Confidence Intervals for the Autocorrelations of the Squares of GARCH Sequences

Piotr Kokoszka[1], Gilles Teyssière[2], and Aonan Zhang[3]

[1] Mathematics and Statistics, Utah State University,
3900 Old Main Hill, Logan UT 84322-3900, USA,
piotr@stat.usu.edu
http://math.usu.edu/~piotr
[2] NBG Bank (Paris) & ASEF,
gilles@ehess.cnrs-mrs.fr
http://www.gillesteyssiere.net
[3] Mathematics and Statistics, Utah State University,
aonanzhang@cc.usu.edu

Abstract. We compare three methods of constructing confidence intervals for sample autocorrelations of squared returns modeled by models from the GARCH family. We compare the residual bootstrap, block bootstrap and subsampling methods. The residual bootstrap based on the standard GARCH(1,1) model is seen to perform best.

1 Introduction

The paper is concerned with assessing finite sample performance of several methods of finding confidence intervals for autocorrelations of squared returns on speculative assets. While the returns themselves are essentially uncorrelated and most econometric and financial models explicitly imply that they are so, their squares exhibit a rich dependence structure. The sample autocorrelation of squared returns $\hat{\rho}_{n,X^2}(\cdot)$ is a measure of volatility clustering. A large value for $\hat{\rho}_{n,X^2}(\cdot)$ is the evidence of either the presence of long memory in the volatility process, or the inadequacy of a GARCH(1,1) process to fit the data under investigation as the true process might be a non–homogeneous GARCH(1,1) process; see [6].

We compare the performance of the various methods by means of their empirical coverage probability (ECP). Suppose we have a method of constructing, say, a 95% confidence interval $(\hat{l}_n, \hat{u}_n)$ from an observed realization $X_1, X_2, \ldots, X_n$. We simulate a large number R of realizations from a specific GARCH type model from which we construct R confidence intervals $(\hat{l}_n^{(r)}, \hat{u}_n^{(r)})$, $r = 1, 2, \ldots, R$. The percentage of these confidence intervals that contain the population autocorrelation is the ECP, which we want to be as close as possible to the nominal coverage probability of 95%. Our objective is to provide answers to the following questions: Does any method have better ECP than the others? If not, what is the range of optimal applicability of each method? Is it better to use equal-tailed or symmetric confidence intervals (see Section 2.1)? How does the coverage depend

on the value of γ_{c2} ? For a given series length n, how should one choose the block length b for the block bootstrap and subsampling? For what lengths n do these methods yield useful confidence intervals?

The ultimate goal is to recommend a practical procedure for finding confidence intervals for squared autocorrelations which assumes minimal prior knowledge of the stochastic mechanism generating the returns.

For ease of reference, recall that the sample autocovariances of the squared returns are

$$\hat{\gamma}_{n,X^2}(h) = \frac{1}{n} \sum_{t=1}^{n-h} \left(X_t^2 - \frac{1}{n-h} \sum_{t=1}^{n-h} X_t^2 \right) \left(X_{t+h}^2 - \frac{1}{n-h} \sum_{t=h+1}^{n} X_t^2 \right), \quad (1)$$

the population autocovariances are $\gamma_{X^2}(h) = E\left[(X_0^2 - EX_0^2)(X_h^2 - EX_0^2)\right]$, while the corresponding autocorrelations (ACF) are

$$\hat{\rho}_{n,X^2}(h) = \frac{\hat{\gamma}_{n,X^2}(h)}{\hat{\gamma}_{n,X^2}(0)}, \quad \rho_{X^2}(h) = \frac{\gamma_{X^2}(h)}{\gamma_{X^2}(0)}. \quad (2)$$

In Section 2, we describe the three methods. Section 3 introduces the various GARCH models we use for the comparison. The results of our simulations are presented in Section 4 with broad conclusions summarized in Section 4.3.

2   Confidence Intervals for ACF of Squared Returns

2.1   Residual Bootstrap

To illustrate the idea, we consider the ARCH(1) model given by

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \omega + \alpha X_{t-1}^2. \quad (3)$$

As we will see in Section 3, the method can be readily extended to any parametric model defined by GARCH type equations by computing the residuals $\hat{Z}_t = X_t/\hat{\sigma}_t$. Since the conditional volatility σ_t^2 is a function of the model parameters, past observations and past innovations, $\hat{\sigma}_t^2$ can be computed recursively once parameter estimates are available. We then proceed as follows:

1. Estimate $\hat{\omega}$ and $\hat{\alpha}$ and compute $\hat{Z}_t = [\hat{\omega} + \hat{\alpha} X_{t-1}^2]^{-1/2} X_t$, with $X_0 = \bar{X}_t$. We use the quasi maximum likelihood estimators (QMLE's) of model parameters with the assumption that the innovations $Z_t \sim N(0,1)$.
2. Form B bootstrap realizations $X_t^2(b) = [\hat{\omega} + \hat{\alpha} X_{t-1}^2(b)]\hat{Z}_t^2(b)$, $t = 1, 2, \ldots, n$, where $\hat{Z}_1^2(b), \ldots \hat{Z}_n^2(b)$, $b = 1, 2, \ldots, B$, are the B bootstrap samples selected with replacement from the squared residuals $\hat{Z}_1^2, \ldots \hat{Z}_n^2$.
3. Calculate the bootstrap autocorrelations $\rho_{n,X^2}^{(b)}(1)$, $b = 1, 2, \ldots, B$ and use their empirical quantiles to find a confidence interval for $\rho_{n,X^2}(1)$.

We now enlarge on step 3). Denote by $F^*_{\rho(1)}$ the EDF (empirical distribution function) of the $\rho^{(b)}_{n,X^2}(1)$, $b = 1, 2, \ldots, B$. The $(\alpha/2)$th and $(1 - \alpha/2)$th quantiles of $F^*_{\rho(1)}$ will yield an *equal-tailed* $(1 - \alpha)$ level confidence interval. To construct a *symmetric* confidence interval centered at $\hat{\rho}_{n,X^2}(1)$, we need the empirical distribution $F^*_{\rho(1),|\cdot|}$ of the B values $|\rho^{(b)}_{n,X^2}(1) - \hat{\rho}_{n,X^2}(1)|$. Denote by $q_{|\cdot|}(1 - \alpha)$ the $(1 - \alpha)$ quantile of $F^*_{\rho(1),|\cdot|}$. Then the symmetric confidence interval is

$$\left(\hat{\rho}_{n,X^2}(1) - q_{|\cdot|}(1 - \alpha), \quad \hat{\rho}_{n,X^2}(1) + q_{|\cdot|}(1 - \alpha) \right).$$

A usual criticism of methods based on a parametric model is that misspecification can lead to large biases. In many applications however these biases have only negligible impact on a statistical procedure of interest. In our setting, we will see that the residual bootstrap confidence intervals based on a misspecified model can produce good coverage probabilities.

2.2   Block Bootstrap

In this section we describe how the popular block-bootstrap of [5] can be used to construct confidence intervals for autocorrelations. This method does not require a model specification, but it relies on a choice of the block size b which is often a difficult task. A good account of block bootstrap is given in [1].

Focusing again on lag one sample autocorrelation of the squared observations, we proceed as follows: having observed the sample $X^2_1, \ldots, X^2_n$, form the vectors $\mathbf{Y}_2 = [X^2_1, X^2_2]', \mathbf{Y}_3 = [X^2_2, X^2_3]', \ldots, \mathbf{Y}_n = [X^2_{n-1}, X^2_n]'$. There are $n - 1$ such vectors. Now choose a block length b and compute the number of blocks $k = [(n-1)/b] + 1$ (if $(n-1)/b$ is an integer we take $k = (n-1)/b$). Choose k blocks with replacement to obtain kb vectors. Choosing the k blocks corresponds to generating k observations from the uniform distribution on $\{2, 3, \ldots, n-b+1\}$. Denote these observations $j_1, j_2, \ldots, j_k$. We thus obtained the kb vectors $\mathbf{Y}_{j_1}, \mathbf{Y}_{j_1+1}, \ldots, \mathbf{Y}_{j_1+b-1}, \ldots, \mathbf{Y}_{j_k}, \mathbf{Y}_{j_k+1}, \ldots, \mathbf{Y}_{k_1+b-1}$. If $(n-1)/b$ is not an integer, remove the last few vectors to have exactly $n - 1$ vectors. This gives us the bootstrap vector process

$$\mathbf{Y}^*_2 = [X^{*2}_1, X^{*2}_2]', \mathbf{Y}^*_3 = [X^{*2}_2, X^{*2}_3]', \ldots, \mathbf{Y}^*_n = [X^{*2}_{n-1}, X^{*2}_n]'.$$

The bootstrap sample autocovariances are computed according to (1) with the X_t replaced by the X^*_t defined above. The empirical distribution of $\hat{\rho}^*_{n,X^2}(1)$ is then an approximation to the distribution of $\hat{\rho}_{n,X^2}(1)$. As described in Section 2.1, the quantiles of the empirical distribution of $|\hat{\rho}^*_{n,X^2}(1) - \hat{\rho}_{n,X^2}(1)|$ can be used to construct symmetric confidence intervals.

2.3   Subsampling

The subsampling methodology is described in detail in [7]. [8] investigated subsampling confidence intervals for autocorrelations of linear time series models like ARMA. We adapt their methodology to the squares of GARCH processes.

To lighten the notations, denote $U_t = X_t^2 - \frac{1}{n}\sum_{j=1}^{n} X_j^2$ and suppress the subscript X^2 in the following formulas in which use definitions (1) and (2). Set

$$s_n^2(h) = \frac{1}{n}\sum_{j=1}^{n-h}\left(U_{j+h} - \hat{\rho}_n(h)U_j\right)^2, \quad \hat{\sigma}_n^2(h) = \frac{s_n^2(h)}{\sum_{j=h}^{n} U_j^2} \qquad (4)$$

and consider the studentized statistic $\hat{\xi}_n = \frac{\hat{\rho}_n(h) - \rho_n(h)}{\hat{\sigma}_n(h)}$. To construct equal-tailed and symmetric confidence intervals, we would need to know the sampling distribution of $\hat{\xi}_n$ and $|\hat{\xi}_n|$, respectively. We use subsampling to approximate these distributions: Consider an integer $b < n$ and the $n - b + 1$ blocks of data $X_t, \ldots, X_{t+b-1}$, $t = 1, \ldots, n - b + 1$. From each of these blocks compute $\hat{\rho}_{b,t}(h)$ and $\hat{\sigma}_{b,t}(h)$ according to respectively (1), (2) and (4), but replacing the original data $X_1, \ldots, X_n$ by $X_t, \ldots, X_{t+b-1}$. Next, compute the subsampling counterpart of the studentized statistic $\hat{\xi}_{b,t}(h) = \dfrac{\hat{\rho}_{b,t}(h) - \hat{\rho}_n(h)}{\hat{\sigma}_{b,t}(h)}$ and construct the EDF

$$L_b(x) = \mathcal{N}_b^{-1}\sum_{t=1}^{n-b+1}\mathbf{1}\left\{\hat{\xi}_{b,t}(h) \le x\right\}, \quad L_{b,|\cdot|}(x) = \mathcal{N}_b^{-1}\sum_{t=1}^{n-b+1}\mathbf{1}\left\{|\hat{\xi}_{b,t}(h)| \le x\right\},$$

with $\mathcal{N}_b = n - b + 1$. The empirical quantiles of L_b and $L_{b,|\cdot|}$ allow us to construct, respectively, equal-tailed and symmetric confidence intervals. For example, denoting by $q_{b,|\cdot|}(1 - \alpha)$ the $(1 - \alpha)$th quantile of $L_{b,|\cdot|}$, a subsampling symmetric $1 - \alpha$ level confidence interval for $\rho_n(h)$ is

$$\left(\hat{\rho}_n(h) - \hat{\sigma}_n(h)q_{b,|\cdot|}(1 - \alpha), \quad \hat{\rho}_n(h) + \hat{\sigma}_n(h)q_{b,|\cdot|}(1 - \alpha)\right).$$

3   GARCH Models

We consider the general framework for GARCH models proposed and studied by [4]; see also references quoted therein. The observations X_t are thus assumed to satisfy $X_t = Z_t\sigma_t$, where Z_t is a sequence of independent identically distributed random variables with zero mean and $\sigma_t^2 = g(Z_{t-1}) + c(Z_{t-1})\sigma_{t-1}^2$. We considered only specifications in which the function $g(\cdot)$ is a constant and the Z_t are standard normal. Denoting $\gamma_{ci} = Ec^i(Z_t)$, [4] proved that under the above assumptions a sufficient and necessary condition for the existence of the $2m$th unconditional moment of X_t is $\gamma_{cm} = Ec_t^m < 1$. Thus, the fourth unconditional moment of X_t exits if and only if $\gamma_{c2} = Ec_t^2 \in [0, 1)$. We considered the following three specific models:

1. The standard GARCH$(1, 1)$ model, for which

$$c_{t-1} = \beta + \alpha Z_{t-1}^2, \quad \sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta\sigma_{t-1}^2. \qquad (5)$$

2. The GJR-GARCH$(1, 1)$ model, see [3], with

$$c_{t-1} = \beta + (\alpha + \phi I(Z_{t-1}))Z_{t-1}^2, \quad \sigma_t^2 = \omega + (\alpha + \phi I(Z_{t-1}))X_{t-1}^2 + \beta\sigma_{t-1}^2, \quad (6)$$

where $I(Z_{t-1}) = 1$ if $Z_{t-1} < 0$, and $I(Z_{t-1}) = 0$ otherwise.

3. The nonlinear GARCH(1,1) model (NL GARCH(1,1,2), see [2], with

$$c_{t-1} = \beta + \alpha(1 - 2\eta \, \text{sign}(Z_{t-1}) + \eta^2)Z_{t-1}^2;$$

$$\sigma_t^2 = \omega + \alpha(1 - 2\eta \, \text{sign}(Z_{t-1}) + \eta^2)X_{t-1}^2 + \beta\sigma_{t-1}^2. \tag{7}$$

Assuming the errors $Z_t \sim N(0,1)$, the values of γ_{c2} and $\rho_{X^2}(1)$ can be computed in a closed form. If we know the model parameters, we can calculate precisely the population autocorrelation $\rho_{X^2}(1)$ and the value of γ_{c2}.

For each of the three models, we considered five parameter choices, which we labeled as models 1 through 5. The lag one autocorrelations for these choices are, respectively, approximately .15, .22, .31, .4, .5. The corresponding values of γ_{c2} are respectively, approximately .1, .3, .5, .7, .9. To facilitate comparison, models with the same index have similar values of γ_{c2} and $\rho_{X^2}(1)$, e.g. standard GARCH and GJR-GARCH with index 3 both have $\gamma_{c2} \approx .5$ and $\rho_{X^2}(1) \approx .31$.

4 Simulation Results

We investigate the performance of the three methods described in Section 2 by comparing the empirical coverage probabilities (ECP's) for the fifteen data generating processes (DGP's) introduced in Section 3. We generated one thousand replications of each DGP and considered realizations of length $n = 100, 250, 500, 1000$. We focused on the most commonly used confidence level of 95%. The standard errors in all tables are about 0.5% and are always smaller than 1%.

4.1 Residual Bootstrap

Table 4.1 presents the ECP of the symmetric confidence interval for the three GARCH models. To save space the results for the equal-tailed confidence interval are not presented, but are discussed in the following conclusions. Equal-tailed and symmetric confidence intervals perform equally well for the standard GARCH and GJR-GARCH. However, for the NL GARCH, the symmetric interval is better than the equal-tailed. It is thus seen that the symmetric confidence interval is preferred over the equal-tailed. The ECP decreases as the value of γ_{c2} approaches 1. Recall that $\gamma_{c2} < 1$ is required for the population autocovariances to exist. When $\gamma_{c2} \approx 0.9$, at least 250 observations are needed to ensure reasonable ECP for the standard GARCH and the GJR-GARCH. For the NL GARCH, even series length of 1000, does not produce satisfactory results. For the standard GARCH and the GJR-GARCH increasing the sample size from 500 to 1000 does not improve the ECP. For the NL GARCH a sample size of 1000 observations is needed, except when $\gamma_{c2} \leq 0.3$.

The somewhat worse performance of the residual bootstrap method for the GJR-GARCH which becomes markedly worse for the NL GARCH can be attributed to identification problems, which are particularly acute for the NL GARCH: for the latter model biases of parameter estimates are very large when η in equation (7) is large. Large η corresponds to large γ_{c2}, we omit the details

Table 1. ECP of *symmetric* confidence intervals constructed using *residual bootstrap.*

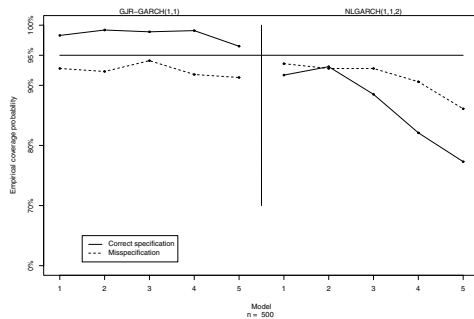|	n	e.c.p. (%)	e.c.p. (%)	e.c.p. (%)	e.c.p. (%)	e.c.p. (%)
STD GARCH	1	2	3	4	5	
100	99.6	85.3	86.0	80.4	77.4	
250	92.9	91.3	92.1	89.4	84.4	
500	93.4	93.4	94.1	93.7	92.7	
1000	95.1	96.8	97.6	97.6	94.4	
GJR GARCH	1	2	3	4	5	
100	97.7	94.8	92.0	89.5	81.5	
250	96.2	96.6	97.0	96.4	92.3	
500	98.3	99.2	98.9	99.1	96.5	
1000	99.0	99.4	99.6	99.8	98.8	
NL GARCH	1	2	3	4	5	
100	95.5	83.8	79.8	74.7	66.0	
250	91.7	87.3	84.3	81.0	73.6	
500	91.7	93.1	88.5	82.1	77.3	
1000	96.4	93.3	92.9	87.0	81.0	

Fig. 1. Comparison of ECP's for symmetric residual bootstrap confidence intervals based on standard GARCH and a correct specification. The nominal coverage of 95% is marked by the solid horizontal line. The series length is $n = 500$.

of the calculation. On the other hand, for the standard GARCH, while they still do exist, the identification problems are much less severe.

Figure 1 shows that estimating the standard GARCH model on all three DGP's might lead to improvements in ECP's, for symmetric confidence intervals and series of length 500. The results for other series lengths look very much the same and are therefore not presented. The residual bootstrap method works best if symmetric confidence intervals are used and the standard GARCH model is estimated. Thus, in our context, misspecifying a model improves the performance of the procedure.

Table 2. ECP of *symmetric* confidence intervals based on the *block bootstrap* method for the five parameter choices in the *GJR-GARCH* model.

Model		1	2	3	4	5
n	b	e.c.p. (%)	e.c.p. (%)	e.c.p. (%)	e.c.p. (%)	e.c.p. (%)
500	3	87.0	82.0	78.4	65.5	61.4
	5	89.1	83.8	73.4	63.0	58.5
	10	87.9	81.8	71.4	60.6	51.9
	15	84.5	78.7	71.8	63.8	52.7
	30	85.6	79.0	69.6	61.3	50.0
1000	5	87.7	84.4	75.2	67.9	59.6
	10	88.6	85.1	70.8	61.0	52.6
	15	89.7	83.0	72.7	63.6	53.3
	30	87.8	80.9	72.7	59.7	51.2

4.2   Block Bootstrap and Subsampling

The implementation of both methods requires a choice of the block length b. We then have a multitude of cases to explore: 15 models, 2 types of confidence intervals (equal-tailed and symmetric), 4 sample sizes and several choices of b. Since we used 10 values of b in our experiments, we obtained 1,200 ECP's. For space constraints, we describe them and present some typical values in Table 2.

The empirical coverage probabilities are generally too low for all choices of n and b and are in the range of 80% to 90% for $\gamma_{c2} \leq 0.3$ and go down to slightly above 50% for $\gamma_{c2} \approx 0.9$. Irrespective of the value of γ_{c2}, choosing smaller b gives higher coverage. However, extremely small b, like 1 or 2, do not work well. We recommend to use $b = 3$ or $b = 5$. The dependence on b is however not substantial, which is very desirable, as in many other applications choosing optimal b is very difficult. There is not much difference of ECPs between equal-tailed and symmetric confidence intervals. The block bootstrap confidence intervals are generally too short and given that the QML estimates underestimate the true value of the autocorrelation, they are shifted too much to the left what causes the under-coverage.

We observed that the subsampling method is very sensitive to the choice of b. Symmetric confidence intervals have a much better ECP than the equal-tailed. By choosing very short b's, such as 3 or 6, we can obtain ECP's that are quite close to 95% for models with $\gamma_{c2} < 0.6$ and fair coverage for models with greater values of γ_{c2}. Such choice of b is somewhat surprising, as autocovariances are then computed from very short sub-series. The ECP's are generally too low for equal-tailed confidence intervals and are typically in the range of 50-70%. As γ_{c2} approaches 1, the empirical coverage decrease and in some cases may be as low as 10%. Complete tables for ECP's are available at the following site: www.gillesteyssiere.net/ktz_iccs2004.
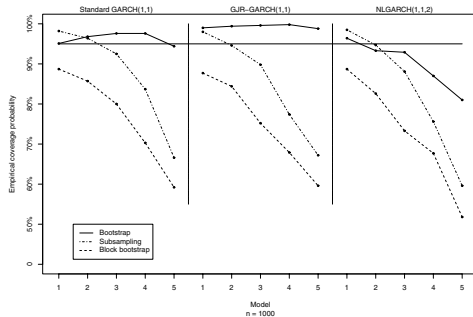
Fig. 2. Comparison of ECP's for symmetric confidence intervals. The nominal coverage 95% is marked by solid horizontal line. The series length is $n = 1000$. For block bootstrap, $b = 5$, for subsampling $b = 3$.

4.3   Conclusions and Practical Recommendations

The best method is residual bootstrap which assumes a standard GARCH(1,1) model. The block bootstrap and subsampling methods do not perform well when γ_{c2} approaches 1. Moreover, these methods require a choice of the block size b. The latter problem is particularly acute for the subsampling method. Except for the NL GARCH, the residual bootstrap method with correct model specification performs reasonably well even for γ_{c2} close to 1. This is probably due to the fact that large values for γ_{c2} correspond to large values of model parameters which are easier to estimate than small values yielding residuals which are close to the unobservable errors. A graphical comparison of symmetric confidence intervals based on the three methods for $n = 1000$ is given below:

References

1. Bühlmann, P.: Bootstrap for time series. Statistical Science **17** (2002) 52–72.
2. Engle, R.F.: Discussion: stock market volatility and the crash of 87. Review of Financial Studies **3** (1990) 103–106.
3. Glosten, L.R., Jagannathan, R. and Runkle, D.: On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance **48** (1993) 1779–1801.
4. He, C. and Teräsvirta, T.: Properties of moments of a family of GARCH processes. Journal of Econometrics **92** (1999) 173–192.
5. Künsch, H.: The jackknife and the bootstrap for general stationary observations. The Annals of Statistics **17** (1989) 1217–1241.
6. Mikosch, T. and Stărică, C.: Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. The Annals of Statistics **28** (2000) 1427–1451.
7. Politis, D.N., Romano, J.P. and Wolf, M.: Subsampling. (1999) Springer Verlag.
8. Politis, D.N., Romano, J.P. and Wolf, M.: Inference for autocorrelations in the possible presence of a unit root. Journal of Time series Analysis (to appear).