Economics Bulletin

Volume 37, Issue 4

Improved two-component tests in Beta-Skew-t-EGARCH models

Fernanda Maria Müller
Federal University of Rio Grande do Sul

Fábio M Bayer
Federal University of Santa Maria

Abstract

This work proposes a likelihood ratio test to assist in the selection of the Beta-Skew-t-EGARCH model with one or two volatility components. To improve the performance of the proposed test in small samples, the bootstrap-based likelihood ratio test and the bootstrap Bartlett correction are considered. The finite sample performance of the tests are assessed using Monte Carlo simulations. The numerical evidence favors the bootstrap-based test. The tests are applied to the DAX log-returns. The results demonstrate the practical usefulness of the proposed two-component tests.

Acknowledgements We gratefully acknowledge partial financial support from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), Brazil.

Citation: Fernanda Maria Müller and Fábio M Bayer, (2017) "Improved two-component tests in Beta-Skew-t-EGARCH models", *Economics Bulletin*, Volume 37, Issue 4, pages 2364-2373

Contact: Fernanda Maria Müller - fernanda.muller@ufrgs.br, Fábio M Bayer - bayer@ufsm.br.

Submitted: April 24, 2017. **Published:** October 26, 2017.

1 Introduction

The Beta-Skew-t-EGARCH model (Harvey and Sucarrat, 2014) was proposed to model the volatility of financial returns. In part, the model is appealing because of its robustness to outliers and volatility jumps in addition to accommodating the conditional asymmetry, the leverage effect, and heavy tails. The model also enables to decompose the volatility into short- and long-term components (Harvey and Sucarrat, 2014, Sucarrat, 2013). The two-component structure of the model improves the adjustment quality, as it is capable of mimicking the long memory pattern present in the autocorrelations of absolute values (Harvey and Sucarrat, 2014).

In this paper, a likelihood ratio (LR) test is proposed to test whether the data series should be modeled through the Beta-Skew-t-EGARCH model with one or two volatility components. The proposed test is referred to as the two-component test. Under the null hypothesis and under sufficient regularity conditions and assumptions, the test statistic has an asymptotic chi-squared distribution (Canepa and Godfrey, 2007). However, in small samples, the approximation of the null distribution of the test statistic by the chi-squared limiting null distribution is limited and can render distorted null rejection rates (Ferrari et al., 2005, Canepa and Godfrey, 2007, Kascha and Trenkler, 2011, Stein et al., 2014, Bayer and Cribari-Neto, 2013). As an alternative, the bootstrap-based LR test (Efron, 1979) and bootstrap Bartlett correction (Rocke, 1989) are considered to improve the performance of the LR test in small samples. At the end of the work, an application to the log-returns of the German stock index DAX is presented. The numerical results confirm the good performance of the proposed tests, thus validating its practical usefulness.

This paper unfolds as follows. Section 2 presents one-component and two-component specifications of Beta-Skew-t-EGARCH model. In Section 3 we introduce the two-component test. We also present the bootstrap-based LR test and bootstrap Bartlett correction of LR statistic. In Section 4 Monte Carlo simulation results are presented and discussed. Section 5 presents an application of the Beta-Skew-t-EGARCH model and of the two-component tests to the log-returns of the DAX index. Section 6 concludes the paper.

2 Beta-Skew-t-EGARCH model

Beta-Skew-t-EGARCH model is a particular case of dynamic conditional score (DCS) models (Harvey, 2013), also known as generalized autoregressive score (GAS) (Creal et al., 2013), wherein the time-varying parameters are based on the score function at time t (Creal et al., 2013). The main advantage of this models is that the score function explores the complete structure of the density, rather than just the mean and higher moments (Harvey, 2013). By utilizing the conditional score it is possible to reduce the prediction error (Creal et al., 2011). In addition, it can insert asymmetry and long memory extensions easier than other class of models (Harvey, 2013).

Let y_t be a financial return at instant t, with $t = 1, \ldots, n$, where n is the sample size, the martingale difference version of the first order one-component Beta-Skew-t-EGARCH model is given by (Harvey and Sucarrat, 2014, Sucarrat, 2013):

$$y_t = \exp(\lambda_t)\varepsilon_t = \sigma_t\varepsilon_t, \quad \varepsilon_t \sim st(0,\sigma_\varepsilon^2,\nu,\gamma), \quad \nu > 2, \quad \gamma \in (0,\infty),$$

$$
\begin{aligned}
\lambda_t &= \omega + \lambda_t^{\dagger}, \\
\lambda_t^{\dagger} &= \phi_1 \lambda_{t-1}^{\dagger} + \kappa_1 u_{t-1} + \kappa^* \text{sgn}(-y_{t-1})(u_{t-1}+1), \quad |\phi_1| < 1,
\end{aligned} \tag{1}
$$

where σ_t is the conditional scale or volatility of y_t, $\text{sgn}(\cdot)$ is the sign function, and ε_t is the conditional error. Further, ω is interpreted as the long-term log-volatility, ϕ_1 is the GARCH parameter, κ_1 is the ARCH parameter, κ^* is the leverage parameter, ν corresponds to the degrees of freedom, γ is the asymmetry, and u_t is the conditional score defined by the derivative of the log-density of y_t in relation to λ_t, see expression (5) (Harvey and Sucarrat, 2014, Sucarrat, 2013). Skewness is introduced into the model using the method developed by Fernández and Steel (1998).

The martingale difference version of the first order two-component Beta-Skew-t-EGARCH model is defined by:

$$
\begin{aligned}
y_t &= \exp(\lambda_t)\varepsilon_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim st(0,\sigma_\varepsilon^2,\nu,\gamma), \quad \nu > 2, \quad \gamma \in (0,\infty), \\
\lambda_t &= \omega + \lambda_{1,t}^{\dagger} + \lambda_{2,t}^{\dagger}, \qquad (2) \\
\lambda_{1,t}^{\dagger} &= \phi_1 \lambda_{1,t-1}^{\dagger} + \kappa_1 u_{t-1} + \kappa^* \text{sgn}(-y_{t-1})(u_{t-1}+1), \quad |\phi_1| < 1, \\
\lambda_{2,t}^{\dagger} &= \phi_2 \lambda_{2,t-1}^{\dagger} + \kappa_2 u_{t-1}, \quad |\phi_2| < 1, \quad \phi_1 \neq \phi_2, \qquad (3)
\end{aligned}
$$

where $\lambda_{1,t}^{\dagger}$ is the short-term component, and $\lambda_{2,t}^{\dagger}$ is the long-term component, ϕ_1 and κ_1 are, respectively, GARCH and ARCH parameters for the short-term component, and ϕ_2 and κ_2 are GARCH and ARCH parameters for the long-term component.

The log-likelihood function of the model is defined by:

$$
\ell(\theta) = \sum_{t=1}^{n} \ln f_y(y_t),
$$

where $\theta = (\omega,\phi_1,\phi_2,\kappa_1,\kappa_2,\kappa^*,\nu,\gamma)$ is the parameter vector of the model and $\ln f_y(y_t)$ is the log-density of y_t, given by (Harvey and Sucarrat, 2014):

$$
\begin{aligned}
\ln f_y(y_t) = {} & \ln 2 - \ln(\gamma + \gamma^{-1}) + \ln \Gamma((\nu+1)/2) - \frac{1}{2}\ln \pi - \ln \Gamma(\nu/2) - \frac{1}{2} \ln \nu \\
& - \lambda_{t|t-1} - \frac{\nu+1}{2} \ln \left(1 + \frac{(y_t-\mu)^2}{\gamma^{2sgn(y_t-\mu)} \nu e^{2\lambda_{t|t-1}}}\right),
\end{aligned} \tag{4}
$$

where μ is a location parameter of y_t. The $\hat{\theta} = (\hat{\omega},\hat{\phi_1},\hat{\phi_2},\hat{\kappa_1},\hat{\kappa_2},\hat{\kappa^*},\hat{\nu},\hat{\gamma})$ values that maximize the log-likelihood function $\ell(\theta)$ are the maximum likelihood estimators (MLEs) of θ. Numerical methods are required to obtain $\hat{\theta}$ (Sucarrat, 2013).

The conditional score of the martingale difference version can be written by (Harvey and Sucarrat, 2014, Sucarrat, 2013):

$$
\begin{aligned}
\frac{\partial \ln f_y(y_t)}{\partial \lambda_t} &= u_t \\
&= \frac{(\nu+1)[y_t^2 + y_t \mu_{\varepsilon*} \exp(\lambda_t)]}{\nu \exp(2\lambda_t) \gamma^{2\text{sgn}(y_t + \mu_{\varepsilon*}\exp(\lambda_t))} + (y_t + \mu_{\varepsilon*} \exp(\lambda_t))^2} - 1,
\end{aligned} \tag{5}
$$

where $\varepsilon*$ is an uncentred skewed variable with mean $\mu_{\varepsilon*}$.

3 Two-component tests

In this section, we introduce the two-component test and its corrected versions for small sample sizes. The issue of interest is to test whether the parameters ϕ_2 and κ_2 in (3) are null; that is:

$$\begin{cases} \mathcal{H}_0 : (\phi_2, \kappa_2) = (0, 0), \\ \mathcal{H}_1 : (\phi_2, \kappa_2) \neq (0, 0). \end{cases} \tag{6}$$

Under $\mathcal{H}_0$, the adequate model is the one-component Beta-Skew-t-EGARCH model of the volatility in (1), and under $\mathcal{H}_1$, the two-component model in (2) must be adjusted.

To perform this test, the likelihood ratio statistic is initially considered, given by:

$$\mathrm{LR} = 2\left[\ell(\widehat{\theta}_1) - \ell(\widehat{\theta}_0)\right], \tag{7}$$

where $\widehat{\theta}_1$ represents the MLEs under an alternative hypothesis, and $\widehat{\theta}_0$ corresponds to the restricted MLE vector (under $\mathcal{H}_0$). The LR statistic under the null hypothesis follows an asymptotic chi-squared distribution with two degrees of freedom (Cordeiro and Cribari-Neto, 2014), subject to appropriate assumptions. However, in small samples, the test can present distorted null rejection rates.

3.1 Bootstrap-based LR test

A common approach to reduce the small size distortion of the LR test is the bootstrap (Efron and Tibshirani, 1993). The steps used to implement the bootstrap-based test (LR^{b}) can be summarized by the following algorithm (Efron and Tibshirani, 1993, Canepa and Godfrey, 2007):

1. Generate, under $\mathcal{H}_0$, B bootstrap resampling $y^{*1}, y^{*2}, \ldots, y^{*B}$ of the model using parametric bootstrapping, i.e, replacing the model parameters by the estimates in $\mathcal{H}_0$ using the original sample;

2. For each resample y^{*b}, where $b = 1, 2, \ldots, B$, the following statistic is computed

$$\mathrm{LR}^{*\mathrm{b}} = 2\left[\ell(\widehat{\theta}_1^{*b}) - \ell(\widehat{\theta}_0^{*b})\right],$$

   where $\widehat{\theta}_0^{*b}$ and $\widehat{\theta}_1^{*b}$ are the maximum likelihood estimators under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively;

3. Repeat steps 1 and 2 a large number of times B;

4. Compute the bootstrap p-value by:

$$p^* = \frac{\#\{LR^{*b} \geq LR\}}{B},$$

   where # denotes the cardinality of the set.

$\mathcal{H}_0$ will be rejected if p^* is smaller than the desired significance level (usually 0.05) (Canepa and Godfrey, 2007).

3.2 Bootstrap Bartlett correction

The performance of the LR test can be also improved in small samples by considering the Bartlett correction of the LR statistic (Bartlett, 1937, Lawley, 1956):

$$\mathrm{LR}_{\mathrm{Bartllet}} = \frac{\mathrm{LR}}{c},$$

where LR is the usual likelihood ratio statistic, $c = E(\mathrm{LR})/g$ is Bartlett's correction factor, and $g = 2$ is the number of restrictions imposed by $\mathcal{H}_0$. The distribution of the corrected $\mathrm{LR}_{\mathrm{Bartllet}}$ test statistic converges faster to the chi-squared limiting null distribution, thereby decreasing the test size distortions introduced by small samples (Bayer and Cribari-Neto, 2013, Rayner, 1990). However, the analytical derivation of Bartlett's correction factor involves cumulants and mixed cumulants up to the fourth order of the log-likelihood function. Such analytical derivation can be cumbersome or even unfeasible in some model classes (Bayer and Cribari-Neto, 2013), particularly in the Beta-Skew-t-EGARCH model, where even the first derivatives, useful in the process of maximizing $\ell(\theta)$, are numerically obtained (Sucarrat, 2013).

Alternatively, Bartlett's correction factor is obtained using the bootstrap method (Rocke, 1989), as follows:

1. Bootstrap resampling $y^{*1}, y^{*2}, \ldots, y^{*B}$ is generated under $\mathcal{H}_0$ using parametric bootstrapping;

2. For each resample y^{*b}, where $b = 1, 2, \ldots, B$, the following statistic is computed

$$\mathrm{LR}^{*\mathrm{b}} = 2\left[\ell(\widehat{\theta}_1^{*b}) - \ell(\widehat{\theta}_0^{*b})\right],$$

   where $\widehat{\theta}_0^{*b}$ and $\widehat{\theta}_1^{*b}$ are the maximum likelihood estimators under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively;

3. The bootstrap Bartlett correction of LR is obtained by:

$$\mathrm{LR}_{\mathrm{B}} = \frac{2\mathrm{LR}}{\overline{\mathrm{LR}}^*},$$

   where $\overline{\mathrm{LR}}^* = \mathrm{B}^{-1}\sum_{b=1}^{B} \mathrm{LR}^{*b}$.

4 Numerical evaluation

We evaluate the finite sample performance of the proposed two-component tests, LR, LR^{b}, and LR_{B}, using Monte Carlo simulations. The number of Monte Carlo replications was set at 1000, and the number of bootstrap resamples was $B = 500$. The same number of Monte Carlo replications was utilized by Omtzigt and Fachin (2002) in bootstrap and Bartlett-corrected tests in cointegrating vectors and by Luger (2012) for hypothesis tests in GARCH models. In the works of Omtzigt and Fachin (2002) and Stein et al. (2014) they also used 500 bootstrap resamples for adjusted LR statistic.

The considered sample sizes are $n = 250$, 500, and 1000. All computational implementations were conducted using `R` programming language (R Core Team, 2017), and the package used to estimate the model parameters was `betategarch` (Sucarrat, 2013).

An R function has been provided at www.ufsm.br/bayer/dois-componentes-boot.zip to perform the proposed tests.

For the analysis of the null rejection rate (size) of the tests, nominal levels equal to 1%, 5%, and 10% were considered. This analysis evaluated the experiments presented in Table 1. The results are listed in Table 2.

Table 1: Considered experiments.

Experiment	ω	ϕ_1	κ_1	κ^*	ν	γ	Characteristic
1	0.1	0.95	0.05	0.02	10	0.8	benchmark
2	0.1	0.98	0.05	0.02	10	0.8	greater persistence
3	0.1	0.95	0.10	0.02	10	0.8	greater response to shocks
4	0.1	0.95	0.05	0.02	5	0.8	greater kurtosis
5	0.1	0.95	0.05	0.02	10	1.2	right-skewed

Table 2: Null rejection rates (%) of the proposed tests.

		1%			5%			10%		
Experiments	n	250	500	1000	250	500	1000	250	500	1000
1	LR	1.9	2.6	1.5	7.2	8.2	6.5	14.0	15.3	12.9
	LR^{b}	1.2	1.1	1.0	4.8	5.3	5.4	9.8	12.7	10.8
	LR_{B}	2.1	1.8	1.2	5.8	7.0	5.9	12.5	13.6	12.0
2	LR	2.0	2.4	1.4	9.0	9.8	7.1	15.5	17.6	13.0
	LR^{b}	0.7	1.5	0.8	4.4	4.9	4.9	10.6	10.8	9.5
	LR_{B}	1.2	1.8	1.1	6.1	6.4	5.7	12.6	13.7	11.0
3	LR	2.6	2.4	1.4	8.7	9.2	6.7	15.7	16.7	13.1
	LR^{b}	0.8	1.0	0.7	5.4	5.6	4.3	10.1	11.5	9.3
	LR_{B}	1.3	1.1	0.8	6.7	6.1	4.6	11.6	11.8	10.1
4	LR	2.8	1.3	1.0	10.2	7.3	5.6	16.9	13.7	11.2
	LR^{b}	1.4	0.6	0.5	6.1	4.4	4.3	11.3	10.5	9.4
	LR_{B}	2.3	0.8	0.6	7.6	6.0	5.1	13.2	11.3	10.3
5	LR	2.3	1.7	1.1	7.8	7.0	6.7	14.4	12.5	13.1
	LR^{b}	1.3	1.1	0.5	5.0	4.7	4.3	10.3	9.8	9.4
	LR_{B}	1.6	1.2	0.6	5.4	5.2	5.4	11.6	10.7	10.6

An analysis of the results in Table 2 verifies that the two-component test based on the LR statistic is more liberal than the tests based on LR^{b} and LR_{B} statistic. The biggest distortion in null rejection rate of LR is observed for the nominal level 1% and $n = 250$ and 500. For example, in Experiment 3 and $n = 250$, the null rejection rate of LR is 2.6%. For $n = 250$ and $n = 500$, the LR^{b} and LR_{B} tests considerably decreases the test size distortions, thereby exhibiting null rejection rates close to the nominal levels. For instance, in Experiment 5 with $n = 250$, the null rejection rates of the LR, LR^{b}, and LR_{B} were equal to 2.3%, 1.3%, and 1.6%, respectively.

We note that generating processes with higher kurtosis (Experiment 4) affect the performance of the LR test in the smaller sample sizes. In this experiment, the biggest distortions are observed for $n = 250$. For example, for the nominal level equals to 5% the null rejection rates of the tests are 10.2 %, 6.1 %, and 7.6 % for the LR, LR^{b}, and LR_{B}, respectively. In general, the bootstrap-based test presents better performance for controlling size distortion. Similar results are identified in Canepa and Godfrey (2007) for the quasi-likelihood ratio test in ARMA models and in Canepa (2016) for the LR test for linear restrictions on the cointegrating vectors. For larger sample sizes, it is clear that tests feature adequate rejection rates.

Our results indicate that the adjusted tests present similar performance, although they do have different features. Despite presenting similar results, the bootstrap Bartlett correction is computationally more efficient, requiring a lower number of bootstrap resamples than the usual bootstrap-based test (Rocke, 1989, Bayer and Cribari-Neto, 2013).

We considered a simulation study to evaluate the non-null rejection rates (power) of the tests. To a fair comparison, we must ensure that the tests have the same size. For this end, we ran a previews simulation in order to calibrate the null rejection rates (size) equal to 5% for all tests. Here we considered the Experiment 1 and $n = 1000$. Than we computed the rejection rates under the alternative hypothesis $\mathcal{H}_1 : (\phi_2,\kappa_2) = (\delta,\delta/30)$, for values of δ ranging from -0.96 to 0.96 by 0.16. We do not show the entire results by brevity. In general, we observed that the LR test is a little more powerful. This result was expected, because it is the most liberal one. Besides, we note that as δ gets away from zero the performance of the tests tends to be equal. For example, when $\delta = -0.96$ and $\delta = 0.16$ the non-null rejection rates were 70.20% and 28.20% for LR, 68.50% and 18.60% for LR^{b}, and 68.90% and 20.30% for LR_{B}, respectively.

Overall, based on the results, it is verified a good performance of the bootstrap-based test in small samples. The use of corrected tests can helps in the choice of the appropriate model, and consequently reduce model risk[1]. Therefore, we recommend its use in empirical analysis.

5 Illustrative example

This section presents an application to actual data of the Beta-Skew-t-EGARCH model and proposed tests. The data used are related to the German stock index DAX. The sampling period ranges from December 14, 2011 to April 2, 2015, totaling 840 daily observations (adjusted closing price). For the analysis, the log-returns were calculated, as follows: $y_t = \ln P_t - \ln P_{t-1}$, for $t = 1, \cdots, 840$, where P_t is the price in t. Diagnostic tests and descriptive statistics of the DAX log-returns indicate that the returns present asymmetry and heavy tails. This can also be observed when analyzing the fitted model (Table 3). These results favor the use of the Beta-Skew-t-EGARCH model.

Considering a 5% significance level, the proposed two-component tests were performed. The uncorrected test (LR) resulted in p-value $\simeq 0.002$, thus rejecting the null hypothesis. On the other hand, the LR^{b} resulted in p-value $\simeq 0.072$ and LR_{B} presented p-value $\simeq 0.904$, indicating that the one-component model (under $\mathcal{H}_0$) is adequate. These results are described in the Table 3. Note that the inference conclusions change when the different tests are considered. Because a moderate sample size for a financial series is being considered, namely $n = 839$, the bootstrap tests are assumed to be more accurate,

[1]Although there is no consensus in the literature to model risk definition, for Hull and Suo (2002), Giannetti et al. (2004), and Barrieu and Ravanelli (2015), model risk refers to the risk of the use of an inadequate or incorrect model.

Table 3: Fitted Beta-Skew-t-EGARCH models applied to the DAX log-returns and results of two-component tests proposed.

One-component Beta-Skew-t-EGARCH model								
	$\hat{\omega}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}^*$	$\hat{\nu}$	$\hat{\gamma}$
Coefficients	-4.607	0.947		0.057		0.063	5.281	0.873
Standard error	0.089	0.017		0.012		0.012	0.969	0.036
BIC								-6.327
Ljung-Box$_{10}$ $\hat{\varepsilon}_t^2$								5.153 (p-value = 0.881)
Lagrange multiplier test$_{10}$ $\hat{\varepsilon}_t$								4.930 (p-value = 0.896)
Two-component Beta-Skew-t-EGARCH model								
	$\hat{\omega}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}^*$	$\hat{\nu}$	$\hat{\gamma}$
Coefficients	-4.568	0.977	0.928	0.066	-0.035	0.078	5.437	0.866
Standard error	0.146	0.023	0.016	0.059	0.064	0.011	1.031	0.036
BIC								-6.326
Ljung-Box$_{10}$ $\hat{\varepsilon}_t^2$								5.423 (p-value = 0.862)
Lagrange multiplier test$_{10}$ $\hat{\varepsilon}_t$								5.461 (p-value = 0.858)
Two-component tests								
LR p-value								$\simeq$ 0.002
LR$^{\text{b}}$ p-value								$\simeq$ 0.072
LR$_{\text{B}}$ p-value								$\simeq$ 0.904

$\hat{\varepsilon}_t$ = standardized residuals

and this is corroborated by diagnostic analysis. The one-component model exhibited a lower Bayesian information criterion (BIC) (Schwarz, 1978), and at the usual significance levels, the $\hat{\kappa}_1$ and $\hat{\kappa}_2$ estimates were not significant in the two-component model.

The estimates and some diagnostic measures of the fitted models, such as the Ljung-Box (Ljung and Box, 1978) test applied to the squared standardized residuals and Lagrange multiplier (Engle, 1982) test applied to the standardized residuals[2], are presented in Table 3. The tests indicated the goodness of the fitted models. Figure 1 presents the estimated conditional deviation and a series of DAX log-returns.

6 Final considerations

We proposed a likelihood ratio test and adjusted versions using the bootstrap Bartlett correction and bootstrap-based test to assist in the selection of the Beta-Skew-t-EGARCH model. For the proposed tests, the numerical results showed that larger sample sizes performed well, although the bootstrap-based test had produced less distorted results with smaller samples. An application to daily DAX log-returns also confirms the adequacy of the adjusted tests. On the basis of these results, one can ascertain that the proposed two-component tests are good alternatives for the selection of the Beta-Skew-t-EGARCH model in practical applications.

[2]Ljung-Box test applied to the squared standardized residuals and Lagrange multiplier test applied to the standardized residuals are recommended to analyze the presence of conditional heteroskedasticity in the residuals. For more details see Tsay (2014).

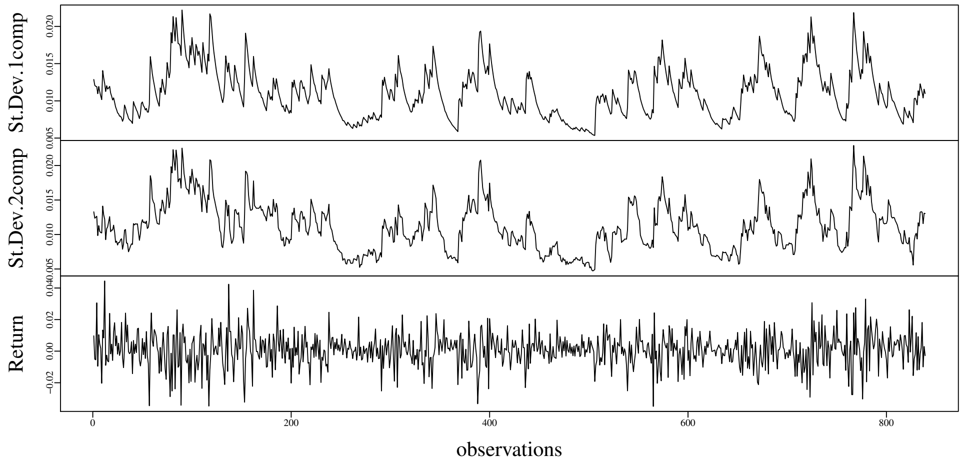

Figure 1: Fitted conditional standard deviation of the one-component model and two-component model and DAX log-returns.

References

Barrieu, P. and Ravanelli, C. (2015), ‘Robust capital requirements with model risk’, *Economic Notes* **44**(1), 1–28.

Bartlett, M. S. (1937), ‘Properties of sufficiency and statistical tests’, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **160**, 268–282.

Bayer, F. M. and Cribari-Neto, F. (2013), ‘Bartlett corrections in beta regression models’, *Journal of Statistical Planning and Inference* **143**, 531–547.

Canepa, A. (2016), ‘A note on Bartlett correction factor for tests on cointegrating relations’, *Statistics & Probability Letters* **110**, 296 – 304.

Canepa, A. and Godfrey, L. G. (2007), ‘Improvement of the quasi-likelihood ratio test in ARMA models: some results for bootstrap methods’, *Journal of Time Series Analysis* **28**(3), 434–453.

Cordeiro, G. M. and Cribari-Neto, F. (2014), *An introduction to Bartlett correction and bias reduction*, Springer Briefs in Statistics.

Creal, D., Koopman, S. J. and Lucas, A. (2011), ‘A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations’, *Journal of Business & Economic Statistics* **29**(4), 552–563.

Creal, D., Koopman, S. J. and Lucas, A. (2013), ‘Generalized autoregressive score models with applications’, *Journal of Applied Econometrics* **28**, 777–795.

Efron, B. (1979), ‘Bootstrap methods: another look at the Jackknife’, *The Annals of Statistics* **7**, 1–26.

Efron, B. and Tibshirani, R. J. (1993), *An introduction to the bootstrap*, Chapman and Hall, New York.

Engle, R. F. (1982), 'Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation', *Econometrica* **50**, 987–1008.

Fernández, C. and Steel, M. (1998), 'On Bayesian modelling of fat tails and skewness', *Journal of the American Statistical Association* **93**(441), 359–371.

Ferrari, S. L. P., Silva, M. F. D. and Cribari-Neto, F. (2005), 'Adjusted profile likelihood for two-parameter exponential family models', *Communications in Statistics - Theory and Methods* **34**(2), 257–276.

Giannetti, A., Clark, J. M. and Anderson, R. I. (2004), 'Model risk and option hedging', *The Quarterly Review of Economics and Finance* **44**(5), 659–677.

Harvey, A. C. (2013), *Dynamic models for volatility and heavy tails*, Cambridge University Press, New York.

Harvey, A. and Sucarrat, G. (2014), 'EGARCH models with fat tails, skewness and leverage', *Computational Statistics & Data Analysis* **76**, 320–338.

Hull, J. and Suo, W. (2002), 'A methodology for assessing model risk and its application to the implied volatility function model', *Journal of Financial and Quantitative Analysis* **37**(02), 297–318.

Kascha, C. and Trenkler, C. (2011), 'Bootstrapping the likelihood ratio cointegration test in error correction models with unknown lag order', *Computational Statistics and Data Analysis* **55**, 1008–1017.

Lawley, D. N. (1956), 'A general method for approximating to the distribution of likelihood ratio criteria', *Biometrika* **43**, 295–303.

Ljung, G. M. and Box, G. E. P. (1978), 'On a measure of a lack of fit in time series models', *Biometrika* **65**(2), 297–303.

Luger, R. (2012), 'Finite-sample bootstrap inference in GARCH models with heavy-tailed innovations', *Computational Statistics & Data Analysis* **56**(11), 3198 – 3211.

Omtzigt, P. and Fachin, S. (2002), Bootstrapping and Bartlett corrections in the cointegrated VAR model, Economics and quantitative methods, Department of Economics, University of Insubria.
URL: *http://EconPapers.repec.org/RePEc:ins:quaeco:qf0212*

R Core Team (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *http://www.R-project.org/*

Rayner, R. K. (1990), 'Bartlett's correction and the bootstrap in normal linear regression models', *Economics Letters* **33**, 255–258.

Rocke, D. M. (1989), 'Bootstrap Bartlett adjustment in seemingly unrelated regression', *Journal of the American Statistical Association* **84**, 598–601.

Schwarz, G. (1978), 'Estimating the dimensional of a model', *Annals of Statistics* **6**(2), 461–464.

Stein, M. C., da Silva, M. F. and Duczmal, L. H. (2014), 'Alternatives to the usual likelihood ratio test in mixed linear models', *Computational Statistics and Data Analysis* **69**, 184–197.

Sucarrat, G. (2013), 'Betategarch: Simulation, estimation and forecasting of Beta-Skew-t-EGARCH models', *The R Journal* **5**(2), 137–147.

Tsay, R. S. (2014), *An introduction to analysis of financial data with R*, John Wiley & Sons.