Simulated Evidence on the Distribution of the Standardized One-Step-Ahead Prediction Errors in ARCH Processes

Degiannakis, Stavros and Xekalaki, Evdokia

Department of Statistics, Athens University of Economics and Business, Greece, Department of Statistics, University of California, Berkeley CA, USA

2007

Simulated Evidence on the Distribution of the Standardized One-Step-Ahead Prediction Errors in ARCH Processes

Stavros Degiannakis
Adjunct Lecturer
Department of Statistics, Athens University of Economics and Business, Greece.
Tel.: +30-210-8203-120
FAX.: +30-210-2614-644
E-mail address: sdegia@aueb.gr
and
Adjunct Assistant Professor
Department of Informatics and Biomedicine, University of Central Greece, Lamia.
Tel.: + 30-223-1066-700
Fax.: +30-223-1066-715
Corresponding Author

Evdokia Xekalaki
Professor
Department of Statistics, Athens University of Economics and Business, Greece.
Tel.: +30-210-8203-269
Fax.: +30-210-8238-798
E-mail address: exek@aueb.gr
and
Department of Statistics, University of California, Berkeley CA, USA.
Tel.: +510-642-2781
Fax.: +510-642-7892

ABSTRACT

In statistical modeling contexts, the use of one-step-ahead prediction errors for testing hypotheses on the forecasting ability of an assumed model has been widely considered. Quite often, the testing procedure requires independence in a sequence of recursive standardized prediction errors, which cannot always be readily deduced particularly in the case of econometric modeling. In this paper, the results of a series of Monte Carlo simulations reveal that independence can be assumed to hold.

Index terms: ARCH models, Monte Carlo Simulation, One-step-ahead Prediction Errors, Predictability, Standardized Prediction Error Criterion.

* An earlier version of the paper was presented at the 7th Hellenic-European conference on computer mathematics and its applications (Degiannakis and Xekalaki 2005c).

I. INTRODUCTION

Defining a standardized prediction error criterion (SPEC), Degiannakis and Xekalaki (2005a) proposed a model selection algorithm for ARCH models. The algorithm allows switching from the model used at time $t-1$ for forecasting volatility to another model for use at time t and, in particular, to the model with the minimum value of the average squared standardized prediction error. As indicated by the results obtained by Degiannakis and Xekalaki (2005b), the SPEC model selection procedure appears to have a satisfactory performance in selecting the model that generates *better* volatility predictions. Moreover, the SPEC algorithm exhibited a satisfactory performance on a simulated options market (Xekalaki and Degiannakis 2005) as well as on trading S&P500 options on a daily basis (Degiannakis and Xekalaki 2001). The general finding is that the prediction performance improves if one switches models over time. In particular, switching from one model to another governed by the SPEC model selection rule appears to lead to a superior predictive performance. The reason might be traced in that jumping from one model to the other according to SPEC reflects a sort of a procedure adapting to the changes of the marketplace. However, model selection procedures based on standardized one-step-ahead prediction errors often require independence in a sequence of recursive standardized prediction errors, which cannot always be readily deduced particularly in the case of econometric modeling. In this paper, on the basis of the results of a series of Monte Carlo simulations, it is conjectured that independence holds. A theoretical justification can be found in Degiannakis and Xekalaki (2005a).

II. THE ARCH PROCESS

An ARCH process, ε_t, is presented as:

$$\begin{aligned} \varepsilon_t &= z_t \sigma_t \\ z_t &\overset{i.i.d.}{\sim} N(0,1) \\ \sigma_t^2 &= g(\sigma_{t-1}, \sigma_{t-2}, \ldots \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) \end{aligned} \quad (1)$$

where z_t is a sequence of independently and identically distributed random variables, with autocorrelation, $Cor(z_t, z_{t+\tau})$, approximately $N(0, T^{-1})$ distributed, σ_t is a time-varying, positive measurable function of the information set at time $t-1$ and $g(.)$ could be a functional form that has been presented in the ARCH literature.

Since very few financial time series have a constant conditional mean of zero, an ARCH model can be presented in a κ^{th} order autoregressive form by letting ε_t be the innovation process in a linear regression:

$$
\begin{aligned}
y_t &= \sum_{i=1}^{\kappa}\left(c_i y_{t-i}\right) + \varepsilon_t \\
\varepsilon_t \mid I_{t-1} &\sim N\left(0, \sigma_t^2\right) \\
\varepsilon_t &= z_t \sigma_t \\
z_t &\overset{i.i.d.}{\sim} N(0,1) \\
\sigma_t^2 &= g\left(\sigma_{t-1}, \sigma_{t-2}, \ldots, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\right)
\end{aligned}
\qquad (2)
$$

The most commonly used conditional variance function is the GARCH(1,1) model: $\sigma_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + b_1\sigma_{t-1}^2$.

A wide range of proposed ARCH models is covered in surveys such as Bollerslev et al. (1994) and Degiannakis and Xekalaki (2004).

III. Simulation of the AR(1)GARCH(1,1) Process

In the sequel, a Monte Carlo simulation is used to provide evidence for the assumption of independently and identically distributed standardized one-step-ahead prediction errors. The procedure consists of three stages:

1. *Generate data from the AR(1)GARCH(1,1) process*

- Generate a series of 32000 values from the standard normal distribution, i.e. $z_t \overset{i.i.d.}{\sim} N(0,1)$.
- Generate an equal number of values $\{\varepsilon_t\}_{t=1}^{32000}$ of the innovation ARCH process, by multiplying the collection $\{z_t\}_{t=1}^{32000}$ by a specific conditional variance form, or $\varepsilon_t = z_t\sqrt{\sigma_t^2}$, for $\sigma_t^2 = 0.0001 + 0.12\varepsilon_{t-1}^2 + 0.8\sigma_{t-1}^2$.
- Generate a first order autoregressive processes, $y_t = 0.06y_{t-1} + \varepsilon_t$, for the conditional mean, based on the values $\{\varepsilon_t\}_{t=1}^{32000}$ of the innovation process.

Panel A of Figure 1 plots the simulated processes while panel A of Table 1 presents the relevant descriptive statistics. According to results obtained in literature (e.g. Engle and Mustafa 1992), the shocks to the variance, $E_t\left(\varepsilon_t^2\right) - E_{t-1}\left(\varepsilon_t^2\right) = \varepsilon_t^2 - \sigma_t^2 \equiv v_t$, generate a martingale difference sequence. These shocks are neither serially independent nor identically distributed. According to the Brock et al.'s (1996) BDS test for independence only the process defined by z_t is independently distributed. The test is presented for two correlated dimensions but it has been computed for higher values and the results are qualitatively unchanged. Panel A of Figure 2 presents the autocorrelation of transformations of the processes defined by z_t, v_t, ε_t. The half-length of the 95% confidence interval for the estimated sample autocorrelation equals $1.96/\sqrt{T} = 0.0113$, in the case of a process with independently and identically normally distributed components. On the other hand, the processes defined by v_t and ε_t are autocorrelated in half of the cases. Ding and Ganger (1996) give the autocorrelation function of the squared errors for the

GARCH(1,1) process and Karanasos (1999) extends the results to the GARCH(p,q) model. He and Teräsvirta (1999) derive the autocorrelation function of the squared and absolute errors for a family of first order ARCH processes.

2. *Estimate the parameters of the AR(1)GARCH(1,1) model*

- The AR(1)GARCH(1,1) model is applied, for the data produced from the AR(1)GARCH(1,1) process. Dropping out the first 1000 data, maximum likelihood estimates of the parameters are obtained by numerical maximization of the log-likelihood function, using a rolling sample of constant size equal to 1000. At each of a sequence of points in time, the maximum likelihood parameter vector, $\hat{\theta}_t \equiv \left(\hat{c}_{1,t}, \hat{a}_{0,t}, \hat{a}_{1,t}, \hat{b}_{1,t}\right)$, is being estimated in order to compute the conditional mean and variance:

$$\begin{aligned} \hat{y}_{t+1|t} &= \hat{c}_{1,t} y_t \\ \hat{\sigma}^2_{t+1|t} &= \hat{a}_{0,t} + \hat{a}_{1,t}\varepsilon^2_{t|t} + \hat{b}_{1,t}\sigma^2_{t|t}. \end{aligned} \tag{3}$$

3. *Compute the standardized one-step-ahead prediction errors,* $\hat{z}_{t+1|t} = \left(y_{t+1} - \hat{y}_{t+1|t}\right)\hat{\sigma}^{-1}_{t+1|t}$

According to Degiannakis and Xekalaki (2005a), under the assumption of constancy of parameters over time, $\left(\hat{\theta}_t\right) = \left(\hat{\theta}_{t+1}\right) = \ldots = \left(\hat{\theta}_T\right) = \left(\hat{\theta}\right)$, the estimated standardized one-step-ahead prediction errors $\hat{z}_{t+1|t}, \hat{z}_{t+2|t+1}, \ldots, \hat{z}_{T+1|T}$ are asymptotically independently standard normally distributed.

- The one-step-ahead estimated processes are presented in Panel B of Figure 1, while Panel B of Table 1 presents the relevant descriptive statistics. According to the tests of normality and independence, the one-step-ahead standardized prediction error process, $\hat{z}_{t+1|t} = \left(y_{t+1} - \hat{y}_{t+1|t}\right)\hat{\sigma}^{-1}_{t+1|t}$, is independently normal distributed. Moreover, if $\hat{z}_{t+1|t} \overset{i.i.d.}{\sim} N(0,1)$, then $\sum_{t=1}^{T} \hat{z}^2_{t+1|t}$ should be chi-square distributed with T degrees of freedom, and mean and variance:

$$E\left[\sum_{t=1}^{T} \hat{z}^2_{t+1|t}\right] = T \text{ and } V\left[\sum_{t=1}^{T} \hat{z}^2_{t+1|t}\right] = 2T. \tag{4}$$

According to Table 2, which presents the descriptive statistics of $\sum_{t=1}^{T} \hat{z}^2_{t+1|t}$, the processes are chi-square distributed in all the cases. Moreover, if $\hat{z}_{t+1|t}$ is a sequence of i.i.d. variables then the autocorrelation of any transformation of $\hat{z}_{t+1|t}$, $Cor\left(\left|\hat{z}_{t+1|t}\right|^d, \left|\hat{z}_{t+\tau+1|t+\tau}\right|^d\right)$, $\forall d > 0$, is $N\left(0, T^{-1}\right)$ distributed. Panel B of Figure 2 presents the autocorrelation of transformations of the processes $\hat{z}_{t+1|t}$, $\hat{\varepsilon}_{t+1|t}$, $\hat{v}_{t+1|t}$. Since the sum of squared standardized one-step-ahead prediction errors is chi-square distributed, and the transformations of $\hat{z}_{t+1|t}$ are not autocorrelated, our findings point towards the independence of the standardized one-step-ahead innovations, $\hat{z}_{t+1|t}$.

IV. Simulation of the GARCH, EGARCH and TARCH Processes

In the sequel, the assumption that the standardized one-step-ahead prediction errors are independently and identically distributed is investigated for higher order of autoregressive processes for the conditional mean and conditional variance functions of the following types:

The GARCH(p,q) model, Bollerslev (1986):

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2\right) + \sum_{i=1}^{p}\left(b_i \sigma_{t-i}^2\right). \tag{5}$$

The EGARCH(p,q) model, Nelson (1991):

$$\ln\left(\sigma_t^2\right) = a_0 + \sum_{i=1}^{q}\left(a_i \left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| + \gamma_i \left(\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right)\right) + \sum_{i=1}^{p}\left(b_i \ln\left(\sigma_{t-i}^2\right)\right). \tag{6}$$

The TARCH(p,q) model, Glosten et al. (1993):

$$\sigma_t^2 = a_0 + \sum_{i=1}^{q}\left(a_i \varepsilon_{t-i}^2\right) + \gamma_1 \varepsilon_{t-1}^2 d_{t-1} + \sum_{i=1}^{p}\left(b_i \sigma_{t-i}^2\right), \tag{7}$$

where $d_t = 1$ if $\varepsilon_t < 0$, and $d_t = 0$ otherwise.

The procedure followed is comprised of the following steps:

1. *Eight processes have been generated with coefficients presented in Table 3*

2. *Estimate the parameters of the simulated processes*

- At each of a sequence of points in time, the maximum likelihood parameter vector $\hat{\theta}_t \equiv \left(\hat{c}_{1,t}, \hat{c}_{2,t}, \hat{c}_{3,t}, \hat{a}_{0,t}, \hat{a}_{1,t}, \hat{a}_{2,t}, \hat{b}_{1,t}, \hat{\gamma}_{1,t}\right)$ is being estimated. The models are estimated 30000 times and the conditional mean and variance are computed in (8)-(11):

The κ^{th} order Autoregressive process:

$$\hat{y}_{t+1|t} = \sum_{i=1}^{\kappa}\left(\hat{c}_{i,t} y_{t+1-i}\right). \tag{8}$$

The GARCH(1,q) model:

$$\hat{\sigma}_{t+1|t}^2 = \hat{a}_{0,t} + \sum_{i=1}^{q}\left(\hat{a}_{i,t} \varepsilon_{t-i+1|t}^2\right) + \hat{b}_{1,t} \sigma_{t|t}^2. \tag{9}$$

The EGARCH(1,1) model:

$$\hat{\sigma}_{t+1|t}^2 = \exp\left(\hat{a}_{0,t} + \hat{a}_{1,t}\left|\frac{\varepsilon_{t|t}}{\sigma_{t|t}}\right| + \hat{\gamma}_{1,t}\left(\frac{\varepsilon_{t|t}}{\sigma_{t|t}}\right) + \hat{b}_{1,t} \ln\left(\sigma_{t|t}^2\right)\right). \tag{10}$$

The TARCH(1,q) model:

$$\hat{\sigma}_{t+1|t}^2 = \hat{a}_{0,t} + \sum_{i=1}^{q}\left(\hat{a}_{i,t} \varepsilon_{t-i+1|t}^2\right) + \hat{\gamma}_{1,t} \varepsilon_{t|t}^2 d_t + \hat{b}_{1,t} \sigma_{t|t}^2. \tag{11}$$

3. *Compute the standardized one-step-ahead prediction errors* $\hat{z}_{t+1|t} = \left(y_{t+1} - \hat{y}_{t+1|t}\right)\hat{\sigma}_{t+1|t}^{-1}$

Due to space limitations all the relative information for each of the eight generated processes are available upon request. The evidence from our findings is in support of the hypothesis of independently and identically distributed standardized one-step-ahead prediction errors in this case too.

Finally, one more set of GARCH(1,1) processes is simulated in order to investigate if changes in the coefficients affect the distribution of $\hat{z}_{t+1|t}$. We generate a series $\{\varepsilon_t\}_{t=1}^{20000}$ of 20000 values for each of 18 innovation GARCH(1,1) processes by multiplying the generated values of z_t by σ_t from $\sigma_t^2 = 0.002 + 0.05\varepsilon_{t-1}^2 + b_1^{(k)}\sigma_{t-1}^2$, where $b_1^{(k)} = 0.05 * k$ for $k = 1,2,\ldots,18$ .There is no evidence against the property of independently distributed standardized prediction errors.

V. CONCLUSION

The findings are in support of the hypothesis of independence of the $\hat{z}_{t+1|t}$. Moreover, changes in the types of conditional variance function, the order of the autoregressive process of the conditional mean and as well as the values of the coefficients considered do not appear to affect these findings.

REFERENCES

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, **31**, 307-327.

Bollerslev, T., R.F., Engle and D., Nelson (1994). ARCH Models. Ch.49 in R. F. Engle and D. L. McFadden eds. *Handbook of Econometrics, IV Elsevier.*

Brock, W., D., Dechert, J., Sheinkman and B., LeBaron (1996). A Test for Independence Based on the Correlation Dimension, *Econometric Reviews*, **15(3)**, 197-235.

Degiannakis, S. and E., Xekalaki, (2001). Using a Prediction Error Criterion for Model Selection in Forecasting Option Prices, *Athens University of Economics and Business, Department of Statistics*, Technical Report, **131**.

Degiannakis, S. and E., Xekalaki (2004). Autoregressive Conditional Heteroscedasticity Models: A Review. *Quality Technology and Quantitative Management*, **1(2)**, 271-324.

Degiannakis, S. and E., Xekalaki (2005a). Predictability and Model Selection in the Context of ARCH Models, *Journal of Applied Stochastic Models in Business and Industry*, **21**, 55-82.

Degiannakis, S. and E., Xekalaki (2005b). Assessing the Performance of a Prediction Error Criterion Model Selection Algorithm in the Context of ARCH Models, *Applied Financial Economics*, forthcoming.

Degiannakis, S. and E., Xekalaki (2005c). On the Independence of the Standardized One-Step-Ahead Prediction Errors in ARCH Models. Paper presented at the *7th Hellenic-European Conference on Computer Mathematics and its Application*s, Athens, Greece.

Ding, Z. and C.W.J., Granger (1996). Modeling Volatility Persistence of Speculative Returns: A New Approach. *Journal of Econometrics*, **73**, 185-215.

Engle, R.F. and C., Mustafa (1992). Implied ARCH models from options prices, *Journal of Econometrics,* **52**, 289-311.

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

He, C. and T., Teräsvirta (1999). Properties of Moments of a Family of GARCH Processes. *Journal of Econometrics*, **92**, 173-192.

Karanasos, M. (1999). The Second Moment and the Autoregressive Function of the Squared Errors of the GARCH Model. *Journal of Econometrics*, **90**, 63-76.

Nelson, D.B. (1991). Conditional Heteroscedasticity in Asset Returns: A New Approach. *Econometrica*, **59**, 347-370.

Xekalaki, E. and S., Degiannakis (2005). Evaluating Volatility Forecasts in Option Pricing in the Context of a Simulated Options Market. *Computational Statistics and Data Analysis*. Special Issue on Computational Econometrics, **49(2)**, 611-629.

FIGURES & TABLES

Table 1. Descriptive statistics of the simulated processes (Panel A) and the one-step-ahead estimated processes (Panel B). The Crámer-Von Misses and Anderson-Darling statistics test the null hypothesis that the process is normally distributed. The BDS statistic tests the null hypothesis that the process is independently and identically distributed.

	Panel A			Panel B		
	$\{z_t\}_{t=2000}^{32000}$	$\{\varepsilon_t\}_{t=2000}^{32000}$	$\{y_t\}_{t=2000}^{32000}$	$\{\hat{z}_{t+1\|t}\}_{t=1}^{30000}$	$\{\hat{\varepsilon}_{t+1\|t}\}_{t=1}^{30000}$	$\{\hat{y}_{t+1\|t}\}_{t=1}^{30000}$
Mean	0.006139	0.000218	0.000232	0.006218	0.000214	1.73E-05
Median	0.001398	3.81E-05	6.16E-05	0.003156	9.65E-05	4.14E-07
Std. Dev.	0.999047	0.035500	0.035562	1.004477	0.035514	0.002390
Skewness	0.015896	0.017497	0.023013	0.016695	0.014996	0.086885
Kurtosis	3.004320	3.703915	3.712639	3.027268	3.699045	5.895584
Crámer-Von Misses	0.039066	2.175911	2.205750	0.046491	2.177987	68.12699
[p-value]	0.6992	0.00	0.00	0.5643	0.00	0.00
Anderson-Darling	0.241252	14.30484	14.52898	0.315299	14.29365	361.4363
[p-value]	0.7728	0.00	0.00	0.5430	0.00	0.00
BDS	8.13E-05	0.011808	0.012084	2.19E-05	0.011849	0.033372
[p-value]	0.8244	0.00	0.00	0.9526	0.00	0.00

Table 2. Descriptive statistics of $\left\{\sum_{j=t-T+1}^{t} \hat{z}_{j+1|j}^{2}\right\}$, for $t = T(T)30000$. The Crámer-Von Misses and Anderson-Darling statistics test the null hypothesis that the process is chi-squared distributed.

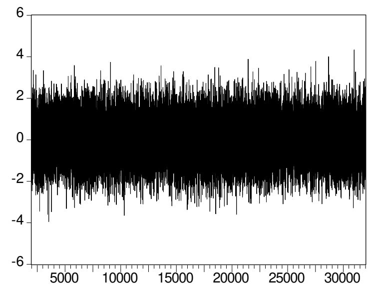
	$T=2$	$T=4$	$T=10$	$T=20$
Mean	2.017960	4.035919	10.08980	20.17960
Variance	4.128895	8.334519	21.34030	41.31167
Crámer-Von Misses	0.110346	0.080465	0.121311	0.112239
[p-value]	0.5364	0.6889	0.4898	0.5274
Anderson-Darling	1.057137	0.590629	1.112229	0.686104
[p-value]	0.3286	0.6569	0.3034	0.5705
Observations	15000	7500	3000	1500

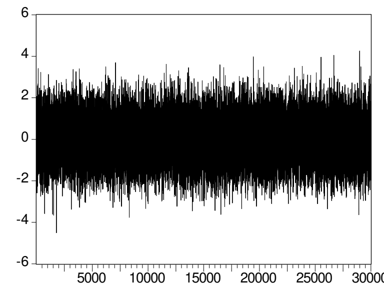
Table 3. Coefficients of the simulated processes.

Model	Parameters							
	c_1	c_2	c_3	a_0	a_1	a_2	b_1	γ_1
AR(1)GARCH(1,1)	0.05	-	-	0.002	0.05	-	0.91	-
AR(1)EGARCH(1,1)	0.05	-	-	0.2	0.05	-	0.2	0.1
AR(1)TARCH(1,1)	0.05	-	-	0.002	0.15	-	0.7	-0.08
AR(1)GARCH(1,2)	0.05	-	-	0.002	0.05	0.08	0.8	-
AR(1)TARCH(1,2)	0.05	-	-	0.002	0.15	0.05	0.7	-0.08
AR(3)GARCH(1,1)	0.1	0.03	-0.02	0.002	0.05	-	0.91	-
AR(3)EGARCH(1,1)	0.12	0.07	-0.03	0.001	0.05	-	0.2	0.1
AR(3)TARCH(1,1)	0.1	0.03	-0.02	0.002	0.15	-	0.7	-0.08

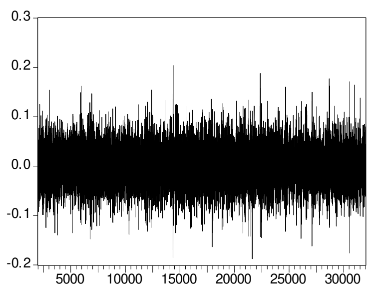
Figure 1. The simulated processes (Panel A) and the one-step-ahead estimated processes (Panel B)

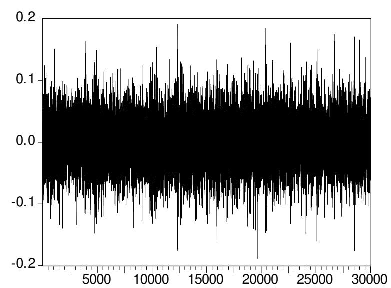
Panel A	Panel B
$\{z_t\}_{t=2000}^{32000}$ process	$\{\hat{z}_{t+1\|t}\}_{t=1}^{30000}$ process



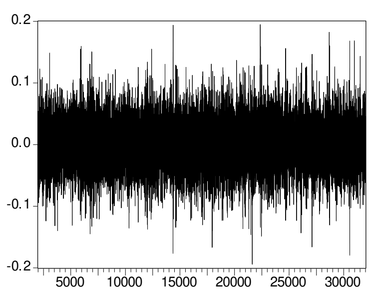


Panel A	Panel B
$\{\varepsilon_t\}_{t=2000}^{32000}$ process	$\{\hat{\varepsilon}_{t+1\|t}\}_{t=1}^{30000}$ process





Panel A	Panel B
$\{y_t\}_{t=2000}^{32000}$ process	$\{\hat{y}_{t+1\|t}\}_{t=1}^{30000}$ process



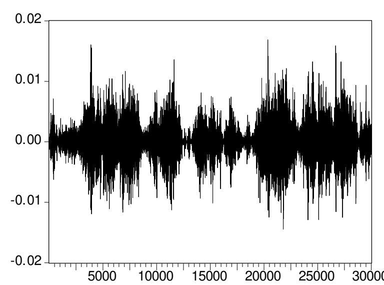
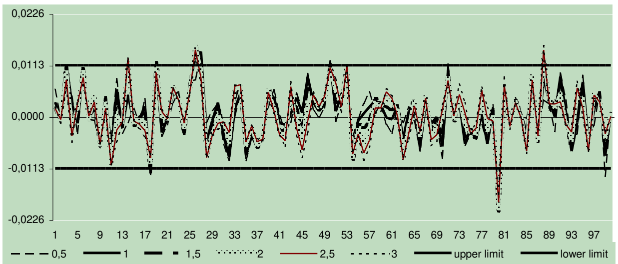


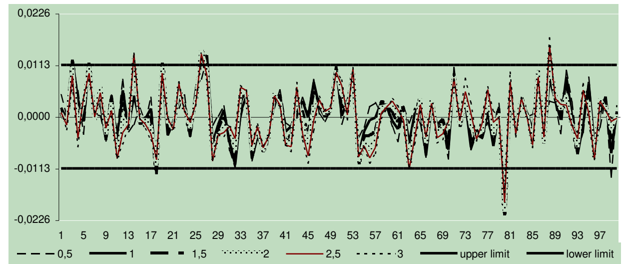
Figure 2. Autocorrelation of transformations of the processes z_t, ε_t, v_t (Panel A) and $\hat{z}_{t+1|t}$, $\hat{\varepsilon}_{t+1|t}$, $\hat{v}_{t+1|t}$ (Panel B)

Panel A	Panel B

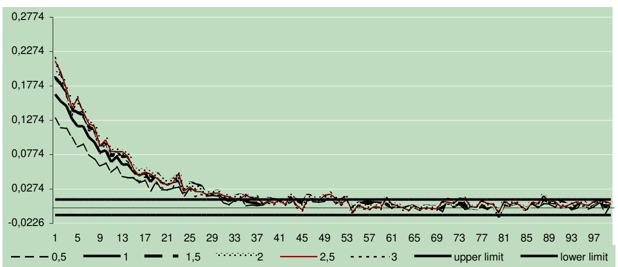
$Cor\left(\left|z_t\right|^d, \left|z_{t+\tau}\right|^d\right)$, $d = 0.5(0.5)3$, $\tau = 1(1)100$



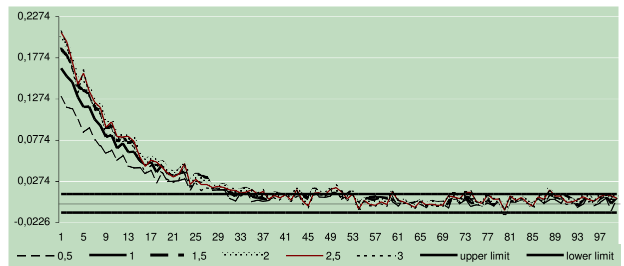
$Cor\left(\left|\hat{z}_{t+1|t}\right|^d, \left|\hat{z}_{t+\tau+1|t+\tau}\right|^d\right)$, $d = 0.5(0.5)3$, $\tau = 1(1)100$



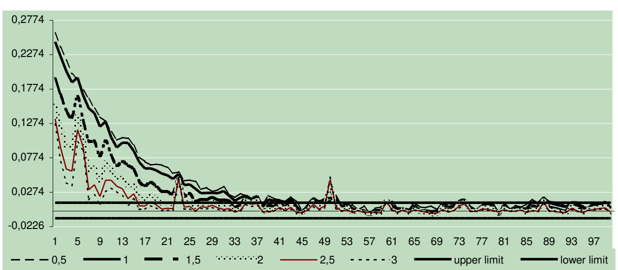
$Cor\left(\left|\varepsilon_t\right|^d, \left|\varepsilon_{t+\tau}\right|^d\right)$, $d = 0.5(0.5)3$, $\tau = 1(1)100$



$Cor\left(\left|\hat{\varepsilon}_{t+1|t}\right|^d, \left|\hat{\varepsilon}_{t+\tau+1|t+\tau}\right|^d\right)$, $d = 0.5(0.5)3$, $\tau = 1(1)100$



$Cor\left(\left|v_t\right|^d, \left|v_{t+\tau}\right|^d\right)$, $d = 0.5(0.5)3$, $\tau = 1(1)100$



$Cor\left(\left|\hat{v}_{t+1|t}\right|^d, \left|\hat{v}_{t+\tau+1|t+\tau}\right|^d\right)$, $d = 0.5(0.5)3$, $\tau = 1(1)100$