10th International Scientific Conference
"Business and Management 2018"
May 3–4, 2018, Vilnius, LITHUANIA
Section: Financial Engineering
http://www.bm.vgtu.lt

ISSN 2029-4441 / eISSN 2029-929X
ISBN 978-609-476-119-5
eISBN 978-609-476-118-8
doi: 10.3846/bm.2018.51
https://doi.org/10.3846/bm.2018.51

FORECASTING CONSUMER PRICE INDEX (CPI) USING TIME SERIES MODELS AND MULTI REGRESSION MODELS (ALBANIA CASE STUDY)

Eralda Gjika (Dhamo)[1], Llukan Puka[2], Oriana Zaçaj[3]

[1, 2]Department of Applied Mathematics, Faculty of Natural Science, University of Tirana, Tirana, Albania
[3]Department of Mathematics, Faculty of Mathematics and Physics Engineering, Polytechnic University, Tirana, Albania
E-mails: [1]eralda.dhamo@fshn.edu.al (corresponding author); [2]llukan.puka@fshn.edu.al; [3]ana_25al@yahoo.com

Abstract. In this work we analyse the CPI index as the official index to measure inflation in Albania, Harmonized Indices of Consumer Prices (HICPs) as the bases for comparative measurement of inflation in European countries and other financial indicators that may affect CPI. This study is an attempt to model CPI based on combination of multiple regression model with time series forecasting models. In the first approach, time series models were used directly on the CPI time series index to obtain the forecast. In the second approach, the time series models (SARIMA, ETS) were used to model and simulate forecast for each subcomponent with significant correlation to CPI and then use the multiple regression model to obtain CPI forecast. The projection of this indicator is important for understanding the country's economic and social development. This study helps researchers in the field of time series modeling, economic analysis and investments.

Keywords: CPI, HICPs, time series, regression, forecast, R software.

JEL Classification: C, C53.

1. Introduction

It is now known that most of Albania's foreign trade is developing with the Eurozone countries, and most of the financial and credit grants accorded to Albania by the EU are in the Euro. Likewise, foreign currency earnings from emigration, as long as their shippers are mostly located in countries that are already part of the Eurozone (mostly in Greece, Italy, Germany, etc.) are dominated by Euro currency. Albania is one of the Western Balkan countries aspiring to become part of the European Union countries. Looking at Albania's developments in the fields of economy, justice, education, social policies and not only, many opinion leaders see Albania's candidacy as one of the top favourites to join the eurozone countries. Albania expects the decision of the EU Council of Ministers to open access negotiations in the first six months of 2018.

The existence of inflation in Albania for the first time was declared in 1990 when it was publicly admitted that the Albanian economy had long been eroded by hidden inflation accumulated over the years. Officially, inflation began to decline since July 1992. In Albania, the consumer price index (CPI) is used for inflation measurement. Throughout the last decade, the basket of goods – 215 total items divided into eight main groups were based on the survey conducted by INSTAT (Institute of Statistics in Albania) in 1993. The family budget survey in December 1999 – November 2000 created the possibility of building a new basket, which, in addition to a more realistic distribution of group weights, contains more items, with a total of 262 pieces. The period 1999–2000, recorded unusual inflation rates for an economy like that of Albania. Although during 1999 Albania was forced to withstand the exodus of Kosovo's population, inflation remained at low rates.

However, the last month of 2000 appears to have given the first signal to the end of the deflationary period in Albania as it recorded a monthly rate of 4.6 percent. The ongoing period up to our days has been characterized by positive percentages, which, though within the boundary, again have a growing trend. This trend became particularly strong in July 2001, when annual inflation surpassed the upper boundary limit by 5.6 percent. Inflation at the end of 1996 was almost three times

higher than a year before due to pyramid schemes, and the premise for further growth was more and more present. Then the economic crisis that included the country in 1997–1999 shock the values of the inflation in Albania. After 1999, inflation rates in Albania started to stabilize in trend and seasonality patterns (see Figure 1).

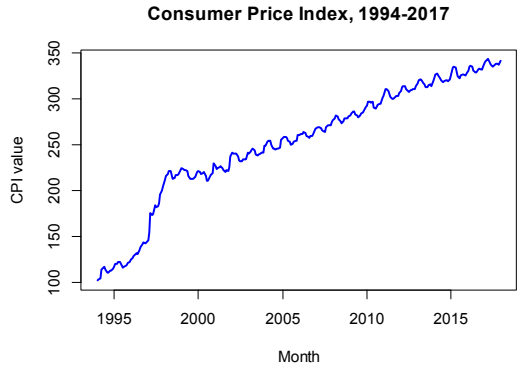

Figure 1. Consumer Price Index in Albania (from January 1994 to December 2017) (source: authors)

The Consumer Price Index (CPI) is one of the main economic indicators constructed to measure the changes over time in the prices of consumer goods and services acquired, used or paid for by households. Consumer Price Index (CPI) is a key variable for macroeconomic analysis of a country. Albanian has adopted the Harmonised Index of Consumer Prices (HICP) in December 2015. It is a consumer price index, which is compiled according to a methodology that has been harmonised across EU countries.

This study is an attempt to model the Albanian CPI indice using multivariate methods combined with time series forecasting models. In the first approach, time series models were used directly on the CPI time series index to obtain the forecast. In the second approach the Seasonal Autoregressive Integrated Moving Average time series models proposed by Box and Jenkins (1970) and Exponential Smoothing Methods presented by Ord, Koehler, and Snyder (1997), Hyndman (2001), Hyndman, Koehler, Snyder, and Grose (2002) and Hyndman, Koehler, Ord, and Snyder (2005). Hyndman et al. (2002) were used to model and simulate forecasts for each subcomponent with significant correlation with CPI and then used for the multi regression model to obtain CPI forecasts. The data were obtained from the Institute of Statistics in Albania, Bank of Albania and the time series were analyzed in all their components (error, trend, seasonality) through R software.

2. Previous research and data analysis

Many authors have studied the dependency of Consumer Price Index (CPI) from other economic variables and time series models (Aka & Pieretti, 2008; Norbert, Wanjoya, & Waititu, 2016).

Multiple regression models are often used to construct explanatory models for economic variables (Bowerman, O'Connell, & Koehler, 2005). In her paper, Popescu (2017) has built a multi regression model to explain total CPI from CPI of food commodities, CPI for non-food goods and CPI for services (case study: Romania). Falnita and Sipos (2007) also propose a multi regression model to explain CPI based on 7 independent variables (exchange rate, interest rate, production price etc.). Other authors work on inflation forecasting models are: Marcellino (2007), Takatoshi and Kiyotaka (2006), Fat Codruta and Dezsi (2011), MacDonald, Menkhoff, and Rebitzky (2009), Gjika, Gjecka, and Zacaj (2016), Jonathan, D. Church (2016).

In their work Lufi and Sinaj (2015) studied the relationship between inflation and the exchange rate and showed that there is a dependency between them. According to a VAR model and the numerical tests it was shown that the CPI is explained by the exchange rate. Also, Zhang, Che, Xu, and Xu (2013) have presented a forecast model for CPI in China for the period 1995–2008 and showed that ARMA model has a forecast accuracy relatively high. Other authors who propose time series model for CPI are Ngailo, Luvanda, and Massawe (2014), Junior, Salomon and Pamplona (2014), Kharimah, Usman, Widiarti, and Elfaki (2015), Kalluci and Gjika (2015).

Among the main indicators influencing inflation measurement are gross domestic product, unemployment rate, exchange rate, economic wellbeing, monetary capability to travel abroad (here we have considered the number of Albanian people traveling abroad). The time series of changes in CPI values, changes in GDP, exchange rate, unemployment rate and the number of Albanian people traveling abroad are reflected in the Figure 2 and Figure 3.

The time series of CPI change rate and GDP change rate values after 2012 have a similar trend of values in quarter and similar periodic behavior over the years. Otherwise, it has occurred with the unemployment rate which has a noticeable change in the second half of 2015. The downward trend of this rate is quite rapid and almost linear.

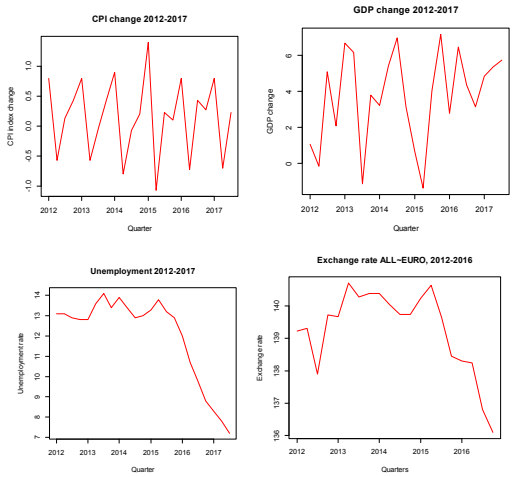

Figure 2. CPI change, GDP change, Unemployment and exchange rate in Albania (quarterly from January 1994 to December 2017) (source: authors)

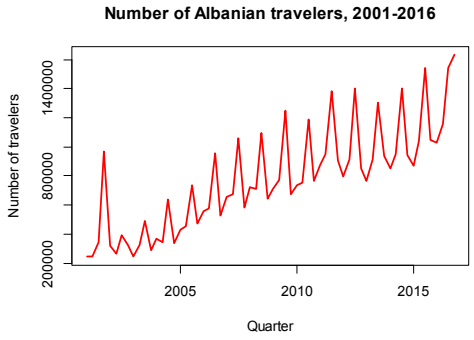

Figure 3a. Time series of number of Albanian people traveling abroad (quarterly from 2001 to 2016) (source: authors)

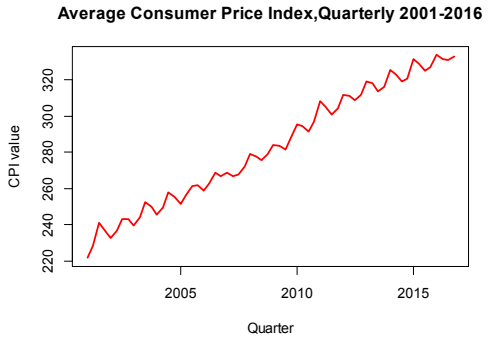

Figure 3b. Time series of Consumer Price Index (quarterly from 2001 to 2016) (source: authors)

Calculating the correlation coefficient (Pearson coefficient) for the quarterly time series we have a value of –0.062 between CPI change rate and GDP change rate; a value of –0.0045 between CPI change rate and unemployment rate; a value of –0.3427 between GDP change rate and unemployment rate. The values show no significant correlation between these indices. It results that the linear model of CPI change with explanation variables GDP change and unemployment rate is not appropriate for Albanian data, it has a value of Multiple R-squared equal to 0.00461.

If we look at Figure 3a the number of Albanians peoples traveling abroad (here are considered air travel, sea and land routes) we notice a growing trend associated with seasonal and periodic behavior. As we can notice on time series of Albanians peoples traveling outside the country and CPI time series for the period 2001–2016 we see a growth trend in both time series and a similar seasonality. The similar behavior of the two time series give us signals that they can be correlated with each other. Number of Albanian travelers outside the country is a quarterly observed time series and the CPI time series is monthly this makes difficult to obtain a mathematical model of CPI explained by number of Albanian travelers outside the country. In our study, we have used as a CPI time series the time series of the average value for the quarters of each year.

3. Multiple linear regression and time series, methodology

The common model to explain inflation is Philip's curve which states there is an inverse correlation between the unemployment rate and inflation. But in the case of Albanian data this model seems to be unappropriated to explain inflation. The unemployment rate as it can be seen from Figure 2 has a linear descending trend

The models we have choose for comparative analysis are: the linear multi regression model and Seasonal Autoregressive Moving Average model. The first model is a linear regression model which is given in Eq. (1), but in our case, it has two explanatory variables:

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n + \varepsilon\,, \qquad (1)$$

where: the dependent variable is CPI and the independent variable is the number of Albanian travelers abroad (by air, sea and land). The correlation coefficient between the number of Albanian people traveling abroad every quarter and the average value of CPI in every quarter has a significant value of 0.79 (Pearson coefficient) which means the two variables are positively correlated. The correlation coefficient between CPI and exchange rate is 0.62. And the correlation coefficient between the two independent variables, the number of Albanian people traveling abroad and the exchange rate is 0.38.

The linear regression model of CPI and Albanian people traveling abroad every quarter (Model 1) is given in Eq. (2):

$$CPI_{quarterly} = 2.271e+02 + 7.207e-05 \cdot OUT_{quarterly} \quad (2)$$

and, $R^2 = 0.6238$.

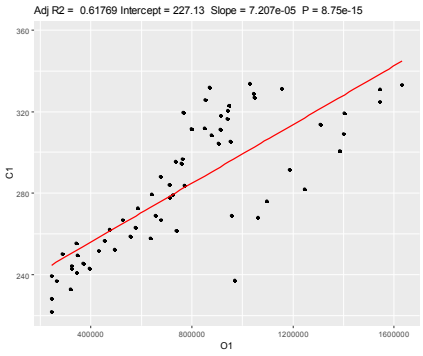

Figure 4. CPI versus number of Albanian people traveling abroad, (quarterly from 2001 to 2016) (source: authors)

Figure 4 is the scatterplot of CPI and number of Albanian people traveling abroad. It is seen that there a correlation exists between these two variables (refer to the value of the correlation coefficient between these variables 0.79)

The linear model of CPI explained by the exchange rate (Model 2) is given in Eq. (3):

$$CPI_{quarterly} = -97.49 + 2.86 \cdot EX_{quarterly} \quad (3)$$

and, $R^2 = 0.3852$.

The third regression model is a multiple linear regression model with two independent variables: Exchange rate (EXquarterly) and number of Albanian travelers abroad (OUTquarterly). The model in this case (Model 3) is:

$$CPI_{quarterly} = 5.839 + 1.741 \cdot EX_{quarterly} + 5.914e-05 \cdot OUT_{quarterly} \quad (4)$$

and, $\mathbf{R^2 = 0.7459}$. A *t-test* for model 3 show that $EX_{quarterly}$ and $OUT_{quarterly}$ have statistically significant predictive capability for CPI, with *p-values* respectively: *1.10e-06* and *2.61e-13*.

Figure 5 shows a clear view of the actual values versus the fitted values for the three models. We can state from the graphical representation that the most consistent model from the three models is the two-variable explanatory model.

The QQ-plot for the two first models (Model 1 and Model 2) show a non regular trend of the quantiles compared to the third model. This is a sign of a good fit of Model 3 for the CPI data.

The goodness of fit from Model 3 is also seen in the residual plot and ACF of the residuals.

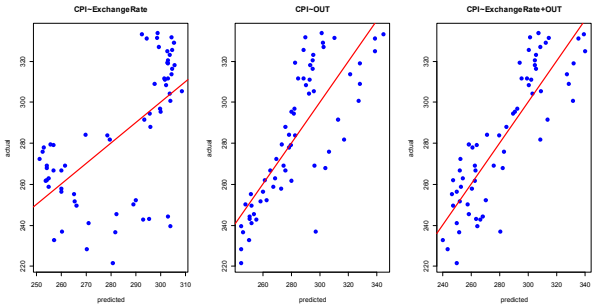

Figure 5. Actual values versus fitted values from each model (source: authors)

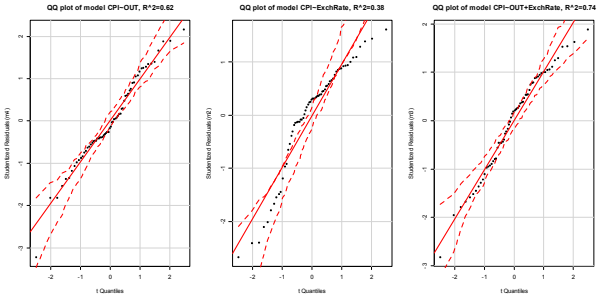

Figure 6. The QQ-plot for the three models (source: authors)

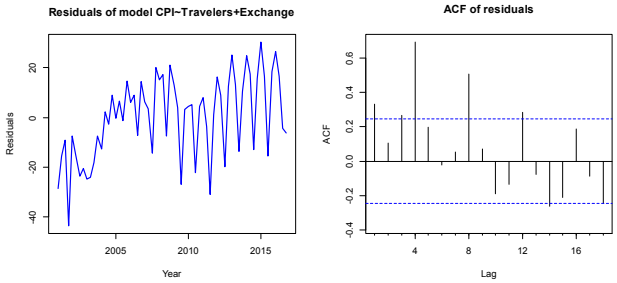

Figure 7. The residuals plot and the ACF of the residuals for Model 3 (source: authors)

The residual plot and the ACF of the residuals for Model 3 (explanatory variables: number of travelers and exchange rate) show that the model has captured the patterns in the data quite well, although there is a small amount of autocorrelation left in the residuals (seen in first part of the residual plot and the significant spikes in the ACF plot). This suggests that the model can be slightly improved, although it is unlikely to make much difference to the resulting forecasts.

The histogram of the residuals (see Figure 8) for Model 3 show a normal distribution which makes the forecast reliable.

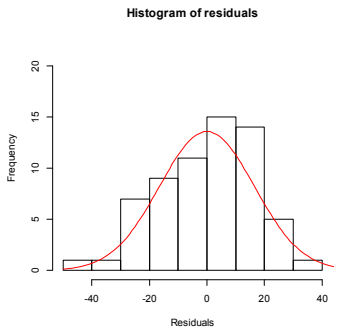

Figure 8. Histogram of the residuals for Model 3 (source: authors)

Concerning the multicollinearity we have used the Variance Inflation Factor (VIF): VIF = 1/(1-R2). The calculations for Model 3 give us a value VIF = 3.84. A value of VIFs lower than 4 shows no multicollinearity. The calculated values of VIFs are quite satisfactory meaning that we may consider the proposed models to forecast CPI.

The next model taken under consideration for CPI is a time series model (Seasonal Autoregressive Integrated Moving Average) which deals with observations collected over equally spaced, discrete time intervals and is given by Eq. (5):

$$\Phi_p(B^s)\phi(B)\nabla_s^D\nabla^d X_t = \alpha + \Theta_Q(B^s)\theta(B)w_t, \quad (5)$$

where:

s = seasonal lag,
ϕ = coefficient for AR process,
Φ = coefficient for seasonal AR process,
θ = coefficient for MA process,
Θ = coefficient for seasonal MA process.

B is the backward shift operator, $\nabla_s^D = (1 - B^s)^D$ and $\nabla^d = (1 - B)^d$, w_t is an uncorrelated random variable with mean zero and constant variance (Box & Jenkins, 1970).

We have used the SARIMA model to forecast the values for each of the time series: Consumer Price Index, Albanian travelers outside the country and Exchange rate. There are in total 64 observations for each time series, from year 2001 to 2016 (quarterly). We have also taken into consideration the Exponential smoothing model presented by Hyndman and Khandakar (2008) and calculations are made through **forecast** package in R software (Hyndman, 2008a, 2008b, 2008c, 2008d; Hyndman & Khandakar, 2008). We have based our "best model selection" procedure on Akaike Information criterion (AIC, AICc), Bayesian Information criterion (BIC).

Akaike Information Criterion is defined to be:

$$AIC = -2\log(Likelihood) + 2p, \quad (6)$$

where p is the number of estimated parameters in the model. Minimizing the *AIC* gives the best model for prediction. It is recommended to consider several models (on same data set) with AIC values close to the minimum. (A difference in *AIC* values of 2 or less is not regarded as substantial and we may choose the simpler but non-optimal model.).

It is recommended to consider several models (on same data set) with AIC values close to the minimum. Shumway and Stoffer (2011). (A difference in *AIC* values of 2 or less is not regarded as substantial and we may choose the simpler but non-optimal model).

AIC corrected (for small sample bias) is defined to be:

$$AIC_C = AIC + \frac{2(p+1)(p+2)}{n-p} \quad (7)$$

Schwartz' Bayesian criterion is defined by:

$$BIC = AIC + p(\log(n) - 2) \quad (8)$$

Table 1 shows the values of the information criteria for the two time series models. The lowest value of the information criteria indicates the "best" predictive model.

Table 1. The accuracy of the models based on the information criteria (source: authors)

Model	Time Series	AIC	AICc	BIC
Sarima	CPI	**268.8**	269.6	277.2
	Traveler	**1590.3**	1590	1594
	Exchange	**277.7**	277.9	282
Exponential	CPI	398.4	401.7	417.8
	Traveler	1765	1768	1784
	Exchange	374.7	375.1	381.2

The SARIMA proposed models are:

Time series:	**Albanian Travelers** ARIMA(0,0,0)(0,1,0)[4] with drift
CPI	ARIMA(1,0,1)(0,1,0)[4] with drift
Exchange rate	ARIMA(0,1,1)

The SARIMA models for the three time series have a lower value of AIC compared to the ETS model for each time series. So, based on the minimum value of information criteria (AIC) we have choose to use the SARIMA forecasted values of the two time series (exchange rate and

number of Albanian people travelling abroad) for predicting through the multiple regression model the values of CPI in 2017 and 2018.

4. Forecast of the Consumer Price Index

The next step of the study is to choose the model which best describes the CPI and obtain the forecasted values of Consumer Price Index (CPI). We have used the predictions of $EX_{quarterly}$ and $OUT_{quarterly}$ from the SARIMA model. The values obtained from the SARIMA model for the two time series are used as input values to the multiple regression model 3.

In Figure 9 are shown the forecasted values and the confidence intervals from the SARIMA model for the quarters of 2017 and 2018 for each time series.

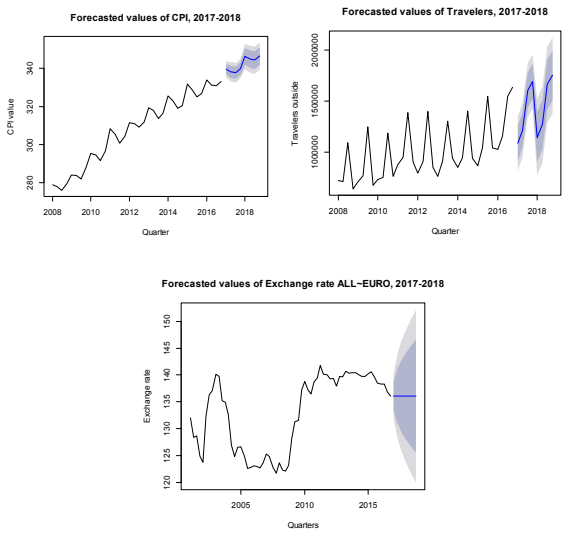

Figure 9. SARIMA model forecast for CPI, number of Albanian people traveling abroad and exchange rate. (quarterly forecast for 2017–2018) (source: authors)

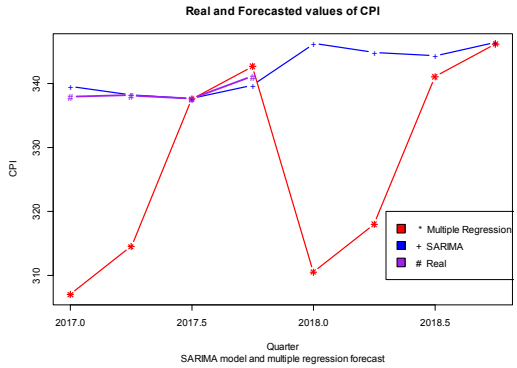

Figure 10. Real values and forecasted values of CPI for the quarters of 2017–2018

We have used the real values of CPI for the quarters of 2017 to discuss about the accuracy of the proposed models. In Figure 10 are shown graphically the real values of CPI for the quarters of 2017 and the forecasted values from the SARIMA model on the CPI time series; the forecasted values of CPI from the multiple regression model where we have used the forecasted values obtained from the SARIMA model for the two variables.

From Figure 10 it is noticed that time series model (SARIMA) for CPI is more suitable for predictions compared to multiple regression model. The confidence intervals for predictions are narrower in the SARIMA model compared to multiple regression model although this one has capture the seasonal behaviour of CPI.

5. Conclusions

Due to the difference in structure and evolution of a country economy there are no universal models to explain inflation.

In this work was studied the dependence of Consumer Price Index (CPI) in Albania from other economic indicators such as exchange rate and number of people traveling abroad. The study found that the two variables: the exchange rate and the number of Albanian people traveling abroad are important for the CPI's performance, they are positively related to CPI and together they explain it at 75%. Time series models combined with multiple regression were used to increase the quality of the forecast. The forecasts of the two main models were compared with the real observations for the 2017 quarters. Graphically, the forecasts obtained from the SARIMA models were satisfactory for a short-term prediction compared to the multiple regression model forecast.

The rising trend of the number of people traveling abroad and their seasonal behavior shows an increase in the level of the country's well-being. This is also seen in the growth of CPI. Considering this study in the econometric viewpoint we can conclude that the Albanian economy is growing satisfactory, which is clearly expressed by the ability of people to travel more abroad and the stability of the European currency (EURO) which positively affects the economic situation in the country, in the Balkan region and the premising position in Europe as well.

Disclosure Statement

The author(s) declare(s) that there is no conflict of interest.

References

Aka B. F., & Pieretii P. (2008). Consumer Price Index dynamics in a small open economy: A structural time series model for Luxembourg. *International Journal of Applied Economics*, *5*(1), 1-13.

Bowerman, B. L., O'Connell, R. T., & Koehler, A. B. (2005). *Forecasting, time series and regression: An applied approach*. Thomson Brooks/Cole, Belmont CA.

Box, G. E. P., & Jenkins, G. (1970). Time series analysis, forecasting and control, *Holden-Day*. San Francisco, CA.

Falnita, E., & Sipos C. (2007). A multiple regression model for inflation rate in Romania in the enlarged EU. *Munich Personal RePEc Archive*, Paper No. 11473, posted 11 November 2008. Retrieved from http://mpra.ub.uni-muenchen.de/11473/

Fat Codruta, M., & Dezsi, E. (2011). Exhange-rates forecasting: Exponential smoothing techniques and ARIMA models. *The Annals of the University of Oradea. Economic Sciences*, 499-508, Romania. Retrieved from http://steconomiceuoradea.ro/anale/volume/2011/n1/046.pdf

Gjika, E., Gjecka, A., & Zacaj, O. (2016). Projection of CPI through time series models (Albanian case), *Workshop: Aspects of Research in Applied Mathematics*. Faculty of Natural Science, University of Tirana. Retrieved from http://akma.morix.al/

Hyndman, R. J. (2001). It's time to move from what to why. *International Journal of Forecasting*, 17, 567-570.

Hyndman, R. J. (2008a). *Expsmooth: Data sets from "Forecasting with exponential smoothing"* by Hyndman, Koehler, Ord & Snyder (2008). R package version 1.11. Retrieved from http://CRAN.R-project.org/ package=forecasting

Hyndman, R. J. (2008b). fma: Data sets from "Forecasting: Methods and applications" By Makridakis, Wheelwright & Hyndman (1998). R package version 1.11. Retrieved from http://CRAN.R-project.org/package=forecasting

Hyndman, R. J. (2008c). Forecast: Forecasting functions for time series. R package version 1.11. Retrieved from http://CRAN.R-project.org/package=forecasting

Hyndman, R. J. (2008d). *Mcomp: Data from the M-Competitions*. R package version 1.11. Retrieved from http://CRAN.R-project.org/package=forecasting

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R, Monash University. *Journal of Statistical Software*, *27*(3). Retrieved from http://www.jstatsoft.org.
https://doi.org/10.18637/jss.v027.i03

Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, *18*(3), 439-454.
https://doi.org/10.1016/S0169-2070(01)00110-8

Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2005). Prediction intervals for exponential smoothing using two new classes of state space models. *International Journal of Forecasting*, *24*(1), 17-37.
https://doi.org/10.1002/for.938

Jonathan, D. Church. (2016, March,). Comparing the Consumer Price Index with the gross domestic product price index and gross domestic product implicit price deflator. *Monthly Labor Review*. U.S. Bureau of Labor Statistics. https://doi.org/10.21916/mlr.2016.13

Junior, P. R., Salomon, F. L. R., & Pamplona, E. O. (2014). ARIMA: An applied time series forecasting model for the Bovespa stock index. *Applied Mathematics*, *5*(21), 3383-3391. https://doi.org/10.4236/am.2014.521315

Kalluci, E., & Gjika, E. (2015). Difference equations and time series analysis in approximating the index prices (for Albania). *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, *2*(4). http://www.jmest.org/wp-content/uploads/JMESTN42350659.pdf

Kharimah, F., Usman, M., Widiarti, W., & Elfaki, F. A. M. (2015). Time series modeling and forecasting of the consumer price index Bandar Lampug, *Science International (Lahore)*, *27*(5), 4619-4624.

Lufi, B., & Sinaj, V. (2015). Exchange rate and inflation rate a casuality analysis: Case of Albania. *Journal of Multidisciplinary Engineering Science and Technology (JMEST)*, *2*(2).

MacDonald, R., Menkhoff, L., & Rebitzky, R. R. (2009). Exchange rate forecaster's performance: Evidence of skill?. *Cesifo Working Paper*, Nr. 2615.

Marcellino, M. (2007). *A comparison of time series models for forecasting GDP growth and inflation*. IEP-Università Bocconi, IGIER and CEPR.

Ngailo, E., Luvanda, E., & Massawe, E. S. (2014). Time series modeling with application to Tanzania inflation data. *Journal of Data Analysis and Information Processing*, *2*(2), 49-59.
https://doi.org/10.4236/jdaip.2014.22007

Norbert, H., Wanjoya, A., & Waititu, A. (2016). Modeling and forecasting consumer price index (case of Rwanda). *American Journal of Theoretical and Applied Statistics*, *5*(3), 101-107.
https://doi.org/10.11648/j.ajtas.20160503.14

Ord, J. K., Koehler, A. B., & Snyder, R. D. (1997). Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of the American Statistical Association*, *92*(440), 1621-1629.
https://doi.org/10.1080/01621459.1997.10473684

Popescu, A. (2017). Correlation analysis of consumer price indeces by multiple regression. *TJMM*, *9*(1), 43-50.

Shumway, R. H., & Stoffer, D. S. (2011). *Time series analysis and its applications With R examples* (Third edition). Springer Science+Business Media, LLC 2011.
https://doi.org/10.1007/978-1-4419-7865-3

Takatoshi, I., & Kiyotaka, S. (2006). Exchange rate changes and inflation in post-crisis asian economies: var analysis of the exchange rate pass-through. *Journal of Money, Credit and Banking*, *40*(7, October 2008). 1407–1438. http://www.nber.org/papers/w12395

Zhang, F., Che, W., Xu, B., & Xu, J. (2013). The Research of ARMA Model in CPI Time Series. In *Proceedings of the 2nd International Symposium on Computer, Communication, Control and Automation (ISCCCA-13)*. Atlantis Press, Paris, France.
https://doi.org/10.2991/isccca.2013.8