The Forecast of Credit Risk of Romania

Simona Andreea DOSPINESCU[1]

Abstract: In this paper we propose to forecast the credit risk in Romania using Box-Jenkins metholody. In this purpose, we used as variable the rate of credit risk for the period 4th quarter of 1995 - 1st quarter of 2014. The treatment of the data is done with the EViews software. The results obtained in this paper show what will happen in the future and can be used as arguments in taking the adequate decisions for prevent or diminuate the consequences of this phenomenon.

Keywords: time series; ARIMA; Box-Jenkins metholody.

JEL Classification: B22, C22, C53, C87, E51.

1 Introduction

The forecast of the credit risk is based on using the indicator the rate of credit risk. In this paper we propose using Box-Jenkins metholody for forecasting the credit risk in Romania for a period of 3 quarters.

The random nature of the credit risk requires as forecasting method the Box-Jenkins methodology. The Box-Jenkins methodology refers to a systematic method for identify, estimate, test and use of models for time series integrated autoregressive and moving average (ARIMA). It is suitable for medium and long time series length (more than 50 observations).

The data used in this analysis have been extracted from the official data of National Bank of Romania. The analysis was carried out for the period 4th quarter of 1995 – 1st quarter of 2014.

2. Method

For realizing the forecast of the analyzed time series we will use classical methods, but also modern methods, such as Box-Jenkins methodology.

2.1. General Elements

Any time series X_t can be, according to Box-Jenkins methodology, as a combination of values and / or errors from the past, X_t and / or e_t.

$$X_t = \emptyset_0 + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q}$$

In order to model the real time series using the equation above are necessary four steps. First, the original series X_t must be transformed to become stationary around its own mean and variance. Second, the values of p and q must be correctly calculated. Third, the values of the parameters $\emptyset_1, \emptyset_2, \ldots, \emptyset_p$ și / sau $\theta_1, \theta_2, \ldots, \theta_q$ must be estimated using non-linear optimization procedures that

[1] Alexandru Ioan Cuza University of Iași, Romania, simona_andreea3@yahoo.com

minimizes the sum of squared errors. Finally, must find practical ways of modeling the seasonal series and calculate the values of the corresponding parameters.

2.2. Steps of Box&Jenkins Methodology

For applying the Box-Jenkins methodology we have to follow the next steps:

- analysis of the series;
- identification of the model. This step has the purpose to detect seasonality and to identify the order of seasonal autoregressive terms and seasonal moving average terms. In this stage is calculated the estimated autocorrelation function (FAC) and the estimated partial autocorrelation function (FACP). These functions measure the statistical dependence between observations of data outputs;
- estimation of ARIMA's parameters. The estimation of ARIMA's parameters is achieved by nonlinear least squares method. The values of the model coefficients are determined in relation to a particular criterion, one of this may be the maximum likelihood criterion. It can be shown that the likelihood function associated with a correct ARIMA model, used to determine the estimates of maximum likelihood of the parameters, contains all the useful information from data series about the model's parameters (Popescu T. and Demetriu S., (1991));
- diagnostic checking. In this stage it is assumed that the errors represent a stationary process and the residues are an white noise type (or independent if the distribution is normal) with a normal distribution with mean and variance constant. The tests used to validate the model are based on the estimated residues. It is checked that the components of this vector are autocorrelated. If there is autocorrelation, the checked model is not correctly specified. In this case the dependencies between the components series are specified in an incomplete manner and we have to return to the model identification step and try another model. Otherwise, the model is good and can be used to make predictions for a given time horizon (Tudorel, A. (2003)).
- forecasting.

3. The empirical analysis

We consider the time series provided by the National Bank of Romania (www.bnro.ro) for the rate of credit risk in Romania (RRC), between 4th quarter of 1995 and 1st quarter of 2014.

3.1. Analysis of The Series

The rate of credit risk is presented in the following figure. During the analyzed period, 4th quarter of 1995 – 1st quarter of 2014, the rate of credit risk presents an increasing trend. This evolution is due to the economic crisis established in July 2007. In 2008 the crisis worsened as stock markets arround the world collapsed and became unstable and the credit risk began to increase.

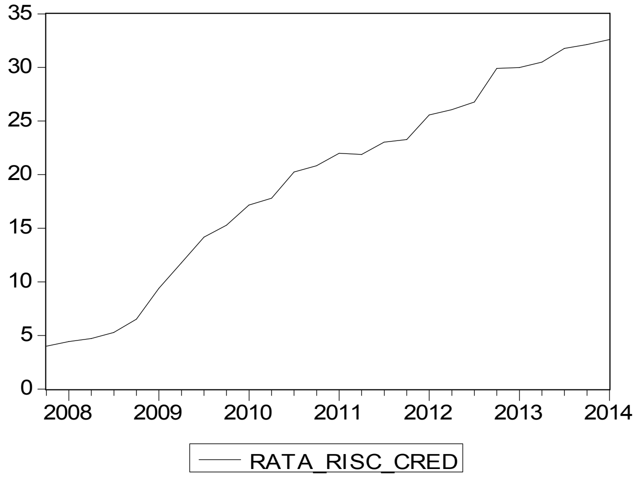


Figure 1 The evolution of credit risk in Romania, between 2007 and 2014

As we see in the 1st figure, the series admits a deterministic trend and may be non-stationary. Numerically, this is demonstrated by the results using Augmented Dickey – Fuller test.

The results obtained in the following table confirm that the analyzed variable is not stationary and is, at least, integrated of order 1.

Table 1 Testing the rate of credit risk stationarity using Augmented Dickey-Fuller test

Variable\Tested model	Intercept	Trend and intercept	None
ADF	-1.424964	-1.068683	1.954179
Probability	0.5531	0.9138	0.9850
Akaike	2.851005	2.897140	3.196326
Schwarz	2.998262	3.093482	3.294497

Source: Data processed using the statistical program EView

To transform the variable in a stationary one we have to differentiate for order 1 the variable the rate of credit risk (DRRC) (Chirilă V. (2013)) and then we test its stationarity with the Augmented Dickey-Fuller test.

Table 2 Testing differentiated variables stationarity using Augmented Dickey-Fuller test

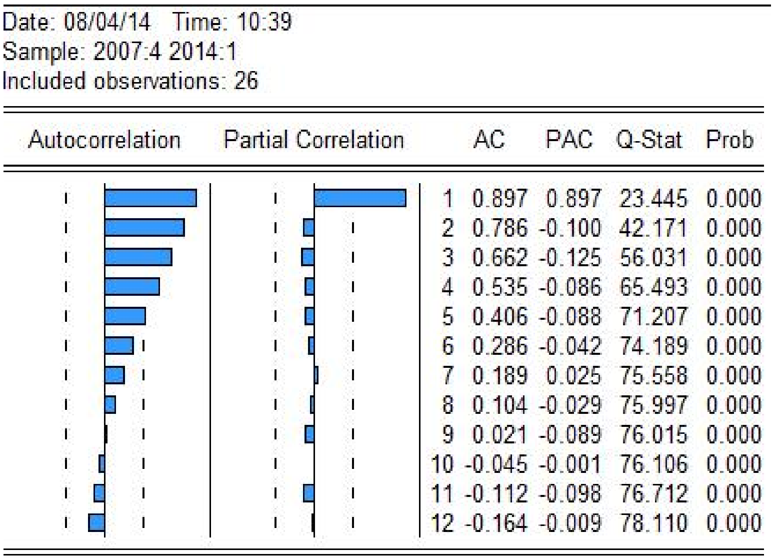
Variable\Tested Model	Intercept	Trend and intercept	None
ADF	-2.711668	-3.165746	-1.131128
Probability	0.0873	0.1157	0.2268
Akaike	2.917406	2.857373	3.084438
Schwarz	3.065514	3.054850	3.183177

Source: Data processed using the statistical program EViews

Because for the analyzed variable, DRRC, the Augmented Dickey-Fuller test don't offer the results that confirm us its stationarity, we include another 2 tests, Philips-Perron test and KPSS test (Chirilă, V. (2012)). The results obtained are presented in the Annex 1 and confirm the stationarity of the variable.

3.2. Identification of The Model

To determine whether the rate of credit risk in Romania (RRC) is the autocorrelated is performed the correloram of this variable.



Date: 08/04/14 Time: 10:39
Sample: 2007:4 2014:1
Included observations: 26

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.897	0.897	23.445	0.000
		2	0.786	-0.100	42.171	0.000
		3	0.662	-0.125	56.031	0.000
		4	0.535	-0.086	65.493	0.000
		5	0.406	-0.088	71.207	0.000
		6	0.286	-0.042	74.189	0.000
		7	0.189	0.025	75.558	0.000
		8	0.104	-0.029	75.997	0.000
		9	0.021	-0.089	76.015	0.000
		10	-0.045	-0.001	76.106	0.000
		11	-0.112	-0.098	76.712	0.000
		12	-0.164	-0.009	78.110	0.000

Figure 1 The credit risk correlogram

The probabilities associated with Ljung-Box test (Q-Stat) show that the variable analyzed is autocorrelated.

3.3. Estimation of ARIMA's Parameters

The graphical representation of the rate of credit risk indicate the existence of a deterministic trend, that can be linear or parabolic. So, we have to determine which is the best model of deterministic trend to use for forecast. Also, the rate of credit risk is an autocorrelated variable and we have to estimate the parameters of the autoregressive model.

The model with linear trend (Jaba, E. (2002)) takes the form:

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

The model with parabolic trend takes the form:

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t$$

The estimation results of the two types of trends are shown in the Annex 2.

The estimated model which include the linear trend, but also the offset variable, is:

$$Y_t = 6{,}112 + 1{,}097t + 0{,}846Y_{t-1} + e_t$$

Because the estimation of the parameter, β_1, is greater than zero, we deduce that the credit risk has an increasing trend.

The estimated model which include the parabolic trend, but also the offset variable, is:

$$Y_t = 0,008 + 2,039t - 0,029t^2 + 0,516Y_{t-1} + e_t$$

Because the estimation of the parameters, β_2, from the parabolic trend is negative indicate us that this curve admits a peak.

3.4. Diagnostic Checking

The estimated regression models respect the assumptions regarding the errors. The results of testing the hypothesis of errors normality, the lack of autocorrelation, homoscedasticy and the error mean is zero.

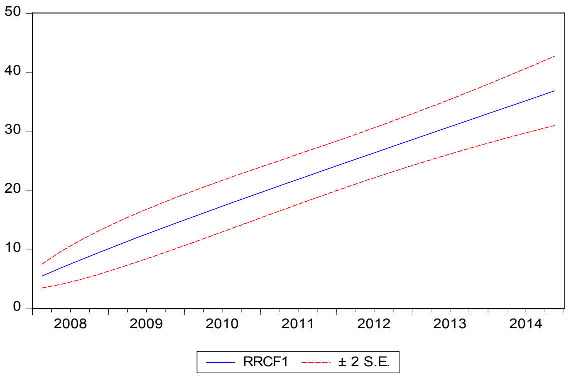
The estimated models are statistically significant because the correlation ratio of each model is significantly different from zero. The estimators models are also significantly different from zero.

3.5. Forecast

The series was modeled taking into account, as suggested schedule a linear trend and a parabolic trend. Therefore we should choose the best model to achieve forecast.

The results obtained (Eviews program) to forecast the model including linear trend are shown in the table below.

Table 5 The forecast of credit risk rate using the linear trend



Forecast: RRCF1	
Actual: RRC	
Forecast sample: 2007Q4 2014Q4	
Adjusted sample: 2008Q1 2014Q4	
Included observations: 25	
Root Mean Squared Error	1.406597
Mean Absolute Error	1.146372
Mean Abs. Percent Error	10.64083
Theil Inequality Coefficient	0.032231
Bias Proportion	0.020416
Variance Proportion	0.201249
Covariance Proportion	0.778335

Source: Data processed using the statistical program EViews

In summary, the table bellow presents some statistics which allow us to choose the best model of the rate of credit risk evolution, that will be used for the forecast. Since there are two models of the same variable we can use to compare the models also the indicators built on Akaike and Schwartz information theory. Since the two models have different numbers of parameters, we use the ratio of adjusted determination.

Table 6 Choosing the model of the rate of credit risk evolution

The Statistic Indicator	Model - Linear trend	Model - Parabolic trend
Root Mean Squared Error	1.406597	**0.883161**
Mean Absolute Error	1.146372	**0.751370**
Mean Abs. Percent Error	10.64083	**5.367873**
Theil Inequality Coefficient	0.032231	**0.020040**
Akaike	2.827134	**2.536816**
Schwartz	2.973399	**2.731836**
Adjusted R-squared	0.989483	**0.992392**

Source: Data processed using the statistical program EViews

The minimum values of the mean square error of adjustment, medium absolute error, medium error as percent, Theil's inequality coefficient, Akaike criterion and Schwartz criterion indicates the best model. The maximum ratio of adjusted determination indicates the best model.

All indicators selected to determine the best forecasting model indicates the model with parabolic trend. The expected values that takes into account the parabolic trend (RRCF2) obtained using Eviews are shown in Figure 3 and Annex 3.

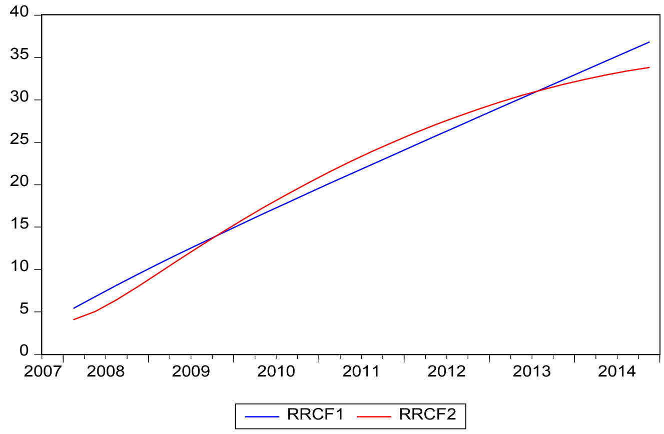


Figure 3 The forecasted values of the rate of credit risk using the model with linear trend (RRCF1) and parabolic trend (RRCF2)

The figure of the forecasted values for the rate of credit risk reveals that the model with parabolic trend shows an increasing evolution, but lower than the model with linear trend.

The predicted values for the rate of credit risk are: for the 2nd quarter of 2014 - 34.61 (model with linear trend), respectively, 32.95 (parabolic trend model), for the 3rd quarter of 2014 - 35.72 (model with linear trend), respectively, 33.42 (parabolic trend model) and for the 4th quarter - 36.82 (model with linear trend), respectively, 33.82 (parabolic trend model).

4. Conclusions

The Box-Jenkins methodology gives the possibility to forecast the rate of the credit risk in Romania and, by this, helps taking the adequate decisions for prevent or diminuate the consequences of this phenomenon.

During the period 4th quarter of 2007 and 1st quarter 2014, the rate of credit risk shows an increasing trend. This evolution is mainly due the economic crisis that was established in July 2007. In 2008 crisis has worsened as stock markets around the world collapsed and became unstable, and credit risk began to rise.

The predicted values for the model with linear trend and parabolic trend model are higher than those registered, which means that the rate of credit risk continue to rise.

5. References

Andrei, T. (2003). *Statistică şi econometrie/Statistics and econometrics*, Bucharest, Economic Publishing

Chirilă, V. & Chirilă C. (2012), Relation Between Expected Return and Volatility at Bucharest Stock Exchange, on Business Cycle Stages", *Annales Universitatis Apulensis, Series Oeconomica,* Nr. 14, Vol. 1, pp. 149-163.

Chirilă, V. (2013) *Analiza econometrică a burselor de valori şi a ciclurilor de afaceri/Econometric analysis of stock markets and business cycles*, ASE Publishing, Bucureşti

Jaba, E. (2002) Statistica, ediţia a 3-a, Editura Economică, Bucureşti

Popescu, S., (1991), *Practica modelării şi predicţiei seriilor de timp/Practical modeling and prediction of time series*, Bucharest, Technical Publishing

Raport Statistic BNR (2014). Web page. Retrieved from http://www.bnr.ro/Raport-statistic.html, date: 09.09.2014

Annex 1. Testing stationarity of the variable DRRC using Phillips-Perron test and KPSS test

Table 1 Testing stationarity of the variable using Phillips-Perron test

Variable\Tested model	Intercept	Trend and intercept	None
PP	-4.795952	-4.996850	-2.222521
Probability	0.0009	0.0027	0.0280
Akaike	2.859970	2.869339	3.273054
Schwarz	2.958141	3.016596	3.322139

Table 2 Testing stationarity of the variable using KPSS test

Variable\Tested model	Intercept	Trend and intercept
KPSS	0.173794	0.085376
Critical value	0.463000	0.146000
Akaike	2.758863	2.802907
Schwarz	2.807618	2.900417

Annex 2. The estimation of the models with linear trend and parabolic trend for the rate of credit risk

Table 1 Estimation of linear deterministic trend

Dependent Variable: RRC
Method: Least Squares
Sample (adjusted): 2008Q1 2014Q1
Included observations: 25 after adjustments
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.111996	4.393740	1.391069	0.1781
@TREND	1.097452	0.218863	5.014336	0.0001
AR(1)	0.845773	0.137295	6.160265	0.0000

R-squared	0.990359	Mean dependent var	20.12360
Adjusted R-squared	0.989483	S.D. dependent var	9.169730
S.E. of regression	0.940395	Akaike info criterion	2.827134
Sum squared resid	19.45555	Schwarz criterion	2.973399
Log likelihood	-32.33917	Hannan-Quinn criter.	2.867702
F-statistic	1129.968	Durbin-Watson stat	1.872430
Prob(F-statistic)	0.000000		

Source: Data processed using the statistical program EViews

Table 2 Estimation of parabolic deterministic trend

Dependent Variable: RRC
Method: Least Squares
Sample (adjusted): 2008Q1 2014Q1
Included observations: 25 after adjustments
Convergence achieved after 4 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.008902	1.575119	0.005652	0.9955
@TREND	2.039890	0.235335	8.668024	0.0000
@TREND^2	-0.029717	0.007817	-3.801547	0.0010
AR(1)	0.516693	0.158561	3.258632	0.0038

R-squared	0.993343	Mean dependent var	20.12360
Adjusted R-squared	0.992392	S.D. dependent var	9.169730
S.E. of regression	0.799832	Akaike info criterion	2.536816
Sum squared resid	13.43435	Schwarz criterion	2.731836
Log likelihood	-27.71020	Hannan-Quinn criter.	2.590906
F-statistic	1044.492	Durbin-Watson stat	1.943140
Prob(F-statistic)	0.000000		

Source: Data processed using the statistical program EViews

Annex 3 The forecasted values for the rate of credit risk

Table 1. The forecasted values for the rate of credit risk using the model with linear trend (RRCF1) and the model with parabolic trend (RRCF2)

Quarters	RRCF1	RRCF2
2007Q4	NA	NA
2008Q1	5.423179	4.081249
2008Q2	6.796121	5.035326
2008Q3	8.126576	6.411664
2008Q4	9.421095	7.977456
2009Q1	10.68522	9.612413
2009Q2	11.92364	11.25438
2009Q3	13.14032	12.87125
2009Q4	14.33861	14.44642
2010Q1	15.52135	15.97133
2010Q2	16.69094	17.44154
2010Q3	17.84940	18.85476
2010Q4	18.99845	20.20981
2011Q1	20.13954	21.50608
2011Q2	21.27390	22.74326
2011Q3	22.40257	23.92118
2011Q4	23.52643	25.03975
2012Q1	24.64621	26.09893
2012Q2	25.76255	27.09871
2012Q3	26.87598	28.03907
2012Q4	27.98694	28.91999

2013Q1	29.09582	29.74149
2013Q2	30.20293	30.50356
2013Q3	31.30856	31.20619
2013Q4	32.41292	31.84939
2014Q1	33.51622	32.43316
2014Q2	**34.61862**	**32.95749**
2014Q3	**35.72025**	**33.42239**
2014Q4	**36.82124**	**33.82786**