Joanna Górka[*]

How many observations should be taken to obtain appropriate VaR measure using a family Sign RCA models?

Accurate modeling of risk is important in finance. A family Sign RCA models could be used to obtain Value-at-Risk (VaR) measure. In alternative approach to obtain VaR measure, different sizes of sample for rolling estimation of models parameters have been used. In the case when one of the Sign RCA models is used, one does not know what size of sample should be taken to obtain appropriate VaR forecasts. In this paper, to choose the size of sample, a Monte Carlo experiment was used. The properties of the time series generated by one of the Sign RCA models were analyzed. The results obtained from Monte Carlo experiment and empirical examples were compared.

Introduction

In the literature, non-linear dynamics of financial time series has generally been described by the class of AR-GARCH models [Bollerslev, 1986; Engle, 1982]. A different, alternative approach to the description of financial time series represents random coefficient autoregressive models (RCA) [Nicholls, Quinn, 1982]. Thavaneswaran and Appadoo (2006) proposed to add the sign function to RCA models. The sign function may describe an asymmetric reaction of changes in returns to good or bad news. For this reason using RCA model with the sign function to obtain VaR forecast seems to be a good idea.

The aim of this paper is to answer the question: How many observations should be taken to obtain the appropriate VaR measure for family Sign RCA models?

[*] Department of Econometrics and Statistics, Nicolaus Copernicus University in Toruń, Poland. e-mail: jgora@umk.pl. This work was financed from the Polish science budget resources in the years 2008-2010 as the research project N N111 434034.

1. The family Sign RCA models

Random coefficient autoregressive models (RCA) are straightforward generalization of the constant coefficient autoregressive models. A full description of this class of models including their properties, estimation methods and some application can be found in Nicholls and Quinn (1982).

The classical random coefficient autoregressive model of first order for stationary univariate time series can be written as:

$$y_t = (\phi + \delta_t) y_{t-1} + \varepsilon_t, \tag{1}$$

where:

$$\begin{bmatrix} \delta_t \\ \varepsilon_t \end{bmatrix} \sim^{iid} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\delta^2 & 0 \\ 0 & \sigma_\varepsilon^2 \end{bmatrix} \right), \tag{2}$$

$$\phi^2 + \sigma_\delta^2 < 1. \tag{3}$$

Condition (3) is necessary and sufficient for the second-order stationarity of y_t. Conditions (2)-(3) ensure the strict stationarity of the process.

The process (1) has zero mean, constant unconditional variance and kurtosis. The constant unconditional variance of RCA(1) is bigger than the unconditional variance of AR(1). If $\sigma_\delta^2 = 0$, the value of kurtosis converges to 3 (similarly to AR(1) models).

A stationary Sign RCA models is given by (Thavaneswaran, Appadoo, 2006):

$$y_t = (\phi + \delta_t + \Phi s_{t-1}) y_{t-1} + \varepsilon_t, \tag{4}$$

where the conditions (2)-(3) are satisfied, and

$$s_t = \begin{cases} 1 & \text{for } y_t > 0 \\ 0 & \text{for } y_t = 0 \\ -1 & \text{for } y_t < 0 \end{cases}. \tag{5}$$

If $\phi + \delta_t > |\Phi|$, the negative value of Φ means, that the negative (positive) observation values at time $t-1$ corresponds to a decrease (increase) of observation values at time t. In the case of stock returns it would suggest (for returns) that after a decrease of stock returns the decrease of stock returns occurs higher than expected, and in the case of the increase of stock returns the increase in stock returns occurs lower than expected.

If conditions (2)-(3) are satisfied, the process (4) has zero mean, constant unconditional variance and kurtosis. When $\sigma_\delta^2 = 0$ and $\Phi = 0$, the kurtosis converges to 3.

Comparing the properties of the RCA model with properties of the sign RCA model one can see that introducing the sign function to the RCA model causes an increase of variance and kurtosis with relation to the variance and kurtosis obtained for the process described by the RCA model without sign function.

Thavaneswaran, Appadoo and Bector (2006) proposed to add MA(1) errors to the RCA or the Sign RCA models, that called a RCA(1)-MA(1) and a Sign RCA(1)-MA(1) respectively.

A sequence of y_t could be described by the RCA(1)-MA(1) model provided it satisfies the equations:

$$y_t = (\phi + \delta_t) y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}, \tag{6}$$

where the conditions (2)-(3) are satisfied. The process described by equation (6) has zero mean, constant unconditional variance and kurtosis. In this case, the kurtosis is exactly the same as in the RCA model. Therefore, the only property which has changed is the value of variance, i.e. bigger.

The Sign RCA(1)-MA(1)[1] is given by:

$$y_t = (\phi + \delta_t + \Phi s_{t-1}) y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}, \tag{7}$$

where the conditions (2)-(3) are satisfied. If conditions (2)-(3) are satisfied, the process (7) has zero mean, constant unconditional variance and kurtosis. From the comparison of properties of the Sign RCA and the Sign RCA-MA models one notice that only the variance of the Sign RCA-MA model is bigger than in the Sign RCA model. Other properties remained the same.

Another modification of the RCA and Sign RCA models is the assumption of the errors. In this case we assume that having GARCH errors.

The RCA model with GARCH errors (RCA(1)-GARCH(p,q)) is given by:

$$y_t = (\phi + \delta_t) y_{t-1} + \varepsilon_t, \tag{8}$$

$$\varepsilon_t = \sqrt{h_t} z_t$$

$$h_t = \alpha_0 + \sum_{t=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}$$

where $z_t \sim N(0, \sigma_z^2)$, $\alpha_0 > 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$.

The Sign RCA(1)-GARCH(p,q) model for the time series y_t can be written as:

$$y_t = (\phi + \delta_t + \Phi s_{t-1}) y_{t-1} + \varepsilon_t, \tag{9}$$

[1] Sign RCA(1) with MA(1) errors.

$$\varepsilon_t = \sqrt{h_t}\, z_t$$

$$h_t = \alpha_0 + \sum_{t=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}$$

where $z_t \sim N\left(0, \sigma_z^2\right)$, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $|\Phi| \leq \alpha_0$.

The assumptions of the errors change both variance and kurtosis of the RCA and the Sign RCA models, respectively.

2. A Monte Carlo Experiment

The term "Monte Carlo" refers to the procedures in which quantities of interest are approximated by generating many random realisations of a stochastic process and averaging them in some way. In economic theory Monte Carlo techniques are used to explore the quantitative properties of models with stochastic elements.

In this research the Monte Carlo experiment involves the following steps:

1. Assume values for the deterministic parts of the model.
2. Generate a (pseudo) random sample of size 1100 for the stochastic elements of the stochastic model from normal distribution function.
3. Generate data (time series).
4. Reduce the sample size to the last 1000 observation.
5. Estimate parameters of the one model from the family Sign RCA models.
6. Calculate one-step-ahead forecasts using rolling estimation and selected forecast of errors.
7. Calculate the Value-at-Risk measures and the traditional VaR test and loss functions.
8. Repeat step 2 to 7 1000 times.

The calculations were done in the Gauss program. All model parameters were estimated using maximum likelihood (MLE)**.** The assumptions of parameters value for the Data Generating Process (GDP) (step 1) come from the empirical application of daily prices of Polish firms' shares on the Warsaw Stock Exchange.

The model parameters were estimated for the different sizes of samples and for all these models the same quantity one-step-ahead forecasts were calculated. Next, the forecast errors such as mean error, mean square error, root mean square error, average absolute error and direction quality measures were calculated.

Using obtained Sign RCA models, VaR measures were calculated. It is used as a tool for measuring market risk and it is defined as the money-loss in a portfolio that is expected to occur over a pre-determined horizon and with pre-determined significance level. To examine accuracy VaR forecast statistics of the Proportion of Failures Test, the Regulatory Loss Function and the Firm's Loss Function were calculated.

If the empirical time series is generated by the appropriate model from the family Sign RCA models as results of Monte Carlo experiment one can obtain expected values of the forecasts errors and accuracy VaR measures.

The Data Generating Process for one of the experiments is described by the RCA model:

$$y_t = (0.15 + \delta_t) y_{t-1} + \varepsilon_t, \tag{10}$$

where $\begin{bmatrix} \delta_t \\ \varepsilon_t \end{bmatrix} \sim^{ind} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0.12 & 0 \\ 0 & 3.95 \end{bmatrix} \right)$.

The Monte Carlo experiment results for this GDP are presented in the Table 1. Notice that the differences between individual forecast errors for different size of sample are small. The conclusions from 100 predictions and 351 predictions are similar. The minimum values for the individual forecast errors were obtained for the sample 400 (for 351 predictions) and sample 250-400 (for 100 predictions). In all cases the signs direction quality measures (Q1) are bigger than 50% and the turning points quality measures (Q2) are smaller than 9%.

Table 1 The forecast errors for the RCA models described by the equation (10)

Sample	ME	MSE	RMSE	MAE	Q1	Q2
100 predictions						
500	0,02329	4,56608	2,12890	1,70285	0,54800	0,00000
400	0,02340	**4,56543**	**2,12877**	1,70275	0,54720	0,00333
300	**0,02277**	4,57006	2,12989	1,70290	**0,54830**	0,00863
250	0,02295	4,56945	2,12976	**1,70263**	0,54470	0,01544
200	0,02333	4,57593	2,13127	1,70433	0,54230	0,03159
150	0,02529	4,58595	2,13344	1,70672	0,54260	0,04565
100	0,02331	4,60090	2,13688	1,71001	0,53840	**0,08899**
351 predictions						
500	0,00926	4,54334	2,12956	1,69335	**0,54778**	0,00000
400	**0,00920**	**4,54258**	**2,12941**	**1,69331**	0,54764	0,00188
300	0,00930	4,54664	2,13035	1,69384	0,54630	0,00761
250	0,00964	4,55061	2,13128	1,69445	0,54530	0,01429
200	0,00928	4,55407	2,13207	1,69490	0,54462	0,02648
150	0,00986	4,56283	2,13409	1,69639	0,54274	0,04083
100	0,00953	4,58464	2,13920	1,70014	0,53943	**0,08240**

Note: ME – mean error. MSE – mean square error. RMSE – root mean square error. MAE – average absolute error. Q1, Q2 – direction quality measures. Bold value means min/max value of appropriate measure.

For the RCA models generated by the equation (10) the VaR measure and backtesting of VaR were calculated. The results are presented in the Table 2. One can see that the Proportion of Failures Test (LR_{POF}) does not reject the null hypothesis for all windows. Almost all the VaR forecasts are underestimated and their values are almost independent of number of calculated forecasts. The number of forecasts has not significant influence on the accuracy. For all significance level and all cases of number predictions the Regulatory Loss Function is the smallest for the biggest size of the sample. On the other hand, the Firm's Loss Function is the smallest for the smallest size of the sample.

Similar results for different values of the RCA model parameters were obtained.

For others models the results of the forecast errors and loss functions led to similar conclusions. Both the Proportion of Failures test and the empirical significant level for different models led to different conclusions.

Table 2 The results of the VaR tests and the loss function for Monte Carlo experiment (the RCA model)

Sample	100 predictions				351 predictions			
	α_emp.	LR_{POF}	RL	FL	α_emp.	LR_{POF}	RL	FL
α= 5%								
500	5,18%	0,897	12,68	342,15	5,12%	0,719	45,64	1200,64
400	5,21%	0,827	12,73	342,20	5,17%	0,684	45,76	1200,13
300	5,18%	0,818	12,89	342,63	5,15%	0,696	46,29	1201,19
250	5,17%	0,858	12,90	342,66	5,17%	0,615	46,54	1201,62
200	5,35%	0,981	13,03	341,90	5,25%	0,641	46,97	1200,67
150	5,39%	0,884	13,11	341,57	5,27%	0,674	47,50	1199,64
100	5,54%	0,864	13,43	340,72	5,40%	0,626	49,22	1197,57
α= 2.5%								
500	2,48%	1,347	5,71	409,47	2,63%	1,025	21,47	1433,85
400	2,51%	1,248	5,75	409,54	2,63%	0,852	21,43	1433,70
300	2,57%	1,319	5,92	409,54	2,64%	0,815	21,82	1434,05
250	2,58%	1,322	5,94	409,51	2,67%	0,914	22,05	1434,31
200	2,49%	1,311	5,80	409,36	2,65%	0,867	22,12	1434,05
150	2,73%	1,305	6,04	408,20	2,75%	0,974	22,68	1431,42
100	2,72%	1,186	6,12	407,39	2,83%	0,985	23,79	1427,94
α= 1%								
500	1,03%	1,311	2,15	488,54	1,09%	1,050	8,30	1710,88
400	1,00%	1,252	2,11	488,79	1,08%	0,985	8,23	1710,92
300	1,02%	1,254	2,20	488,92	1,13%	1,051	8,58	1710,39
250	1,07%	1,320	2,25	488,66	1,15%	1,145	8,70	1710,58
200	1,12%	1,226	2,24	487,92	1,19%	1,075	8,91	1709,25
150	1,16%	1,279	2,28	487,15	1,26%	1,193	9,24	1706,11
100	1,17%	1,271	2,36	486,08	1,29%	1,228	9,77	1701,86

Note: LR_{POF} – estimates of the Proportion of Failures Test statistics. RL – Regulatory Loss Function. FL – Firm's Loss Function.

3. The empirical application

The data of PBG shares (PBG Capital Group) were used from bossa.pl for the period from 23-rd September 2005 to 18-th February 2009, which yields 852 observations. For the analyze daily percentage log returns of share were used.

For the first 500 observations of returns series the descriptive statistics and some tests were calculated. The percentages log stock returns have: the mean 0.363, the standard deviation 2.094, the kurtosis 5.344 and the skewnes 0.095. This returns series has autocorrelation. The LBI test rejects the null hypothesis of non random coefficient to returns for 10% significance level.

Next, the parameters of six models from the family Sign RCA models for the first 500 observations for time series were estimated. Models with statistically significant parameters are presented in the Table 4.

Table 4 Models from family Sign RCA for PBG

Model	ϕ	Φ	σ_ε^2	σ_δ^2	α_0	α_1	β_1	ARCH(3)	Q(3)	AIC	BIC
RCA	0.143 (0.052)		3.946	0.120				5.616	1.271	2161.262	2173.906
Sign RCA	0.133 (0.052)	0.085 (0.051)	3.963	0.108				4.798	0.839	2160.768	2177.626
Sign RCA-GARCH(1,1)	0.113 (0.050)	0.094 (0.051)		0.023	0.602	0.047	0.813	4.513	0.540	2158.402	2187.905

Note: In () is standard deviation parameter. Q(3) – estimates of the Box-Ljung test statistics of order 3. ARCH(3) – estimates of the Engel ARCH test statistics of order 3. AIC – Akaike's information criterion. BIC – Bayesian information criterion.

Table 5 The forecast errors for PBG

Sample	ME	MSE	RMSE	MAE	Q1	Q2
100 predictions						
500	-0,18791	7,78262	2,78973	2,13466	**0,51724**	0,00000
400	-0,20985	**7,75433**	**2,78466**	2,11965	**0,51724**	0,00000
300	-0,21751	7,77598	2,78854	2,12075	0,50575	0,00000
250	-0,21784	7,77416	2,78822	2,11668	0,49425	0,04762
200	-0,21865	7,78725	2,79056	2,12091	0,44828	0,14286
150	**-0,22403**	7,76373	2,78635	**2,11521**	0,50575	**0,35714**
100	-0,20999	7,82018	2,79646	2,12824	0,50575	0,02381
351 predictions						
500	-0,21629	8,18328	2,86064	2,20126	0,46688	0,47771
400	-0,22434	8,15999	2,85657	2,19584	0,47950	0,59873
300	-0,22933	8,16591	2,85761	2,19722	0,48265	0,58599
250	-0,22971	**8,14298**	**2,85359**	**2,19342**	0,48265	0,63057
200	-0,23014	8,15517	2,85573	2,19599	0,46057	0,66879
150	-0,23751	8,16789	2,85795	2,19762	0,47634	**0,70701**
100	**-0,23959**	8,22862	2,86856	2,21174	**0,48580**	0,59236

Note: ME – mean error. MSE – mean square error. RMSE – root mean square error. MAE – average absolute error. Q1, Q2 – direction quality measures. Bold value means min/max value of appropriate measure.

The autoregressive parameter for all models is almost the same. Only the variance of autoregressive parameter for the Sign RCA-GARCH model was significantly changed. The residuals of all models are not correlated and they have not ARCH effect. From AIC point of view the best model is the Sign RCA-GARCH(1,1) model. On the other hand, the value of Bayesian information criterion is the smallest for the RCA model. For next calculation, the RCA model is chosen. Let us that the GDP described by equation (10) has similar values of parameters as the RCA model for PBG.

The MSE and the RMSE (see the Table 5) for the 100 forecasts lead to the same conclusions as for GDP for the 100 forecasts. However, the other forecast errors give different conclusions, especially for the Q2 measure. In this case, the 70% cases of the turning points by the RCA models for the 150 size of sample were predicted. It has never happened to the model from the family Sign RCA models.

Table 6 Results of the VaR tests and the loss function for PBG

Sample	100 predictions				351 predictions			
	α_emp.	LR_{POF}	RL	FL	α_emp.	LR_{POF}	RL	FL
α= 5%								
500	11%	5.733**	54,69	384,42	8,83%	8,924***	187,36	1442,89
400	13%	9,537***	57,60	367,03	8,83%	8,924***	174,86	1456,28
300	13%	9,537***	52,76	369,31	7,69%	4,633**	155,83	1507,41
250	11%	5,733**	48,36	380,88	7,12%	2,959*	146,42	1538,14
200	8%	1,616	41,52	395,69	6,27%	1,103	132,82	1561,54
150	10%	4,131*	43,18	395,33	6,27%	1,103	135,92	1602,58
100	8%	1,616	41,05	400,40	5,70%	0,345	127,83	1629,97
α= 2.5%								
500	6%	3,633*	33,70	446,92	4,84%	6,234**	115,61	1674,19
400	8%	7,927***	36,07	425,21	5,41%	9,215***	107,25	1689,23
300	7%	5,626**	31,06	433,52	4,56%	4,925**	92,48	1754,06
250	5%	1,996	27,57	449,78	3,99%	2,710	85,43	1796,51
200	5%	1,996	24,67	459,96	3,99%	2,710	77,20	1819,65
150	4%	0,783	22,94	469,07	3,70%	1,821	78,49	1869,09
100	5%	1,996	23,24	464,27	3,99%	2,710	73,20	1891,85
α= 1%								
500	3%	2,632	19,49	523,61	3,13%	10,313***	67,06	1947,01
400	4%	5,182**	20,75	501,55	3,13%	10,313***	59,64	1979,78
300	3%	2,632	16,44	513,85	2,56%	6,056**	49,08	2060,13
250	2%	0,783	14,54	531,41	2,28%	4,259**	44,70	2111,14
200	2%	0,783	12,21	545,15	1,99%	2,719*	38,01	2148,73
150	3%	2,632	12,30	547,04	1,99%	2,719*	39,43	2202,47
100	3%	2,632	11,48	544,78	1,71%	1,472	33,28	2241,44

Note: *, **, *** indicate rejection of H_0 at the 10%, 5% and 1% significant level, respectively. LR_{POF} – estimates of the Proportion of Failures Test statistics. RL – Regulatory Loss Function. FL – Firm's Loss Function.

The Proportion of Failures Test and loss functions for PBG and 5%, 2.5% and 1% significance levels for 100 and 351 forecasts are presented in the Table 6. One can see that the accuracy test rejects the null hypothesis for windows size 500, 400 and 300 observations for the 5% and 2.5% significance levels and for each number of predictions. All the VaR forecasts are underestimated and they are independent of number of calculated forecasts. For 1% significance level the results depend on the number of predictions. The VaR predictions are accurate and underestimated for almost all cases for 100 predictions but they are not accurate for window size 500-250 for 351 predictions. The Regulatory Loss Function is the smallest for the biggest size of sample and the Firm's Loss Function is the smallest for the smallest size of sample.

The results of loss functions are similar to the results of loss functions for the GDP (see Table 2). In all cases, the empirical significance levels are bigger than for the empirical significance levels for the GDP (see Table 2), respectively. The properties of the empirical time series from period to period could be changed. In this case, the RCA model is good enough to predict turning points for the window size 150 and 351 forecasts.

Conclusions

In this paper, the fundamental properties of the family Sign RCA models to obtain accurate VaR measures were presented. The GDP described by any models form the family Sign RCA models, has:

- the smallest value of the Regulatory Loss Function for the biggest size of sample,
- the smallest value of the Firm's Loss Function for the smallest size of sample,
- accuracy the VaR measures (the accuracy test rejects the null hypothesis).

If the GDP is one of the RCA, the RCA-MA and the Sign RCA-MA models the 200-400 sample are preferred. For these models, for each sample, VaR forecasts are underestimated. For the RCA-MA and the Sign RCA-MA models the smallest sample is for the smallest significant level. This dependence does not occur for the RCA model and the choice of the size of sample depends on the parameters values. When the GDP is the Sign RCA model then the smallest sizes of samples (i.e. 100 observations) are preferred. Overestimated VaR forecasts by the Sign RCA models for all sample are obtained.

References

Appadoo, S.S., Thavaneswaran, A., Singh, J. (2006), RCA Models with Correlated Errors, *Applied Mathematics Letters*, 19, 824–829.

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

Brzeszczyński, J., Kelm, R. (2002), *Ekonometryczne modele rynków finansowych.* WIG-Press, Warszawa.

Engle, R. F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50, 987–1006.

Doman, M., Doman, R. (2004), *Ekonometryczne modelowanie dynamiki polskiego rynku finansowego,* Wyd. AE w Poznaniu, Poznań.

Lee, S. (1998), Coefficient Constancy Test in a Random Coefficient Autoregressive Model, *Journal of Statistical Planning and Inference*, 74, 93–101.

Lopez J.A., *Methods for Evaluating Value-at-Risk Estimates*, „FRBNY Economic Policy Review", 1998.

Nicholls, D.F., Quinn, B.G. (1982), *Random Coefficient Autoregressive Models: An Introduction*, Springer, New York.

Sarma M, Thomas S., Shah A. (2003), Selection of Value-at-Risk Models, *Journal of Forecasting*, 22, 337-358.

Thavaneswaran, A., Appadoo, S.S., Bector C.R. (2006), Recent Developments in Volatility Modeling and Applications, *Journal of Applied Mathematics and Decision Sciences*, 2006, 1–23.

Thavaneswaran, A., Appadoo, S.S. (2006), Properties of a New Family of Volatility Sing Models, *Computers and Mathematics with Applications*, 52, 809–818.