IWQW

Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

Diskussionspapier
Discussion Papers

No. 05/2012

Detecting Outliers in Time Series

Vlad Ardelean
University of Erlangen-Nuremberg

ISSN 1867-6707

Friedrich-Alexander-Universität
IWQW
Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung

Detecting Outliers in Time Series

Working paper. August 29, 2012

Vlad Ardelean[1]
University of Erlangen-Nuremberg
Vlad.Ardelean@wiso.uni-erlangen.de

Keywords and phrases: GARCH processes; Detection of outliers; CUSUM-type test

Abstract

In parametric time series analysis there is the implicit assumption of no aberrant observations, so-called outliers. Outliers are observations that seem to be inconsistent with the assumed model. When these observations are included to estimate the model parameters, the resulting estimates are biased.
The fact that markets have been affected by shocks (i.e. East Asian crisis, Dot-com bubble, subprime mortgage crisis) make the assumption that no outlier is present questionable.
This paper addresses the problem of detecting outlying observations in time series. Outliers can be understood as a short transient change of the underlying parameters. Unfortunately tests designed to detect structural breaks cannot be used to find outlying observations. To overcome this problem a test normally used to detect structural breaks is modified. This test is based on the cumulative sum (CUSUM) of the squared observations. In comparison to a likelihood-ratio test neither the underlying model nor the functional form of the outliers have to be specified.
In a simulation study the finite sample behaviour of the proposed test is analysed. The simulation study shows that the test has reasonable power against a variety of alternatives.
Moreover, to illustrate the behaviour of the proposed test we analyse the returns of the Volkswagen stock.

[1]Corresponding author: Vlad Ardelean. Department of Statistics and Econometrics. University of Erlangen-Nuremberg. Lange Gasse 20. D-90403 Nuremberg. E-Mail: Vlad.Ardelean@wiso.uni-erlangen.de

1 Introduction

Returns of financial time series are often modelled by a generalised autoregressive conditional heteroscedasticity (GARCH) process, proposed by Bollerslev (1986). As the GARCH model is able to capture many stylised facts of financial returns.

The time-varying volatility of returns can be forecasted when the parameters of the process are known. Such a forecast can be used as an input for pricing options, other derivatives, trading and hedging strategies. Furthermore, the risk of an asset can be measured by the predicted volatility. The usefulness of the forecast is questionable when the estimated parameter are biased. In a simulation study, Ardelean (2009) investigates the effect on the estimated parameters when outlying observations are present and finds that the resulting estimates are biased upwards, especially the offset parameter.

Outliers are not a theoretical problem but occur more often than usually assumed in financial data, see for example Charles and Darné (2005) and the references therein. Even though many stylised facts of financial time series can be modelled with a GARCH process, the residuals of the fitted time series show excess kurtosis. According to Carnero et al. (2007) this indicates the presence of outliers.

This paper focuses on the question whether an observation is outlying or not. It is organized as follows: Section 2 gives a short introduction on GARCH models and their estimation. Section 3 gives a short review on the identification of outliers in GARCH processes. Section 4 introduces a test based on the cumulated sums of the squared observations to identify outlying observations. In Section 5 the finite-sample behaviour of the proposed test is analysed via a simulation study. Finally the tests are applied to financial data to check whether they contain outlying observations.

2 GARCH Processes

GARCH(p,q) processes are nowadays a standard tool to model financial asset returns. This is due to the fact that many of the stylised facts of financial assets are captured by this model, see Rama (2001). This chapter gives a short overview of GARCH(p,q) models and their estimation via the maximum likelihood method. In the following chapters, the likelihood is used to identify outlying observations. A more rigorous treatment of GARCH(p,q) models and their properties can be found, for example, in Berkes et al. (2003) or Lindner (2009).

Definition 1 (Bollerslev (1986)).
A stochastic process $(X_t)_{t\in\mathbb{Z}}$ is said to be a GARCH(p,q) process, if:

$$X_t|\mathcal{F}_{t-1} = \sigma_t \nu_t,$$

$$\sigma_t^2 = (\sigma_t(\theta))^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2,\ t \in \mathbb{Z}$$

with $\theta = (\alpha_0, \alpha_1, \cdots, \alpha_p, \beta_1, \cdots, \beta_q)$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $i = 1, \cdots, p$ and $\beta_j \geq 0$, $j = 1, \cdots, q$. $\mathcal{F}_t$ denotes the information set of the process up to time t and (ν_t) is an i.i.d. sequence of real-valued random variables, independent of σ_0, with $E_F(\nu_t) = 0$ and $E_F(\nu_t^2) = 1$. The probability distribution F of ν_t has a continuous density (with respect to Lebesgue measure on the real line), and its density f is positive on $(-\infty, +\infty)$

The parameters can be estimated via the maximum likelihood(ML) method. Let $x_1, \cdots, x_n$ be the observed sample, then the likelihood $\mathcal{L}$ is given by:

$$\begin{aligned}
\mathcal{L}(\theta) &= f_{X_1,\ldots,X_n}(x_1,\ldots,x_n;\theta) \\
&= f_{X_1}(x_1;\theta) f_{X_2|X_1}(x_2|x_1;\theta) \cdots f_{X_n|X_{n-1}\cdots X_1}(x_n|x_{n-1}\cdots x_1;\theta) \\
&= \prod_{i=\max(p,q)+1}^{n} f_{X_i|\mathcal{F}_{i-1}}(x_i|\mathcal{F}_{i-1};\theta).
\end{aligned}$$

The initial values $x_1, \cdots, x_{i-1}$ can be set to zero. Let $\Theta \in \mathcal{C}$ and let $\mathcal{C}$ be a compact subspace of $\mathbb{R}^{1+p+q}$, i.e.

$$C_\varepsilon := \left\{ a \in \mathbb{R}_+^{p+1}, b \in \mathbb{R}_+^q \;\middle|\; a_0 \in [\varepsilon, 1/\varepsilon],\ \sum_{i=1}^{p} a_i \geq \varepsilon,\ \sum_{i=1}^{p} a_i + \sum_{j=1}^{q} b_j \leq 1-\varepsilon \right\},$$

where $\varepsilon \in \mathbb{R}^+$. Then the maximum likelihood estimator is defined as:

$$\hat{\theta} = \arg\max_{\theta\in\Theta} \mathcal{L}(\theta).$$

Under general assumptions the maximum likelihood estimator is consistent and asymptotically normal even when the true distribution F of the innovations is not known, see for example Francq and Zakoïan (2004, Theorem 2.2). Furthermore, nested hypotheses can be tested on the basis of the difference between the maximum likelihood under the null hypothesis and under the alternative hypothesis, using the likelihood ratio test. Under general assumptions the likelihood ratio test statistic has a limiting χ^2 distribution under the null hypothesis, see for example Engle (1984). In the next chapter the likelihood ratio test by Doornik and Ooms (2005) to detect outlying observations is reviewed.

3 Identifying outliers

The detection of an outlier is simple when there is no dependency structure within the observations. In that case, the notion of which observations are possibly outlying simplifies to the question of which observations are the largest and the smallest, receptively. These observations can be tested for outlyingness, since they may be not consistent with the assumed distribution. A more rigorous analysis of outliers can be found in Barnett and Lewis (1994) and the references therein.
For dependent processes, especially autoregressive processes, Fox (1972) introduced two types of outliers namely additive (type-I) and innovational (type-II) outliers.
Type-I outliers only effect a single observation, while type-II outliers also affect the following observations.
Intervention analysis introduced by Box and Tiao (1975) can be embedded in the definition of Fox (1972) if more than one outlier can occur.
Let $(X_t)_{t\in\mathbb{Z}}$ be the underlying and unobserved GARCH(1,1) process, and let $(Y_t)_{t\in\mathbb{Z}}$ be the observable process which contains outliers. Then the two types of outliers can be modelled as follows:

additive outliers

$$
\begin{aligned}
&Y_t = X_t + \nu 1_\tau(t)\\
&X_t|\mathcal{F}_{t-1} \sim N(0, \sigma_{t-1}^2),\\
&\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2,
\end{aligned}
$$

innovational outliers

$$
\begin{aligned}
&Y_t = X_t + \nu 1_\tau(t)\\
&X_t|\mathcal{F}_{t-1} \sim N(0, \sigma_{t-1}^2),\\
&\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i Y_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2,
\end{aligned}
$$

where $\nu \in \mathbb{R}$ denotes the size of the outlier, $\tau \in \mathbb{Z}$ represents the time of occurrence of the outlier and $1_t(\tau)$ is the indicator function, which is one if $\tau = t$ and zero if $\tau \neq t$. As time dependent volatility is based on the disturbed processes for an innovational outlier, the outlier affects subsequent observations. Similar definitions for outliers in GARCH models can be found in Hotta and Tsay (2012), Doornik and Ooms (2005) and Carnero et al. (2007).

3.1 Outlier detection with the likelihood ratio test

Outlier detection in autoregressive processes was first proposed by Fox (1972). He proposed to use a likelihood ratio test if the type and the location of the outlier is known. If the location is unknown, each time point has to be tested individually. Box and Tiao (1975) extended the method in order to find outliers in autoregressive-moving-average (ARMA) processes. Abraham and Yatawara (1988) proposed a score test instead of a likelihood ratio test. Iterative procedures were introduced, among others, by Chang et al. (1988), Chen and Liu (1993) and Tsay (1986). A different approach was introduced by Bruce and Martin (1989), who proposed to omit observations and test the changes in the estimated parameters. The literature for outlier detection in GARCH processes is rather sparse. Franses and Ghijsels (1999) used the method proposed by Chang et al. (1988) by using the fact that a squared GARCH(1,1) process is an ARMA process. Recently, Hotta and Tsay (2012) proposed a Lagrange multiplier test to detect outliers. Lately the likelihood displacement approach proposed by Cook (1986) and extended by Billor and Loynes (1993) was applied to GARCH processes. Liu (2004) used elliptical error while Zhang and King (2005) used Gaussian errors.
The likelihood approach is more flexible since not all GARCH extensions have an ARMA representation. Mimicking the approach by Fox (1972) to detect outliers in autoregressive processes, Doornik and Ooms (2005) propose a likelihood ratio test to detect outliers. With the likelihood ratio test the occurrence of both types of outliers can be tested simultaneously. It is assumed that the functional form of the outlier is known, for example the form defined above. To test whether an outlier of any type occurred at a specific time τ one estimates the parameters of an extended

model:

$$Y_t = \nu_1 1_\tau + X_t,$$
$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 + \sum_{i=1}^{p} \nu_{1+i} 1_{\tau-i},$$

where ν_1 is the size of an additive outlier occurring at time point τ and $\nu_2, \ldots, \nu_{p+1}$ is the size of each innovational outlier. The parameters of this extended model can be estimated via the ML-method. Since the model with no outlier is nested ($\nu_1 = \nu_2 = \cdots = \nu_{p+1} = 0$) one can test the null hypothesis 'no outlier occurred at time τ' with a likelihood ratio test. Let $\hat{\theta}_0 = (\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_p, \hat{\beta}_1, \ldots, \hat{\beta}_q)$ be the restricted ML estimate and
$\hat{\theta}_1 = (\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_p, \hat{\beta}_1, \ldots, \hat{\beta}_q, \hat{\nu}_1, \ldots, \hat{\nu}_{p+1})$ be the unrestricted ML estimate. The test statistic λ_τ for an outlier at time τ is defined as:

$$\lambda_\tau = 2(\log L(\hat{\theta}_0) - \log L(\hat{\theta}_1)) \overset{a}{\sim} \chi^2(p+1)$$

Since the time of an outlier is not known, the test statistic is computed for every $\tau \leq n$

$$M_n = \max_{t \leq n} \lambda_t.$$

The distribution of M_n is non-standard. The parameters that estimate the size of the outliers appear only in the alternative. This problem was first discussed by Davies (1977) and Davies (1987). Andrews and Ploberger (1994) introduced optimal tests for this problem, which have exponential form, but have the disadvantage that they cannot detect the location of the outlier. Andrews and Ploberger (1995) showed that it is a best test if the alternative hypothesis is sufficiently distant from the null hypothesis. Under certain regularity conditions, Andrews and Ploberger (1995) showed that the distribution of the max LR test statistic converges under the null hypothesis. Doornik and Ooms (2005) approximated the distribution of M_n by:

$$P(M_n \leq x) = \exp\left(\exp\left(-\frac{x + 1.283 - 1.88 \log n(1 + \frac{12}{n})}{2.223}\right)\right).$$

When more than one outlier is present the outliers can be detected via an iterative procedure, see for example Rosner (1975).

3.2 Outlier detection based on online analysis

To save computational time we propose a method that reduces the number of observations that need to be tested. Instead of estimating the time-varying scale from the given process by a parametric model, one can estimate the variance model-free.

Rousseeuw and Hubert (1996) proposed a affine-invariant estimate for the scale in a bivariate regression. Any three points in $\mathbb{R}^2$ determine a (possibly degenerated) triangle. The height of such a triangle can be used as a measure of volatility.

The volatility modelled by a GARCH process is clustered, so it is reasonable to assume that the volatility of any datum can be approximated by its past volatilities. Gelper et al. (2009) extend the methodology of Rousseeuw and Hubert (1996) to the time series setting. Let y_t be the observed process, then the height of a triangle formed by three successive observations is:

$$h_i = \left| y_{i+1} - \frac{y_i + y_{i+2}}{2} \right|.$$

Different scale estimators can be based on the h_i. Let n be the length of a window, and for each time t the observations $y_{t-n+1}, \ldots, y_t$ are part of the window. From the observations within the window, $n-2$ such triangles can be constructed. Furthermore, let $h_{(i)}$ be the ordered sequence of the heights in the current window. A scale estimate for the time t is given by:

$$\hat{\sigma}_t = Q^{\alpha}_{\text{adj}}(y_1, \ldots, y_n),$$

where α satisfies $0 < \alpha < 1$ and Q^{α}_{adj} is a function that depends on the heights in the current window and α. The parameter α regulates the trade off between robustness and efficiency. Gelper et al. (2009) discuss different choices of the scale estimate and propose

$$Q^{\alpha}_{\text{adj}}(y_1, \ldots, y_n) = c_q h_{\langle \alpha(n-2) \rangle},$$

where c_q is a consistency factor and depends on the distribution of the innovations. If the underlying distribution is the normal distribution then the consistency factor is given by:

$$c_q = \left(\sqrt{\frac{3}{2}} \Phi^{-1} \left(\frac{\alpha + 1}{2} \right) \right),$$

where Φ^{-1} is the quantile function of the normal distribution.

This allows the identification of the time points which are possible candidates for outliers:

- Estimate the variance with the above mentioned procedure.
- Calculate the estimated residuals $\hat{r}_t = y_t / \sigma_t \; t = 1, \cdots, n$.
- The likelihood ratio test is only computed for those observations with $|r_t| > 1.65 \;\; t = 1, \cdots, n$.

The bound for the observations that are considered is somewhat arbitrary, but in the case of a GARCH(p,q) with normal innovations this boundary corresponds to the 97.5% confidence interval for the properly standardized observations. With this pre-procedure the number of observations for which a test for outlyingness has to be computed decrease drastically.

4 Outlier detection based on the increments of a Brownian motion

Outliers can be interpreted as an external shock to the system and can be tested via the likelihood ratio test introduced in the last section. But outliers can also be interpreted as short change of parameters of the underlying process. A change in the parameters of the underlying process is called a structural break. In order to test for structural breaks Inclan and Tiao (1994) propose to use the iterated cumulative sums of the squared processes. Under suitable regularity conditions the cumulative sums of the squared processes converges to a Brownian Motion. This test is applicable to GARCH processes but fails to detect outliers when they are present. This can be explained by the fact that outliers correspond to very short structural breaks. We modify the test in order to detect outliers.

The detection procedure is based on the fact that the correctly standardized cumulative sum of the squared GARCH process converges to a Brownian motion. Furthermore the distribution of the increments of a Brownian motion is known. In Theorem 2 the test is proposed. The following Theorem gives sufficient conditions under which the correctly standardized cumulative sum of random variables converges to a Brownian Motion.

Theorem 1 (Theorem 11 from Merlevède et al. (2006))**.** *Let $\mathcal{F}_0$ be a $\sigma-$Algebra on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying $\mathcal{F}_0 \subseteq \mathbb{T}(\mathcal{F}_0)$, where $\mathbb{T} : \Omega \to \Omega$.Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary sequence with $E(X_0) = 0$ and $E(X_0)^2 < \infty$ Assume that the following holds:*

$$\sum_{i=1}^{\infty} \frac{||E(S_n|\mathcal{F}_0)||_2}{n^{\frac{3}{2}}} < \infty,$$

where $S_n = \sum_{i=1}^{n} X_i$ and $||X||_p = (E(|X|^p))^{\frac{1}{p}}$. Then,

$$\left\{ max_{1\leq k\leq n} \frac{S_k^2}{n} \ : n \geq 1 \right\}$$

is uniformly integrable and

$$W_n \overset{D}{\to} \sqrt{\eta} W,$$

where $W_n(r) = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{\lfloor nr \rfloor} X_i$ and W a Brownian Motion, ν is a non-negative random variable with finite mean $E[\eta] = \sigma^2$ and independent of $\{W(r); r \geq 0\}$. Moreover, η is determined by the limit

$$\lim_{n\to\infty} E\left(\frac{S_n^2|\mathcal{I}}{n}\right) = \eta \ in \ L_1,$$

where $\mathcal{I}$ is the invariant sigma field. In particular, $\lim_{n\to\infty} E\left(\frac{S_n^2}{n}\right) = \sigma$.

The following Theorem is the basis for the outlier detection method.

Theorem 2.
Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary process that fulfils the assumptions of the previous theorem and let $\xi_t = X_t^2 - Var(X_t)$, then

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{\lfloor rn \rfloor} \xi_i \overset{D}{\to} W(r),$$

where $\sigma = E\left((\sum_{i=1}^{n} X_i^2/n)^2\right)$.
It holds furthermore that:

$$\max_{1\leq i\leq n} \xi_i - \xi_{i-1} \to G,$$

where G is a Gumbel distribution with suitable normalizing constants $\mu_g(n)$ and $\sigma_g(n)$.

Proof (Theorem 2) From Theorem 1 we have that

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{\lfloor rn \rfloor} \xi_i \overset{D}{\to} W(r).$$

By definition the following holds for a Brownian motion:

1. W(0)=0.

2. Let $t_1, t_2, t_3, t_4 \in [0, 1]$ with $t_1 < t_2, t_3 < t_4$. Then
   $W(t_2) - W(t_1)$ and $W(t_4) - W(t_3)$ are stochastically independent.

3. $\forall \ t_1, t_2 \in [0, 1]$ with $t_1 \leq t_2$ it holds: $W(t_2) - W(t_1) \ N(0, t_2 - t_1)$.

From this follows that $\xi_t - \xi_{t-1}$ is i.i.d. normal with $\mu_\xi = 0$ and $\sigma_\xi^2 = \frac{1}{n}$ Furthermore, the maximum of i.i.d. normal distributed random variables lies in the domain of attraction of a Gumbel distribution. The location parameter $\mu_g(n)$ and the scale parameter $\sigma_g(n)$ are given in Takahashi (1987)

$$\mu_g(n) = \left((2\log(n))^{\frac{1}{2}} - \left(\log(\log(n) + \log(4\pi) \right) / \left(2/(2\log(n))^{\frac{1}{2}} \right) \right) \sqrt{\frac{1}{n}} \tag{1}$$

$$\sigma_g(n) = \left(2\log(n) \right)^{-\frac{1}{2}} \sqrt{\frac{1}{n}} \tag{2}$$

□

Under certain conditions on the parameters the assumptions for Theorem 2 hold for GARCH processes and various extensions. The conditions for a GARCH(1,1) and the TGARCH process proposed by Zakoïan (1994) can be found in **?**.

With this computationally inexpensive test not only the presence but also the occurrence time of an outlier can be identified. σ^2 can be consistently estimated by :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(x_i^2 - \hat{\sigma}_x^2 \right)^2 + \frac{2}{n} \sum_{l=1}^{m} w(l,m) \sum_{i=1}^{n} \left(x_i^2 - \hat{\sigma}_x^2 \right)^2,$$

where $\hat{\sigma}_x^2$ is a consistent estimator of the unconditional variance of the process X_t and $w(l,m)$ is a lag window, i.e. the Bartlett window defined by

$$w(l,m) = 1 - \frac{l}{m+1}.$$

The bandwidth m can be chosen according to an automatic procedure proposed by Newey and West (1994).

In the next section the finite sample properties of the proposed test are evaluated.

5 Simulation study

In the following chapter the finite samples properties of the proposed test are investigated.

5.1 Setting

For the studies we will consider the following four data generating process (DGP)

- DGP-1: GARCH(1,1) $\alpha_0 = 0.001,\ \alpha_1 = 0.1,\ \beta_1 = 0.75$
- DGP-2: GARCH(1,1) $\alpha_0 = 0.1,\ \alpha_1 = 0.1,\ \beta_1 = 0.75$
- DGP-3: GARCH(1,1) $\alpha_0 = 0.01,\ \alpha_1 = 0.3,\ \beta_1 = 0.3$
- DGP-4: T-GARCH(1,1,1) $\alpha_0 = 0.01,\ \alpha_1 = 0.1,\ \alpha_2 - = 0.2,\ \beta = 0.7$ proposed by Zakoïan (1994) with the following equation for the standard deviation:

  $$\sigma_t = \alpha_0 + \alpha_1 x_{t-1}^+ - \alpha_2 x_{t-1}^- + \beta \sigma_{t-1}.$$

  where $x_t^+ = \max(x_t, 0)$ and $x_t^- = \min(x_t, 0)$.

The sample size is $n = 500$ or $n = 1000$ and the innovations are standard normal. All four processes fulfil assumptions for Theorem 1 (see Example 1 and 2 in the appendix).
The proposed test is non-parametric in the sense that the underlying process does not need to be known or specified. In order to investigate the capability of the likelihood ration test to detect outliers in a misspecified model, we simulate DGP-4 but for the likelihood ratio test it is assumed that the data follow a GARCH(1,1) which will be denoted by DGP-4*.

5.2 Small sample parameters

5.2.1 Convergence

In a first step the convergence of the parameters of the Gumbel distribution is evaluated for the four DGP's. Each process is simulated 1000 times with different sample size n. For each simulation we record the maximal value of the test statistic. This is repeated 1000 times. First, we test if the distribution of the maxima is indeed the Gumbel distribution. This is tested with a chi-squared goodness of fit test with 20 classes. Table 1 shows the rejection frequencies of the null hypothesis at the significance level $\alpha = 0.01$.

DGP	Observations		
	250	500	1000
DGP-1	0.015	0.011	0.009
DGP-2	0.018	0.013	0.011
DGP-3	0.019	0.015	0.012
DGP-4	0.021	0.016	0.014

Table 1: Rejection frequencies

In a next step, a Gumbel distribution is fitted to the maxima. The estimated parameters and the standard deviation are shown in table 8 and 9 in the appendix. As can be seen, the convergence is rather slow, the estimated parameters are higher than the asymptotic parameters from equation (1) and (2) and the estimated parameters differ for each DGP. The critical values implied by asymptotic parameters are lower then the simulated values.

5.2.2 Small sample parameters

As a consequence, the parameters for the Gumbel distribution are estimated from the largest k Observations.

Theorem 3 (Theorem 5 of Weissman (1978)). *Let $x_1, \cdots, x_n$ be the observed sample from a distribution F that belongs to the domain of attraction of a Gumbel distribution. Furthermore, let $x_{(1)} \geq \cdots \geq x_{(n)}$ be the ordered observations. The MLE estimator of μ_g and σ_g based on the k-largest observations are given by:*

$$\hat{\sigma}_g = \frac{1}{k}\sum_{i=1}^{k} x_{(i)} + x_{(k)}$$

$$\hat{\mu}_g = \hat{\sigma}_g \ln(k) + x_{(k)}$$

The procedure is sensitive to the choice of k. Finding the 'optimal' k is equivalent to determining the beginning of the tail of the distribution. Different possibilities to determine k exist, see

Strawderman and Zelterman (1998).

$$k = \underset{m}{\operatorname{argmin}} \left\{ \left(\hat{\gamma}_m - R_m^{(1)} \right)^2 \right\}, \text{ with} \tag{3}$$

$$\hat{\gamma}_m = R_m^{(1)} + 1 - \frac{1}{2} \left(1 - \frac{(R_m^{(1)})^2}{R_m^{(2)}} \right)^{-1}$$

$$R_m^{(j)} = k^{-1} \sum_{i=1}^{k} \left(\log(x_i) - \log(x_{k+1}) \right)^j .$$

This choice is motivated by the result of Davis and Resnick (1984):

$$\hat{\gamma}_k - R_k^{(1)} \xrightarrow{P} 0.$$

The choice for this criterion has another justification. The distribution function F of a random variable X is in the domain of attraction of a Gumbel distribution iff

$$\lim_{t \to t_0} \frac{E[(X-t)^2 | X > t]}{(E[X-t | X > t])^2} = 2, \tag{4}$$

with $t_0 = \sup\{t | \ F(t) < 1\}$.
The empirical analogue to equation (4) for the log-transformed random variable is $\frac{R_m^{(2)}}{(R_m^{(1)})^2}$. It also converges in probability to two, when F is in the domain of attraction of a Gumbel distribution. Using equation (3) to find the value of k corresponds to choosing the value of k 'most consistent' with the data coming from a Gumbel distribution cf. Strawderman and Zelterman (1998, p. 448).

5.3 Detecting outliers

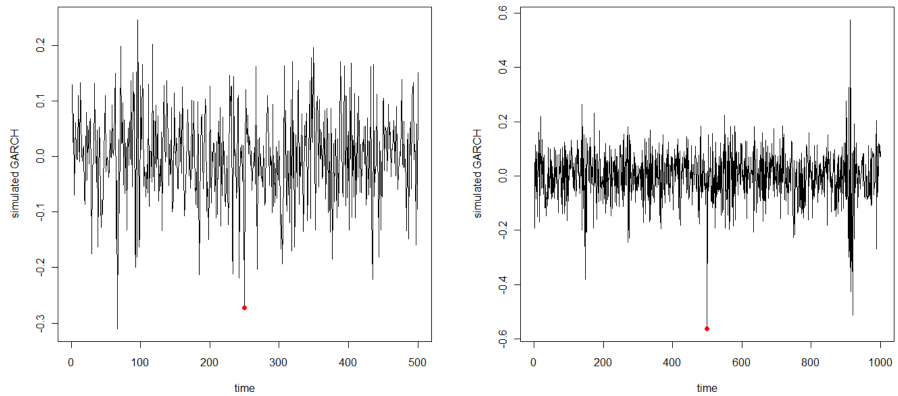
The following tests were used to detect one or more outliers.

- $CUSUM - type$: Test based on the increments of a Brownian motion .
- LR : Likelihood-ratio test proposed by Doornik and Ooms (2005).
- $LR - OSE$: Likelihood-ratio test proposed by Doornik and Ooms (2005) but only on those time points t with estimated residuals $|y_t/\hat{\sigma}_t| > 1.66$ are considered. To obtain a maximal finite sample breakdown point, α is set to 0.2625, see Gelper et al. (2009).

We modify the GARCH processes by adding one additive outlier or one innovational outlier at time $\lfloor n/2 \rfloor$ of size $3, 5, 7$ or adding two additive outliers, respectively two innovational outliers at time $\lfloor n/2 \rfloor$, $\lfloor n/2+50 \rfloor$. The size of every outlier is either relative or fixed. The size of an relative outlier is the current standard deviation at time t σ_t times 3,5 or 7, while the size of an fixed outlier is the unconditional standard deviation σ times 3,5 or 7. Figure 1 shows DGP-1 with relative additive outlier of size 3 respectively size 5.

5.4 Results

The results are summarized in the following tables. The first number is the relative number of detections of an outlier, while the number in parenthesis is the relative number of correctly



(a) (b)

Figure 1: Simulated GARCH process (DGP-1) with one relative additive outlier of size 3 at time 250 (left) and with one relative additive outlier of size 5 at time 500 (right)

identified outliers. These two measures are proposed for example by David (1981), Hawkins (1980) and Barnett and Lewis (1994) to measure the performance of an outlier test. The results for method $LR - OSE$ are not shown since they are the same as for LR. The advantage lies in the reduced number of the likelihood ratio tests that have to be performed.

Table 2 shows the results when no outlier is present. The likelihood-ratio approach does not hold its significance level in small samples. This effect is ample when a wrong underlying model for the data generating process is assumed. The CUSUM-type test holds its significance level.

Table 3 indicates, that both tests are consistent and the test proposed by Doornik and Ooms (2005) is able to detect small outliers better (size=3). For larger outliers both tests give similar results. Interestingly, when using a GARCH(1,1) model to detect outliers in an TGARCH process outliers are still detected reasonably well. This can be due to the fact that outliers in a TGARCH process are extreme outliers in a GARCH process. The results for other types of outliers are given in the Appendix. The type of outliers seem to influence the detection rate for the CUSUM-type test. Especially, relative outliers are detected better than fixed outliers. This can be explained by the fact that adding a fixed value in a period with high volatility doesn't seem too unlikely.

When more than one small outlier is present, an iterative procedure is used. The significance level of each test is adjusted such that the overall significance level of detecting an outlier when no outlier is present stays at $\alpha = 0.01$ or respectively $\alpha = 0.05$.

The iterative procedure used here is the forward method proposed by Hawkins (1980):

- Test the time series for an occurrence of an outlier
- If the test detects an outlier, split the series in two and test each half (CUSUM-type) or adjust the series for the outlier (LR-Test)
- Repeat until no more outlier is detected.

DGP	Observations	LR		$CUSUM-type$	
		0.95	0.99	0.95	0.99
DGP 1	n=500	0.086	0.016	0.066	0.014
	n=1000	0.068	0.012	0.058	0.012
DGP 2	n=500	0.092	0.022	0.06	0.016
	n=1000	0.072	0.008	0.04	0.008
DGP 3	n=500	0.054	0.014	0.058	0.018
	n=1000	0.04	0.008	0.052	0.008
DGP 4	n=500	0.066	0.022	0.062	0.02
	n=1000	0.05	0.014	0.058	0.014
DGP 4*	n=500	0.132	0.038	-	-
	n=1000	0.102	0.032	-	-

Table 2: Detection rate when no outlier is present

DGP	Observations		LR		$CUSUM-type$	
			0.95	0.99	0.95	0.99
DGP 1	n=500	3	0.324 (0.272)	0.136 (0.13)	0.204 (0.168)	0.082 (0.072)
		5	0.962 (0.95)	0.94 (0.938)	0.904 (0.898)	0.666 (0.666)
		7	0.988 (0.988	0.982 (0.982)	1 (1)	0.994 (0.994)
	n=1000	3	0.4 (0.332)	0.212 (0.188)	0.21 (0.19)	0.092 (0.09)
		5	0.968 (0.95)	0.944 (0.94)	0.922 (0.912)	0.708 (0.706)
		7	0.998 (0.982	0.99 (0.976)	0.998 (0.998)	0.988 (0.988)
DGP 2	n=500	3	0.39 (0.326)	0.176 (0.164)	0.222 (0.2)	0.072 (0.072)
		5	0.836 (0.816)	0.812 (0.798)	0.906 (0.906)	0.682 (0.682)
		7	0.712 (0.694)	0.682 (0.668)	1 (1)	0.996 (0.996)
	n=1000	3	0.41 (0.348)	0.18 (0.166)	0.256 (0.212)	0.106 (0.098)
		5	0.95 (0.94)	0.932 (0.926)	0.926 (0.914)	0.73 (0.726)
		7	0.926 (0.92)	0.902 (0.9)	1 (1)	0.994 (0.994)
DGP 3	n=500	3	0.402 (0.36)	0.182 (0.178)	0.194 (0.11)	0.07 (0.052)
		5	0.984 (0.978)	0.966 (0.96)	0.636 (0.61)	0.366 (0.362)
		7	0.988 (0.98)	0.974 (0.968)	0.948 (0.932)	0.802 (0.798)
	n=1000	3	0.346 (0.3)	0.154 (0.152)	0.268 (0.1)	0.114 (0.058)
		5	0.994 (0.992)	0.986 (0.986)	0.792 (0.764)	0.402 (0.378)
		7	0.996 (0.996)	0.99 (0.99)	0.962 (0.956)	0.834 (0.834)
DGP 4	n=500	3	0.548 (0.52)	0.348 (0.334)	0.302 (0.272)	0.118 (0.114)
		5	0.866 (0.86)	0.858 (0.856)	0.9 (0.894)	0.724 (0.724)
		7	0.932 (0.912)	0.916 (0.906)	0.99 (0.99)	0.958 (0.958)
	n=1000	3	0.758 (0.732)	0.572 (0.564)	0.316 (0.272)	0.122 (0.106
		5	0.98 (0.976)	0.974 (0.972)	0.914 (0.898)	0.762 (0.762)
		7	0.1 (0.95)	0.988 (0.978)	0.992 (0.992)	0.974 (0.974)
DGP 4*	n=500	3	0.826 (0.798)	0.67 (0.66)	-	-
		5	0.904 (0.876)	0.884 (0.866)	-	-
		7	0.914 (0.902)	0.908 (0.902)	-	-
	n=1000	3	0.826 (0.786)	0.65 (0.632)	-	-
		5	0.932 (0.924)	0.922 (0.892)	-	-
		7	0.95 (0.942)	0.944 (0.938)	-	-

Table 3: Detection rate when one relative additive outlier is present.

DGP	Observations		LR		$CUSUM-type$	
			0.95	0.99	0.95	0.99
DGP 1	n=500	3	0.556 (0.088)	0.312 (0.036)	0.246 (0.052)	0.066 (0)
		5	0.978 (0.894)	0.956 (0.818)	0.77 (0.644)	0.382 (0.254)
		7	0.984 (0.942)	0.97 (0.924)	0.98 (0.94)	0.772 (0.706)
	n=1000	3	0.608 (0.116)	0.354 (0.03)	0.27 (0.052)	0.11 (0.016)
		5	0.984 (0.898)	0.978 (0.85)	0.848 (0.73)	0.574 (0.418)
		7	0.988 (0.95)	0.968 (0.924)	0.998 (0.976)	0.948 (0.918)
DGP 2	n=500	3	0.524 (0.118)	0.3 (0.03)	0.234 (0.032)	0.078 (0)
		5	0.778 (0.692)	0.752 (0.656)	0.78 (0.608)	0.462 (0.256)
		7	0.604 (0.51)	0.572 (0.482)	0.988 (0.95)	0.82 (0.752)
	n=1000	3	0.55 (0.108)	0.296 (0.026)	0.282 (0.032)	0.11 (0.002)
		5	0.948 (0.91)	0.934 (0.888)	0.892 (0.718)	0.65 (0.404)
		7	0.876 (0.828)	0.866 (0.808)	1 (0.98)	0.972 (0.94)
DGP 3	n=500	3	0.534 (0.094)	0.284 (0.026)	0.284 (0.026)	0.134 (0.002)
		5	0.986 (0.946)	0.984 (0.928)	0.656 (0.372)	0.368 (0.118)
		7	0.998 (0.968)	0.996 (0.96)	0.896 (0.772)	0.67 (0.474)
	n=1000	3	0.554 (0.126)	0.322 (0.03)	0.318 (0.014)	0.17 (0.006)
		5	1 (0.986)	1 (0.97)	0.684 (0.312)	0.438 (0.136)
		7	1 (0.99)	0.998 (0.986)	0.956 (0.832)	0.792 (0.622)
DGP 4	n=500	3	0.796 (0.574)	0.332 (0.238)	0.308 (0.018)	0.12 (0.002)
		5	0.646 (0.612)	0.492 (0.372)	0.67 (0.298)	0.424 (0.086)
		7	0.86 (0.824)	0.624 (0.496)	0.908 (0.762)	0.712 (0.454)
	n=1000	3	1 (0.574)	1 (0.574)	0.266 (0.014)	0.122 (0.004)
		5	1 (0.95)	1 (0.95)	0.72 (0.332)	0.486 (0.142)
		7	1 (0.966)	1 (0.966)	0.938 (0.76)	0.808 (0.576)
DGP 4*	n=500	3	0.602 (0.096)	0.384 (0.034)	-	-
		5	0.936 (0.902)	0.926 (0.884)	-	-
		7	0.722 (0.648)	0.708 (0.636)	-	-
	n=1000	3	0.686 (0.16)	0.408 (0.052)	-	-
		5	0.996 (0.98)	0.996 (0.95)	-	-
		7	0.938 (0.928)	0.938 (0.928)	-	-

Table 4: Detection rate when two relative additive outlier are present.

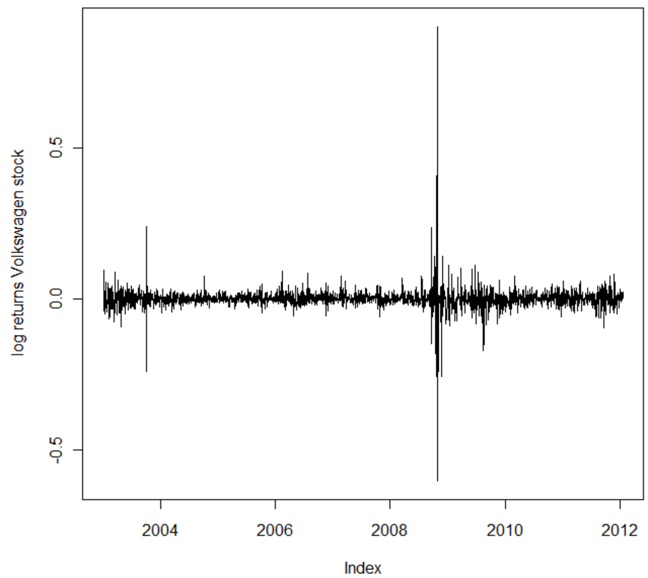
6 Empirical analysis

We analyse the daily log returns of the Volkswagen (VW) stock between 01.01.2003 and 08.11.2011. To obtain a short overview of the data some descriptive statistics are given.The ADF-test on stationarity and the Jarque-Bera test on normality are also carried out. As can be seen from table 5 the log returns are leptokurtic, stationary and not normally distributed.

Asset	mean	std. dev.	skewness	kurtosis	p-value of ADF	p-value of JB
Volkswagen	0.0005	0.0371	5.9008	221.2313	< 0.01	2.2e-16

Table 5: Descriptive Statistics

To check whether a GARCH model is suitable for the given data a lagrange multiplier test to detect arch effects is carried out, Engle (1982). Since there may be outliers in the data a robust



version of the test as proposed by van Dijk et al. (1999) is applied. Table 6 shows the results for the Lagrange multiplier test. For all lags the null of no (g)arch effects can be rejected.

	Lag			
Asset	1	5	10	15
VW	2.2e-16	2.2e-16	2.2e-16	2.2e-16

Table 6: p-value of the robust LM-test for arch effects.

The table 7 summarizes the results. A date indicates that an outlier was detected at the particular date. Dates that are bold indicate that only one of the tests found an outlier

$\alpha = 0.05$	LR		CUSUM-type	
VW	30.09.03	16.09.08	30.09.03	16.09.08
	17.10.08	24.10.08	24.10.08	27.10.08
	27.10.08	28.10.08	28.10.08	**25.11.08**

Table 7: Location of outliers detected at $\alpha = 0.05$

Since it is not clear whether an observation is outlying or not, news regarding the VW stock are used as a proxy. If extreme market movements correspond to news, an outlier is assumed, otherwise the detected outlier is regarded as false alarm. A similar approach was used in Charles and Darné (2006). The following news correspond to detected dates:

30.09.03: Compensation for the economic disadvantages caused by the division of Germany was dismissed by European Court of Justice.

16.09.08: Porsche announced that the company had increased its stake in Volkswagen AG to 35 percent.

27.10.08: Porsche announced that they effectively held over 74 percent of the Volkswagen AG stock.

25.11.08: The Volkswagen stock lost index weight in the MSCI World market index .

However, news were not found for all detected dates. This could indicate that there were no outliers or extreme market movements. For the dates close to the 28.10.2008 the outliers can be explained with the turmoil and shortage of freely traded Volkswagen stock, while for the 17.10.2008 no news could be found.

7 Conclusion

In this paper a new test to detect outlying observations was proposed. For a small outlier a likelihood ratio test has more accurate results. This changes when there is more than one outlier present or when the size of the outlier increases. Another advantage of the proposed test is that it is non-parametric, so no exact model for the time series has to be assumed. We can reduce the number of observations that have to be tested with the LR-test by first applying the online analysis approach.
The CUSUM type method can be combined with non-parametric estimation. Since every detected outlier splits the sample into two parts, the number of outliers that can be detected is rather small. Instead of splitting the sample one can use a non-parametric method to estimate the value of the observation that is detected as an outlier. Possible non-parametric methods are proposed by Biau et al. (2010).

References

Abraham, B. and Yatawara, N. (1988). A score test for detection of time series outliers. *Journal of time series analysis*, 9(2):109–119.

Andrews, D. W. K. and Ploberger, W. (1994). Optimal Tests when a Nuisance Parameter is Present Only Under the Alternative. *Econometrica*, 62(6):1383–1414.

Andrews, D. W. K. and Ploberger, W. (1995). Admissibility of the Likelihood Ratio Test When a Nuisance Parameter is Present Only Under the Alternative. *The Annals of Statistics*, 23(5):1609–1629.

Ardelean, V. (2009). The impact of outliers on different estimators for GARCH processes: an empirical study. *American Statsitical Association*.

Barnett, V. and Lewis, T. (1994). *Outliers in statistical data*. Wiley & Sons, 3 edition.

Berkes, I., Horvát, K., and Kokoszka, P. (2003). GARCH processes: structure and estimation. *Bernoulli*, 9:201–227.

Biau, G., Bleakley, K., Györfi, L., and Ottucsák, G. (2010). Nonparametric sequential prediction of time series. *Journal of Nonparametric Statistics*, 22(3):297–317.

Billor, N. and Loynes, R. (1993). Local influence: a new approach. *Communications in Statistics - Theory and Methods*, 22(6):1595–1611.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31:307–322.

Box, G. E. P. and Tiao, G. C. (1975). Intervention Analysis with Applications to Economic and Environmental Problems. *Journal of the American Statistical Association*, 70(349):70–79.

Bruce, A. G. and Martin, R. D. (1989). Leave-k-Out Diagnostics for Time Series. *Journal of the Royal Statistical Society. Series B (Methodological)*, 51(3):363–424.

Carnero, M. A., Peña, D., and Ruiz, E. (2007). Effects of outliers on the identification and estimation of GARCH Models. *Journal of time series analysis*, 28(4):471–497.

Carrasco, M. and Chen, X. (2002). Mixing and moment properties if various garch and stochastic volatility models. *Econometric Theory*, 18(01):17–39.

Chang, I., Tiao, G. C., and Chen, C. (1988). Estimation of Time Series Parameters in the Presence of Outliers. *Technometrics*, 30(2):193–204.

Charles, A. and Darné, O. (2005). Relevance of detecting outliers in GARCH models for modelling and forecasting financial data: L'intérêt de détecter les outliers dans les modèles GARCH pour modéliser et prévoir les données financières. *Finance*, 26(1):33–71.

Charles, A. and Darné, O. (2006). Large shocks and the September 11th terrorist attacks on international stock markets. *Economic Modelling*, 23(4):683–698.

Chen, C. and Liu, L.-M. (1993). Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association*, 88(421):284–297.

Cook, R. D. (1986). Assessment of Local Influence. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(2):p 133–169.

David, H. A. (1981). *Order statistics. 2nd edition*. John Wiley & Sons, New York and N.Y.

Davidson, J. (1994). *Stochastic Limit Theory*. Oxford University Press.

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 64(2):247–254.

Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*, 74(1):33–43.

Davis, R. and Resnick, S. (1984). Tail Estimates Motivated by Extreme Value Theory. *The Annals of Statistics*, 12(4):1467–1487.

Doornik, J. A. and Ooms, M. (2005). Outlier Detection in GARCH Models.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007.

Engle, R. F. (1984). *Handbook of Econometrics*. Elsevier Science B.V.

Fox, A. J. (1972). Outliers in Time Series. *Journal of the Royal Statistical Society. Series B (Methodological)*, 34(3):350–363.

Francq, C. and Zakoïan, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10:605–637.

Franses, P. H. and Ghijsels, H. (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting*, 15(1):1–9.

Gelper, S., Schettlinger, K., Croux, C., and Gather, U. (2009). Robust online scale estimation in time series: A model-free approach. *Journal of Statistical Planning and Inference*, 139(2):335–349.

Hawkins, D. M. (1980). *Identification of outliers*. Chapman and Hall, London and and New York.

Hotta, L. K. and Tsay, R. S. (2012). Outliers in GARCH Processes. In Bell, W. R., Holan, S. H., and McElroy, T., editors, *Economic time series*, pages 337–358. CRC Press, Boca Raton and FL.

Inclan, C. and Tiao, G. C. (1994). Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance. *Journal of the American Statistical Association*, 89(427):913–923.

Lindner, A. M. (2009). Stationarity, Mixing, Distributional Properties and Moments of GARCH(p,q)-processes. In Mikosch, T., Kreiß, J.-P., Davis, R. A., and Andersen, T. G., editors, *Handbook of Financial Time Series*, pages 43–69. Springer Berlin Heidelberg.

Liu, S. (2004). On Diagnostics in Conditionally Heteroskedastic Time Series Models under Elliptical Distributions. *Journal of Applied Probability*, 41:p 393–405.

Merlevède, F., Peligrad, M., and Utev, S. (2006). Recent advances in invariance principles for stationary sequences. *Probability Surveys*, 3:1–35.

Newey, W. K. and West, K. D. (1994). Automatic Lag Selection in Covariance Matrix Estimation. *The Review of Economic Studies*, 61(4):631–653.

Oodaira, H. and Yoshihara, K.-i. (1972). Functional central limit theorems for strictly stationary processes satisfying the strong mixing condition. *Kodai Mathematical Seminar Reports*, 24(3):259–269.

Rama, C. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.

Rosner, B. (1975). On the Detection of Many Outliers. *Technometrics*, 17(2):221–227.

Rousseeuw, P. and Hubert, M. (1996). Regression-free and robust estimation of scale for bivariate data. *Computational Statistics & Data Analysis*, 21(1):67–85.

Strawderman, R. L. and Zelterman, D. (1998). A semiparametric bootstrap for simulating extreme order statistics. In N. Balakrishnan and C.R. Rao, editors, *Order Statistics: Theory & Methods*, volume 16 of *Handbook of Statistics*, pages 441–462. Elsevier.

Takahashi, R. (1987). Normalizing constants of a distribution which belongs to the domain of attraction of the Gumbel distribution. *Statistics & Probability Letters*, 5(3):197–200.

Tsay, R. S. (1986). Time Series Model Specification in the Presence of Outliers. *Journal of the American Statistical Association*, 81(393):p 132–141.

van Dijk, D., Franses, P., and Lucas, A. (1999). Testing for ARCH in the presence of additive outliers. *Journal of Applied Econometrics*, 14:539–562.

Weissman, I. (1978). Estimation of Parameters and Larger Quantiles Based on the k Largest Observations. *Journal of the American Statistical Association*, 73(364):812–815.

Zakoïan, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955.

Zhang, X. and King, M. L. (2005). Influence Diagnostics in Generalized Autoregressive Conditional Heteroscedasticity Processes. *Journal of Business & Economic Statistics*, 23(1):p 118–129.

8 Appendix

8.1 Examples

In order to show under which parameter restrictions the conditons of Theorem 1 hold we need the next two results.
The next Corollary makes it easier to check if the conditions for Theorem 1 hold.

Corollary 1 (Corollary 12 Merlevède et al. (2006))**.** *Let everything be defined as in Theorem 1. Assume that*

$$\sum_{n=1}^{\infty} \frac{||E(X|\mathcal{F}_0)||_2}{n^{\frac{1}{2}}} < \infty.$$

Then the conclusion of Theorem 1 hold.

To show that the conditions of Theorem 1 hold for a GARCH(1,1) we need a result form Carrasco and Chen (2002)

Corollary 2 (Corollary 6 from Carrasco and Chen (2002))**.** *Let the processes X_t and the process of the innovations (ν_t) be defined as in Definition 1. Assume that there is an integer $s \geq 1$ such that either (a) or (b) is fulfilled*

$$(a)\ E\bigg(\left((\beta_1 + \alpha_1 \cdot \nu_t)^2 \right)^s \bigg) < 1$$
$$(b) E(|\nu_t|)^{2s} < \infty \quad \beta_1 + \alpha_1 < \frac{1}{E(\nu_t)^{\frac{1}{s}}}.$$

Then

- *If σ_0 is initialized from its invariant measure, then (σ_t) and (X_t) are strictly stationary and β-mixing with exponential decay*
- *$E(\sigma_t)^2 < \infty$ and $E(|X_t|^{2s}) < \infty$. Condition (a) with $s = 2$ is also necessary to simultaneously have (σ_t) geometric ergodic and $E(|\sigma_t|^{2s}) < \infty$.*

Example GARCH(1,1) Let the processes X_t be an GARCH(1,1) then the assumptions for Theorem 1 hold for the process $\xi_t = X_t^2 - \sigma$ under certain regularity conditions.
The process is stationary and has finite second moments when $E(X_t^4)$ is finite.
We use condition (a) of Corollary 2 with $s = 2$:

$$E\bigg(\left((\beta_1 + \alpha_1 \cdot \nu_t)^2 \right)^s \bigg) = \beta_1^2 + 2 \cdot \alpha_1 \beta E(\nu_t^2) + \alpha_1^2 E(\nu_t^4) < 1$$

The assumption that the fourth moment of the underlying GARCH exists is not uncommon, especially when the parameters are estimated with the QML-approach.
Every squared GARCH(p,q) X_t process can be represented as an ARMA$(\max(p, q), q)$ process and every ARMA-Process can be represented as a stationary $MA(\infty)$ process.

$$\xi_k = \sum_{j=-\infty}^{k} a_{k-j} v_j = \sum_{i=1}^{\infty} a_i v_{k-i}$$

with $\sum_{j=0}^{\infty} a_j^2 < \infty$ and v_t is a martingale difference regarding the filtration $\mathcal{F}_{t-1}$ with finite second moment $\sigma_v < \infty$.

$$||E(\xi_k|\mathcal{F}_0)||_2 = E\left(E(\xi_k|\mathcal{F}_0)^2\right) = E\left(\left(\sum_{i=1}^{\infty} a_i v_{k-i}\right)^2\right)) = \sigma_v^2 \sum_{i=k}^{\infty} a_i^2 \to 0 \text{ for } k \to \infty,$$

since $E(\xi_k|\mathcal{F}) = \sum_{j=-\infty}^{0} a_{k-j} v_j = \sum_{i=k}^{\infty} a_i v_{k-i}$ and the $v_i' s$ are uncorrelated.
A GARCH process fulfils the assumptions for Theorem 1 when $\beta^2 + 2 \cdot \alpha\beta E(\nu_t^2) + \alpha^2 E(\nu_t^4) < 1$.
The process $\xi_t^2 - \sigma$ fullfils the assumptions for Theorem 1 when X_t^2 is stationary and $\beta - mixing$.
The process X_t^2 is stationary and $\beta-$mixing with exponential rate and therefore $\alpha-$mixing. From Davidson (1994) Theorem 14.1 we have that when the process $(X_t)_{t\in\mathbb{Z}}$ is $\alpha-$mixing then

$$\xi_t = g(X_t, X_{t-1}, \cdots, X_{t-\tau})$$

is also $\alpha-$mixing when g is a measurable function and τ is finite.
In our case the function g is:

$$g(X_t, X_{t-1}, \cdots, X_{t-\tau}) = X_t^2 - \frac{\sum_{i=1}^{\tau} X_{t-i}{}^2}{\tau - 1},$$

which is continuous. The second part of the function is a consistent estimate of the variance.

To show that the conditions of Theorem 1 hold for a TGARCH(1,1) we need the following two results

Theorem 4 (Theorem 2 from Oodaira and Yoshihara (1972)). *Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary process with $E(X_0) = 0$ and $E(X_0)^2 < \infty$ that is alpha-mixing with mixing coefficients $\alpha(t)$. If*

$$E|X_j|^{2+\delta} < \infty \text{ and } \sum_{i=1}^{\infty} \alpha(t)^{\frac{\delta}{2}+\delta},$$

then $\sigma < \infty$. If $\sigma > 0$ then:

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{\lfloor nt \rfloor} X_t \overset{D}{\to} B,$$

where and B a Brownian Motion.

The following results gives necessary the conditions for a TGARCH(1,1) to have existing higher moments ergodicity and stationarity.

Corollary 3 (Corollary 11 from Carrasco and Chen (2002)). *Let the processes X_t and the process of the innovations (ν_t) be defined as in Definition 1. Assume that there is an integer $s \geq 1$ such that either (a) or (b) is fulfilled*

$$(a)\ E\Big(\left(\beta_1 + \alpha_1|n_t| + \alpha_2 \max(0, -\nu_t)\right)^s\Big) < 1$$

$$(b) E(|\nu_t|)^{2s} < \infty \quad \beta + \alpha < \frac{1}{E(\nu_t)^{\frac{1}{s}}}.$$

Then

- *If σ_0 is initialized from its invariant measure, then (σ_t) and (X_t) are strictly stationary and β-mixing with exponential decay*
- *$E(\sigma_t)^2 < \infty$ and $E(|X_t|^{2s}) < \infty$. Condition (a) with $s = 2$ is also necessary to simultaneously have (σ_t) geometric ergodic and $E(|\sigma_t|^{2s}) < \infty$.*

Example TGARCH(1,1) Let $(X_t)_{t\in\mathbb{Z}}$ be a TGARCH(1,1) process. A functional limit theorem hold for the process $\xi_t = X_t^2 - \sigma$, where $\sigma = E(X_t^2)$. From Theorem 4 we have a sufficient conditions is the existence of the $E|X_j|^{2+\delta}$ moment with $\delta > 0$ and the fact that the process is $\alpha-$mixing. We can use Theorem 3 with $\delta = \frac{1}{2}$. The 5^{th} Moment of an TGARCH exists when:

$$E[(\beta + \alpha_1|\nu_t| + \alpha_2 \max(0, -\nu_t))^5] < 1.$$

$$\begin{aligned}
E\Bigg(& \alpha_1^5|\nu_t|^5 + 5\alpha_1^4\beta_1|\nu_t|^4 + 5\alpha_1^4\alpha_2 \max(0, -\nu_t)|\nu_t|^4 + 10\alpha_1^3\beta_1^2|\nu_t|^3 + 20\alpha_1^3\beta_1(\alpha_2 \max(0, -\nu_t))|\nu_t|^3 + \\
&+ 10\alpha_1^3(\alpha_2 \max(0, -\nu_t))^2|\nu_t|^3 + 10\alpha_1^2\beta_1^3|\nu_t|^2 + 30\alpha_1^2\beta_1^2(\alpha_2 \max(0, -\nu_t))|\nu_t|^2 + \\
&+ 30\alpha_1^2\beta_1(\alpha_2 \max(0, -\nu_t))^2|\nu_t|^2 + 10\alpha_1^2(\alpha_2 \max(0, -\nu_t))^3|\nu_t|^2 + 5\alpha_1\beta_1^4|\nu_t| + \\
&+ 20\alpha_1\beta_1^3(\alpha_2 \max(0, -\nu_t))|\nu_t| + 30\alpha_1\beta_1^2(\alpha_2 \max(0, -\nu_t))^2|\nu_t| + 20\alpha_1\beta_1(\alpha_2 \max(0, -\nu_t))^3|\nu_t| + \\
&+ 5\alpha_1(\alpha_2 \max(0, -\nu_t))^4|\nu_t| + \beta_1^5 + 5\beta_1^4(\alpha_2 \max(0, -\nu_t)) + 10\beta_1^3(\alpha_2 \max(0, -\nu_t))^2 + \\
&+ 10\beta_1^2(\alpha_2 \max(0, -\nu_t))^3 + 5\beta_1(\alpha_2 \max(0, -\nu_t))^4 + (\alpha_2 \max(0, -\nu_t))^5\Bigg) < 1
\end{aligned}$$

When the distribution of ν_t is symmetric $E(\max(0, -\nu_t)) = E(|\nu_t|)/2$. This equation simplifies to:

$$\begin{aligned}
& 6.38\alpha_1^5 + 15\alpha_1^4\beta_1 + 31.9\alpha_1^4\alpha_2 + 25.6\alpha_1^3\beta_1^2 + 30\alpha_1^3\alpha_2\beta_1 + 31.4\alpha_1^3\alpha_2^2 + 10\alpha_1^2\beta_1^3 + 24\alpha_1^2\beta_1^2\alpha_2 + \\
& + 45\alpha_1^2\beta_1\alpha_2 + 31.9\alpha_1^2\alpha_2^3 + 4\alpha_1\beta_1^4 + 10\alpha_1\beta_1^3\alpha_2 + 24\alpha_1\beta_1^2\alpha_2^2 + 30\alpha_1\beta_1\alpha_2^3 + 15.95\alpha_1\alpha_2^4 + \beta_1^5 + \\
& + 4\beta_1^4\alpha_2 + 5\beta_1^3\alpha_2^2 + 8\beta_1^2\alpha_2^3 + 7.5\beta_1\alpha_2^4 + 3.14\alpha_2^5 < 1
\end{aligned}$$

The parameter constellations that fullfill this equation are not as strict as it seems. The following table gives the largest parameter β_1 that is allowed for given α_1 and α_2

α_1	α_2	β_1
0.05	0.1	0.87
0.1	0.15	0.78
0.1	0.2	0.73

with $E(|\nu_t|) = \sqrt{2}/\sqrt{\pi} \sim 0.8$ $E(|\nu_t|^3) = 2\sqrt{2}/\sqrt{\pi} \sim 1.6$ and $E(|\nu_t|^5) = 8\sqrt{2}/\sqrt{\pi} \sim 6.38$
Furthermore the process X_t is stationary and $\beta-$mixing with exponential rate and therefore $\alpha-$mixing. From Davidson (1994) Theorem 14.1 we have that when the process $(X_t)_{t\in\mathbb{Z}}$ is $\alpha-$mixing then

$$\xi_t = g(X_t, X_{t-1}, \cdots, X_{t-\tau})$$

is also $\alpha-$mixing when g is a measurable function and τ is finite.
In our case the function g is:

$$g(X_t, X_{t-1}, \cdots, X_{t-\tau}) = X_t^2 - \frac{\sum_{i=1}^{\tau} {X_{t-i}}^2}{\tau - 1},$$

which is continuous. The second part of the function is a consistent estimate of the variance. This fact together with the parameter restriction such that the 5^{th} Moment exist are sufficient assumptions of Theorem 3.

8.2 Convergence

T	True Param	DGP-1	DGP-2	DGP-3	DGP-4
50	0.348	1.084 (0.023)	1.106 (0.023)	1.069 (0.022)	1.061 (0.022)
100	0.271	0.890 (0.018)	0.899 (0.018)	0.745 (0.014)	0.809 (0.016)
250	0.190	0.574 (0.010)	0.568 (0.009)	0.434 (0.005)	0.509 (0.007)
500	0.144	0.399 (0.004)	0.393 (0.004)	0.327 (0.003)	0.372 (0.003)
1000	0.108	0.287 (0.002)	0.292 (0.002)	0.250 (0.002)	0.281 (0.002)
2000	0.081	0.214 (0.001)	0.215 (0.001)	0.190 (0.001)	0.215 (0.002)
3000	0.068	0.185 (0.001)	0.186 (0.001)	0.164 (0.001)	0.186 (0.001)
4000	0.059	0.164 (0.001)	0.164 (0.001)	0.149 (0.001)	0.168 (0.001)
5000	0.054	0.150 (0.001)	0.151 (0.001)	0.139 (0.001)	0.155 (0.001)
10000	0.040	0.116 (0.001)	0.114 (0.001)	0.109 (0.001)	0.123 (0.001)
15000	0.033	0.100 (0.001)	0.099 (0.001)	0.096 (0.001)	0.108 (0.001)
20000	0.029	0.089 (0.001)	0.089 (0.001)	0.087 (0.001)	0.097 (0.001)

Table 8: Asymptotic and estimated location parameter for the resulting Gumbel for finite samples

T	True Param	DGP-1	DGP-2	DGP-3	DGP-4
50	0.050	0.714 (0.020)	0.709 (0.020)	0.673 (0.019)	0.689 (0.016)
100	0.033	0.554 (0.017)	0.568 (0.016)	0.430 (0.012)	0.508 (0.014)
250	0.019	0.310 (0.009)	0.275 (0.008)	0.142 (0.004)	0.220 (0.006)
500	0.0127	0.127 (0.003)	0.129 (0.003)	0.085 (0.002)	0.105 (0.003)
1000	0.0090	0.067 (0.002)	0.070 (0.002)	0.061 (0.002)	0.069 (0.002)
2000	0.0057	0.043 (0.001)	0.044 (0.001)	0.043 (0.001)	0.048 (0.001)
3000	0.0046	0.036 (0.001)	0.036 (0.001)	0.039 (0.001)	0.038 (0.001)
4000	0.0039	0.031 (0.001)	0.030 (0.001)	0.034 (0.001)	0.035 (0.001)
5000	0.0034	0.028 (0.001)	0.028 (0.001)	0.031 (0.001)	0.033 (0.001)
10000	0.0023	0.020 (0.001)	0.021 (0.001)	0.026 (0.001)	0.026 (0.001)
15000	0.0019	0.018 (0.001)	0.018 (0.001)	0.023 (0.001)	0.022 (0.001)
20000	0.0016	0.016 (0.001)	0.015 (0.001)	0.021 (0.001)	0.019 (0.001)

Table 9: Asymptotic and estimated scale parameter for the resulting Gumbel distribution for finite samples

8.3 Results

8.3.1 One outlier

DGP	Observations		LR		$CUSUM-type$	
			0.95	0.99	0.95	0.99
DGP 1	n=500	3	0.394 (0.338)	0.172 (0.164)	0.184 (0.168)	0.06 (0.06)
		5	0.956 (0.948)	0.908 (0.902)	0.932 (0.93)	0.686 (0.686)
		7	0.996 (0.992)	0.988 (0.984)	1 (1)	0.99 (0.99)
	n=1000	3	0.454 (0.402)	0.24 (0.22)	0.194 (0.158)	0.082 (0.072)
		5	0.954 (0.942)	0.91 (0.904)	0.966 (0.96)	0.738 (0.736)
		7	0.994 (0.99)	0.992 (0.988)	1 (1)	1 (1)
DGP 2	n=500	3	0.43 (0.366)	0.216 (0.204)	0.188 (0.172)	0.072 (0.066)
		5	0.828 (0.808)	0.77 (0.762)	0.938 (0.934)	0.654 (0.652)
		7	0.734 (0.72)	0.716 (0.706)	0.998 (0.998)	0.984 (0.984)
	n=1000	3	0.45 (0.388)	0.22 (0.21)	0.196 (0.184)	0.084 (0.076)
		5	0.94 (0.932)	0.918 (0.912)	0.954 (0.948)	0.792 (0.788)
		7	0.956 (0.94)	0.934 (0.918)	1 (1)	0.996 (0.996)
DGP 3	n=500	3	0.482 (0.454)	0.296 (0.284)	0.162 (0.108)	0.05 (0.034)
		5	0.966 (0.962)	0.922 (0.922)	0.77 (0.758)	0.454 (0.45)
		7	0.984 (0.98)	0.972 (0.968)	0.986 (0.986)	0.902 (0.902)
	n=1000	3	0.57 (0.544)	0.322 (0.314)	0.22 (0.078)	0.074 (0.026)
		5	0.956 (0.95)	0.93 (0.926)	0.786 (0.764)	0.468 (0.446)
		7	0.986 (0.98)	0.976 (0.972)	0.988 (0.974)	0.94 (0.934)
DGP 4	n=500	3	0.122 (0.05)	0.044 (0.03)	0.144 (0.098)	0.034 (0.03)
		5	0.482 (0.444)	0.306 (0.3)	0.708 (0.658)	0.416 (0.386)
		7	0.784 (0.776)	0.688 (0.684)	0.972 (0.958)	0.898 (0.888)
	n=1000	3	0.134 (0.072)	0.042 (0.028)	0.202 (0.166)	0.064 (0.044)
		5	0.536 (0.496)	0.328 (0.316)	0.798 (0.79)	0.48 (0.478)
		7	0.854 (0.842)	0.758 (0.754)	0.972 (0.97)	0.908 (0.906)
DGP 4*	n=500	3	0.594 (0.53)	0.41 (0.39)	-	-
		5	0.886 (0.874)	0.828 (0.82)	-	-
		7	0.82 (0.792)	0.778 (0.762)	-	-
	n=1000	3	0.664 (0.61)	0.482 (0.454)	-	-
		5	0.954 (0.95)	0.928 (0.924)	-	-
		7	0.976 (0.97)	0.97 (0.966)	-	-

Table 10: Detection rate when one fixed additive outlier is present.

DGP	Observations		LR		$CUSUM - type$	
			0.95	0.99	0.95	0.99
DGP 1	n=500	3	0.356 (0.29)	0.164 (0.15)	0.154 (0.132)	0.048 (0.042)
		5	0.964 (0.942)	0.904 (0.894)	0.858 (0.846)	0.554 (0.55)
		7	1 (0.978)	0.994 (0.972)	0.978 (0.966)	0.9 (0.894)
	n=1000	3	0.372 (0.322)	0.202 (0.184)	0.178 (0.14)	0.056 (0.048)
		5	0.976 (0.966)	0.924 (0.918)	0.924 (0.912)	0.672 (0.664)
		7	0.998 (0.992)	0.992 (0.988)	0.996 (0.988)	0.968 (0.962)
DGP 2	n=500	3	0.398 (0.34)	0.158 (0.126)	0.102 (0.902)	0.046 (0.02)
		5	0.792 (0.774)	0.636 (0.618)	0.76 (0.724)	0.424 (0.402)
		7	0.93 (0.908)	0.918 (0.9)	0.908 (0.908)	0.782 (0.782)
	n=1000	3	0.398 (0.34)	0.188 (0.176)	0.17 (0.15)	0.086 (0.07)
		5	0.892 (0.874)	0.836 (0.818)	0.86 (0.854)	0.524 (0.522)
		7	0.962 (0.958)	0.948 (0.924)	0.978 (0.978)	0.882 (0.882)
DGP 3	n=500	3	0.436 (0.402)	0.256 (0.246)	0.16 (0.084)	0.042 (0.026)
		5	0.948 (0.94)	0.896 (0.892)	0.644 (0.572)	0.346 (0.318)
		7	0.99 (0.982)	0.968 (0.96)	0.922 (0.842)	0.73 (0.67)
	n=1000	3	0.488 (0.46)	0.266 (0.258)	0.228 (0.058)	0.078 (0.018)
		5	0.956 (0.952)	0.908 (0.904)	0.72 (0.59)	0.416 (0.338)
		7	1 (0.988)	0.98 (0.976)	0.96 (0.846)	0.866 (0.774)
DGP 4	n=500	3	0.124 (0.05)	0.04 (0.024)	0.118 (0.062)	0.038 (0.028)
		5	0.45 (0.408)	0.258 (0.252)	0.606 (0.692)	0.338 (0.324)
		7	0.78 (0.766)	0.67 (0.662)	0.904 (0.884)	0.76 (0.76)
	n=1000	3	0.132 (0.066)	0.042 (0.028)	0.196 (0.058)	0.064 (0.012)
		5	0.474 (0.432)	0.278 (0.266)	0.638 (0.582)	0.358 (0.342)
		7	0.836 (0.826)	0.72 (0.716)	0.952 (0.936)	0.836 (0.83)
DGP 4*	n=500	3	0.54 (0.462)	0.328 (0.3)	-	
		5	0.892 (0.868)	0.846 (0.83)	-	
		7	0.884 (0.864)	0.852 (0.838)	-	
	n=1000	3	0.578 (0.512)	0.374 (0.342)	-	
		5	0.938 (0.932)	0.892 (0.888)	-	
		7	0.982 (0.976)	0.972 (0.966)	-	-

Table 11: Detection rate when one fixed innovational outlier is present.

DGP	Observations		LR		$CUSUM-type$	
			0.95	0.99	0.95	0.99
DGP 1	n=500	3	0.296 (0.252)	0.144 (0.128)	0.182 (0.166)	0.07 (0.064)
		5	0.996 (0.966)	0.976 (0.948)	0.874 (0.868)	0.598 (0.598)
		7	1 (0.978)	0.998 (0.978)	0.988 (0.986)	0.932 (0.932)
	n=1000	3	0.34 (0.278)	0.168 (0.156)	0.234 (0.176)	0.086 (0.074)
		5	0.998 (0.976)	0.982 (0.962)	0.892 (0.886)	0.634 (0.626)
		7	1 (0.99)	0.998 (0.99)	0.994 (0.978)	0.97 (0.956)
DGP 2	n=500	3	0.306 (0.232)	0.138 (0.126)	0.2 (0.17)	0.052 (0.044)
		5	0.906 (0.888)	0.894 (0.878)	0.834 (0.822)	0.558 (0.552)
		7	0.926 (0.914)	0.898 (0.888)	0.996 (0.984)	0.918 (0.912)
	n=1000	3	0.322 (0.254)	0.14 (0.124)	0.242 (0.182)	0.09 (0.076)
		5	0.984 (0.97)	0.962 (0.95)	0.898 (0.87)	0.674 (0.656)
		7	0.978 (0.966)	0.976 (0.964)	0.998 (0.986)	0.984 (0.974)
DGP 3	n=500	3	0.282 (0.24)	0.122 (0.118)	0.18 (0.084)	0.074 (0.04)
		5	0.988 (0.984)	0.976 (0.972)	0.578 (0.498)	0.302 (0.274)
		7	0.996 (0.992)	0.994 (0.99)	0.876 (0.804)	0.696 (0.656)
	n=1000	3	0.348 (0.302)	0.144 (0.142)	0.274 (0.086)	0.118 (0.046)
		5	0.998 (0.988)	0.98 (0.972)	0.598 (0.526)	0.38 (0.316)
		7	0.998 (0.992)	0.998 (0.992)	0.89 (0.796)	0.728 (0.656)
DGP 4	n=500	3	0.678 (0.65)	0.472 (0.46)	0.248 (0.218)	0.096 (0.092)
		5	0.804 (0.792)	0.802 (0.786)	0.844 (0.836)	0.67 (0.67)
		7	0.888 (0.868)	0.88 (0.866)	0.978 (0.976)	0.926 (0.926)
	n=1000	3	0.706 (0.676)	0.502 (0.494)	0.286 (0.206)	0.128 (0.106)
		5	0.988 (0.982)	0.982 (0.976)	0.878 (0.856)	0.68 (0.676)
		7	0.994 (0.988)	0.990 (0.982)	0.992 (0.988)	0.962 (0.958)
DGP 4*	n=500	3	0.722 (0.722)	0.546 (0.532)	-	
		5	0.894 (0.872)	0.888 (0.872)	-	
		7	0.948 (0.932)	0.932 (0.912)	-	
	n=1000	3	0.774 (0.73)	0.55 (0.534)	-	
		5	0.986 (0.978)	0.978 (0.97)	-	
		7	1 (1)	1 (1)	-	

Table 12: Detection rate when one relative innovational outlier is present.

8.3.2 Two outliers

DGP	Observations		LR		CUSUM − type	
			0.95	0.99	0.95	0.99
DGP 1	n=500	3	0.54 (0.088)	0.294 (0.024)	0.26 (0.024)	0.078 (0.004)
		5	0.972 (0.88)	0.942 (0.814)	0.79 (0.696)	0.396 (0.258)
		7	986 (0.958)0.992 (0.968)	0.974 (0.93)	0.974 (0.95)	0.81 (0.776)
	n=1000	3	0.588 (0.154)	0.344 (0.048)	0.23 (0.028)	0.068 (0.004)
		5	0.982 (0.882)	0.966 (0.812)	0.914 (0.828)	0.59 (0.42)
		7	0.992 (0.968)	0.978 (0.944)	0.998 (0.978)	0.984 (0.962)
DGP 2	n=500	3	0.582 (0.128)	0.308 (0.038)	0.23 (0.016)	0.06 (0)
		5	0.786 (0.698)	0.746 (0.636)	0.808 (0.72)	0.428 (0.278)
		7	0.974 (0.970)	0.968 (0.958)	0.992 (0.97)	0.812 (0.774)
	n=1000	3	0.588 (0.138)	0.37 (0.036)	0.208 (0.026)	0.066 (0.002)
		5	0.912 (0.862)	0.886 (0.806)	0.904 (0.798)	0.598 (0.444)
		7	0.988 (0.982)	0.982 (0.972)	1 (0.988)	0.98 (0.962)
DGP 3	n=500	3	0.696 (0.24)	0.474 (0.084)	0.236 (0.004)	0.042 (0)
		5	0.99 (0.912)	0.972 (0.846)	0.558 (0.428)	0.212 (0.124)
		7	0.996 (0.962)	0.992 (0.94)	0.912 (0.864)	0.596 (0.524)
	n=1000	3	0.742 (0.242)	0.47 (0.068)	0.144 (0.006)	0.082 (0)
		5	0.998 (0.9)	0.992 (0.836)	0.718 (0.498)	0.354 (0.17)
		7	1 (0.976)	1 (0.958)	0.976 (0.93)	0.852 (0.782)
DGP 4	n=500	3	0.122 (0.05)	0.042 (0.028)	0.146 (0.028)	0.038 (0.004)
		5	0.482 (0.444)	0.306 (0.3)	0.664 (0.528)	0.292 (0.17)
		7	0.784 (0.776)	0.688 (0.684)	0.932 (0.89)	0.676 (0.608)
	n=1000	3	0.134 (0.072)	0.12 (0.008)	0.188 (0.04)	0.06 (0.02)
		5	0.536 (0.496)	0.328 (0.316)	0.696 (0.52)	0.356 (0.166)
		7	0.854 (0.842)	0.758 (0.754)	0.972 (0.948)	0.824 (0.768)
DGP 4*	n=500	3	0.822 (0.322)	0.634 (0.158)	-	-
		5	0.922 (0.818)	0.91 (0.764)	-	-
		7	0.698 (0.63)	0.688 (0.602)	-	-
	n=1000	3	0.844 (0.38)	0.694 (0.204)	-	-
		5	0.992 (0.874)	0.984 (0.818)	-	-
		7	0.958 (0.928)	0.958 (0.91)	-	-

Table 13: Detection rate when two fixed additive outlier are present.

DGP	Observations		LR		$CUSUM-type$	
			0.95	0.99	0.95	0.99
DGP 1	n=500	3	0.494 (0.062)	0.248 (0.016)	0.198 (0.024)	0.052 (0.002)
		5	0.988 (0.858)	0.964 (0.784)	0.62 (0.432)	0.274 (0.122)
		7	1 (0.94)	0.992 (0.924)	0.84 (0.718)	0.536 (0.366)
	n=1000	3	0.508 (0.1)	0.278 (0.03)	0.226 (0.02)	0.062 (0.004)
		5	0.994 (0.864)	0.986 (0.752)	0.832 (0.636)	0.438 (0.242)
		7	1 (0.95)	0.998 (0.946)	0.982 (0.902)	0.848 (0.712)
DGP 2	n=500	3	0.51 (0.076)	0.252 (0.012)	0.16 (0.006)	0.036 (0.002)
		5	0.958 (0.82)	0.93 (0.736)	0.642 (0.466)	0.284 (0.114)
		7	0.974 (0.878)	0.956 (0.856)	0.868 (0.724)	0.542 (0.352)
	n=1000	3	0.542 (0.076)	0.296 (0.02)	0.198 (0.01)	0.054 (0)
		5	0.986 (0.894)	0.968 (0.806)	0.832 (0.634)	0.474 (0.258)
		7	0.986 (0.938)	0.98 (0.926)	0.986 (0.91)	0.864 (0.734)
DGP 3	n=500	3	0.638 (0.132)	0.394 (0.048)	0.168 (0)	0.056 (0)
		5	0.992 (0.714)	0.98 (0.642)	0.466 (0.206)	0.18 (0.046)
		7	1 (0.788)	1 (0.782)	0.746 (0.476)	0.454 (0.23)
	n=1000	3	0.66 (0.158)	0.432 (0.088)	0.248 (0.004)	0.102 (0)
		5	0.994 (0.712)	0.99 (0.662)	0.664 (0.29)	0.322 (0.068)
		7	1 (0.794)	0.998 (0.774)	0.904 (0.584)	0.714 (0.412)
DGP 4	n=500	3	0.132 (0.066)	0.042 (0.028)	0.176 (0)	0.038 (0)
		5	0.45 (0.408)	0.258 (0.252)	0.364 (0)	0.246 (0)
		7	0.78 (0.766)	0.67 (0.662)	0.744 (0.744)	0.588 (0.628)
	n=1000	3	0.124 (0.05)	0.04 (0.024)	0.216 (0)	0.086 (0)
		5	0.474 (0.432)	0.278 (0.266)	0.458 (0.432)	0.23 (0.224)
		7	0.836 (0.826)	0.72 (0.716)	0.832 (0.822)	0.626 (0.624)
DGP 4*	n=500	3	0.586 (0)	0.348 (0)	-	-
		5	0.906 (0.014)	0.852 (0.008)	-	-
		7	0.868 (0.038)	0.842 (0.026)	-	-
	n=1000	3	0.568 (0.002)	0.374 (0.002)	-	-
		5	0.95 (0.004)	0.882 (0.002)	-	-
		7	0.988 (0.002)	0.98 (0)	-	-

Table 14: Detection rate when two fixed innovational outlier are present.

DGP	Observations		LR		$CUSUM-type$	
			0.95	0.99	0.95	0.99
DGP 1	n=500	3	0.544 (0.064)	0.22 (0.016)	0.21 (0.022)	0.042 (0.002)
		5	0.98 (0.868)	0.962 (0.782)	0.628 (0.434)	0.29 (0.134)
		7	0.988 (0.936)	0.98 (0.922)	0.85 (0.694)	0.538 (0.364)
	n=1000	3	0.572 (0.098)	0.262 (0.024)	0.264 (0.038)	0.096 (0.012)
		5	0.998 (0.892)	0.996 (0.828)	0.768 (0.584)	0.47 (0.26)
		7	1 (0.96)	1 (0.952)	0.976 (0.898)	0.818 (0.71)
DGP 2	n=500	3	0.402 (0.037)	0.206 (0.012)	0.204 (0.018)	0.056 (0)
		5	0.942 (0.83)	0.932 (0.786)	0.63 (0.378)	0.296 (0.114)
		7	0.962 (0.878)	0.952 (0.856)	0.866 (0.684)	0.598 (0.354)
	n=1000	3	0.456 (0.066)	0.232 (0.016)	0.264 (0.018)	0.092 (0.002)
		5	0.98 (0.902)	0.976 (0.89)	0.82 (0.566)	0.522 (0.248)
		7	0.982 (0.934)	0.976 (0.926)	0.984 (0.848)	0.856 (0.68)
DGP 3	n=500	3	0.49 (0.068)	0.234 (0.016)	0.3 (0.01)	0.136 (0)
		5	0.996 (0.786)	0.988 (0.764)	0.578 (0.188)	0.306 (0.044)
		7	1 (0.826)	1 (0.822)	0.772 (0.418)	0.528 (0.172)
	n=1000	3	0.668 (0.098)	0.286 (0.032)	0.31 (0.008)	0.162 (0.004)
		5	0.998 (0.804)	0.994 (0.78)	0.626 (0.158)	0.396 (0.05)
		7	1 (0.838)	0.998 (0.836)	0.898 (0.482)	0.692 (0.276)
DGP 4	n=500	3	0.816 (0.41)	0.368 (0.186)	0.218 (0)	0.092 (0)
		5	0.63 (0.588)	0.18 (0.146)	0.622 (0)	0.358 (0)
		7	0.734 (0.666)	0.284 (0.254)	0.928 (0)	0.764 (0)
	n=1000	3	1 (0.348)	1 (0.348)	0.236 (0)	0.106 (0)
		5	1 (0.836)	1 (0.836)	0.668 (0)	0.4 (0)
		7	0.998 (0.93)	0.998 (0.93)	0.882 (0)	0.774 (0)
DGP 4*	n=500	3	0.414 (0)	0.186 (0)	-	-
		5	0.948 (0.006)	0.908 (0.004)	-	-
		7	0.862 (0.016)	0.858 (0.008)	-	-
	n=1000	3	0.46 (0.002)	0.262 (0)	-	
		5	0.984 (0.002)	0.954 (0)	-	-
		7	0.966 (0.008)	0.964 (0.002)	-	-

Table 15: Detection rate when two relative innovational outlier are present.

Diskussionspapiere 2012
Discussion Papers 2012

01/2012 **Wrede, Matthias**: Wages, Rents, Unemployment, and the Quality of Life

02/2012 **Schild, Christopher-Johannes:** Trust and Innovation Activity in European Regions - A Geographic Instrumental Variables Approach

03/2012 **Fischer, Matthias:** A skew and leptokurtic distribution with polynomial tails and characterizing functions in closed form

04/2012 **Wrede, Matthias**: Heterogeneous Skills and Homogeneous Land: Segmentation and Agglomeration

Diskussionspapiere 2011
Discussion Papers 2011

01/2011 **Klein, Ingo, Fischer, Matthias and Pleier, Thomas:** Weighted Power Mean Copulas: Theory and Application

02/2011 **Kiss, David:** The Impact of Peer Ability and Heterogeneity on Student Achievement: Evidence from a Natural Experiment

03/2011 **Zibrowius, Michael:** Convergence or divergence? Immigrant wage assimilation patterns in Germany

04/2011 **Klein, Ingo and Christa, Florian:** Families of Copulas closed under the Construction of Generalized Linear Means

05/2011 **Schnitzlein, Daniel**: How important is the family? Evidence from sibling correlations in permanent earnings in the US, Germany and Denmark

06/2011 **Schnitzlein, Daniel**: How important is cultural background for the level of intergenerational mobility?

07/2011 **Steffen Mueller**: Teacher Experience and the Class Size Effect - Experimental Evidence

08/2011 **Klein, Ingo:** Van Zwet Ordering for Fechner Asymmetry

09/2011 **Tinkl, Fabian and Reichert Katja:** Dynamic copula-based Markov chains at work: Theory, testing and performance in modeling daily stock returns

10/2011 **Hirsch, Boris and Schnabel, Claus:** Let's Take Bargaining Models Seriously: The Decline in Union Power in Germany, 1992 – 2009

11/2011 **Lechmann, Daniel S.J. and Schnabel, Claus** : Are the self-employed really jacks-of-all-trades? Testing the assumptions and implications of Lazear's theory of entrepreneurship with German data

12/2011 **Wrede, Matthias**: Unemployment, Commuting, and Search Intensity

13/2011 **Klein, Ingo**: Van Zwet Ordering and the Ferreira-Steel Family of Skewed Distributions

Diskussionspapiere 2010
Discussion Papers 2010

01/2010 **Mosthaf, Alexander, Schnabel, Claus and Stephani, Jens:** Low-wage careers: Are there dead-end firms and dead-end jobs?

02/2010 **Schlüter, Stephan and Matt Davison:** Pricing an European Gas Storage Facility using a Continuous-Time Spot Price Model with GARCH Diffusion

03/2010 **Fischer, Matthias, Gao, Yang and Herrmann, Klaus:** Volatility Models with Innovations from New Maximum Entropy Densities at Work

04/2010 **Schlüter, Stephan and Deuschle, Carola:** Using Wavelets for Time Series Forecasting – Does it Pay Off?

05/2010 **Feicht, Robert and Stummer, Wolfgang:** Complete closed-form solution to a stochastic growth model and corresponding speed of economic recovery.

06/2010 **Hirsch, Boris and Schnabel, Claus:** Women Move Differently: Job Separations and Gender.

07/2010 **Gartner, Hermann, Schank, Thorsten and Schnabel, Claus:** Wage cyclicality under different regimes of industrial relations.

08/2010 **Tinkl, Fabian**: A note on Hadamard differentiability and differentiability in quadratic mean.

Diskussionspapiere 2009
Discussion Papers 2009

Diskussionspapiere 2008
Discussion Papers 2008

01/2008 **Grimm, Veronika and Gregor Zoettl**: Strategic Capacity Choice under Uncertainty: The Impact of Market Structure on Investment and Welfare

02/2008 **Grimm, Veronika and Gregor Zoettl**: Production under Uncertainty: A Characterization of Welfare Enhancing and Optimal Price Caps

03/2008 **Engelmann, Dirk and Veronika Grimm**: Mechanisms for Efficient Voting with Private Information about Preferences

04/2008 **Schnabel, Claus and Joachim Wagner**: The Aging of the Unions in West Germany, 1980-2006

05/2008 **Wenzel, Tobias**: On the Incentives to Form Strategic Coalitions in ATM Markets

06/2008 **Herrmann, Klaus**: Models for Time-varying Moments Using Maximum Entropy Applied to a Generalized Measure of Volatility

07/2008 **Klein, Ingo and Michael Grottke:** On J.M. Keynes' “The Principal Averages and the Laws of Error which Lead to Them” - Refinement and Generalisation