Robust bootstrap forecast densities for GARCH models: returns, volatilities and Value-at-Risk.

Carlos Trucios*
University of Campinas, Campinas, Brazil - ctrucios@gmail.com

Luiz K. Hotta
University of Campinas, Campinas, Brazil - hotta@ime.unicamp.br

Esther Ruiz
University Carlos III de Madrid, Madrid, Spain - ortega@est-econ.uc3m.es

Abstract

GARCH models are widely used to forecast the volatility and can be used to construct forecast densities for financial returns. These densities can be useful to obtain forecast intervals and quantiles of interest as, for example, the Value-at-Risk. In this context, bootstrap procedures can be useful as they allow obtaining forecast densities for returns and volatilities that incorporate the parameter uncertainty without assuming any particular error distribution. In this work, we analyze the effect of outliers on the construction of bootstrap forecast densities for returns and volatilities when they are based on both standard Maximum Likelihood and robust procedures. The results have implications on the construction of forecast intervals for returns and volatilities and on VaR forecasts. Finally, we propose a robust modification with good finite sample properties.

Keywords: Outliers; Robust estimator; Smooth bootstrap; Winsorized bootstrap.

1. Introduction

In the context of financial time series, density forecast of future returns focus the interest of many researches and practitioners as it is fundamental to obtain risk measures as, for example, Value-at-Risk (VaR). Furthermore measuring the uncertainty around future volatilities is also of interest in many financial models. It is well known that conditional variances of financial returns evolve over time and, in order to construct forecast densities for future returns and volatilities, one should take into account this evolution. One of the most popular models for the evolution of daily conditional variances of financial returns is the GARCH model. After estimating the GARCH parameters and assuming that the standardized returns have a conditional Normal distribution, the forecast densities of future returns are usually approximated using this distribution; see, for example, Linsmeier and Pearson (2000) and Kuester et al. (2006) among many others. However, approximating the forecast densities of future returns in this way has two main drawbacks. First, the corresponding densities do not take into account the parameter uncertainty and, second, the distribution of standardized returns may have heavy tails and can even be asymmetric; see, for instance, Bollerslev (1987), Pagan (1996) and Cont (2001) among many others. To solve these problems, Pascual et al. (2006) propose a bootstrap procedure for GARCH models that not only allows the construction of forecast densities for future returns but also for future volatilities incorporating the parameter uncertainty without assuming any particular error distribution.

When dealing with long time series of financial returns, it is not unusual to find extreme observations that cannot be explained by the GARCH model even when assuming heavy tailed standardized returns; see, for example, Franses and Ghijsels (1999), Charles and Darné (2005) and Hotta and Zevallos (2013). Outliers may cause biases on the usual Maximum Likelihood (ML) estimators of the parameters of GARCH models and, consequently, on the estimated volatilities; see, for example, Catalán and Trívez (2007) and Carnero et al. (2012). As a consequence, it is expected that bootstrap forecast densities for returns and volatilities based on ML estimators of parameters and standard filters for volatilities could be distorted giving a misleading picture of what can be expected in the future.

The first objective of this work is to analyze the effect of outliers on the forecast densities of returns and volatilities constructed using the bootstrap procedure proposed by Pascual et al. (2006). We show that the performance of the bootstrap procedure, even if based on robust estimates; see, Muler and Yohai (2008), Carnero et al. (2012) and Boudt et al. (2013), is not adequate when constructing forecast densities for future volatilities. Furthermore, the bootstrap forecast densities are highly influenced by outliers mainly when they appear at the end of the series. Second, we propose a robust bootstrap procedure to obtain forecast densities for returns and volatilities. We carry out Monte Carlo experiments to compare the performance of both alternatives and show that the new algorithm has good finite sample properties when forecasting returns and volatilities in the context of GARCH models.

2. Contaminated GARCH models Generalized Autoregressive Conditional Heteroscedastic (GARCH) models were proposed by Bollerslev (1986), extending the model proposed by Engle (1982), to represent the dynamic dependence often observed in the second order moments of some economic time series. Given its popularity in empirical applications, in this paper, we focus on the GARCH(1,1) model. The GARCH(1,1) model contaminated by observational or additive outliers is defined by Hotta and Tsay (2012) as follows

$$y_t = z_t + sign(z_t) w_t I_t(t \in A) \tag{1a}$$

$$z_t = \sigma_t \epsilon_t \tag{1b}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{1c}$$

for $t = 1, ..., T$, where y_t is the return observed at time t and z_t is the uncontaminated GARCH(1,1) process with σ_t being the volatility that depends on past uncontaminated returns. The disturbance ϵ_t is an independent white noise sequence with variance one and w_t is the size of the outlier at time t. $I_t(\cdot)$ is the indicator function and A is the set of contaminated observations. Finally, the parameters are assumed to satisfy the usual positivity and stationary conditions, namely $\alpha_0 > 0$, $\alpha_1, \beta \geq 0$ and $\alpha_1 + \beta < 1$.

3. Bootstrap Forecast Intervals

Pascual et al. (2006) propose a bootstrap procedure to construct forecast densities of returns and volatilities that incorporate the parameter uncertainty without relying on particular assumptions about the error distribution.

The PRR algorithm, is described next for the sake of clarity.

- **Step 1:** Estimate the parameters θ of model (1) without outliers by quasi maximum likelihood with Normal distribution (QMLn), obtaining $\hat{\theta}^{(N)} = (\hat{\alpha_0}^{(N)}, \hat{\alpha_1}^{(N)}, \hat{\beta}^{(N)})$ and obtain the corresponding standardized residuals $\hat{\epsilon}_t = \frac{y_t}{\hat{s}_t}$, $t = 1, ..., T$ where $\hat{s}_t = \alpha_0^{(N)} + \alpha_1^{(N)} y_{t-1}^2 + \beta^{(N)} \hat{s}_{t-1}^2$, $t = 2, ..., T$, is evaluated with the parameters substituted by the corresponding QMLn estimates, and $\hat{s}_1^2$ being the sample variance of y_t. Denote by $\hat{F}_\epsilon$ the empirical distribution of the centered standardized residuals.
- **Step 2:** Generate a bootstrap series y_t^*, $t = 1, ..., T$, as follows

$$\begin{aligned} s_t^{*2} &= \hat{\alpha_0}^{(N)} + \hat{\alpha_1}^{(N)} y_{t-1}^{*2} + \hat{\beta}^{(N)} s_{t-1}^{*2} \\ y_t^* &= s_t^* \epsilon_t^*, \end{aligned} \tag{2}$$

  where ϵ_t^* is a random draw with replacement from $\hat{F}_\epsilon$ and $s_1^{*2} = \hat{s}_1^2$. Compute the bootstrap estimates $\hat{\theta}^{*(N)} = (\hat{\alpha_0}^{*(N)}, \hat{\alpha_1}^{*(N)}, \hat{\beta}^{*(N)})$
- **Step 3:** For $h = 1, ..., H$ construct $\hat{y}_{T+h|T}^*$ and $\hat{s}_{T+h|T}^*$ using the bootstrap estimates $\hat{\theta}^{*(N)}$ and the original series y_t, $t = 1, ..., T$ as follows

$$\begin{aligned} \hat{s}_{T+h|T}^{*2} &= \hat{\alpha_0}^{*(N)} + \hat{\alpha_1}^{*(N)} \hat{y}_{T+h-1|T}^{*2} + \hat{\beta}^{*(N)} \hat{s}_{T+h-1|T}^{*2} \\ \hat{y}_{T+h|T}^* &= \epsilon_{T+h|T}^* \hat{s}_{T+h|T}^*, \end{aligned} \tag{3}$$

  where $y_{T|T}^* = y_T$, $\epsilon_{T+h|T}^*$ are random draws with replacement from $\hat{F}_\epsilon$ and

$$\hat{s}_{T|T}^{*2} = \frac{\hat{\alpha_0}^{*(N)}}{1 - \hat{\alpha_1}^{*(N)} - \hat{\beta}^{*(N)}} + \hat{\alpha_1}^{*(N)} \sum_{j=0}^{T-2} \hat{\beta}^{*j(N)} \left(y_{T-j-1}^2 - \frac{\hat{\alpha_0}^{*(N)}}{1 - \hat{\alpha_1}^{*(N)} - \hat{\beta}^{*(N)}} \right). \tag{4}$$

- **Step 4:** Repeat steps 2 and 3 B times obtaining B bootstrap replicates $(\hat{y}^{*(1)}_{T+h|T}, ..., \hat{y}^{*(B)}_{T+h|T})$ of y_{T+h} and $(\hat{s}^{*(1)}_{T+h|T}, ..., \hat{s}^{*(B)}_{T+h|T})$ of s_{T+h}.

Denote by $\hat{G}^*_y(x) = \#\left(y^*_{T+h|T} \leq x\right)/B$, the empirical bootstrap distribution function of the B bootstrap replicates $\left(y^{*(1)}_{T+h|T}, ..., y^{*(B)}_{T+h|T}\right)$. Using $\hat{G}^*_y(x)$, one can compute the VaR of returns as the required quantile of $\hat{G}_y(x)$ and also $100(1-\delta)\%$ forecast intervals as follows

$$[L^*_Y(h), U^*_Y(h)] = [Q^*_Y(\frac{\delta}{2}), Q^*_Y(1-\frac{\delta}{2})], \tag{5}$$

where $Q^*_x = \hat{G}^{*-1}_x$. Bootstrap forecast intervals of future volatilities can be obtained in a similar way. This procedure work very well when the series is not contaminated by outliers, as showed in Pascual et al. (2006), however, when the series is contaminated by atypical observations the performance is pour, even using robust estimators of the parameters.
Figure 1, shows the results for volatilities using the QMLn estimator and some robust estimators, as BM of Muler and Yohai (2008), Bs of Carnero et al. (2012) and BVT of Boudt et al. (2013). Using any estimator the result when the series is not contaminated shows good performance, and when the series is contaminated in the middle the robust estimator improve the performance of the algorithm. However, when the series is contaminated near to the end of the series the algorithm has not good performance, even, when robust estimators was used. As a consequence of the high error coverage below, the returns estimate coverage is

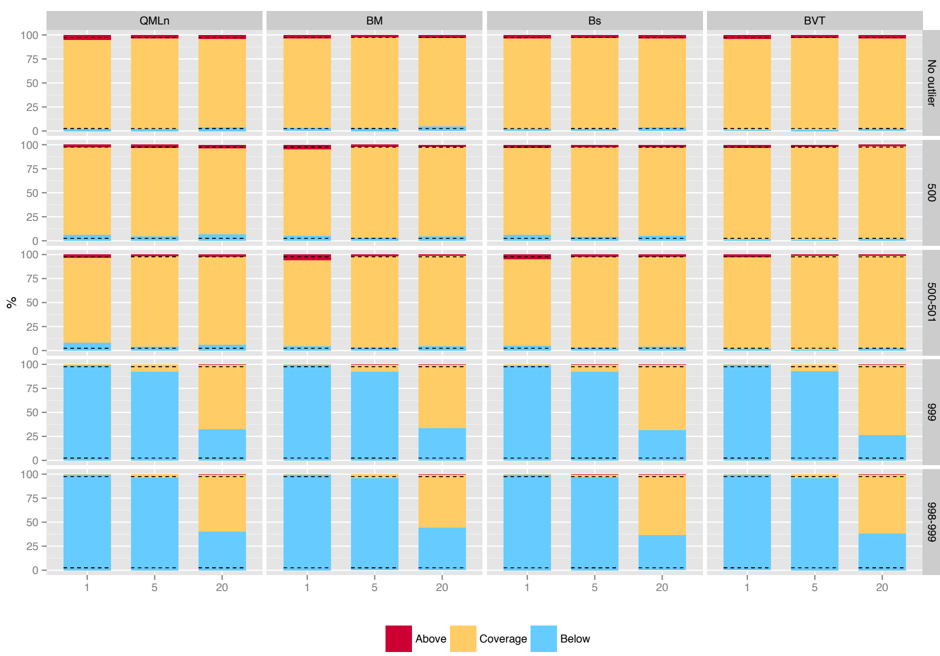

Figure 1: Estimated coverage and coverages above and below of the 95% bootstrap forecast interval for volatilities using the PRR algorithm with QMLn, BM, Bs and BVT estimators. Sample size $T = 1000$. Based on 500 Monte Carlo replications.

overestimated, because the length of intervals in larger than necessary.

4. Robust bootstrap forecast intervals

As we have seen in the previous section, bootstrap forecast densities of returns and volatilities may be distorted when the data contains outliers even if the bootstrap forecast densities are based on robust estimators. These difficulties could be attributed to the important negative effect of bootstrap replicates where the proportion

of outliers could be even much higher that in the original series. To deal with this problem, we propose a robust bootstrap procedure to obtain forecast intervals that consider two alternative bootstrap procedures instead the traditional bootstrap to sample the residuals. First, we consider using smoothing bootstrap that calls for replacing the empirical distribution of the residuals by a smoothed version. Second, we consider a winsorized bootstrap procedure in which the bootstrap samples are constructed so that the distribution of the residuals in each bootstrap replicate reflects that observed in the original series. These alternative bootstrap methods are combined with a bounded robust form to obtain the volatilities.
In the new algorithm, denoted as PRR_R, steps 1, 2 and 3 are substituted, respectively by:

- Step 1: Estimate the parameters using the robust BVT estimator and estimate the underlying volatility. Obtain the corresponding centered residuals, $\hat{\epsilon}_t^R$, and their empirical distribution function, $\hat{F}_{\hat{\epsilon}^R}$.
- Step 2: Generate a bootstrap series, y_t^*, $t = 1, ..., T$, as follows

$$\begin{aligned} s_t^{*2} &= \hat{\alpha_0}^{(BVT)} + \hat{\alpha_1}^{(BVT)} s_{t-1}^{*2} r_c \left(\frac{y_{t-1}^{*2}}{s_{t-1}^{*2}} \right) + \hat{\beta}^{(BVT)} s_{t-1}^{*2} \\ y_t^* &= s_t^* \varepsilon_t^*, \end{aligned} \tag{6}$$

  where ε_t^* are bootstrap extractions of $\hat{F}_{\hat{\epsilon}^R}$ and $s_1^{*2} = \hat{s}_1^2$ where $r_c(\cdot)$ is defined in Boudt et al. (2013) as

$$r_c(x) = \begin{cases} 1.0465 \times x, & \text{if } |x| \leq c \\ 1.0465 \times c, & \text{if } |x| > c. \end{cases} \tag{7}$$

- Step 3: Obtain h-steps-ahead forecast for returns and volatilities. For $h = 1$

$$\begin{aligned} \hat{s}_{T+1|T}^{*2} &= \hat{\alpha_0}^{*(BVT)} + \hat{\alpha_1}^{*(BVT)} \hat{s}_{T|T}^{*2} r_c \left(\frac{\hat{y}_{T|T}^{*2}}{\hat{s}_{T|T}^{*2}} \right) + \hat{\beta}^{*(BVT)} \hat{s}_{T|T}^{*2} \\ \hat{y}_{T+1|T}^* &= \varepsilon_{T+1|T}^* \hat{s}_{T|T}^*, \end{aligned} \tag{8}$$

  where $\hat{y}_{T|T} = y_T$, $\varepsilon_{T+1|T}^*$ is a bootstrap extraction and $s_{T|T}^{*2}$ is obtained from the following recursion for $t = 2, ..., T$,

$$\hat{s}_{t|T}^{*2} = \hat{\alpha_0}^{*(BVT)} + \hat{\alpha_1}^{*(BVT)} \hat{s}_{t-1|T}^{*2} r_c \left(\frac{y_{t-1}^2}{\hat{s}_{t-1|T}^{*2}} \right) + \hat{\beta}^{*(BVT)} \hat{s}_{t-1|T}^{*2}, \tag{9}$$

  with

$$r_c(x) = \begin{cases} x, & \text{if } |x| \leq 9 \\ \varepsilon_t^{2*}, & \text{if } |x| > 9, \end{cases} \tag{10}$$

  where ε_t^* is a bootstrap extraction and $s_{1|T}^{*2} = \frac{\hat{\alpha_0}^{*(BVT)}}{1 - \hat{\alpha_1}^{*(BVT)} - \hat{\beta}^{*(BVT)}}$.

For $h = 2, ..., H$

$$\begin{aligned} \hat{s}_{T+h|T}^{*2} &= \hat{\alpha_0}^{*(BVT)} + \hat{\alpha_1}^{*(BVT)} \hat{y}_{T+h-1|T}^{*2} + \hat{\beta}^{*(BVT)} \hat{s}_{T+h-1|T}^{*2} \\ \hat{y}_{T+h|T}^* &= \varepsilon_{T+h|T}^* \hat{s}_{T+h|T}^*. \end{aligned} \tag{11}$$

5. Simulation

To evaluate the performance of the PRR_R algorithm, we carry out Monte Carlo experiments, we evaluate our algorithm using three alternatives for re-sampling residuals in the PRR_R algorithm: classic bootstrap (re-sampling with replacement), smooth bootstrap and winsorized bootstrap.
Figure 2 shows the average estimate coverage forecast intervals for volatilities for uncontaminated and contaminated series with 95% nominal coverage. The new algorithm with $\xi = 0.01$ shows a good performance in all cases, even when outliers are at the end of the series.

For returns, the estimate coverages in all cases are very close to the nominal coverage and the results for uncontaminated and contaminated series do not present strong differences. The new algorithm also corrected the overestimation of the return coverage presented in the PRR algorithm when working with contaminated series.

For volatilities, the PRR_R algorithm bounds the influence of outliers in the constructions of the forecast densities. Regardless whether outliers are in the middle or at the end of the series, or the series is contaminated with one isolated or two consecutive outliers, the estimated coverages are close to the nominal coverages robustifying the PRR algorithm. For uncontaminated series the results are also good, showing that the PRR_R algorithm works well in all situations.

Figure 2: Estimated coverage and coverages above and below of the 95% bootstrap forecast interval for volatilities using the PRR_R algorithm with classic, smooth and winsorized bootstrap. Sample size $T = 1000$. Based on 500 Monte Carlo replications.

The new algorithm has a good performance regardless the bootstrap procedure to sample the innovations, with slightly better performance when classic and winsorized bootstrap are used compared with the smooth bootstrap.

5. Conclusions

In this work, we modified the PRR algorithm and propose a robust bootstrap procedure to obtain forecast densities for returns and volatilities in GARCH models. Our procedure is based on the PRR algorithm and hold the features of incorporating the uncertainty due to parameter estimation and it does not require any assumption on the error distribution. Our approach uses a robust form to estimate the parameters and the volatility in the GARCH model and bounds the influence of atypical observation in the construction of the forecast densities. We analyze the finite sample behavior of the proposed bootstrap procedure by means of extensive Monte Carlo experiments and it is shown that our method has a good performance in both of cases with and without outliers. Finally, the new bootstrap procedure provides a new tool for practitioners to compute forecast densities for returns and volatilities with the same features of PRR algorithm but with the guarantee of obtaining forecast intervals robust to additive outliers.

Acknowledgements
The first two authors acknowledge financial support from São Paulo Research Foundation (FAPESP) grants 2012/09596-0 and 2013/00506-1, and Laboratory EPIFISMA. The third author is grateful for financial support from Project ECO2012-32401 by the Spanish Government.

References

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307-327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of returns. *The Review of Economics and Statistics*, 69(3):542-547.

Boudt, K., Danielsson, J., and Laurent, S. (2013). Robust forecasting of dynamic conditional correlation GARCH models. *International Journal of Forecasting,* 29(2):244-257.

Carnero, M. A., Peña, D. and Ruiz, E. (2007). Estimating GARCH volatility in the presence of outliers. E*conomics Letters*, 114(1):86-90.

Catalán, B. and Trívez, F. J. (2007). Forecasting volatility in GARCH model with additive outliers. *Quantitative Finance*, 7(6):591-596.

Charles, A. and Darné, O. (2005). Outliers and GARCH models in financial data. *Economics Letters*, 86(3):347-352.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223-236.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987-1007.

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

Hotta, L. K. and Tsay, R. S. (2012). Outliers in GARCH processes. In W. Bell, T. M. and Holan, S. H., editors, *Economic Time Series*: Modeling and Seasonality, pages 337-358. Chapman and Hall/CRC Press.

Hotta, L. K. and Zevallos, M. (2013). Test of outliers and influential observations in GARCH models: A review. *Estadística*, 65(184):65-85.

Kuester, K., Mittnik, S., and Paolella, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53-89.

Linsmeier, T. J. and Pearson, N. D. (2000). Value at risk. *Financial Analysts Journal*, 56(2):47-67.

Muler, N. and Yohai, V. J. (2008). Robust estimates for GARCH models. *Journal of Statistical Planning and Inference*, 138(10):2918-2940.

Pagan, A. (1996). The econometrics of financial markets. *Journal of Empirical Finance*, 3(1):15-102.

Pascual, L., Romo, J., and Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics & Data Analysis*, 50(9):2293-2312.