm{Cov}[\hat{\alpha}_{1_{nT}}, \hat{\beta}_{1_{nT}}] = \hat{s}^2(\hat{\eta}_{S,1_{nT}}) \tag{33}$$

$$\begin{aligned}\mathrm{Var}[\hat{\eta}_{A,1_{nT}}] = {} & \mathrm{Var}[\hat{\alpha}_{1_{nT}}] + \mathrm{Var}[\hat{\beta}_{1_{nT}}] + \frac{1}{4}\mathrm{Var}[\hat{\delta}_{1_{nT}}] + 2\mathrm{Cov}[\hat{\alpha}_{1_{nT}}, \hat{\beta}_{1_{nT}}] + \\ & \frac{1}{2}\left(2\mathrm{Cov}[\hat{\delta}_{1_{nT}}, \hat{\alpha}_{1_{nT}}] + 2\mathrm{Cov}[\hat{\delta}_{1_{nT}}, \hat{\beta}_{1_{nT}}]\right) = \hat{s}^2(\hat{\eta}_{S,1_{nT}})\end{aligned} \tag{34}$$

as the VP is a sum of dependent variables. The bias (32) is measured as Euclidean norm. For the variances of the VP (33) and (34), the covariance between each parameter has to be taken into account. To assess the assumption of normality the one-sample Kolmogorov-Smirnov-Test (KS)

$$d^+_{M_{nT}} = \max\left[\max_{m=1,..,M_{nT}} \left(F_{M_{nT}}(\hat{\theta}_{mnT}) - F_0(\hat{\theta}_{mnT})\right); 0\right] \tag{35a}$$

$$d^-_{M_{nT}} = \max\left[\max_{m=1,..,M_{nT}} \left(F_0(\hat{\theta}_{mnT}) - F_{M_{nT}}(\hat{\theta}_{mnT})\right); 0\right] \tag{35b}$$

$$d_{M_{nT}} = \max(d^+_{M_{nT}}; d^-_{M_{nT}}) \tag{35c}$$

is conducted. Here $F_0(\hat{\theta}_{mnT})$ is the theoretical continuous distribution, in this case, the standard normal distribution. $F_{M_{nT}}(\hat{\theta}_{mnT})$ is a step function of the ascending ordered normalised parameter estimators with discontinuities at M_{nT} points. So for each replication with the particular estimator $\hat{\theta}_{mnT}$, $F_{M_{nT}}(\hat{\theta}_{mnT})$ gives the fraction of values smaller than $\hat{\theta}_{mnT}$ over all replications. The maximum distance between the theoretical and the empirical distribution function is measured by $d_{M_{nT}}$. The quantiles are approximated by

$$d_{M_{nT;1-\alpha}} \approx \sqrt{-\frac{1}{2M_{nT}}\ln\frac{\alpha}{2}} \tag{36}$$

in this case, $\alpha = 0.05$ to be consistent with assumptions above. The mean and the standard deviation for normalisation are received by (28), respectively (29). For further detailed description see Rao (2002, p. 420ff).

3.1.2 Asymptotic Statistics

Although every time series is generated by a $z_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$ standardised innovation process, for each replication and each related converged estimator the robust asymptotic covariance matrix $\mathbf{\Sigma}(\hat{\boldsymbol{\theta}}_{mnT})$ and the resulting robust-standard error $\mathbf{se}(\hat{\boldsymbol{\theta}}_{mnT})$ (14) were estimated, to force $\mathbf{\Sigma}(\hat{\boldsymbol{\theta}}_{mnT})$ beeing p.s.d. for the demanding estimation process with up to 19 parameters. With the $\mathbf{se}(\hat{\boldsymbol{\theta}}_{mnT})$ confidence intervals (CI) around each estimated parameter with a significance level of 95% was constructed. The fraction of $\tilde{\boldsymbol{\theta}}_n$ covered by each CI around $\hat{\boldsymbol{\theta}}_{mnT}$ is measured by the so-called coverage probability (p_c)

$$\text{95\% CI:} \quad \hat{\boldsymbol{\theta}}_{mnT} \pm 1.96 \cdot \mathbf{se}(\hat{\boldsymbol{\theta}}_{mnT}) \tag{37}$$

$$\hat{p}_c = \frac{\#\{\tilde{\boldsymbol{\theta}}_n \in \text{CI}\}}{M_{nT}} \tag{38}$$

$$\text{Bernoulli Standard Error:} \quad \sqrt{\frac{\hat{p}_c(1-\hat{p}_c)}{M_{nT}}}, \tag{39}$$

following Lumsdaine (1995). The bias (30) is applied to test consistency $\bar{\hat{\boldsymbol{\theta}}}_{nT} \xrightarrow{p} \tilde{\boldsymbol{\theta}}_n$ assumption. To examine asymptotically normality $\sqrt{T}(\bar{\hat{\boldsymbol{\theta}}}_{nT} - \tilde{\boldsymbol{\theta}}_n) \xrightarrow{d} \mathcal{N}(\tilde{\boldsymbol{\theta}}_n, \mathbf{\Sigma}(\tilde{\boldsymbol{\theta}}_n))$ a Gaussian kernel estimator of $\bar{\hat{\boldsymbol{\theta}}}_{nT}$ will be compared with related Gaussian distribution. The kernel bandwidth is chosen by $b = 1.06\hat{\boldsymbol{s}}(\hat{\boldsymbol{\theta}}_{nT})T^{-1/5}$ according to the proposals by Silverman (1998, Equation (3.28)). To keep the presentation concise, only a few of the total amount of 150 different parameter distributions will be depicted in figures 14, 15, 19, 20 and 21.

3.2 Results

Before considering the results of the simulation study, the DGP will be described. In the original paper by Engle and Rangel (2008), 48 real equity return time series with realised volatilities as exogenous variable x_t in (21a) were examined. In the global view no appreciable difference between the standard GARCH and the spline-GARCH model for the ARCH effect (α_1) were apparent. For the GARCH effect (β_1) a decrease from GARCH to spline-GARCH was observed, but independent to the number of knots. This independence is due to the global view over all 48 time series. For the purposes of this paper, just one real equity return time series is considered. Just to keep in mind, every DGP process was originally estimated for a $T = 9835$ time series. A look into the estimations of DGP in tables 4 and 5, reveals some of the general patterns of the spline-GARCH model. In contrast to the gobal view of the original paper, a dependence with the number of knots seem obvious, even if within this study, the spline-GJR-GARCH model is also analysed. Referring to the GARCH case includes the standard GARCH(1,1) model as well as the spline(K)-GARCH(1,1) model. The same holds for the GJR-GARCH case. So unlike in the GARCH case, where initially a slight increase between 1 and 4 knots appears, for the GJR-models the VP decline continuously. As mentioned in section 2.1.3 , it is in particular $\tilde{\beta}_1$ driving the VP. So regarding just the $\tilde{\beta}_1$ parameters, there is a continuous decline from approximately 0.9 to 0.87 observable for both cases. So a lowering of the VP is an immanent pattern within the the spline-GARCH and the spline-GJR-Garch model. Regarding the other single parameters some differences appear. In line with the findings in the original paper, the $\tilde{\alpha}_1$ parameter increase slightly in the GARCH case, but decrease substantially in the GJR-GARCH case from 0.02 to 0.007. For the spline-GJR-GARCH models, the $\tilde{\delta}_1$ parameters increase from zero knot case to the 14 knots case from 0.12 to 0.16. So a time-varying pattern of the unconditional variances seems to strengthen the leverage effect. The black curves in figure 8 represent the values of the DGP.

In regard to the results of the simulation study, for $T \in \{100, 500, 1000\}$, referred to as the small samples, the replications are very compressed compared to the original process. So some of the $M = 1000$ ML optimisations didn't converge. In particular, the small samples with $K = 14$ had convergence rates only between $55\% - 80\%$. For $T \in \{5000, 10000, 25000\}$, referred to as the large samples, the convergence rates are on or just below 100%. Additionally, as discussed above, there were only parameter constraints in the τ_t equation imposed. Subsequently the time series lengths $T \in \{100, 500\}$ had a considerable fraction of explosive parameter constellations and negative estimators, in particular for $\hat{\alpha}_{1_{mnT}}$ in GJR models. Nevertheless, for all GARCH models, the condition $h_t > 0 \quad \forall t$ holds. The GJR-GARCH models, in contrast, had a significant number of negative h_t for some t in different model setups. So, in particular, all GJR-GARCH models with $K = 9$ over all time series lengths occurred negative h_t for some t. For GJR-GARCH models with $K \in \{0, 1, 4, 15\}$ for $T \in \{10000, 25000\}$ no further negative values for h_t occur. Within the time series lengths $T \in \{1000, 5000, 10000, 25000\}$, every single estimation met the recommended covariance stationary constraints (cf. (8) and (17)). In advance, it can be noted, that the results of the large samples are more robust than the small samples. As some GJR-GARCH models generated negative conditional variances, the results for the GJR-GARCH models have to analysed selectively. As mentioned above, de Boor (2001, p. 84ff) recommended not using truncated spline functions, when adjacent knots are very close. In the small samples with a high knots/observation ratio, this problem occurs and could indicate some of the occurred problems with related estimators. All estimators and related statistics can be found in tables (6)-(9).

In figures 8 - 10 the behaviour of the the VP in presence of different spline models is depicted. In figure 8 the estimated VP are presented. As the black curve provides the VP of the true values, it can be seen at the first glance, that bias declines with increasing time series length, even if there is a rise with an increasing number of knots (figure 9). Also in line with asymptotic maximum-likelihood-theory are the standard-deviations $\hat{s}^2(\hat{\eta}_{1_{nT}})$ of the

VP. As figure 10 reveals, the variance also declines with increasing time series length, but rises with an increasing number of knots, as one might expect. There are no noteworthy differences between the values and the global pattern between the sample statistics of the symmetric and asymmetric GARCH models, in the case of large samples.
The $\hat{p}_c$ results are listed in table 10 and depicted in figures 11 - 13. For the zero knot case the results are in line with those by Lumsdaine (1995) [5] and approach the determined significance level with rising time series length for the $\tilde{\alpha}_1$ and $\tilde{\beta}_1$ parameters. The advertised significance level is clearly understated for all three parameters in the GJR-GARCH zero knot case. For the spline-GJR-GARCH model the coverage probabilities for $\tilde{\alpha}_1$ and $\tilde{\delta}_1$ rise continuously from 1 to 14 knots and matches the determined significance level for $T \in \{5000, 25000\}$. For the $\tilde{\beta}_1$ parameters, a decline between zero and 9 knots and an increase for 14 knots to a fraction clearly below 95% is evident. Considering the spline-GARCH models another picture emerges, as there is no stringent pattern throughout all parameters within all CI apparent. With the occurrence of 1 knot, the $\hat{p}_c$ of both parameters break down, for $\tilde{\beta}_1$ even to nearly zero. With increasing number of knots, the coverage probability for the $\tilde{\beta}_1$ parameters rise throughout, whereas for the $\tilde{\alpha}_1$ the coverage probability has no such clear pattern. On the one hand, as the CI is computed with normality assumption, there is evidence, that for some model orders, the normality assumption could be violated. As also for the large samples, the advertised significance level is mostly understated, the assumption of asymptotic normality could not be met by all parameters. On the other hand, the CI is also computed with robust-standard-errors and here the same pattern as for the $\hat{p}_c$ is apparent for the $\mathbf{se}(\hat{\boldsymbol{\theta}}_{mnT})$ [6]. So in some cases, the CI is very narrow and accordingly the $\hat{p}_c$ is low.
Comparing the results of the coverage probability with the KS-statistics, the picture differs considerably, as can be seen in tables 11 and 12. Throughout all zero-knot GARCH and GJR-GARCH maximum-likelihood-estimators for $T \in \{1000, 5000, 10000, 25000\}$ are significantly normally distributed. For the spline-GARCH models, all ML-estimators for $T \in \{5000, 10000, 25000\}$ are significantly normally distributed. Regarding the depicted distributions of the GARCH parameters in figures 14 and 15 and for the GJR-GARCH parameters in figures 19, 20 and 21 there is evidence, that the approximation to the normal distribution decelerate with increasing number of knots. For the small samples the distributions of the $\hat{\beta}_{1_{mnT}}$ parameters are heavily skewed, whereas the $\hat{\alpha}_{1_{mnT}}$ and $\hat{\delta}_{1_{mnT}}$ parameters are symmetrically distributed. In the GARCH case the distributions of the $\hat{\beta}_{1_{mnT}}$ parameters are bimodal, with major mode around the mean. The minor mode rises with an increasing number of knots. In the GJR-GARCH case, the minor knot is clearly smaller, than in the GARCH case for all number of knots. In both zero knot cases, there is even in the small samples no bimodal distribution for the $\hat{\beta}_{1_{mnT}}$ parameters. The same applies to all $\hat{\alpha}_{1_{mnT}}$ and $\hat{\delta}_{1_{mnT}}$ parameters throughout all models. The tails of the small samples are fatter than normal, more pronounced in the GARCH case. When parameters are restricted, this issue is often observed, also for large samples, and called pile-up effect. Within this study, no restrictions for short-term volatility parameters are imposed. The fat tails of the small sample distributions are therefore the result of the data.

[5] even if Lumsdaine (1995) just examine CI for $T \in \{200, 500\}$ and with different DGP.
[6] $\mathbf{se}(\hat{\boldsymbol{\theta}}_{mnT})$ results on request.

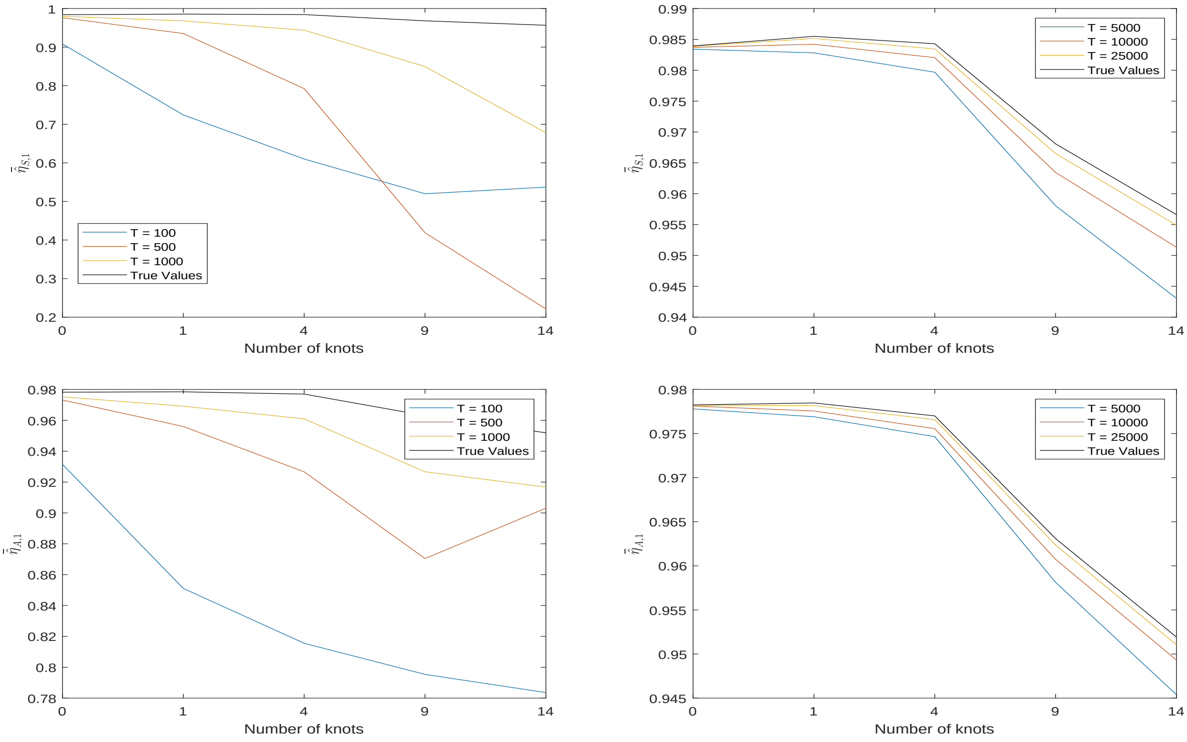

Figure 8: VP $\hat{\eta}_{1_{nT}}$. GARCH(1,1)(top row) and GJR-GARCH(1,1) (bottom row). Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). The black curve represents the true values $\tilde{\eta}_{S,1_{nT}}$ and $\tilde{\eta}_{A,1_{nT}}$. Some GJR-GARCH replications and related ML estimators result to some negative h_t

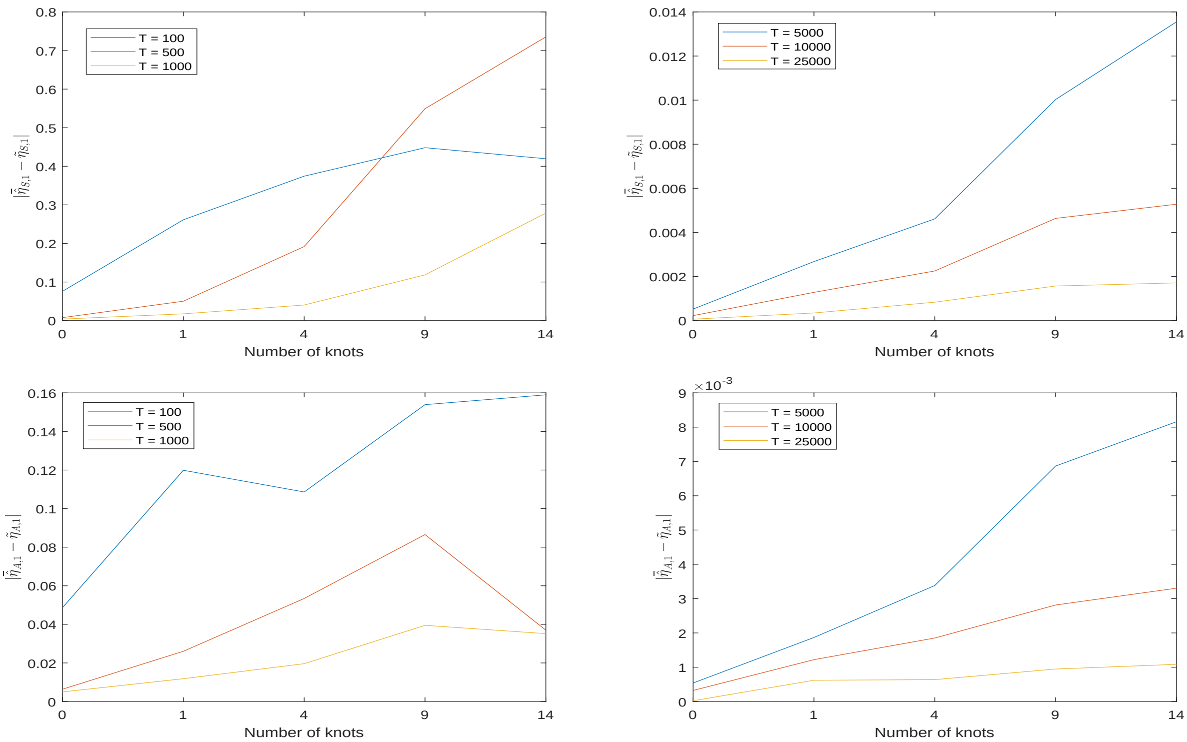

Figure 9: Bias $\|\hat{\eta}_{1_{nT}} - \tilde{\eta}_{1_n}\|_2$. GARCH(1,1)(top row) and GJR-GARCH(1,1) (bottom row). Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). Some GJR-GARCH replications and related ML estimators result to some negative h_t.

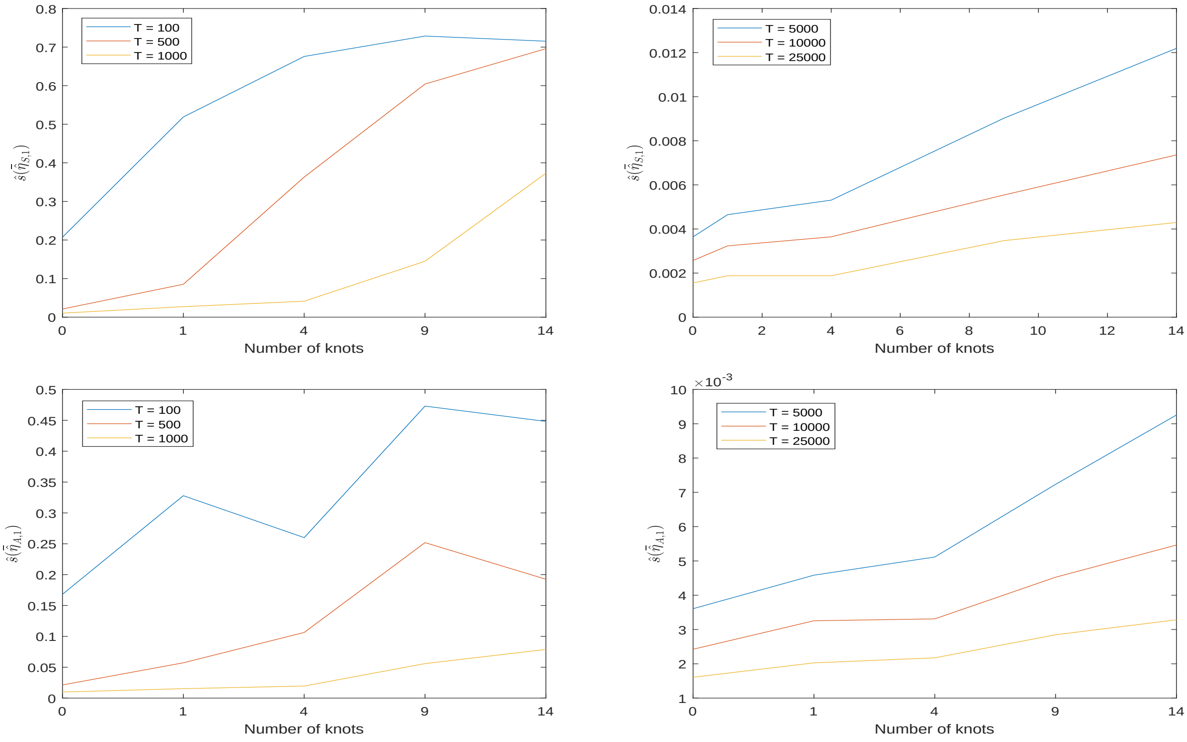

Figure 10: Standard deviation $\hat{s}(\hat{\eta}_{1_{nT}})$. GARCH(1,1)(top row) and GJR-GARCH(1,1) (bottom row). Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). Some GJR-GARCH replications and related ML estimators result to some negative h_t.

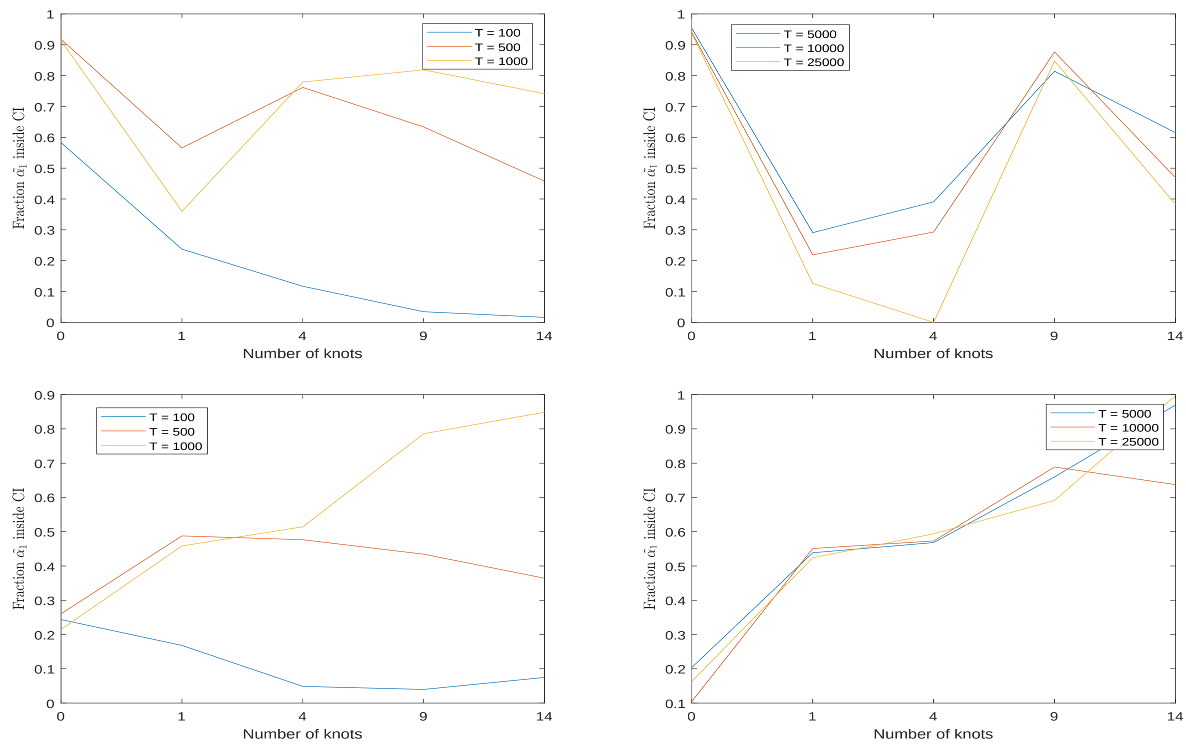

Figure 11: Coverage Probability of $\tilde{\alpha}_1$.GARCH(1,1)(top row) and GJR-GARCH(1,1) (bottom row). Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). Some GJR-GARCH replications and related ML estimators result to some negative h_t.

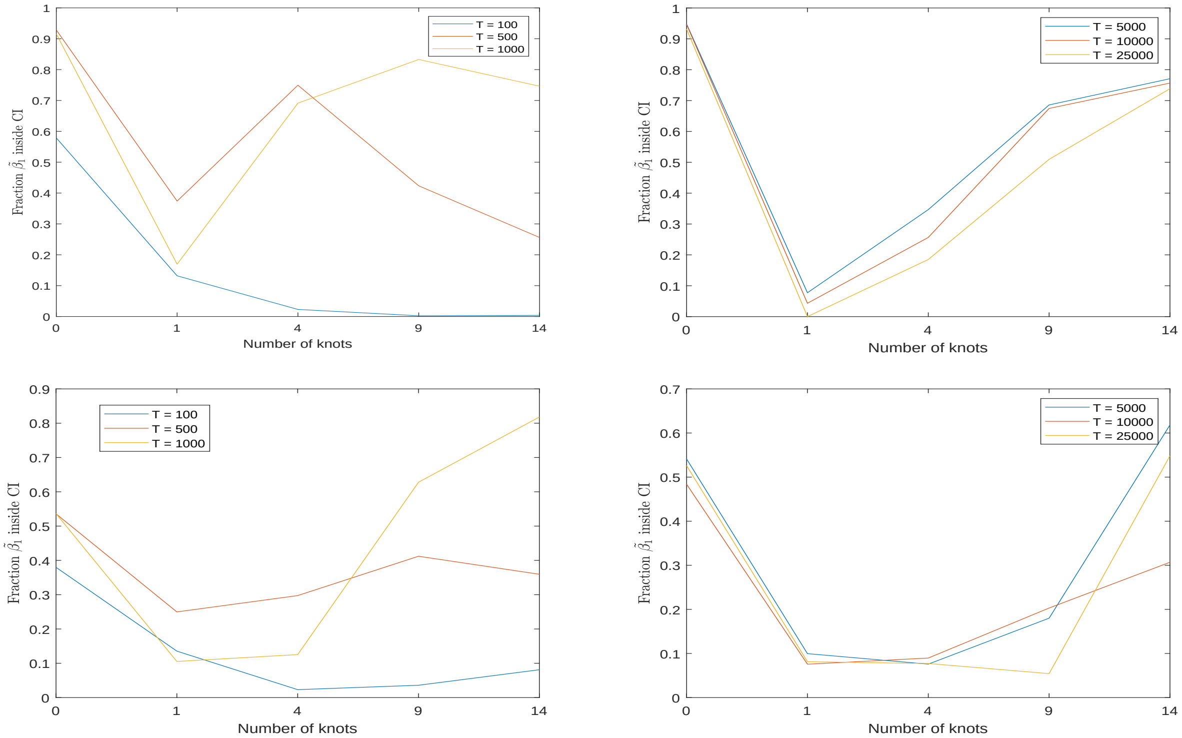

Figure 12: Coverage Probability of $\tilde{\beta}_1$Garch(1,1)(top row) and GJR-Garch(1,1) (bottom row).Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). Some GJR-GARCH replications and related ML estimators result to some negative h_t.

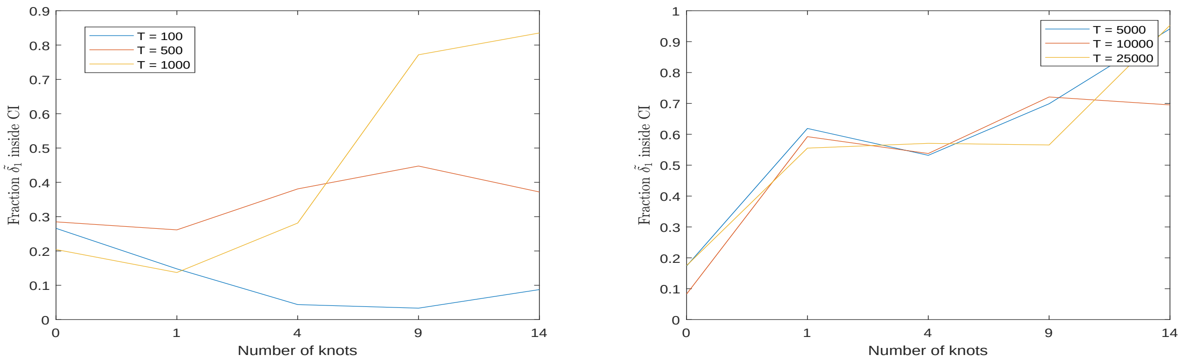

Figure 13: Coverage Probability of $\tilde{\delta}_1$ GJR-Garch(1,1).Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). Some GJR-GARCH replications and related ML estimators result to some negative h_t.

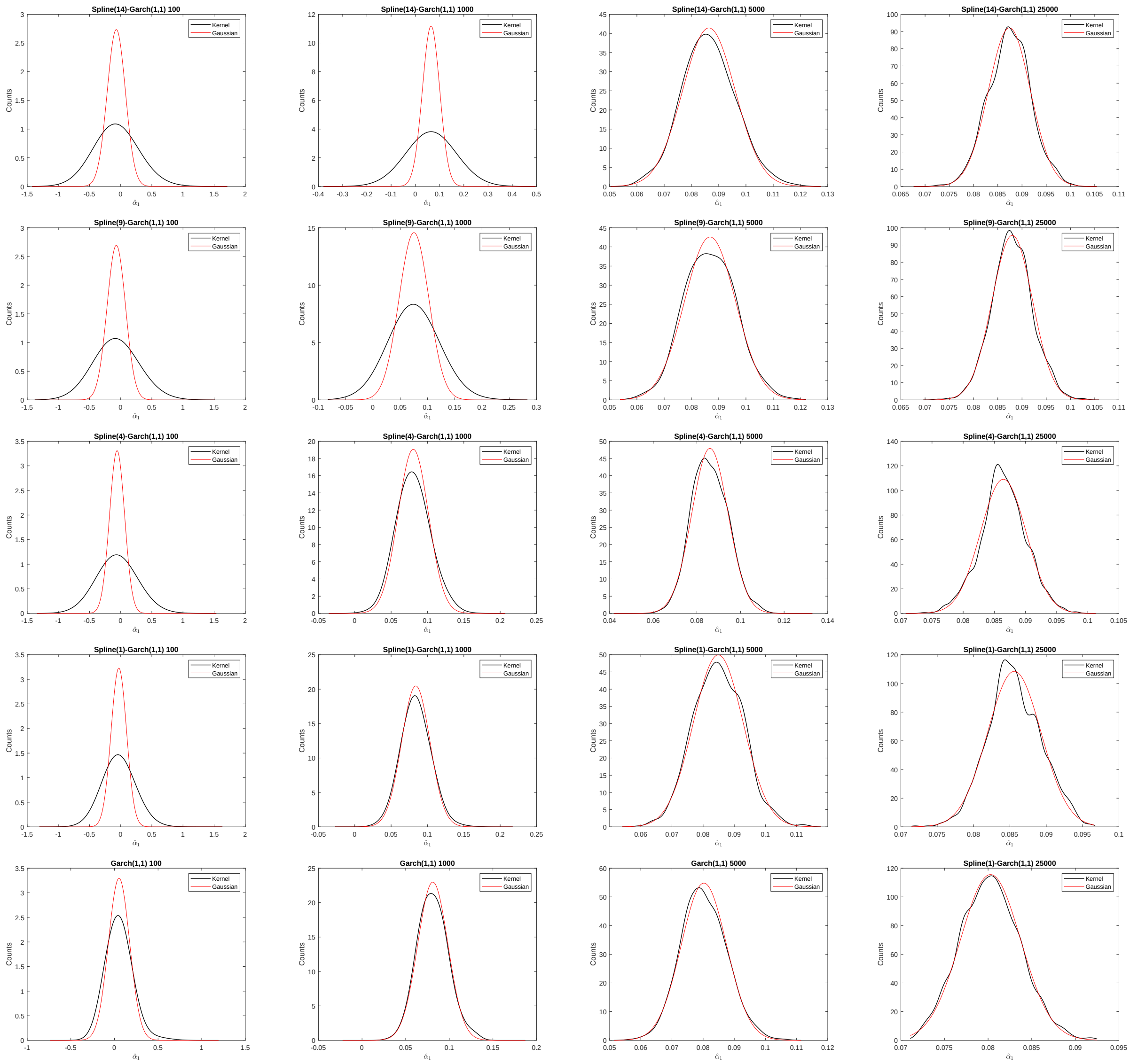

Figure 14: Asymptotic Normality of $\hat{\alpha}_1$ in GARCH(1,1) case. The black curves represent the kernel density estimator for $\hat{\alpha}$, the red curve the related normal distribution. In the top row the spline(14)-GARCH(1,1) distribution for $T \in \{100, 1000, 5000, 25000\}$ is depicted. The subjacent rows are the spline(K)-GARCH(1,1) models with $K \in \{0, 1, 4, 9\}$ and the same time series lengths.

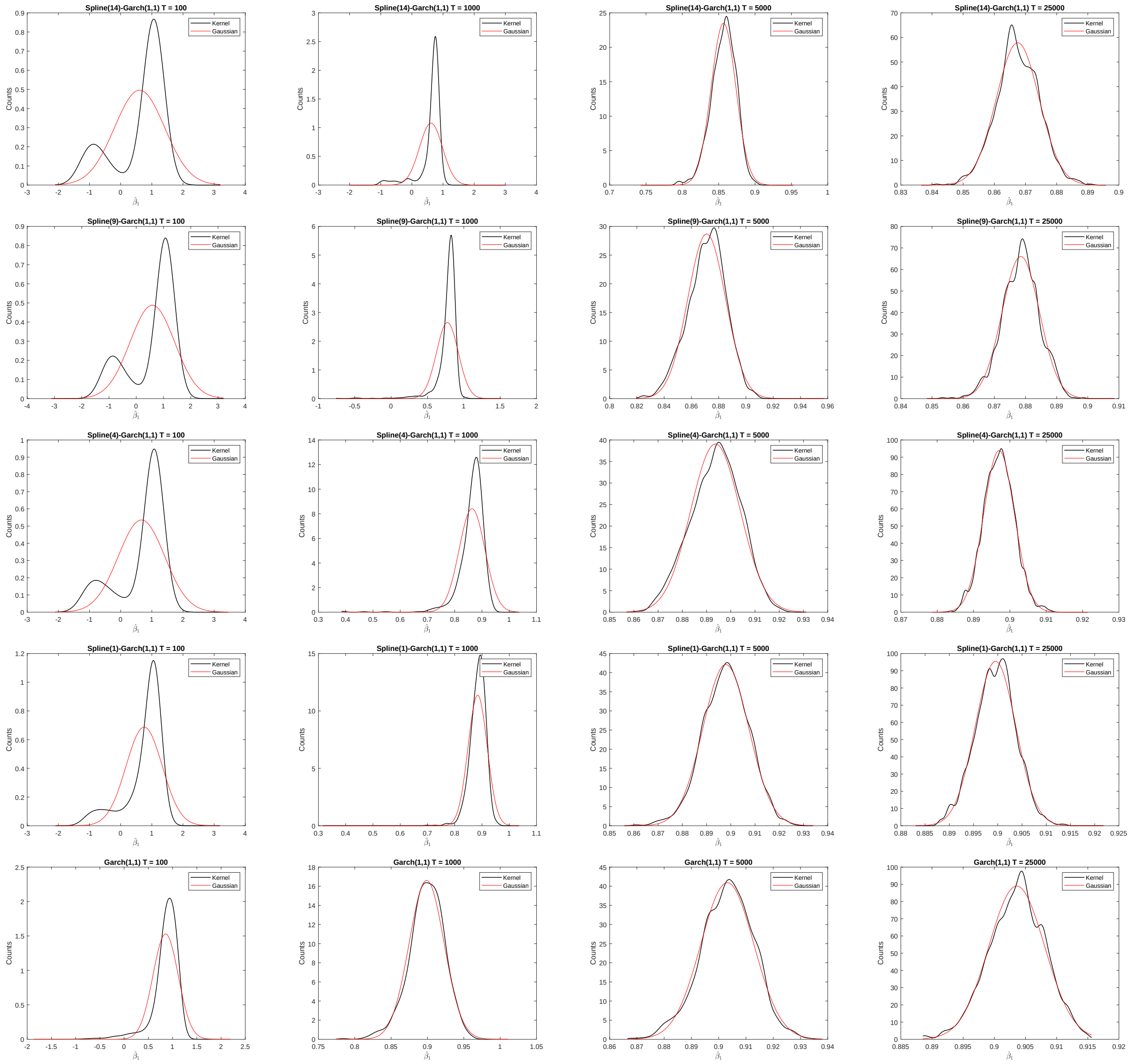

Figure 15: Asymptotic Normality of $\hat{\beta}_1$ in GARCH(1,1) case. The black curves represent the kernel density estimator for $\hat{\beta}$, the red curve the related normal distribution. In the top row the spline(14)-GARCH(1,1) distribution for $T \in \{100, 1000, 5000, 25000\}$ is depicted. The subjacent rows are the spline(K)-GARCH(1,1) models with $K \in \{0, 1, 4, 9\}$ and the same time series lengths.

4 Conclusion

The intention of the spline-GARCH model of Engle and Rangel (2008) was to relax the assumption of constant long-term variance, on which standard GARCH models are based. This is achieved by decomposing the variance into a short horizon component, the standard GARCH model and a long horizon component, the exponential spline function. One objective of the latter is to smooth the innovation process and so lowering the geometric decay rate of the VP. This may reduce the doubt on a spurious long-memory effect, but may not detect every break in the structure of the process, as stated by Engle et al. (2013). In this paper the Data Generating Processes (DGP) were estimated by real equity data and confirmed some of the results within the original paper. With these DGP, 60000 different time series were simulated and 60 different models were estimated to stress the finite-sample properties of the related symmetric and asymmetric GARCH parameters $(\alpha_1, \beta_1, \delta_1)$ when the unconditional variance is time-varying. The results for the parameters of the spline function were not discussed within this paper. They are available on request.
For the spline-GARCH model more parameters have to be estimated, than in standard GARCH models. Therefore the optimisation routine for the ML method is demanding and very sensitive for good starting values. In this paper, an ordinary-least-squares approach to receive appropriate starting values was introduced. The optimisation routine could easily end up in a local maximum or a saddle point far away from the global maximum, starting with bad values. For the simulation study, the estimated parameters of the DGP process have been used as starting values.
First of all, the volatility persistence decreases with an increasing number of knots. This behaviour could already be observed with the estimations of the DGP. For the replications, a similar picture arises, but with a tremendous lowering of the volatility persistence for the small samples. But these results have to be treated with caution, as there could occur problems with parameter estimators of the truncated spline function. The estimations get closer to the true values with longer time series. So for all large samples, the bias is in a reasonable range of 1%. The same holds for the zero knot cases, but here also for the small samples. On the whole, the bias of the single parameters shows the same behaviour. Also in line with asymptotic ML theory are the variances of the volatility persistences and the single parameters. Such a clear pattern is not apparent regarding the coverage probability. Here some irregularities rise to doubts on the assumption of asymptotic normality for some parameters in some model setups. This behaviour of the coverage probability could be explained by narrow confidence intervals due to small robust-standard-errors for some model setups. Unlike the coverage probability, the Kolmogorov-Smirnov-statistics are based on the sample statistics. For all large samples, the assumption of asymptotic normal distributed parameters following the Kolmogorov-Smirnov-statistic holds. It can be emphasised that the ML-estimators of the GARCH and GJR-GARCH parameters in the spline-GARCH model are consistent and asymptotically normally distributed, even though some GJR-GARCH replications had problematic parameter constellations. But empirical researchers should be suspicious using the spline-GARCH model for short horizons, even if there is a vast reduction of the VP apparent.
It should be stressed, that the applied model setups rely on only one arbitrarily chosen sample, the S&P500 equity index from 1980-1999. Some of the simulated time series are heavily compressed or stretched, as the original sample contains $T = 9835$ data points. Whether the events happened within the original sample period, are applicable to each and every simulation is not evident. But as a typical developed market long time series, with the S&P500 sample, some of the intentions of the spline-GARCH model could be examined. Based on

the results of this article, further studies should examine different parameter constellations. Another issue is the omitted restrictions on the estimation of the short-term volatility. There may be different results imposing parameter restriction, in particular in small samples and for the most of the considered GJR-GARCH models.

Acknowledgements

Parts of this paper were presented at the annual meeting of the German Statistical Society (DStatG) in Trier 2019. I would like to thank my supervisor Prof. Dr. Hermann Singer for constructive criticism and many insightful conversations.

References

Amado, C., Silvennoinen, A. and Teräsvirta, T. (2018), 'Models with multiplicative decomposition of conditional variances and correlations'.
URL: *https://econpapers.repec.org/RePEc:aah:create:2018-14*

Audrino, F. and Bühlmann, P. (2009), 'Splines for financial volatility', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **71**(3), 655–670.

Beran, J., Feng, Y. and Ghosh, S. (2013), *Long-Memory Processes: Probabilistic Properties and Statistical Methods*.
URL: *http://dx.doi.org/10.1007/978-3-642-35512-7*

Black, F. (1976), 'Studies of stock price volatility changes', *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, American Statistical Association* pp. 177–181.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**(3), 307–327.

Bollerslev, T. and Wooldridge, J. M. (1992), 'Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances', *Econometric Reviews* **11**(2), 143–172.

Brownlees, C. and Gallo, G. M. (2010), 'Comparison of volatility measures: A risk management perspective', *Journal of financial econometrics : official journal of the Society for Financial Econometrics* **8**(1), 29–56.

Cai, J. (1994), 'A markov model of switching-regime arch', *Journal of Business & Economic Statistics* **12**(3), 309–316.

Čižek, P. and Spokoiny, V. (2009), Varying coefficient garch models, *in* 'Handbook of Financial Time Series', Springer, pp. 169–185.

Conrad, C. and Kleen, O. (2018), 'Two are better than one: Volatility forecasting using multiplicative component garch models', *SSRN Electronic Journal* .

de Boor, C. (2001), *A practical guide to splines*, Vol. 27 of *Applied mathematical sciences*, rev. ed., 1. hardcover print edn, Springer, New York, NY.
URL: *http://www.loc.gov/catdir/enhancements/fy0817/2001049644-d.html*

Dennis, J. E. and Schnabel, R. B. (1983), *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice-Hall series in computational mathematics.

Diebold, F. X. (1986), 'Modeling the persistence of conditional variances: A comment', *Econometric Reviews* **5**(1), 51–56.

Ding, Z., Granger, C. W. and Engle, R. F. (1993), 'A long memory property of stock market returns and a new model', *Journal of Empirical Finance* **1**(1), 83–106.

Drost, F. C. and Nijman, T. E. (1993), 'Temporal aggregation of garch processes', *Econometrica* **61**(4), 909.

Eilers, P. H. C. and Marx, B. D. (1996), 'Flexible smoothing with b -splines and penalties', *Statistical Science* **11**(2), 89–121.

Engle, R. F. (1982), ‘Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation’, *Econometrica* **50**(4), 987–1007.

Engle, R. F. and Bollerslev, T. (1986), ‘Modelling the persistence of conditional variances’, *Econometric Reviews* **5**(1), 1–50.

Engle, R. F. and Lee, G. G. (1999), A long-run and short-run component model of stock return volatility, *in* C. Engle, H. W. Robert F. Engle, H. White, R. F. Engle, R. ENGLE and C. Granger, eds, ‘Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W.J. Granger’, Oxford University Press, pp. 475–497.

Engle, R. F. and Mezrich, J. (1996), ‘Garch for groups’, *RISK* **10**(9), 36–40.

Engle, R. F. and Rangel, J. G. (2008), ‘The spline-garch model for low-frequency volatility and its global macroeconomic causes’, *Review of Financial Studies* **21**(3), 1187–1222.

Engle, R., Ghysels, E. and Sohn, B. (2013), ‘Stock market volatility and macroeconomic fundamentals’, *Review of Economics and Statistics* **95**(3), 776–797.

Fama, E. F. (1970), ‘Efficient capital markets: A review of theory and empirical work’, *The Journal of Finance* **25**(2), 383.

Feng, Y. (2004), ‘Simultaneously modeling conditional heteroskedasticity and scale change’, *Econometric Theory* **20**(03), 173.

Fiorentini, G., Calzolari, G. and Panattoni, L. (1996), ‘Analytic derivatives and the computation of garch estimates’, *Journal of Applied Econometrics* **11**(4), 399–417.

Ghysels, E., Sinko, A. and Valkanov, R. (2007), ‘Midas regressions: Further results and new directions’, *Econometric Reviews* **26**(1), 53–90.

Glosten, L. R., Jagannathan, R. and Runkle, D. E. (1993), ‘On the relation between the expected value and the volatility of the nominal excess return on stocks’, *The Journal of Finance* **48**(5), 1779–1801.

Goldman, E. and Shen, X. (2017), ‘Analysis of asymmetric garch volatility models with applications to margin measurement’.

Goldman, E. and Wang, T. (2015), ‘The spline-threshold-garch volatility model and tail risk’.

González-Rivera, G. (1998), ‘Smooth-transition garch models’, *Studies in Nonlinear Dynamics and Econometrics* (3).

Hamilton, J.D., Susmel, R. (1994), ‘Autoregressiv conditional heteroskedasticity and changes in regime’, *Journal of Econometrics* (64), 307–333.

Hansen, P. R. and Lunde, A. (2001), ‘A comparison of volatility models: Does anything beat a garch(1,1)?’.
URL: *http://www.cls.dk/caf/wp/wp-84.pdf*

Hansen, P. R. and Lunde, A. (2005), ‘A forecast comparison of volatility models: does anything beat a garch(1,1)?’, *Journal of Applied Econometrics* **20**(7), 873–889.

He, C. and Teräsvirta, T. (1999), 'Properties of moments of a family of garch processes', *Journal of Econometrics* **92**(1), 173–192.

Hillebrand, E. (2005), 'Neglecting parameter changes in garch models', *Journal of Econometrics* **129**(1-2), 121–138.

Hillebrand, E. and Medeiros, M. C. (2008), Estimating and forecasting garch models in the presence of structural breaks and regime switches, *in* D. E. Rapach and M. E. Wohar, eds, 'Forecasting in the presence of structural breaks and model uncertainty', Vol. 3 of *Frontiers of Economics and Globalization*, Emerald, Bingley, pp. 303–327.

Hillebrand, E. T. (2004), 'Neglecting parameter changes in autoregressive models', *SSRN Electronic Journal* .

Lamoureux, C. G. and Lastrapes, W. D. (1990), 'Persistence in variance, structural change, and the garch model', *Journal of Business & Economic Statistics* **8**(2), 225.

Laurent, S. (2013), *Estimating and forecasting ARCH models using G@rch 7*, Timberlake Consultants, London and Union, NJ.

Lumsdaine, R. L. (1995), 'Finite-sample properties of the maximum likelihood estimator in garch(1,1) and igarch(1,1) models: A monte carlo investigation', *Journal of Business & Economic Statistics* **13**(1), 1.

Medeiros, M. C. and Veiga, A. (2009), 'Modeling multiple regimes in financial volatility with a flexible coefficient garch(1, 1) model', *Econometric Theory* **25**(1), 117–161.
URL: *http://www.jstor.org/stable/20532434*

Mercurio, D. and Spokoiny, V. (2004), 'Statistical inference for time-inhomogeneous volatility models', *The Annals of Statistics* **32**(2), 577–602.

Mikosch, T. and Starica, C. (2004), 'Nonstationarities in financial time series, the long-range dependence, and the igarch effects', *Review of Economics and Statistics* **86**(1), 378–390.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: A new approach', *Econometrica* **59**(2), 347–370.

Rangel, J. G. and Engle, R. F. (2012), 'The factor-spline-garch model for high and low frequency correlations', *Journal of business & economic statistics : JBES : a publication of the American Statistical Association* **30**(1), 109–124.

Rao, C. R. (2002), *Linear statistical inference and its applications*, Wiley series in probability and statistics, 2. ed., paperback ed. edn, Wiley, New York.

Silvennoinen, A. (2006), *Essays on autoregressive conditional heteroskedasticity*, Stockholm School of Economics, EFI, Economic Research Institute, Stockholm.

Silverman, B. W. (1998), *Density estimation for statistics and data analysis*, Monographs on statistics and applied probability, 1. crc repr edn.

Zakoian, J. M. (1994), 'Threshold heteroskedastic models', *Journal of Economic Dynamics and Control* **18**, 931–955.

A Tables

	GARCH(1,1)	Spline(1)	Spline(4)	Spline(9)	Spline(14)
$\tilde{\alpha}_1$	**0.0804** (0.0142)	**0.0856** (0.0006)	**0.0866** (0.0010)	0.0881 (0.263)	**0.0876** (0.0014)
$\tilde{\beta}_1$	**0.9036** (0.0183)	**0.8999** (0.0001)	**0.8977** (0.0041)	**0.8800** (0.13)	**0.869** (0.0077)
$\tilde{\alpha}_1 + \tilde{\beta}_1$	0.9840	0.9855	0.9843	0.9681	0.9566
$\tilde{c}'$		**0.4054** (0.0342)	−0.0401 (0.0796)	0.276 (3.89)	**−0.109** (0.0326)
$\tilde{w}_0$		**−0.6443** (0.0854)	**7.31** (0.701)	**−4.08** (0.364)	**19.5** (0.0682)
$\tilde{w}_1$		**0.2259** (0.0876)	**−23.7** (0.221)	**15.4** (1.98)	**−248.0** (0.0905)
$\tilde{w}_2$			**37.4** (2.47)	**16.4** (1.69)	**440.0** (0.0501)
$\tilde{w}_3$			**−20.7** (7.16)	**−153.0** (1.31)	**−257.0** (0.116)
$\tilde{w}_4$			**5.21** (1.83)	**346.0** (0.204)	**−35.7** (0.0893)
$\tilde{w}_5$				**−455.0** (0.66)	**100.0** (1.19)
$\tilde{w}_6$				**369.0** (0.839)	**282.0** (1.49)
$\tilde{w}_7$				**−185.0** (0.218)	**−458.0** (0.604)
$\tilde{w}_8$				**7.98** (0.336)	**84.5** (2.01)
$\tilde{w}_9$				**149.0** (6.36)	**52.2** (0.201)
$\tilde{w}_{10}$					**409.0** (1.57)
$\tilde{w}_{11}$					**−851.0** (1.19)
$\tilde{w}_{12}$					**780.0** (17.3)
$\tilde{w}_{13}$					**−497.0** (55.2)
$\tilde{w}_{14}$					**599.0** (58.9)
BIC	2.6681	2.6700	2.6716	**2.6675**	2.6684

Table 4: GARCH(1,1): Estimated Parameter Values from S&P500 Sample and DGP

	GJR-GARCH(1,1)	Spline(1)	Spline(4)	Spline(9)	Spline(14)
$\tilde{\alpha}_1$	**0.0203** (0.0008)	**0.0155** (0.0019)	**0.015** (0.00216)	**0.0101** (0.00241)	**0.00696** (0.00341)
$\tilde{\beta}_1$	**0.9** (0.00336)	**0.894** (0.0002)	**0.892** (0.0003)	**0.878** (0.0009)	**0.867** (0.0004)
$\tilde{\delta}_1$	**0.116** (0.00106)	**0.137** (0.0005)	**0.14** (0.00343)	**0.149** (0.00131)	**0.155** (0.00521)
$\tilde{\alpha}_1 + \tilde{\beta}_1 + \frac{1}{2}\tilde{\delta}_1$	0.9783	0.978	0.977	0.9626	0.9515
$\tilde{c}'$		**0.521** (0.00757)	**0.18** (0.00937)	**0.357** (0.0103)	−0.00339 (0.00879)
$\tilde{w}_0$		**−0.924** (0.0584)	**5.29** (0.271)	**−3.26** (0.0142)	**19.4** (0.0365)
$\tilde{w}_1$		**0.197** (0.0645)	**−18.8** (0.109)	**16.4** (0.101)	**−229.0** (0.0505)
$\tilde{w}_2$			**30.0** (0.783)	**−1.95** (0.0217)	**386.0** (0.0538)
$\tilde{w}_3$			**−16.1** (2.25)	**−108.0** (0.12)	**−214.0** (0.0288)
$\tilde{w}_4$			0.0683 (0.147)	**284.0** (0.115)	**−25.0** (0.0242)
$\tilde{w}_5$				**−391.0** (0.252)	**71.3** (0.0285)
$\tilde{w}_6$				**322.0** (0.519)	**285.0** (0.3)
$\tilde{w}_7$				**−162.0** (0.0688)	**−481.0** (0.211)
$\tilde{w}_8$				**12.1** (0.471)	**190.0** (0.515)
$\tilde{w}_9$				**97.6** (3.74)	**−80.4** (1.45)
$\tilde{w}_{10}$					**472.0** (1.34)
$\tilde{w}_{11}$					**−810.0** (4.24)
$\tilde{w}_{12}$					**664.0** (18.1)
$\tilde{w}_{13}$					**−331.0** (60.0)
$\tilde{w}_{14}$					**285.0** (45.1)
BIC	2.6544	2.6418	2.6429	**2.6370**	2.6375

Table 5: GJR-GARCH(1,1): Estimated Parameter Values from S&P500 Sample and DGP

	GARCH(1,1)					GJR-GARCH (1,1)				
	$T = 100$					$T = 100$				
knots	$\bar{\hat{\eta}}_{S,1}$	bias	$\hat{s}(\hat{\eta}_{S,1})$	RMSE	M_{nT}	$\bar{\hat{\eta}}_{A,1}$	bias	$\hat{s}(\hat{\eta}_{A,1})$	RMSE	M_{nT}
-	0.9081	0.0759	0.2073	0.2208	973	0.9314	0.0468	0.1682	0.1751	940
1	0.7242	0.2613	0.5190	0.5810	910	0.8511	0.1199	0.3279	0.3491	886
4	0.6099	0.3744	0.6757	0.7725	872	0.8155	0.1486	0.4346	0.4588	823
9	0.5199	0.4482	0.7286	0.8554	841	0.7954	0.1539	0.4731	0.4975	780
14	0.5371	0.4195	0.7153	0.8292	804	0.7836	0.1590	0.4482	0.4756	789
	$T = 500$					$T = 500$				
knots	$\bar{\hat{\eta}}_{S,1}$	bias	$\hat{s}(\hat{\eta}_{S,1})$	RMSE	M_{nT}	$\bar{\hat{\eta}}_{A,1}$	bias	$\hat{s}(\hat{\eta}_{A,1})$	RMSE	M_{nT}
-	0.9762	0.0077	0.0211	0.0225	950	0.9732	0.0064	0.0213	0.0222	948
1	0.9352	0.0502	0.0853	0.0990	944	0.9560	0.0260	0.0572	0.0628	937
4	0.7924	0.1919	0.3631	0.4107	898	0.9266	0.0534	0.1063	0.1190	898
9	0.4190	0.5492	0.6042	0.8165	895	0.8705	0.0865	0.2519	0.2663	762
14	0.2213	0.7354	0.6959	1.0125	660	0.9031	0.0370	0.1926	0.1961	648
	$T = 1000$					$T = 1000$				
knots	$\bar{\hat{\eta}}_{S,1}$	bias	$\hat{s}(\hat{\eta}_{S,1})$	RMSE	M_{nT}	$\bar{\hat{\eta}}_{A,1}$	bias	$\hat{s}(\hat{\eta}_{A,1})$	RMSE	M_{nT}
-	0.9800	0.0040	0.0105	0.0112	937	0.9752	0.0050	0.0098	0.0110	945
1	0.9680	0.0175	0.0272	0.0323	909	0.9692	0.0118	0.0153	0.0193	940
4	0.9439	0.0404	0.0413	0.0578	865	0.9610	0.0196	0.0195	0.0277	910
9	0.8495	0.1185	0.1449	0.1872	927	0.9267	0.0395	0.0559	0.0684	911
14	0.6780	0.2786	0.3730	0.4656	595	0.9168	0.0352	0.0788	0.0863	552
	$T = 5000$					$T = 5000$				
knots	$\bar{\hat{\eta}}_{S,1}$	bias	$\hat{s}(\hat{\eta}_{S,1})$	RMSE	M_{nT}	$\bar{\hat{\eta}}_{A,1}$	bias	$\hat{s}(\hat{\eta}_{A,1})$	RMSE	M_{nT}
-	0.9834	0.0005	0.0036	0.0037	977	0.9778	0.0005	0.0036	0.0036	985
1	0.9829	0.0027	0.0046	0.0054	983	0.9769	0.0019	0.0046	0.0049	984
4	0.9797	0.0046	0.0053	0.0070	975	0.9746	0.0034	0.0051	0.0061	989
9	0.9580	0.0100	0.0090	0.0135	974	0.9581	0.0069	0.0072	0.0100	983
14	0.9430	0.0135	0.0122	0.0182	938	0.9454	0.0082	0.0093	0.0123	888
	$T = 10000$					$T = 10000$				
knots	$\bar{\hat{\eta}}_{S,1}$	bias	$\hat{s}(\hat{\eta}_{S,1})$	RMSE	M_{nT}	$\bar{\hat{\eta}}_{A,1}$	bias	$\hat{s}(\hat{\eta}_{A,1})$	RMSE	M_{nT}
-	0.9837	0.0002	0.0026	0.0026	992	0.9781	0.0003	0.0024	0.0315	990
1	0.9842	0.0013	0.0032	0.0035	996	0.9776	0.0012	0.0035	0.0325	991
4	0.9820	0.0023	0.0036	0.0043	990	0.9756	0.0019	0.0038	0.0330	993
9	0.9634	0.0046	0.0055	0.0072	983	0.9607	0.0028	0.0053	0.0338	985
14	0.9513	0.0053	0.0074	0.0091	978	0.9493	0.0033	0.0055	0.0064	971
	$T = 25000$					$T = 25000$				
knots	$\bar{\hat{\eta}}_{S,1}$	bias	$\hat{s}(\hat{\eta}_{S,1})$	RMSE	M_{nT}	$\bar{\hat{\eta}}_{A,1}$	bias	$\hat{s}(\hat{\eta}_{A,1})$	RMSE	M_{nT}
-	0.9839	0.0001	0.0016	0.0016	995	0.9782	0.0000	0.0016	0.0016	1000
1	0.9851	0.0003	0.0019	0.0019	999	0.9782	0.0006	0.0020	0.0021	996
4	0.9834	0.0008	0.0020	0.0022	1000	0.9766	0.0006	0.0022	0.0023	997
9	0.9665	0.0016	0.0035	0.0038	995	0.9624	0.0009	0.0028	0.0030	994
14	0.9549	0.0017	0.0043	0.0046	990	0.9511	0.0011	0.0033	0.0035	996

Table 6: Volatility Persistence. The sample statistics are computed as given in equations (32)- (34). The red highlighted numbers indicate, that some GJR-GARCH replications and related ML estimators result to some negative h_t.

	GARCH(1,1)					GJR-GARCH (1,1)				
	$T = 100$					$T = 100$				
knots	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}
-	0.0528	0.0276	0.1210	0.1241	973	-0.0269	0.0472	0.1222	0.1310	940
1	-0.0296	0.1153	0.1236	0.1690	910	-0.1009	0.1164	0.1405	0.1824	886
4	-0.0578	0.1444	0.1206	0.1881	872	-0.1209	0.1359	0.1314	0.1891	823
9	-0.0701	0.1581	0.1480	0.2166	841	-0.1439	0.1541	0.1321	0.2030	780
14	-0.0697	0.1573	0.1459	0.2145	804	-0.1421	0.1491	0.1411	0.2053	789
	$T = 500$					$T = 500$				
knots	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}
-	0.0802	0.0001	0.0276	0.0276	950	0.0162	0.0041	0.0304	0.0306	948
1	0.0812	0.0044	0.0332	0.0335	944	-0.0027	0.0182	0.0377	0.0419	937
4	0.0649	0.0217	0.0454	0.0503	898	-0.0210	0.0360	0.0488	0.0607	898
9	0.0393	0.0488	0.0628	0.0796	895	-0.0611	0.0712	0.0508	0.0875	762
14	0.0217	0.0659	0.0651	0.0926	660	-0.0811	0.0880	0.0362	0.0952	648
	$T = 1000$					$T = 1000$				
knots	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}
-	0.0811	0.0008	0.0174	0.0174	937	0.0175	0.0028	0.0190	0.0192	945
1	0.0839	0.0017	0.0195	0.0196	909	0.0083	0.0072	0.0207	0.0219	940
4	0.0805	0.0061	0.0209	0.0218	865	0.0012	0.0138	0.0220	0.0259	910
9	0.0752	0.0129	0.0274	0.0302	927	-0.0187	0.0289	0.0288	0.0408	911
14	0.0649	0.0227	0.0357	0.0423	595	-0.0381	0.0450	0.0312	0.0548	552
	$T = 5000$					$T = 5000$				
knots	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}
-	0.0803	0.0001	0.0073	0.0073	977	0.0197	0.0006	0.0077	0.0077	985
1	0.0849	0.0008	0.0080	0.0080	983	0.0142	0.0013	0.0080	0.0081	984
4	0.0860	0.0006	0.0083	0.0083	975	0.0125	0.0025	0.0085	0.0088	989
9	0.0868	0.0012	0.0094	0.0094	974	0.0057	0.0044	0.0084	0.0095	983
14	0.0864	0.0012	0.0096	0.0097	938	0.0022	0.0048	0.0089	0.0101	888
	$T = 10000$					$T = 10000$				
knots	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}
-	0.0802	0.0002	0.0054	0.0054	992	0.0000	0.0003	0.0055	0.0055	990
1	0.0851	0.0005	0.0059	0.0059	996	0.0006	0.0006	0.0055	0.0055	991
4	0.0866	0.0000	0.0059	0.0059	990	0.0012	0.0007	0.0057	0.0059	993
9	0.0882	0.0001	0.0062	0.0062	983	0.0018	0.0010	0.0059	0.0062	985
14	0.0871	0.0004	0.0067	0.0067	978	0.0048	0.0021	0.0062	0.0066	971
	$T = 25000$					$T = 25000$				
knots	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}	$\bar{\hat{\alpha}}_1$	bias	$\hat{s}(\hat{\alpha}_1)$	RMSE	M_{nT}
-	0.0803	0.0001	0.0035	0.0035	995	0.0201	0.0002	0.0034	0.0034	1000
1	0.0856	0.0000	0.0037	0.0037	999	0.0152	0.0004	0.0035	0.0035	996
4	0.0864	0.0001	0.0037	0.0037	1000	0.0145	0.0005	0.0036	0.0037	997
9	0.0880	0.0001	0.0042	0.0042	995	0.0095	0.0007	0.0037	0.0037	994
14	0.0874	0.0002	0.0043	0.0043	990	0.0062	0.0007	0.0038	0.0038	996

Table 7: α_1. The sample statistics are computed as given in equations (28)- (31). The red highlighted numbers indicate, that some GJR-GARCH replications and related ML estimators result to some negative h_t.

	GARCH(1,1)					**GJR-GARCH (1,1)**				
	$T = 100$					$T = 100$				
knots	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}
-	0.8553	0.0483	0.2608	0.2653	973	0.8983	0.0015	0.2306	0.2306	940
1	0.7538	0.1461	0.5807	0.5988	910	0.8908	0.0035	0.3951	0.3951	886
4	0.6677	0.2300	0.7451	0.7798	872	0.8810	0.0109	0.4917	0.4919	823
9	0.5900	0.2900	0.8174	0.8673	841	0.8786	0.0002	0.5440	0.5440	780
14	0.6068	0.2622	0.8047	0.8464	804	0.8575	0.0099	0.5207	0.5208	789
	$T = 500$					$T = 500$				
knots	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}
-	0.8960	0.0076	0.0422	0.0428	950	0.8975	0.0023	0.0425	0.0425	948
1	0.8540	0.0458	0.0970	0.1073	944	0.8864	0.0078	0.0797	0.0801	937
4	0.7275	0.1703	0.3647	0.4025	898	0.8745	0.0174	0.1346	0.1357	898
9	0.3797	0.5003	0.6058	0.7857	895	0.8631	0.0153	0.2799	0.2804	762
14	0.1996	0.6695	0.7088	0.9750	660	0.9185	0.0511	0.2121	0.2182	648
	$T = 1000$					$T = 1000$				
knots	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}
-	0.8988	0.0047	0.0240	0.0245	937	0.8976	0.0022	0.0244	0.0245	945
1	0.8841	0.0158	0.0351	0.0385	909	0.8896	0.0046	0.0269	0.0273	940
4	0.8634	0.0343	0.0474	0.0585	865	0.8861	0.0058	0.0336	0.0341	910
9	0.7743	0.1057	0.1505	0.1839	927	0.8678	0.0106	0.0744	0.0752	911
14	0.6131	0.2560	0.3694	0.4494	595	0.8772	0.0098	0.1067	0.1071	552
	$T = 5000$					$T = 5000$				
knots	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}
-	0.9032	0.0004	0.0097	0.0098	977	0.8998	0.0000	0.0095	0.0095	985
1	0.8980	0.0019	0.0095	0.0097	983	0.8937	0.0006	0.0092	0.0093	984
4	0.8937	0.0040	0.0102	0.0110	975	0.8910	0.0009	0.0103	0.0103	989
9	0.8712	0.0088	0.0139	0.0165	974	0.8760	0.0024	0.0123	0.0125	983
14	0.8566	0.0124	0.0170	0.0210	938	0.8640	0.0034	0.0146	0.0150	888
	$T = 10000$					$T = 10000$				
knots	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}
-	0.9035	0.0001	0.0070	0.0070	992	0.8995	0.0003	0.0067	0.0067	990
1	0.8991	0.0008	0.0069	0.0070	996	0.8937	0.0006	0.0065	0.0065	991
4	0.8954	0.0023	0.0073	0.0077	990	0.8913	0.0007	0.0069	0.0069	993
9	0.8752	0.0047	0.0089	0.0101	983	0.8774	0.0010	0.0082	0.0082	985
14	0.8642	0.0048	0.0111	0.0121	978	0.8662	0.0012	0.0094	0.0095	971
	$T = 25000$					$T = 25000$				
knots	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}	$\bar{\hat{\beta}}_1$	bias	$\hat{s}(\hat{\beta}_1)$	RMSE	M_{nT}
-	0.9036	0.0000	0.0045	0.0045	995	0.8999	0.0001	0.0043	0.0043	1000
1	0.8995	0.0003	0.0042	0.0042	999	0.8940	0.0003	0.0044	0.0044	996
4	0.8970	0.0007	0.0043	0.0043	1000	0.8918	0.0001	0.0044	0.0044	997
9	0.8785	0.0015	0.0060	0.0062	995	0.8782	0.0003	0.0053	0.0054	994
14	0.8675	0.0015	0.0069	0.0071	990	0.8671	0.0003	0.0058	0.0059	996

Table 8: β_1. The sample statistics are computed as given in equations (28)- (31). The red highlighted numbers indicate, that some GJR-GARCH replications and related ML estimators result to some negative h_t.

GJR-GARCH (1,1)

knots	$\bar{\hat{\delta}}_1$	bias	$\hat{s}(\hat{\delta}_1)$	RMSE	M_{nT}
		$T = 100$			
-	0.1163	0.0037	0.1738	0.1739	940
1	0.1224	0.0151	0.1979	0.1985	886
4	0.1109	0.0293	0.1924	0.1947	823
9	0.1214	0.0276	0.1820	0.1841	780
14	0.1365	0.0187	0.1986	0.1994	789
		$T = 500$			
-	0.1163	0.0026	0.0452	0.0453	948
1	0.1446	0.0072	0.0514	0.0519	937
4	0.1463	0.0061	0.0630	0.0633	898
9	0.1369	0.0122	0.0689	0.0699	762
14	0.1314	0.0237	0.0710	0.0749	648
		$T = 1000$			
-	0.1163	0.0039	0.0306	0.0308	945
1	0.1425	0.0051	0.0311	0.0315	940
4	0.1474	0.0072	0.0341	0.0348	910
9	0.1552	0.0061	0.0416	0.0420	911
14	0.1552	0.0001	0.0481	0.0481	552
		$T = 5000$			
-	0.1163	0.0002	0.0118	0.0118	985
1	0.1380	0.0006	0.0136	0.0136	984
4	0.1423	0.0021	0.0142	0.0144	989
9	0.1529	0.0038	0.0155	0.0160	983
14	0.1584	0.0033	0.0158	0.0161	888
		$T = 10000$			
-	0.1163	0.0004	0.0091	0.0091	990
1	0.1380	0.0006	0.0096	0.0096	991
4	0.1410	0.0008	0.0099	0.0099	993
9	0.1500	0.0009	0.0106	0.0106	985
14	0.1565	0.0014	0.0112	0.0619	971
		$T = 25000$			
-	0.1163	0.0000	0.0053	0.0053	1000
1	0.1380	0.0006	0.0060	0.0060	996
4	0.1407	0.0004	0.0060	0.0060	997
9	0.1495	0.0004	0.0067	0.0067	994
14	0.1555	0.0004	0.0071	0.0071	996

Table 9: δ_1. The sample statistics are computed as given in equations (28)- (31). The red highlighted numbers indicate, that some GJR-GARCH replications and related ML estimators result to some negative h_t.

	GARCH(1,1)			GJR-GARCH (1,1)			
	$T = 100$			$T = 100$			
knots	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	M_{nT}	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	$\tilde{\delta}_1$	M_{nT}
-	0.5827 (0.0158)	0.5786 (0.0158)	973	0.2436 (0.0140)	0.3798 (0.0158)	0.2660 (0.0144)	940
1	0.2374 (0.0141)	0.1319 (0.0112)	910	0.1682 (0.0126)	0.1354 (0.0115)	0.1479 (0.0119)	886
4	0.1170 (0.0109)	0.0229 (0.0051)	872	0.0486 (0.0075)	0.0231 (0.0052)	0.0437 (0.0071)	823
9	0.0345 (0.0063)	0.0024 (0.0017)	841	0.0397 (0.0070)	0.0359 (0.0067)	0.0333 (0.0064)	780
14	0.0162 (0.0044)	0.0037 (0.0022)	804	0.0748 (0.0094)	0.0811 (0.0097)	0.0875 (0.0101)	789
	$T = 500$			$T = 500$			
knots	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	M_{nT}	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	$\tilde{\delta}_1$	M_{nT}
-	0.9179 (0.0089)	0.9284 (0.0084)	950	0.2605 (0.0143)	0.5359 (0.0162)	0.2848 (0.0147)	948
1	0.5657 (0.0161)	0.3739 (0.0157)	944	0.4877 (0.0163)	0.2497 (0.0141)	0.2615 (0.0144)	937
4	0.7617 (0.0142)	0.7494 (0.0145)	898	0.4766 (0.0167)	0.2973 (0.0153)	0.3808 (0.0162)	898
9	0.6335 (0.0161)	0.4235 (0.0165)	895	0.4344 (0.0180)	0.4121 (0.0178)	0.4475 (0.0180)	762
14	0.4576 (0.0194)	0.2561 (0.0170)	660	0.3642 (0.0189)	0.3596 (0.0189)	0.3719 (0.0190)	648
	$T = 1000$			$T = 1000$			
knots	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	M_{nT}	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	$\tilde{\delta}_1$	M_{nT}
-	0.9136 (0.0092)	0.9157 (0.0091)	937	0.2148 (0.0134)	0.5365 (0.0162)	0.2042 (0.0131)	945
1	0.3597 (0.0159)	0.1694 (0.0124)	909	0.4585 (0.0163)	0.1053 (0.0100)	0.1372 (0.0112)	940
4	0.7792 (0.0141)	0.6913 (0.0157)	865	0.5143 (0.0166)	0.1253 (0.0110)	0.2813 (0.0149)	910
9	0.8188 (0.0127)	0.8328 (0.0123)	927	0.7859 (0.0136)	0.6279 (0.0160)	0.7717 (0.0139)	911
14	0.7412 (0.0180)	0.7462 (0.0178)	595	0.8487 (0.0157)	0.8180 (0.0169)	0.8352 (0.0162)	552
	$T = 5000$			$T = 5000$			
knots	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	M_{nT}	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	$\tilde{\delta}_1$	M_{nT}
-	0.9529 (0.0068)	0.9478 (0.0071)	977	0.2041 (0.0128)	0.5411 (0.0159)	0.1736 (0.0121)	985
1	0.2909 (0.0145)	0.0773 (0.0085)	983	0.5386 (0.0159)	0.0996 (0.0095)	0.6189 (0.0155)	984
4	0.3908 (0.0156)	0.3467 (0.0152)	975	0.5683 (0.0158)	0.0758 (0.0084)	0.5319 (0.0159)	989
9	0.8142 (0.0125)	0.6858 (0.0149)	974	0.7599 (0.0136)	0.1801 (0.0123)	0.6989 (0.0146)	983
14	0.6151 (0.0159)	0.7708 (0.0137)	938	0.9696 (0.0058)	0.6182 (0.0163)	0.9414 (0.0079)	888
	$T = 10000$			$T = 10000$			
knots	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	M_{nT}	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	$\tilde{\delta}_1$	M_{nT}
-	0.9395 (0.0076)	0.9466 (0.0071)	992	0.1051 (0.0097)	0.4838 (0.0159)	0.0828 (0.0088)	990
1	0.2189 (0.0131)	0.0432 (0.0064)	996	0.5510 (0.0158)	0.0757 (0.0084)	0.5923 (0.0156)	991
4	0.2929 (0.0145)	0.2566 (0.0139)	990	0.5730 (0.0157)	0.0896 (0.0091)	0.5378 (0.0158)	993
9	0.8769 (0.0105)	0.6745 (0.0149)	983	0.7888 (0.0130)	0.2030 (0.0128)	0.7208 (0.0143)	985
14	0.4693 (0.0160)	0.7566 (0.0137)	978	0.7374 (0.0141)	0.3069 (0.0148)	0.6952 (0.0148)	971
	$T = 25000$			$T = 25000$			
knots	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	M_{nT}	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	$\tilde{\delta}_1$	M_{nT}
-	0.9367 (0.0077)	0.9317 (0.0080)	995	0.1630 (0.0117)	0.5260 (0.0158)	0.1750 (0.0120)	1000
1	0.1261 (0.0105)	0 (0)	999	0.5241 (0.0158)	0.0813 (0.0087)	0.5552 (0.0157)	996
4	0 (0)	0.1850 (0.0123)	1000	0.5938 (0.0156)	0.0772 (0.0085)	0.5707 (0.0157)	997
9	0.7347 (0.0140)	0.3920 (0.0155)	995	0.6911 (0.0147)	0.0543 (0.0072)	0.5654 (0.0157)	994
14	0.3828 (0.0154)	0.7384 (0.0140)	990	0.9970 (0.0017)	0.5482 (0.0158)	0.9528 (0.0067)	996

Table 10: Coverage Probability. Coverage probability and the bernoulli standard errors (in parentheses) are computed as given in equations (37)- (39). The red highlighted numbers indicate, that some GJR-GARCH replications and related ML estimators result to some negative h_t.

knots	$T = 100$	$T = 500$	$T = 1000$	$T = 5000$	$T = 10000$	$T = 25000$
			$\hat{\alpha}_{nT}$			
0	0.001	0.004	**0.545**	**0.346**	**0.485**	**0.924**
1	0.000	0.025	**0.336**	**0.887**	**0.406**	**0.356**
4	0.000	0.000	**0.414**	**0.287**	**0.503**	**0.439**
9	0.000	0.007	**0.213**	**0.752**	**0.928**	**0.833**
14	0.000	0.002	**0.758**	**0.480**	**0.554**	**0.775**
			$\hat{\beta}_{nT}$			
0	0.000	0.000	**0.228**	**0.557**	**0.949**	**0.356**
1	0.000	0.000	0.000	**0.920**	**0.840**	**0.892**
4	0.000	0.000	0.000	**0.174**	**0.295**	**0.998**
9	0.000	0.000	0.000	**0.171**	**0.766**	**0.403**
14	0.000	0.000	0.000	**0.117**	**0.168**	**0.763**

Table 11: Kolmogorov-Smirnov-Test for GARCH (1,1). p-Values of two-tailed Kolmogorov-Smirnov-Test with significance level $\alpha = 0.05$

$H_0 : F_0(\hat{\theta}) = F_{M_{nT}}(\hat{\theta}) \quad H_1 : F_0(\hat{\theta}) \neq F_{M_{nT}}(\hat{\theta})$

knots	$T = 100$	$T = 500$	$T = 1000$	$T = 5000$	$T = 10000$	$T = 25000$
			$\hat{\alpha}_{nT}$			
0	0.000	0.001	**0.410**	**0.318**	**0.917**	**0.398**
1	0.000	**0.355**	**0.779**	**0.871**	**0.414**	**0.780**
4	0.000	**0.459**	**0.195**	**0.797**	**0.950**	**0.828**
9	0.000	0.000	**0.789**	**0.636**	**0.635**	**0.845**
14	0.000	0.000	0.021	**0.609**	**0.525**	**0.957**
			$\hat{\beta}_{nT}$			
0	0.000	0.000	**0.131**	**0.195**	**0.805**	**0.832**
1	0.000	0.000	**0.177**	**0.345**	**0.566**	**0.843**
4	0.000	0.000	0.009	**0.625**	**0.839**	**0.885**
9	0.000	0.000	0.000	**0.436**	**0.952**	**0.974**
14	0.000	0.000	0.000	**0.378**	**0.747**	**0.620**
			$\hat{\delta}_{nT}$			
0	0.000	0.022	**0.108**	**0.551**	**0.119**	**0.693**
1	0.000	**0.193**	**0.928**	**0.901**	**0.803**	**0.907**
4	0.000	**0.561**	**0.612**	**0.651**	**0.940**	**0.955**
9	0.000	0.002	**0.956**	**0.841**	**0.954**	**0.952**
14	0.000	0.000	**0.276**	**0.546**	**0.634**	**0.460**

Table 12: Kolmogorov-Smirnov-Test for GJR-GARCH (1,1). p-Values of two-tailed Kolmogorov-Smirnov-Test with significance level $\alpha = 0.05$

$H_0 : F_0(\hat{\theta}) = F_{M_{nT}}(\hat{\theta}) \quad H_1 : F_0(\hat{\theta}) \neq F_{M_{nT}}(\hat{\theta})$

The red highlighted numbers indicate, that some GJR-GARCH replications and related ML estimators result to some negative h_t.

B Figures

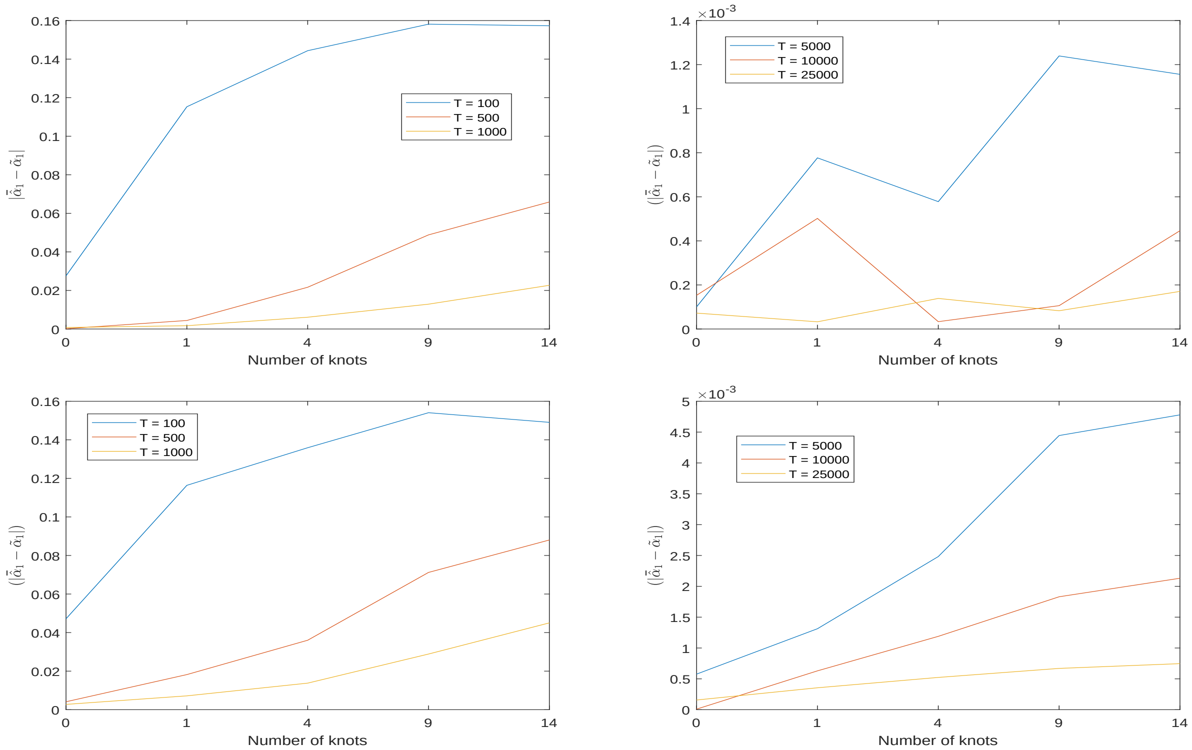

Figure 16: Bias $|\bar{\tilde{\alpha}}_1 - \tilde{\alpha}_1|$.Garch(1,1)(top row) and GJR-Garch(1,1) (bottom row). Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). Some GJR-GARCH replications and related ML estimators result to some negative h_t.

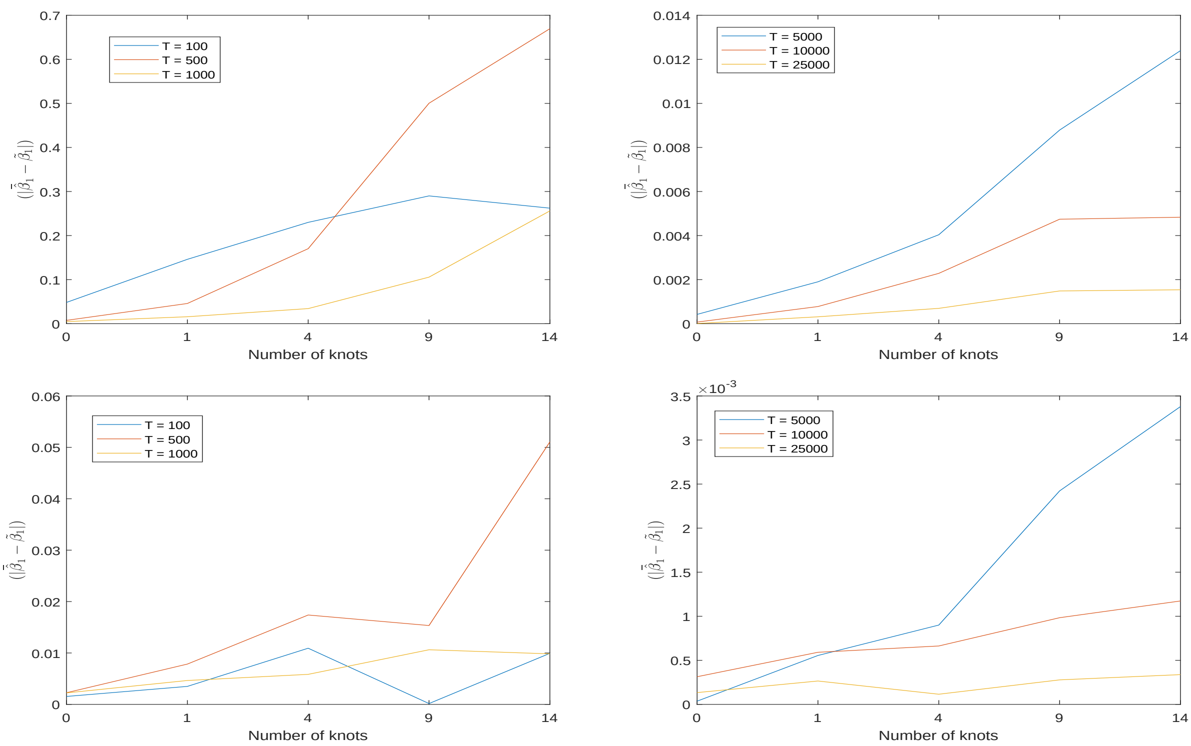

Figure 17: Bias $|\bar{\hat{\beta}}_1 - \tilde{\beta}_1|$. Garch(1,1)(top row) and GJR-Garch(1,1) (bottom row). Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). Some GJR-GARCH replications and related ML estimators result to some negative h_t.

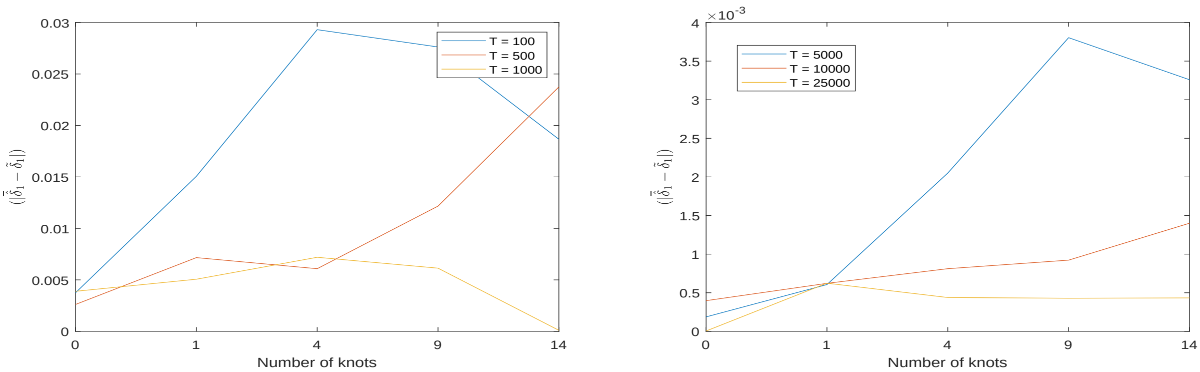

Figure 18: Bias $|\bar{\hat{\delta}}_1 - \tilde{\delta}_1|$. GJR-Garch(1,1). Small samples $T \in \{100, 500, 1000\}$ (left-hand side), large samples $T \in \{5000, 10000, 25000\}$(right-hand side). Some GJR-GARCH replications and related ML estimators result to some negative h_t.

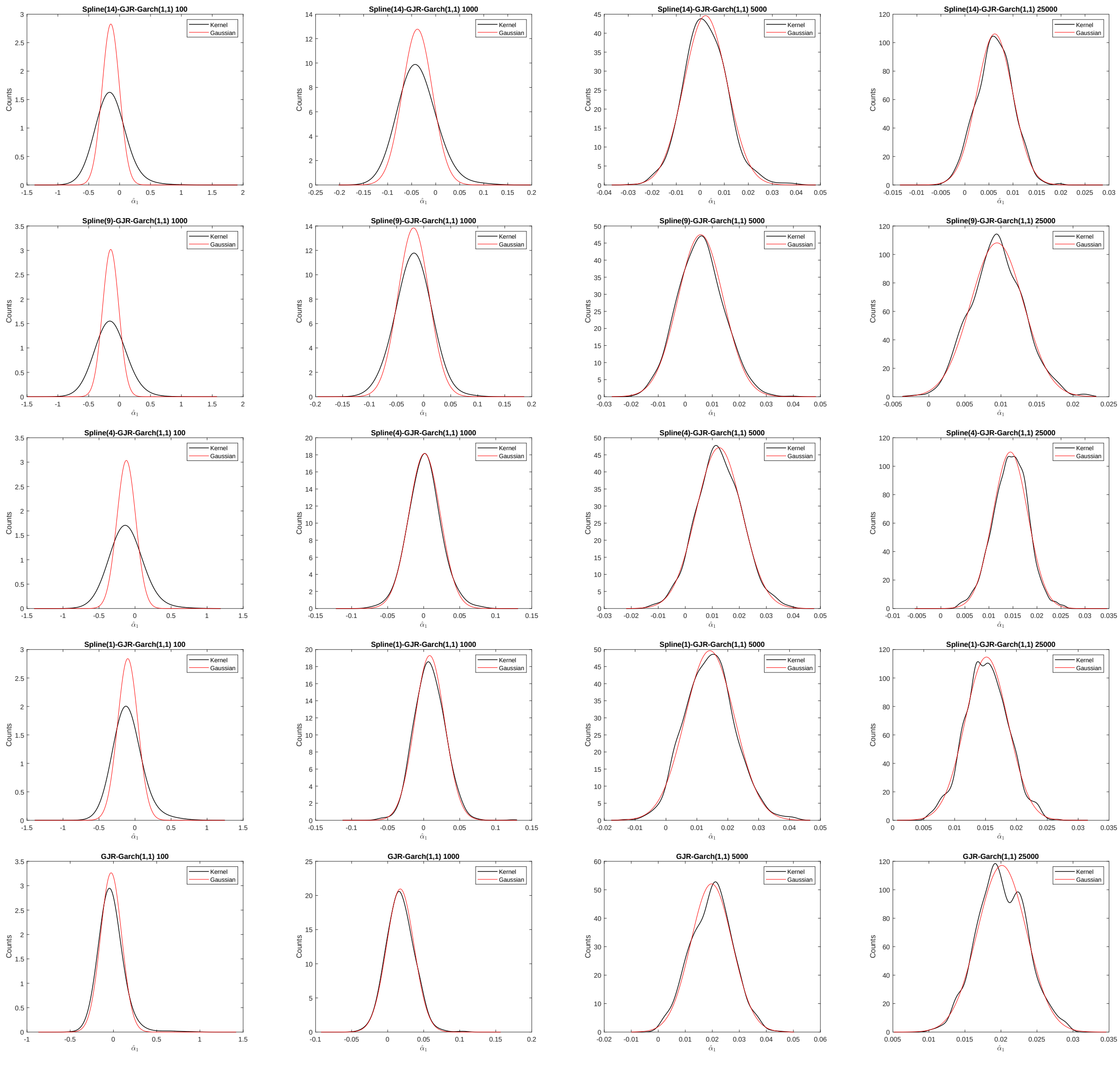

Figure 19: Asymptotic Normality of $\hat{\alpha}_1$ in GJR-GARCH(1,1) case. The black curves represent the kernel density estimator for $\hat{\alpha}$, the red curve the related normal distribution. In the top row the spline(14)-GARCH(1,1) distribution for $T \in \{100, 1000, 5000, 25000\}$ is depicted. The subjacent rows are the spline(K)-GARCH(1,1) models with $K \in \{0, 1, 4, 9\}$ and the same time series lengths.

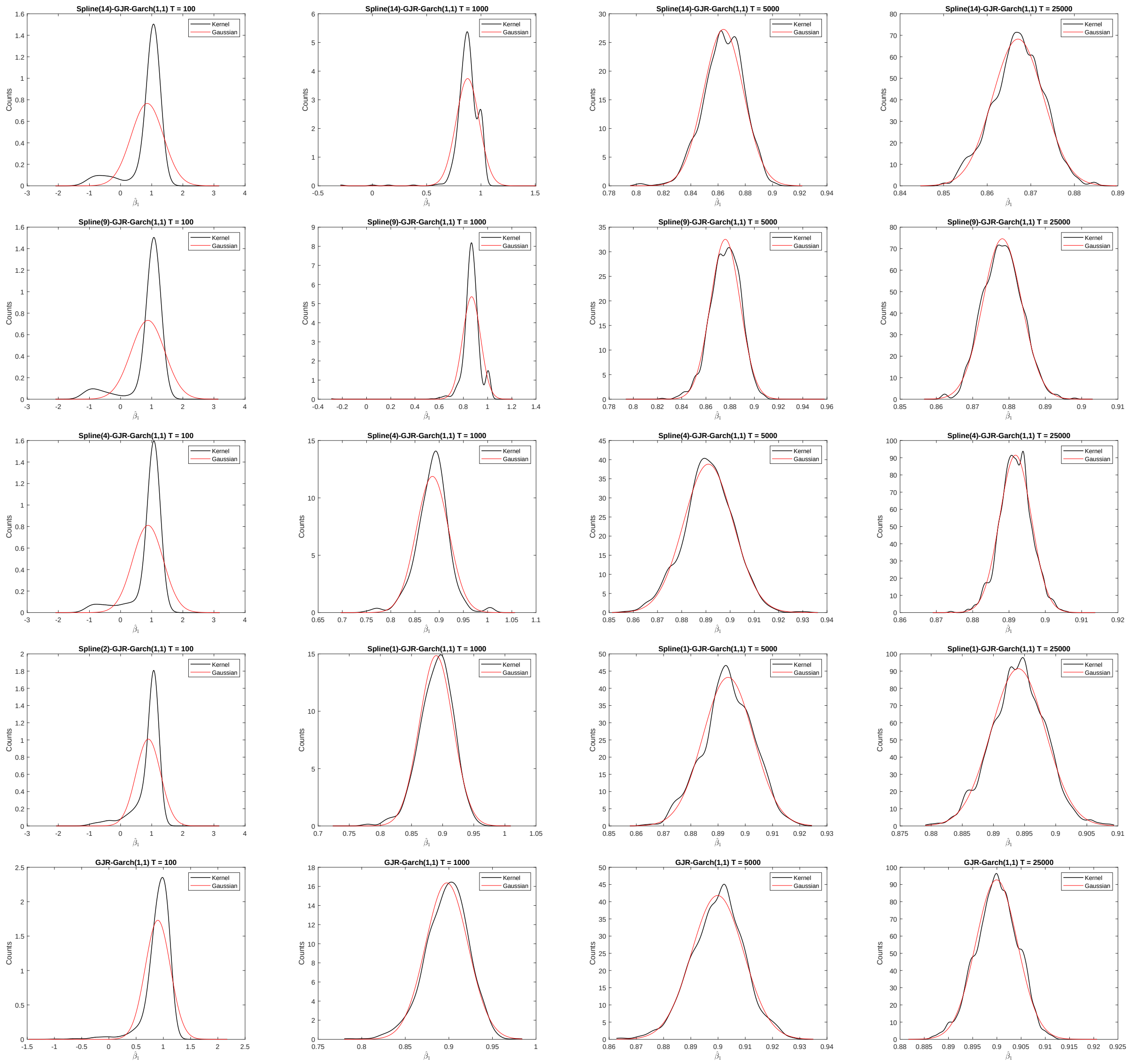

Figure 20: Asymptotic Normality of $\hat{\beta}_1$ in GJR-GARCH(1,1) case. The black curves represent the kernel density estimator for $\hat{\beta}$, the red curve the related normal distribution. In the top row the spline(14)-GARCH(1,1) distribution for $T \in \{100, 1000, 5000, 25000\}$ is depicted. The subjacent rows are the spline(K)-GARCH(1,1) models with $K \in \{0, 1, 4, 9\}$ and the same time series lengths.

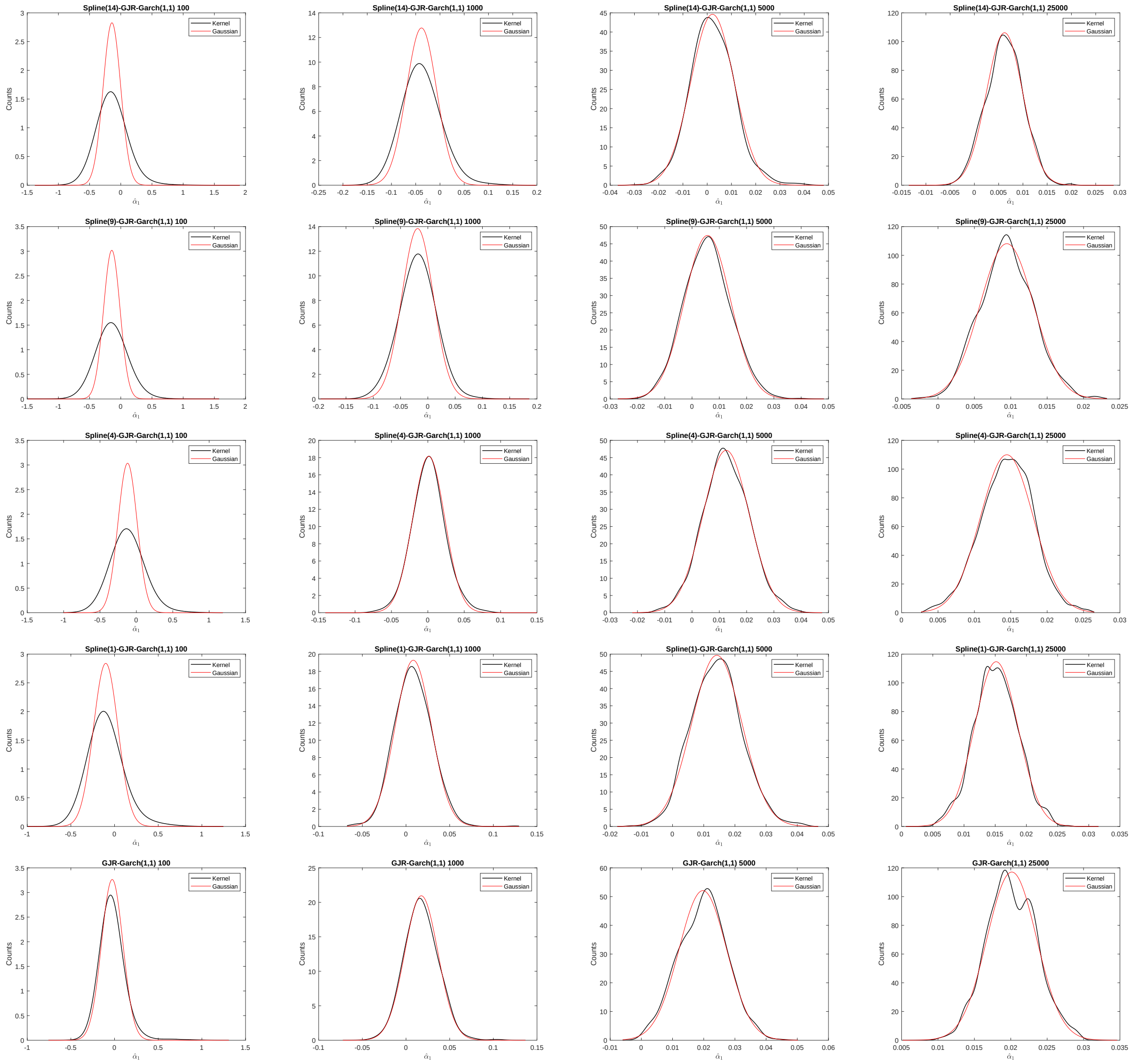

Figure 21: Asymptotic Normality of $\hat{\delta}_1$ in GJR-GARCH(1,1) case. The black curves represent the kernel density estimator for $\hat{\delta}$, the red curve the related normal distribution. In the top row the spline(14)-GARCH(1,1) distribution for $T \in \{100, 1000, 5000, 25000\}$ is depicted. The subjacent rows are the spline(K)-GARCH(1,1) models with $K \in \{0, 1, 4, 9\}$ and the same time series lengths.

Die Diskussionspapiere ab Nr. 403 (2007) bis heute, können Sie im Internet unter
http://www.fernuni-hagen.de/wirtschaftswissenschaft/forschung/beitraege.shtml einsehen
und zum Teil downloaden.
Ältere Diskussionspapiere selber erhalten Sie nur in den Bibliotheken.

Nr	Jahr	Titel	Autor/en
420	2008	Stockkeeping and controlling under game theoretic aspects	Fandel, Günter Trockel, Jan
421	2008	On Overdissipation of Rents in Contests with Endogenous Intrinsic Motivation	Schlepütz, Volker
422	2008	Maximum Entropy Inference for Mixed Continuous-Discrete Variables	Singer, Hermann
423	2008	Eine Heuristik für das mehrdimensionale Bin Packing Problem	Mack, Daniel Bortfeldt, Andreas
424	2008	Expected A Posteriori Estimation in Financial Applications	Mazzoni, Thomas
425	2008	A Genetic Algorithm for the Two-Dimensional Knapsack Problem with Rectangular Pieces	Bortfeldt, Andreas Winter, Tobias
426	2008	A Tree Search Algorithm for Solving the Container Loading Problem	Fanslau, Tobias Bortfeldt, Andreas
427	2008	Dynamic Effects of Offshoring	Stijepic, Denis Wagner, Helmut
428	2008	Der Einfluss von Kostenabweichungen auf das Nash-Gleichgewicht in einem nicht-kooperativen Disponenten-Controller-Spiel	Fandel, Günter Trockel, Jan
429	2008	Fast Analytic Option Valuation with GARCH	Mazzoni, Thomas
430	2008	Conditional Gauss-Hermite Filtering with Application to Volatility Estimation	Singer, Hermann
431	2008	Web 2.0 auf dem Prüfstand: Zur Bewertung von Internet-Unternehmen	Christian Maaß Gotthard Pietsch
432	2008	Zentralbank-Kommunikation und Finanzstabilität – Eine Bestandsaufnahme	Knütter, Rolf Mohr, Benjamin
433	2008	Globalization and Asset Prices: Which Trade-Offs Do Central Banks Face in Small Open Economies?	Knütter, Rolf Wagner, Helmut
434	2008	International Policy Coordination and Simple Monetary Policy Rules	Berger, Wolfram Wagner, Helmut
435	2009	Matchingprozesse auf beruflichen Teilarbeitsmärkten	Stops, Michael Mazzoni, Thomas
436	2009	Wayfindingprozesse in Parksituationen - eine empirische Analyse	Fließ, Sabine Tetzner, Stefan
437	2009	ENTROPY-DRIVEN PORTFOLIO SELECTION a downside and upside risk framework	Rödder, Wilhelm Gartner, Ivan Ricardo Rudolph, Sandra
438	2009	Consulting Incentives in Contests	Schlepütz, Volker

439	2009	A Genetic Algorithm for a Bi-Objective Winner-Determination Problem in a Transportation-Procurement Auction"	Buer, Tobias Pankratz, Giselher
440	2009	Parallel greedy algorithms for packing unequal spheres into a cuboidal strip or a cuboid	Kubach, Timo Bortfeldt, Andreas Tilli, Thomas Gehring, Hermann
441	2009	SEM modeling with singular moment matrices Part I: ML-Estimation of time series	Singer, Hermann
442	2009	SEM modeling with singular moment matrices Part II: ML-Estimation of sampled stochastic differential equations	Singer, Hermann
443	2009	Konsensuale Effizienzbewertung und -verbesserung – Untersuchungen mittels der Data Envelopment Analysis (DEA)	Rödder, Wilhelm Reucher, Elmar
444	2009	Legal Uncertainty – Is Hamonization of Law the Right Answer? A Short Overview	Wagner, Helmut
445	2009	Fast Continuous-Discrete DAF-Filters	Mazzoni, Thomas
446	2010	Quantitative Evaluierung von Multi-Level Marketingsystemen	Lorenz, Marina Mazzoni, Thomas
447	2010	Quasi-Continuous Maximum Entropy Distribution Approximation with Kernel Density	Mazzoni, Thomas Reucher, Elmar
448	2010	Solving a Bi-Objective Winner Determination Problem in a Transportation Procurement Auction	Buer, Tobias Pankratz, Giselher
449	2010	Are Short Term Stock Asset Returns Predictable? An Extended Empirical Analysis	Mazzoni, Thomas
450	2010	Europäische Gesundheitssysteme im Vergleich – Effizienzmessungen von Akutkrankenhäusern mit DEA –	Reucher, Elmar Sartorius, Frank
451	2010	Patterns in Object-Oriented Analysis	Blaimer, Nicolas Bortfeldt, Andreas Pankratz, Giselher
452	2010	The Kuznets-Kaldor-Puzzle and Neutral Cross-Capital-Intensity Structural Change	Stijepic, Denis Wagner, Helmut
453	2010	Monetary Policy and Boom-Bust Cycles: The Role of Communication	Knütter, Rolf Wagner, Helmut
454	2010	Konsensuale Effizienzbewertung und –verbesserung mittels DEA – Output- vs. Inputorientierung –	Reucher, Elmar Rödder, Wilhelm
455	2010	Consistent Modeling of Risk Averse Behavior with Spectral Risk Measures	Wächter, Hans Peter Mazzoni, Thomas

456	2010	Der virtuelle Peer – Eine Anwendung der DEA zur konsensualen Effizienz-bewertung –	Reucher, Elmar
457	2010	A two-stage packing procedure for a Portuguese trading company	Moura, Ana Bortfeldt, Andreas
458	2010	A tree search algorithm for solving the multi-dimensional strip packing problem with guillotine cutting constraint	Bortfeldt, Andreas Jungmann, Sabine
459	2010	Equity and Efficiency in Regional Public Good Supply with Imperfect Labour Mobility – Horizontal versus Vertical Equalization	Arnold, Volker
460	2010	A hybrid algorithm for the capacitated vehicle routing problem with three-dimensional loading constraints	Bortfeldt, Andreas
461	2010	A tree search procedure for the container relocation problem	Forster, Florian Bortfeldt, Andreas
462	2011	Advanced X-Efficiencies for CCR- and BCC-Modell – Towards Peer-based DEA Controlling	Rödder, Wilhelm Reucher, Elmar
463	2011	The Effects of Central Bank Communication on Financial Stability: A Systematization of the Empirical Evidence	Knütter, Rolf Mohr, Benjamin Wagner, Helmut
464	2011	Lösungskonzepte zur Allokation von Kooperationsvorteilen in der kooperativen Transportdisposition	Strangmeier, Reinhard Fiedler, Matthias
465	2011	Grenzen einer Legitimation staatlicher Maßnahmen gegenüber Kreditinstituten zur Verhinderung von Banken- und Wirtschaftskrisen	Merbecks, Ute
466	2011	Controlling im Stadtmarketing – Eine Analyse des Hagener Schaufensterwettbewerbs 2010	Fließ, Sabine Bauer, Katharina
467	2011	A Structural Approach to Financial Stability: On the Beneficial Role of Regulatory Governance	Mohr, Benjamin Wagner, Helmut
468	2011	Data Envelopment Analysis - Skalenerträge und Kreuzskalenerträge	Wilhelm Rödder Andreas Dellnitz
469	2011	Controlling organisatorischer Entscheidungen: Konzeptionelle Überlegungen	Lindner, Florian Scherm, Ewald
470	2011	Orientierung in Dienstleistungsumgebungen – eine explorative Studie am Beispiel des Flughafen Frankfurt am Main	Fließ, Sabine Colaci, Antje Nesper, Jens

471	2011	Inequality aversion, income skewness and the theory of the welfare state	Weinreich, Daniel
472	2011	A tree search procedure for the container retrieval problem	Forster, Florian Bortfeldt, Andreas
473	2011	A Functional Approach to Pricing Complex Barrier Options	Mazzoni, Thomas
474	2011	Bologna-Prozess und neues Steuerungsmodell – auf Konfrontationskurs mit universitären Identitäten	Jost, Tobias Scherm, Ewald
475	2011	A reduction approach for solving the rectangle packing area minimization problem	Bortfeldt, Andreas
476	2011	Trade and Unemployment with Heterogeneous Firms: How Good Jobs Are Lost	Altenburg, Lutz
477	2012	Structural Change Patterns and Development: China in Comparison	Wagner, Helmut
478	2012	Demografische Risiken – Herausforderungen für das finanzwirtschaftliche Risikomanagement im Rahmen der betrieblichen Altersversorgung	Merbecks, Ute
479	2012	"It's all in the Mix!" – Internalizing Externalities with R&D Subsidies and Environmental Liability	Endres, Alfred Friehe, Tim Rundshagen, Bianca
480	2012	Ökonomische Interpretationen der Skalenvariablen u in der DEA	Dellnitz, Andreas Kleine, Andreas Rödder, Wilhelm
481	2012	Entropiebasierte Analyse von Interaktionen in Sozialen Netzwerken	Rödder, Wilhelm Brenner, Dominic Kulmann, Friedhelm
482	2013	Central Bank Independence and Financial Stability: A Tale of Perfect Harmony?	Berger, Wolfram Kißmer, Friedrich
483	2013	Energy generation with Directed Technical Change	Kollenbach, Gilbert
484	2013	Monetary Policy and Asset Prices: When Cleaning Up Hits the Zero Lower Bound	Berger, Wolfram Kißmer, Friedrich
485	2013	Superknoten in Sozialen Netzwerken – eine entropieoptimale Analyse	Brenner, Dominic, Rödder, Wilhelm, Kulmann, Friedhelm
486	2013	Stimmigkeit von Situation, Organisation und Person: Gestaltungsüberlegungen auf Basis des Informationsverarbeitungsansatzes	Julmi, Christian Lindner, Florian Scherm, Ewald
487	2014	Incentives for Advanced Abatement Technology Under National and International Permit Trading	Endres, Alfred Rundshagen, Bianca
488	2014	Dynamische Effizienzbewertung öffentlicher Dreispartentheater mit der Data Envelopment Analysis	Kleine, Andreas Hoffmann, Steffen

489	2015	Konsensuale Peer-Wahl in der DEA -- Effizienz vs. Skalenertrag	Dellnitz, Andreas Reucher, Elmar
490	2015	Makroprudenzielle Regulierung – eine kurze Einführung und ein Überblick	Velauthapillai, Jeyakrishna
491	2015	SEM modeling with singular moment matrices Part III: GLS estimation	Singer, Hermann
492	2015	Die steuerliche Berücksichtigung von Aufwendungen für ein Studium − Eine Darstellung unter besonderer Berücksichtigung des Hörerstatus	Meyering, Stephan Portheine, Kea
493	2016	Ungewissheit versus Unsicherheit in Sozialen Netzwerken	Rödder, Wilhelm Dellnitz, Andreas Gartner, Ivan
494	2016	Investments in supplier-specific economies of scope with two different services and different supplier characters: two specialists	Fandel, Günter Trockel, Jan
495	2016	An application of the put-call-parity to variance reduced Monte-Carlo option pricing	Müller, Armin
496	2016	A joint application of the put-call-parity and importance sampling to variance reduced option pricing	Müller, Armin
497	2016	Simulated Maximum Likelihood for Continuous-Discrete State Space Models using Langevin Importance Sampling	Singer, Hermann
498	2016	A Theory of Affective Communication	Julmi, Christian
499	2016	Approximations of option price elasticities for importance sampling	Müller, Armin
500	2016	Variance reduced Value at Risk Monte-Carlo simulations	Müller, Armin
501	2016	Maximum Likelihood Estimation of Continuous-Discrete State-Space Models: Langevin Path Sampling vs. Numerical Integration	Singer, Hermann
502	2016	Measuring the domain-specificity of creativity	Julmi, Christian Scherm, Ewald
503	2017	Bipartite Strukturen in Sozialen Netzen – klassische versus MaxEnt-Analysen	Rödder, Wilhelm Dellnitz, Andreas Kulmann, Friedhelm Litzinger, Sebastian Reucher, Elmar
504	2017	Langevin and Kalman Importance Sampling for Nonlinear Continuous-Discrete State Space Models	Singer, Hermann
505	2017	Horizontal versus vertical fiscal Equalization	Anetsberger, Georg Arnold, Volker
506	2017	Formative and Reflective Measurement Models	Singer, Hermann
507	2017	Identifizierung von führenden Köpfen in terroristischen Netzwerken – ein entropiebasiertes Verfahren –	Dellnitz, Andreas Litzinger, Sebastian Rödder, Wilhelm
508	2017	Die Bedeutung der steuerlichen Norm § 5 Abs. 2 EStG für die handelsrechtliche Rechnungslegung	Meyering, Stephan
509	2018	Ein erweitertes Effizienzmaß für DMUs im BCC-Modell – eine ökonomiegerechte DEA-Anpassung –	Rödder, Wilhelm Dellnitz, Andreas Litzinger, Sebastian

510	2018	A concise proof of Gaussian smoothing	Singer, Hermann
511	2018	Empirical evidence on the topological properties of structural paths and some notes on its theoretical explanation	Stijepic, Denis
512	2018	On the predictability of economic structural change by the Poincaré-Bendixson theory	Stijepic, Denis
513	2018	On development paths minimizing the aggregate labor-reallocation costs in the three-sector framework and an application to structural policy	Stijepic, Denis
514	2018	Models of Continuous Dynamics on the 2-Simplex and Applications in Economics	Stijepic, Denis
515	2018	A Note on the Ideological Content of Modern Economic Dynamics Models and Ideology-Reducing Meta-Modeling	Stijepic, Denis
516	2018	Logistik für Versand von Studienmaterialien der FernUniversität in Hagen – Optimierte Bereitstellung bei der Kommissionierung	Brenner, Dominic Gädeke, Andre Kulmann, Friedhelm Kleine, Andreas
517	2019	Zur Quantifizierung von Macht und Machtallianzen – ein struktureller Ansatz in sozialen Netzwerken –	Dellnitz, Andreas Rödder, Wilhelm
518	2019	Kolmogorov Backward Equations with Singular Diffusion Matrices	Singer, Hermann
519	2020	Finite-Sample Properties of GARCH Models in the Presence of Time-Varying Unconditional Variance. A Simulation Study	Old, Oliver