Forecasting Foreign Reserves in Bosnia and Herzegovina Using Box-Jenkings Methodology

Dejan Kovacevic
Central Bank of Bosnia and Herzegovina, Marsala Tita 25, 71000 Sarajevo
e-mail: dejan.kovacevic@cbbh.ba Tel. +387-33-278-272

October 2017

Abstract

In this paper we select an appropriate univariate time series model for forecasting monthly time series of foreign reserves. The paper considers both, a standard procedure for seasonal unit root testing as well as extended procedures with endogenous determination of structural breaks, coupled with simulation exercises to obtain critical values for the tests. Empirical evidence lends no support to the hypothesis that the series contains seasonal unit roots. The best model in terms of forecasting performance by various criteria is fitted to the series in first differences and forecasts produced.

Keywords: Currency Board, Foreign Reserves, Time Series Analysis, Box-Jenkins Methodology, Forecasting

JEL Classification: E42, E59, C13, C32, C53

1 Introduction

In this paper we fit an appropriate univariate time series model to CBBH foreign reserves data for forecasting purposes. Namely, the foreign reserves represent one of the most important macroeconomic indicators in Bosnia and Herzegovinas economy, especially in the context of a monetary policy strategy —currency board —implemented in the country. This is one of the key variables for the functioning of automatic adjustment mechanism of the currency board that facilitates constant balancing of the economic processes and restoring external and internal equilibrium. For those reasons, knowing current levels of the foreign reserves and their future developments through forecasting exercises is of essential importance for maintaining macroeconomic and financial stability of the economy, as well as for the process of foreign reserves management in the central bank.

Descriptive analysis, coupled with formal seasonality test, indicate that both deterministic and stochastic seasonality could be present in the data. Depending on the type of seasonality present, various approaches to modelling seasonality could be taken. If seasonal patterns are stable or repetitive, regression approaches with the introduction of seasonal dummy variables may be implemented to describe its behavior. On the other hand, if innovations have cumulative or persistent effects that change or introduce a new seasonal pattern, a different approach that takes stochastic seasonality into account needs to be taken. In that respect, formal investigation whether unit roots are present in the long run (at zero frequency) and/or in each of the seasons is taken by applying a standard battery of unit root tests coupled with the HEGY seasonal unit root test. Possible structural breaks in the series have been treated using unit root tests with endogenous determination of the breaks, both at zero frequency (Zivot-Andrews, 1992 and Clemente, Montanes and Reyes, 1998) and other frequencies of monthly data (modified HEGY tests with endogenously determined structural breaks based on innovational outlier model). Finally, formal steps in the Box-Jenkins (1970) methodology are conducted to determine the structure of the models that best fit the series in terms of its trend, cycle and irregular components.

With the aim of determining basic characteristics of the series in terms of possible presence of the stochastic trend and the characteristic of the stochastic seasonality in the data, the HEGY seasonal unit root test has been applied. Towards that end, the critical values used in the testing procedure for inference have been obtained by Monte Carlo simulations to increase the precision of the test for this particular series. It is observed that resulting critical values are similar to those available in the literature for the series with the similar length (number of observations) and structure with respect to deterministic component. Since visual inspection of the series indicates presence of the structural breaks, the HEGY test has been modified to explicitly take them into consideration. The results offer evidence for the presence of the stochastic trend at zero frequency and no seasonal unit roots (no conclusive evidence for seasonal unit roots at bimonthly and other frequencies) so there is no need for transforming the series by seasonal differencing. Accordingly, the first difference of the series is used in the formal econometrical analysis for obtaining a model with the best forecasting performances.

Finally, a horse race in terms of the forecasting performances of the candidate models is carried out applying a battery of tests and taking into consideration different criteria to obtain the best model. The model is used to produce forecast of the foreign reserves in Bosnia and Herzegovina in the form of a fan chart.

The outline of the paper is following. The first section gives introductory comments on the aim of the paper and the methodology used. The second section presents a brief description of the main characteristics of the data with accompanying descriptive analysis. The unit root testing procedures of the foreign reserves series and its outcomes are explained in the third section. The fourth section considers the forecasting procedure and characteristics of the alternative model specifications and provides (the best) forecasts of the series. The last section presents concluding remarks.

2 The main characteristics of the foreign reserves

The foreign reserves are one of the most important macroeconomic indicators in an economy. IMF (2013) defines foreign reserves as those external assets that are readily available to and controlled by monetary authorities for meeting the balance of payments financing needs, intervention in exchange markets to affect the currency exchange rate and other related purposes (confidence in the currency and the economy, basis for foreign borrowing). However, it needs to be emphasized that the specific form of monetary policy strategy implemented in Bosnia and Herzegovina excludes some of those possible uses.[1]

2.1 The role of foreign reserves in domestic economy and the factors that affect data generating process

This variable lays at the heart of an (orthodox) currency board as a form of monetary policy strategy. Accordingly, most of the variations in the foreign reserves in Bosnia and Herzegovina are related to the functioning of the automatic adjustment mechanism of the currency board, a basic mechanism of the strategy that enables accommodation of the main domestic macroeconomic variables in the economy to developments in its balance of payments. Namely, a balance of payments deficit causes capital outflow from the country and a proportional reduction in the money supply in the economy. The basic principles of the currency board of the full backing of monetary liabilities by the foreign reserves and the full convertibility of domestic currency into foreign and vice versa dictate that the foreign reserves changes are one for one reflected in the changes of the money supply (monetary base). A decrease in the money supply leads to a domestic interest rates increase, reduction in aggregate demand and real exchange rate depreciation. The money supply contraction decreases labor force and other factors of production demand, causing a decline in domestic prices in comparison to foreign prices. All this movements contribute to reaching a new balance. The reverse happens in the case of a balance of payments surplus. This mechanism is expected to essentially generate cyclical or truly stochastic developments in the foreign reserves series that in a formal analysis may be represented as irregular or stochastic seasonality processes.

On the other hand there are one off events that affect foreign reserves series in a semi deterministic way since most of them are known or arranged in advance, before they actually happen. They may be considered as a form of shocks to the series, since they cause its significant increase or decrease at a time of their occurrence. Those events include, for instance, euro conversion, IMF standby arrangement tranches payments, selling off major companies to foreign investors, foreign credits, grants and other major capital inflows, foreign debt servicing etc. From the perspective of one particular point in time those events may be considered deterministic, but from the perspective of the overall economy, they could be treated just as manifestations of the stochastic processes of its functioning. Accordingly, in a regression analysis, these shocks could be treated by the introduction of the dummy variables, or left unchecked if they do not cause undesirable characteristics of the series that

[1] E.g. intervention in the exchange markets to affect the exchange rate of domestic currrency.

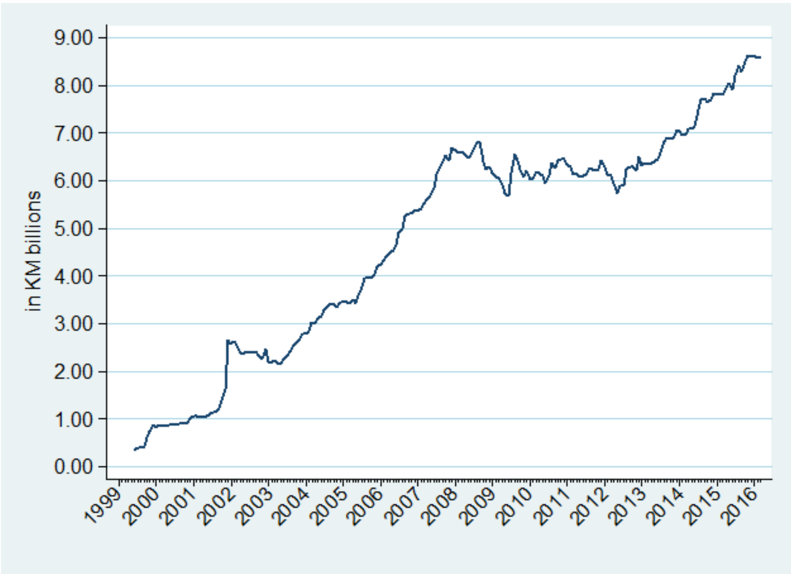
complicate the estimation and inference.

Deterministic seasonality can be expected to explain some part of the variation observed in the series. There are several factors that could generate the deterministic seasonality in the foreign reserves data. For example, an increase in money remittances can be observed during summer months when diaspora comes back to the country for a summer vacation. Additionally, commercial banks close their open foreign exchange positions towards the end of the year through buying foreign currency (mainly euro) from the Central bank in exchange for domestic currency to observe strict banking agencies' rules on the net open foreign currency positions. Normally, the reverse operations are conducted at the beginning of the year.

2.2 The data

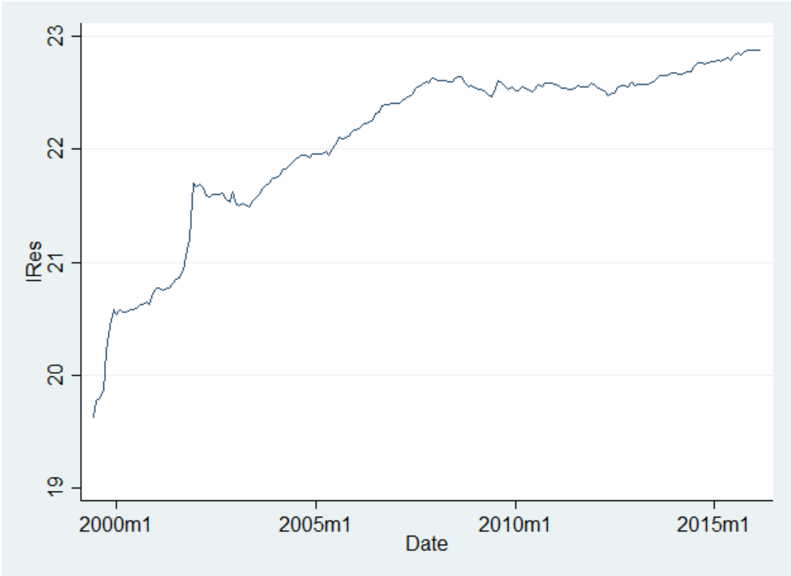
In this paper we analyze and forecast the monthly series of the Central Bank of Bosnia and Herzegovina's (CBBH) foreign reserves. The data are taken from the official statistics of the Bank. The foreign reserves series is presented in figure 1. The time series starts with the June 1999 and ends with the March 2016 making overall 202 observations. The values of individual data are given in the domestic currency (KM) equivalents.

Figure 1: CBBH foreign reserves



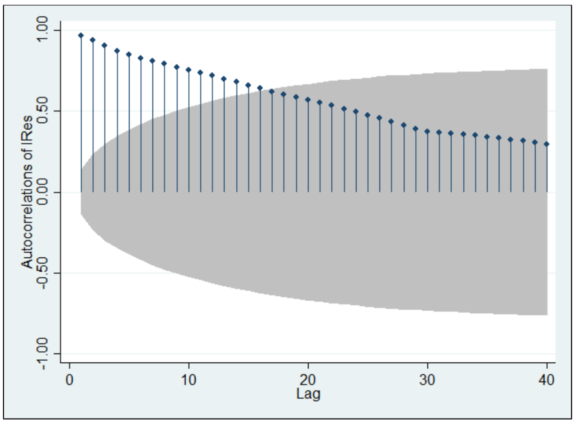
The series exhibits a clear (stochastic or deterministic) upward trend, with a noticeable change in the trend in 2008 when the effects of the financial crisis started to appear in Bosnian economy as well as a subsequent pickup in growth and resultant increase in the slope of the trend from 2012 onwards. The dominance of the upward trend indicates non-stationarity of the series in levels. Figure 2 shows the logarithm of the CBBH foreign reserves series in levels. The logarithmic transformation of the series is performed to diminish its volatility and dampen the magnitude of the fluctuations (Enders, 2010). The figure shows a constant upward movement of the series, suggesting its non-stationarity. The second important feature of the series in this figure is a relatively persistent pattern of short term volatility of the series, which at least partly is expected to be attributable to the deterministic seasonality in the data, though it does not clearly show in the graph line.

Figure 2: Logarithm of the CBBH foreign reserves

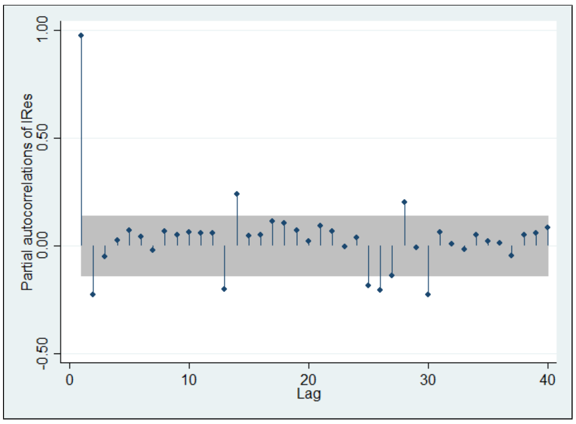


As presented in figure 3, auto correlation function shows exponential decay indicative of an unit root or near unit root process. On the other hand partial auto correlation function is significant up to the second lag (with changing sign) that, in combination with the shape of the auto correlation function, could indicate a second order autoregressive process. Additionally, significance of the partial autocorrelation function (round) every twelve lags could point to a stochastic seasonality present in the data generation process.

Figure 3: ACF and PACF functions of the logarithm of foreign reserves



(a) ACF function

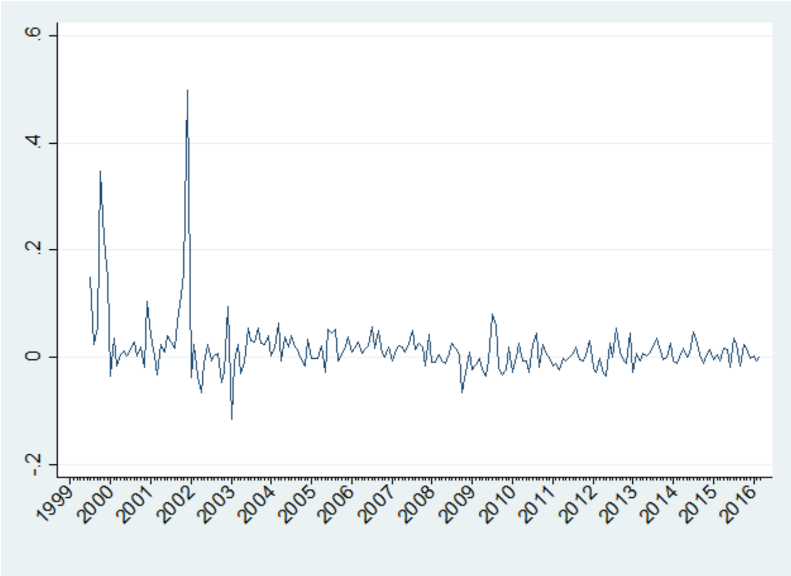


(b) PACF function

Note: 95% confidence bands when calculationg confidence interval bounds $[se = 1/\sqrt{n}]$.

To make the series more amenable to formal empirical analysis the transformation in the form of the first difference seems appropriate. The first difference of the logarithmic series is presented in figure 4, with the aim of removing (stochastic) trend from the data. The differencing appears to yield a series with a relatively stable mean, i.e. a relatively constant mean monthly growth rate. Apart from the beginning period of increased volatility in the series, the variance of the series appears to be quite constant. Thus, the underlying series in levels appears to be difference stationary, with no clear signs of deterministic seasonality present in the data by visual inspection of the figure.

Figure 4: Monthly growth rates of the CBBH foreign reserves



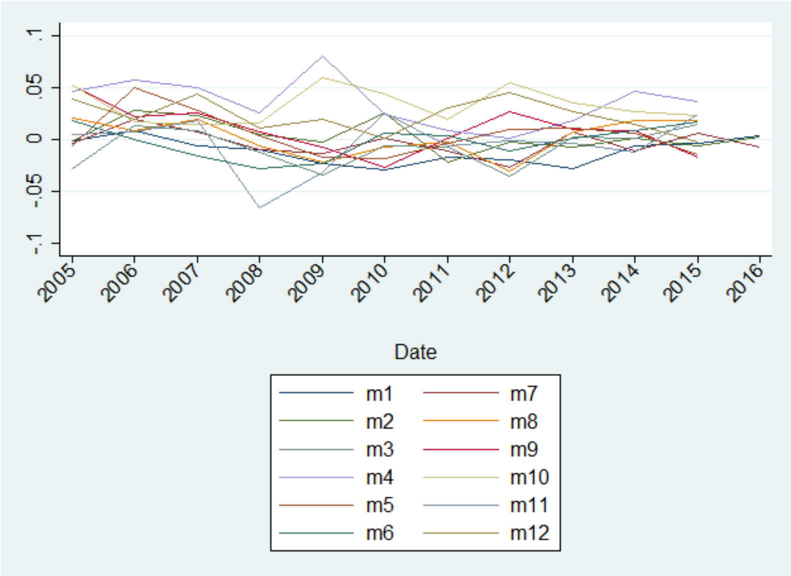
However, the outcome of a seasonality test, where the first difference of the foreign reserves series is regressed on a constant and the set of eleven seasonal dummy variables, clearly shows the presence of seasonality in the data since most of the seasonal dummy variables appear highly statistically significant. It is interesting to note that all the coefficients on seasonal dummy variables have negative signs, because the highest level of theforeign reserves is usually present in December (treated by the constant).

Table 1: Seasonality test of the log of foreign reserves: dummy variable representation

Variables	**Coef.**	**Std. Err.**	**t**	**P> | t |**	**[95% Conf. interval]**	
Const.	0.06861	0.01266	5.419	0.000	0.04363	0.09359
m1	-0.08495	0.01791	-4.744	0.000	-0.12028	-0.04963
m2	-0.06695	0.01791	-3.739	0.000	-0.10228	-0.03163
m3	-0.06584	0.01791	-3.677	0.000	-0.10116	-0.03051
m4	-0.07172	0.01819	-3.944	0.000	-0.10760	-0.03585
m5	-0.07110	0.01819	-3.910	0.000	-0.10697	-0.03522
m6	-0.05331	0.01819	-2.931	0.004	-0.08918	-0.01744
m7	-0.03031	0.01791	-1.693	0.092	-0.06564	0.00501
m8	-0.04002	0.01791	-2.235	0.027	-0.07534	-0.00469
m9	-0.05380	0.01791	-3.004	0.003	-0.08913	-0.01848
m10	-0.04395	0.01791	-2.454	0.015	-0.07928	-0.00863
m11	-0.04999	0.01791	-2.792	0.006	-0.08532	-0.01467

To learn more about the seasonality features of the data, the so called Franses graph for the data in first differences is given in the following figure, where every quarter is presented in a separate line. The initial volatile period from 1999 to 2004 is left out to provide a clearer representation. As it can be clearly seen, there is a lot of intertwining between different quarters' seasonal lines and one cannot easily discern the effects of different seasons. This could be an indication of the stationary stochastic seasonality present in the data that might interact with the deterministic component of the seasonality process.

Figure 5: Franses graph: 2005-2016 period



Since graphical representation of the series in levels and first differences and in the form of the seasonal graph for the data in the first differences and the seasonality test give mixed signals about possible deterministic seasonality present in the data and indicate the possible presence of the stochastic seasonality, both ARIMA models with and without seasonal dummy variables are going to be investigated as the candidate models for forecasting the series of the foreign reserves.

3 Unit root testing

Before proceeding with formal ARIMA modelling, the order of the integration of the series under consideration needs to be determined. Nonstationary variables may have a pronounced deterministic and/or stochastic trend, appearing to meander without a constant long-run mean or variance (Enders, 2010). In that case, the series needs to be transformed to ensure stationarity before any further formal steps in the ARIMA methodology are taken. Shocks to a non-stationary series have a permanent nature and the effects of those shocks are never eliminated. On the other hand, shocks to a stationary time series are temporary and over time the effects of the shocks dissipate and the series reverts to its long-run level. The nature of the trend has important implications for the appropriate transformation necessary to attain a stationary series.

In order to check for (non)stationarity of the foreign reserves series, a battery of the standard unit root tests has been employed: Augmented Dickey Fuller (1979, 1981) test, Phillips-Perron (1988) test, Dickey-Fuller test with GLS de-trending (Elliot, Rothenberg and Stock, 1996) test and Kwiatkowski-Phillips-Schmidt-Shin (1992) test and results presented in table 2.

The standard unit root tests of the foreign reserves series give somewhat mixed results, depending on a specification of the test equation and a form of the test itself, but overall indicate that the series is nonstationary. The Augmented Dickey-Fuller test shows that the

Table 2: Unit root tests

Unit root test	**Specification: Trend (yes/no)**	**Test statistics**	**Critical values**		
Left tailed tests			**1%**	**5%**	**10%**
ADF	No	-3.459	-3.481	-2.884	-2.576
	Yes	-2.762	-4.011	-3.438	-3.138
PP(Z(rho))	No	-5.534	-20.137	-13.902	-11.135
	Yes	-9.565	-28.073	-21.104	-17.837
PP(Z(t))	No	-5.170	-3.467	-2.883	-2.573
	Yes	-4.148	-4.006	-3.437	-3.137
DFGLS	No	1.184	-2.586	-1.946	-1.656
	Yes	-0.460	-3.480	-2.812	-2.534
Right tailed test			**10%**	**5%**	**1%**
KPSS	No	1.230	0.347	0.463	0.739
	Yes	0.355	0.199	0.146	0.216

Note: Lag length for the difference of the dependant variable in the test equation is thirteen for all the tests. With some tests trend variable appears no statistically significant, possibly due to the presence of structural breaks.

series is integrated of order one, but only at 1 percent significance level for the specification without trend.The Phillips-Perron test Z (t) statistics indicates no presence of the unit root for both specifications. However, the Phillips-Perron test Z (rho) test statistics clearly shows the presence of the unit root in both cases, as is the case with the Dickey-Fuller test with GLS de-trending. Additionally, the Kwiatkowski-Phillips-Schmidt-Shin test implies the presence of the unit root for both specifications —the zero hypothesis of stationarity is strongly rejected at the conventional level of significance.

Visual inspection of the foreign reserves series suggests that, as in case of many other macroeconomic time series in Bosnia and Herzegovina, it most likely exhibits one or more structural breaks. A structural break in a time series may be the result of some unique economic event, or can reflect an institutional, legislative or technical change. It can also be the outcome of changes in economic policies or large economic shocks, such as the outset of a crisis etc. In the case of a non-stationary series, shocks such as structural breaks can have permanent effect on the series. Since structural breaks in time series can affect the result of the standard unit root tests, the special form of unit root tests that account for structural breaks has to be implemented. According to Perron (1990), there is a tradeoff between power of the tests and the amount of a priori information one is willing to incorporate with respect to the choice of the break point —less information yields low power and vice versa.

Perron (1989, 1990), Perron and Vogelsang (1992) and Perron (1997) suggested a unit root test that allows for two different forms of structural break called the Additive Outlier (AO) and the Innovational Outlier (IO) models. The AO model changes are assumed to take place rapidly, while in the IO model changes are assumed to take place gradually over a certain period of time. In the first model for testing unit root, the effect of the change on the level of the series is not affected by dynamics of the correlation structure of the series. In the second, it is assumed that the series reacts to the change in the mean in the same way it responds to other shocks, implying that there is a transition period in the adjustment of the series. Both models include the assumption of no break under the null hypothesis of unit root. The main difference between the two procedures is that in the IO case, the estimation is conducted on a single equation, while the AO requires an auxiliary regression.

The structural breaks can be treated as being exogenous or endogenous. When the date of the structural break can be determined, different forms of the Perron (1989) unit root test that treats breaks as being exogenous can be used, i.e. the timing of the break is known in advance. In essence, the testing procedure entails the addition of dummies to capture the different segments in which the series is divided (before and after the break). On the other hand, if the date of the break is uncertain, different types of unit root tests with endogenous breaks should be used in which determination of the timing of the break is part of the estimation procedure. Several studies have been developed using various methodologies for endogenously determining the break date: Banerjee, Lumisdaine and Stock (1992), Zivot and Andrews (1992), Perron and Vogelsang (1992) and Perron (1997), for example, that allow for the possibility of only one structural break. Additionally, other tests have been developed, allowing for the possibility of multiple structural breaks such as Lumsdaine and Papell (1997) and Clemente, Montas and Reyes (1998), for example.

Since the series is characterized by many shocks and outliers, especially in the initial period, it is relevant to consider the possibility of breaks in the series when testing for unit roots (at zero frequency). To this end, Zivot-Andrews (1990) and Clemente, Montanes and Reyes (1998) tests have been used.

Table 3: Unit root tests with structural breaks

Unit root test	**Test stat.**	**Critical values**			**Break date(s)**	
		1%	**5%**	**10%**		
Zivot-Andrews test	-4.630	-5.570	-5.080	-4.820	2008m1	
CMR test						
ONE BREAK						
Additive Outlier	-4.107		-3.560		2005m8	
Innovational Outlier	-3.674		-4.270		2001m7	
TWO BREAKS						
Additive Outlier	-2.023		-5.490		2002m1	2006m4
Innovational Outlier	-4.684		-5.490		2001m7	2005m4

Note: Lag selection criteria for the Zivot-Andrews test is Akaike (AIC) and one lag of the dependent variable is included in the test equation. Breaks in both, the intercept and the trend are allowed. The test is performed by STATA's zandrews command. The Clemente-Montans-Reyes (CMR) (IO, AO) tests are carried out by the clemio1 and clemio2 routines in STATA.

All test results indicate the presence of the unit root in the series, apart from the Clemente-Montanes-Reyes test with Additive Outlier specification. The break dates suggested by the tests have an economic interpretation: the beginning of the effects of the economic crisis in Bosnia and Herzegovina (2008m1), the euro conversion (2001m7 and more precisely, 2002m1) and the VAT tax introduction at the beginning of 2006 (2005m4, 2005m8, 2006m4).

It is well known that in a time series with frequency higher than yearly (quarterly, monthly, etc.) the unit root may be present in the long run (at zero frequency) and/or at other frequences. Hylleberg et al (1990) developed a special technique for testing unit roots at different frequencies —HEGY test —that investigates the possibility of the presence of both long run and seasonal roots at different cycles. Initially developed for quarterly data, in subsequent work their technique was extended to monthly data. In order to increase the precision of the test and to obtain critical values more suitable to specific features of the series under analysis, the critical values for testing monthly unit roots in this paper have been calculated and tabulated (under the hypothesis of the unit root) using Monte Carlo simulations.

The testing for unit roots in monthly time series is equivalent to the testing for the significance of the parameters in the regression:

$$\begin{aligned}\Phi(L)s_{8,t} = \mu_t + \pi_1 s_{1,t-1} + \pi_2 s_{2,t-1} + \pi_3 s_{3,t-1} + \pi_4 s_{3,t-2} + \pi_5 s_{4,t-1} + \pi_6 s_{4,t-2} \\ + \pi_7 s_{5,t-1} + \pi_8 s_{5,t-2} + \pi_9 s_{6,t-1} + \pi_{10} s_{6,t-2} + \pi_{11} s_{7,t-1} + \pi_{12} s_{7,t-2} + e_t \end{aligned} \quad (1)$$

Other auxiliary equations for individual variables in the regression are given in the appendix. The μ_t stands for the deterministic part and may consist of a constant, seasonal dummies and trend.

Table 4: Testing for seasonal unit roots by HEGY test

	Specification without trend	**Specification with trend**
$\boldsymbol{\pi 1}$	-3.09 ***	-2.45 *
$\boldsymbol{\pi 2}$	-3.59	-3.59
$\boldsymbol{\pi 3}$	-5.84	-5.83
$\boldsymbol{\pi 4}$	-4.62	-4.65
$\boldsymbol{\pi 5}$	-6.23	-6.20
$\boldsymbol{\pi 6}$	-5.93	-5.91
$\boldsymbol{\pi 7}$	-3.32	-3.32
$\boldsymbol{\pi 8}$	0.21 ***	0.19 ***
$\boldsymbol{\pi 9}$	-3.82	-3.79
$\boldsymbol{\pi 10}$	-4.53	-4.50
$\boldsymbol{\pi 11}$	-4.75	-4.68
$\boldsymbol{\pi 12}$	0.19 ***	0.16 ***
$\boldsymbol{\pi 3, \pi 4}$	31.30	31.47
$\boldsymbol{\pi 5, \pi 6}$	20.36	20.18
$\boldsymbol{\pi 7, \pi 8}$	17.74	17.88
$\boldsymbol{\pi 9, \pi 10}$	12.31	12.14
$\boldsymbol{\pi 11, \pi 12}$	14.19	13.87
$\boldsymbol{\pi 3, ..., \pi 12}$	56.43	55.91

Note: * Significant at 10% level.** Significant at 5% level. *** Significant at 1% level. The deterministic part of the test equation includes a constant and standard set of centered seasonal dummy variables without or with linear trend.

The results of the test are presented in table 4 above. The non-seasonal unit root statistics corresponds to the t-statistics of π_1 and the seasonal unit root statistics with two months per cycle corresponds to π_2. Since pairs of complex roots are conjugates, these roots are present when pairs of corresponding $\pi's$ are zero simultaneously. Those seasonal unit roots correspond to the F-statistics for the pairs of $\pi's$. Simulated critical values for t-tests for separate $\pi's$ and for F-tests of pair of $\pi's$, as well as for joint test of $\pi_3 = \ldots = \pi_{12}$ are given in the table 1 in the Appendix. Comparing the calculated test statistics with the appropriate critical values, it can be seen that there is evidence for the presence of the unit root at zero frequency, in accordance with the results of the standard unit root tests. However, there is no indication of the unit root at other frequencies suggesting that the series should be first differences in order to obtain stationarity of the data.

Seasonal unit root tests in finite samples suffer from severe size distortions and power reductions when breaks are present in the data (Hassler and Rodrigues, 2002). As in the case of the standard non-seasonal unit root tests, neglecting (seasonal) mean shifts can bias unit root tests towards non-rejection (Perron, 1989) or to spurious rejection of the null hypothesis (Leybourne, Mills and Newbold, 1998). Structural breaks may not only have general effects on the levels and trends of the series, but they may also change the observed pattern of seasonality. Therefore, it is necessary to allow for possible seasonal mean shifts when testing for seasonal unit roots to precisely determine the characteristics of a time series in terms of the unit root presence.

If it is assumed that there is a single break that occurs at time $T_B(1 < T_B < T)$, it is possible to test the null hypothesis of the Innovational Outlier (IO) break in monthly data by estimating the following equation:

$$\Delta_{12}y_t = \mu_t + \beta_t + \sum_{k=2}^{1} 2\delta_k D_{kt} + \sum_{k=1}^{1} 2\pi_k s_{k,t} + \sum_{i=1}^{\rho-1} \psi_i \Delta_{12} y_{t-i} + \sum_{k=2}^{12} \theta_k S_{kt} + \sum_{k=2}^{12} \eta_k \Delta_{12} S_{kt} + \upsilon_t \quad (2)$$

$$S_{kt} = \left\{ \begin{array}{ll} 1 & tT_B \\ 0 & t < T_B \end{array} \right\} \quad (3)$$

This is the standard HEGY test equation with addition of S_{kt} that stands for seasonal dummy variables that start to be active at the time of the break. The standard procedure for conducting this test requires us to estimate the equation above for all the potential break dates in the data. It is advisable to restrict the range of possible breaks to $[T_B^*, , T - T_B^*]$, where $T_B^* = \lambda T$ to assure the results are asymptotically valid, as recommended by Franses and Volgelsang (1998). The value of λ is called the amount of trimming and it practically excludes some observations at the beginning and at the end of the series from the potential break dates.

There are two approaches for selecting the break date. The first method involves minimizing the value of the $t_{\pi_i}(T_B)$ and maximizing the value of the $F_{\pi_{odd},\pi_{even}}(T_B)$ statistics over all possible break dates. De facto, this procedure boils down to selecting the break date when the statistics is least favorable to the null hypothesis. This method can be defined:

$$\hat{T}_{B,\pi_i} = argmin(t_{\pi_i}(T_B)), \qquad i = 1, 2 \quad (4)$$

$$\hat{T}_{B,F_{o,e}} = argmax(F_{odd,even}(T_B)) \quad (5)$$

In this procedure there will be identified as many break dates as there are unit roots considered because, for each frequency, the selection is based on statistics that is least favorable to the null (Franses and Volgelsang, 1998). In this procedure the trimming is allowed, but is not required.

The second method bases the selection of the break date on the maximization of the significance of the seasonal dummy shift variable:

$$\hat{T}_B = argmax(F_\theta(T_B)) \quad (6)$$

This procedure identifies a unique break date and the trimming of the observations is necessary. According to Perron and Volgelsang (1992), the second method has more power than the first. Additionally, as shown in Harvey, Leybourne and Newbold (2001) the second procedure tends to anticipate the break date by one period in the case of quarterly data.

However, Mendez (2015) found no conclusive evidence of the anticipation of the break dates in the case of monthly data.

In order to check the results obtained by the standard HEGY test we employ the HEGY test when structural breaks are determined endogenously. The test equation specifications are the same as before. The critical values for both procedures, the one that minimizes t-statistics and maximizes F-statistics, as well as the one that maximizes the significance of the dummy shift variable, are given in tables 11 and 12 in the Appendix, respectively. The set of the tabulated critical values are calculated assuming that the break size follows a standard normal distribution. The break dates identified are concentrated round two major events —the euro conversion during 2001 and the VAT tax introduction at the beginning of the 2006.

Table 5: Testing for seasonal unit roots by HEGY test with endogenous structural breaks

	Specification without trend	**Break date**	**Specification with trend**	**Break date**
$\boldsymbol{\pi 1}$	−1.55 *	2001m10	−4.08 *	2003m1
$\boldsymbol{\pi 2}$	−4.29 ***	2001m3	−4.21 **	2003m11
$\boldsymbol{\pi 3, \pi 4}$	22.23	2006m1	51.10	2003m12
$\boldsymbol{\pi 5, \pi 6}$	13.61 ***	2002m3	16.73	2002m3
$\boldsymbol{\pi 7, \pi 8}$	59.63	2001m10	80.26	2003m9
$\boldsymbol{\pi 9, \pi 10}$	29.54	2005m8	28.82	2005m8
$\boldsymbol{\pi 11, \pi 12}$	33.40	2001m6	42.30	2001m6

Note: * Significant at 10% level.** Significant at 5% level. *** Significant at 1% level. The test is based on the Innovational Outlier model. Brake dates are determined on the basis of minimum t, maximum F statistics criteria.

In table 5 above we present the results of the test when the break selection is based on the least favorable statistics to the null hypothesis, minimizing t and maximizing F values. For each frequency the particular time of the break found is presented as different break dates are associated with different roots. As can be clearly seen we still cannot reject the null of the unit root at zero frequency. However, in comparison to the results of the standard HEGY test evidence for new roots at 2 months and 2.4 months frequency for the specification without trend and for 2 months only for the specification with trend seems to appear.

In the following table we give the results of the seasonal unit root test when the break date is determined by maximizing the significance of the seasonal break dummies. In this case a single break date is identified. Again, we cannot reject the null of the unit root at zero frequency. But, the evidence for more roots seems to appear at different frequencies, especially in the case of the specification with trend where roots at almost all frequencies appear to be present.

The fact that when the HEGY test with endogenous consideration of break dates is used gives an indication of new seasonal unit roots seems to complicate our judgement. However, it is important to bear in mind that this test has less power than the test without the structural break because more parameters have been estimated using the same number of observations, as explained in Mendez (2015). This could be an explanation for the appearance of the new seasonal roots.

The true purpose of the extended HEGY test is to check for the possibility that structural breaks may disguise otherwise stationary processes as those showing unit root(s). In that

Table 6: Testing for seasonal unit roots by HEGY test with endogenous structural breaks

	Specification without trend	**Specification with trend**
Break date	2008m11	2008m11
$\boldsymbol{\pi 1}$	-0.23 *	-2.03 *
$\boldsymbol{\pi 2}$	-1.87 *	-2.18 *
$\boldsymbol{\pi 3, \pi 4}$	11.05	10.58 ***
$\boldsymbol{\pi 5, \pi 6}$	1.64 *	2.23 *
$\boldsymbol{\pi 7, \pi 8}$	7.47 *	6.83 *
$\boldsymbol{\pi 9, \pi 10}$	15.28	15.19
$\boldsymbol{\pi 11, \pi 12}$	9.90 ***	10.37 ****

Note:* Significant at 10% level.** Significant at 5% level. *** Significant at 1% level. The test is based on the Innovational Outlier model. Brake dates are determined on the basis of the maximum value of t statistics on the dummy shift variable.

respect, only in cases when the HEGY test without structural breaks indicates the presence of the unit root should the extended HEGY test be considered to check the validity of the findings since the latter has less power than the former (Mendez, 2015). Additionally, this procedure identifies break date at 2008m11 round the time when the effects of the large economic crisis started to appear in the economy of Bosnia and Herzegovina.

4 Forecasting

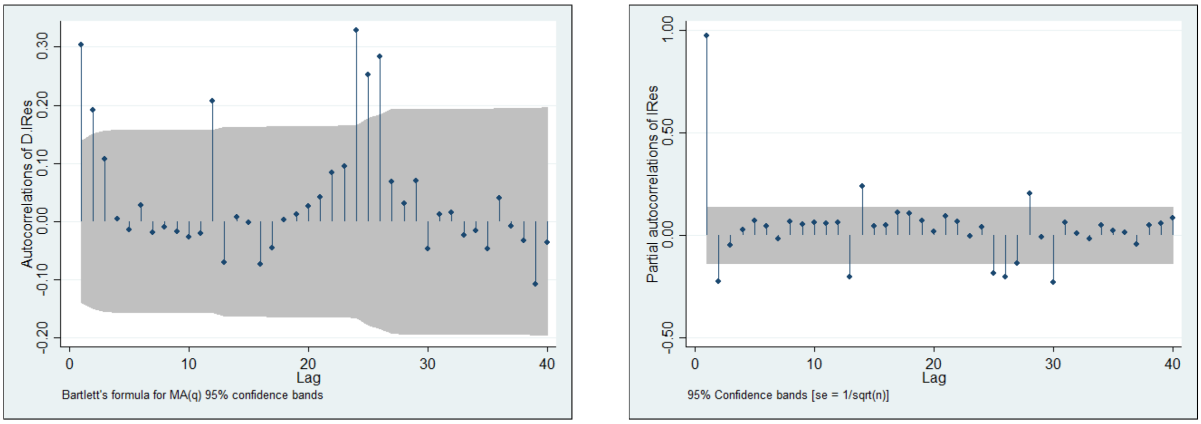
Franses (1991) shows that correctly taking account of the type of seasonality and non-stationarity in monthly data can improve the forecasting performance. The formal unit root tests results here confirm the initial assumption that it is appropriate to first difference the foreign reserves series to obtain stationarity. Since there is no clear evidence of the seasonal unit roots, Δ_{12} seasonal filter will not be applied. Accordingly, the monthly growth rates of the foreign reserves will be analyzed to estimate and fit the best ARIMA model for forecasting purposes. Towards this end, formal Box-Jenkins (1976) strategy for appropriate model selection will be implemented.

Both, the auto correlation function and the partial autocorrelation function are significant up to the second lag, with changing sign in the second case which could serve as a rough indication of the shape of the ARIMA model. Significance of the autocorrelation function and the partial autocorrelation function at approximately every twelve lags suggest that the stochastic seasonality may be present in the data generation process.

Estimation and evaluation results of the competing models for the forecasting exercise are presented in table 7. The competing specifications analyzed are the following:

1. ARIMA((1,3), 1, 1/2),
2. ARIMA((2),1,1),
3. ARIMA(0,1,1/3),
4. ARIMA(0,1,(1/3,12)) + seasonal dummies (additive stochastic seasonality) and

Figure 6: ACF and PACF functions of the first difference of the logarithm of foreign reserves



(a) ACF function

(b) PACF function

Note: 95% confidence bands when calculationg confidence interval bounds [$se = 1/\sqrt{n}$].

5. SARIMA(0,1,1/3)(0,0,1) + seasonal dummies (multiplicative stochastic seasonality).

All models mostly display highly significant autoregressive and moving average estimation parameters. As can be seen in the table, the constant is not statistically significant in the first three specifications without seasonal dummy variables. However, the constant is kept in the model specifications for the forecast evaluation. The Box-Pierce portmanteau statistics does not provide evidence of the autocorrelation present in the residuals.

The first model is not invertible, since all roots of the characteristic polynomial for the MA component are not outside the unit circle and thus will not be further taken into consideration. The Akaike and Schwartz-Bayesian statistics suggest the fourth and the third specification to have the best overall fit, respectively.

To evaluate competing models in terms of their forecasting performance they are estimated on a sample trimmed by 50 observations at the end of the series and forecasts for 50 out-of-sample months are generated from each of these models. Both, rolling (constant) window and expanding window procedures are implemented.

The standard set of forecast accuracy measures for one to four periods ahead, given in tables 13 to 16 in the Appendix, show no clear indication as to what is the best model in terms of the forecasting performances. Generally, they favor simpler model specifications with no deterministic and stochastic seasonality modelled in the case of the expanding window strategy and vice versa in the case of the rolling window strategy. However, the MAPE criterion always gives advantage to more parsimonious models. Since it is known that these statistics can mask an important feature of the forecast evaluation exercises that higher forecast errors may be caused by a few observational errors, they should be supported by other forecast evaluation statistics.

Theil U2 statistics is a measure of the forecasting performance between competing models in the form of a ratio of their mean squared errors. The value of the ratio smaller than one indicates that the first model's forecasting performance (whose mean squared error is given in the numerator) is better than the second's and vice versa. The resulting test statistics presented in table 9 generally indicates that the models with deterministic and stochastic seasonality explicitly accounted for have better forecasting performance than the simpler

Table 7: Estimation results of different ARIMA specifications

Variables	**(1) ar1_3ma1_2_12**	**(2) ar2ma1_12**	**(3) ma3_12**	**(4) ma3_12seas**	**(5) ma3mma12seas**
Constant	0.0175	0.0174	0.0175	0.0694***	0.0687***
	(0.0130)	(0.0108)	(0.0113)	(0.0179)	(0.0182)
m1				-0.0860***	-0.0851***
				(0.0244)	(0.0201)
m2				-0.0673**	-0.0669*
				(0.0307)	(0.0345)
m3				-0.0668*	-0.0661**
				(0.0363)	(0.0329)
m4				-0.0710**	-0.0700*
				(0.0342)	(0.0382)
m5				-0.0691*	-0.0698*
				(0.0403)	(0.0418)
m6				-0.0534	-0.0529
				(0.0330)	(0.0333)
m7				-0.0288	-0.0295
				(0.0253)	(0.0245)
m8				-0.0416	-0.0402
				(0.0352)	(0.0347)
m9				-0.0568**	-0.0538**
				(0.0282)	(0.0269)
m10				-0.0410**	-0.0414**
				(0.0176)	(0.0175)
m11				-0.0480***	-0.0484***
				(0.0186)	(0.0160)
L.ar	-0.139**				
	(0.0550)				
L2.ar		0.226***			
		(0.0672)			
L3.ar	0.246***				
	(0.0551)				
L1.ma	0.496***	0.342***	0.371***	0.393***	0.347***
	(0.0659)	(0.0421)	(0.0498)	(0.0581)	(0.0497)
L2.ma	0.598***		0.340***	0.306***	0.224**
	(0.0671)		(0.0465)	(0.0740)	(0.0954)
L3.ma			0.109***	0.136**	0.155**
			(0.0409)	(0.0685)	(0.0769)
L12.ma	0.502***	0.285***	0.386***	0.283***	
	(0.0648)	(0.0692)	(0.0611)	(0.0695)	
L13.ma					0.140*
					(0.0827)
Sigma	0.0468***	0.0496***	0.0487***	0.0455***	0.0463***
	(0.00164)	(0.00109)	(0.00128)	(0.00171)	(0.00165)
Obs.	201	201	201	201	201
AIC	-625.5	-626.7	-629.6	-635.8	-630.6
BIC	-608.9	-603.6	-609.8	-579.6	-574.4
BP(12)	0.993	0.342	0.948	0.889	0.889
BP(24)	0.456	0.131	0.333	0.775	0.775

Note: Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. The evaluation criteria are the Box-Pierce portmanteau test statistics calculated for m lags. The Akaike (AIK) and Schwartz-Bayesian (BIC) criterions are given as measures of the overall fit of the model.

Table 8: Theil U2 statistics between alternative specifications

	Period 1	Period 2	Period 3	Period 4
Expanding window				
ar2ma1_12 vs ma3_12	1.154	0.996	0.995	0.994
ar2ma1_12 vs ma3_12seas	0.707	1.042	1.538	1.054
ar2ma1_12 vs ma3mma12seas	0.653	1.077	1.632	1.123
ma3_12 vs ma3_12seas	0.613	1.046	1.546	1.060
ma3_12 vs ma3mma12seas	0.566	1.081	1.641	1.129
ma3_12seas vs ma3mma12seas	0.924	1.033	1.061	1.065
Constant window				
ar2ma1_12 vs ma3_12	1.013	1.014	1.002	0.974
ar2ma1_12 vs ma3_12seas	1.033	1.472	4.719	2.014
ar2ma1_12 vs ma3mma12seas	1.131	1.445	5.093	1.657
ma3_12 vs ma3_12seas	1.020	1.453	4.708	2.068
ma3_12 vs ma3mma12seas	1.117	1.426	5.082	1.701
ma3_12seas vs ma3mma12seas	1.095	0.981	1.079	0.823

models where the seasonality effects are left in the residual. However, the Theil U2 statistics favors simpler models for very short term forecasts (one period ahead) in the expanding window strategy case. When models with the seasonality treated are directly compared more evidence is found in favor of the model with multiplicative stochastic seasonality, especially when the expanding window strategy is implemented.

The results of comparing the competing models' forecasting performances in terms of Diebold and Mariano (1995) test are outlined in table 17 in the Appendix. Here, the null of equal forecast accuracy is tested against the alternative of the model with a lower value of the given criterion being better. The test results generally show that more parsimonious model specifications have equal forecasting performance as more complex models.[2] When the two most complex model specifications with the deterministic and stationary stochastic seasonality modelled are directly compared, the test results roughly show that in the case of the rolling window strategy the model with additive seasonality fares better then the one with multiplicative sthochastic seasonality.

Table 9 presents decomposition of the mean squared error of the forecasts to bias proportion (BP), variance proportion (VP) and covariance proportion (CP). The bias proportion shows how far the mean of the forecast is from the mean of the actual series. The variance proportion indicates how far the variation of the forecast is from the variation of the actual series, while the covariance proportion measures the remaining unsystematic forecasting error. By definition, these three measures sum to one. For a good forecast, the bias and variance proportions should be as small as possible, since this suggests the model is providing a good estimate of the underlying data generating process. So, most of the mean square error should be due to the covariance proportion unsystematic component. In the case of the expanding window procedure, the statistics indicates the ARIMA(0,1,(1/3,12)) + seasonal dummies for one period ahead forecasts and the SARIMA(0,1,1/3)(0,0,1) + seasonal dummies for two to four step ahead forecasts to be the best specifications. The rolling window procedure suggests the second to outperform others for all the forecasting periods.

Taking into consideration all the available evidence on the forecasting performances of

[2]Since we are dealing with nested models, only the constant window procedure has been used to compute forecast statistics (Giacomini and White, 2006).

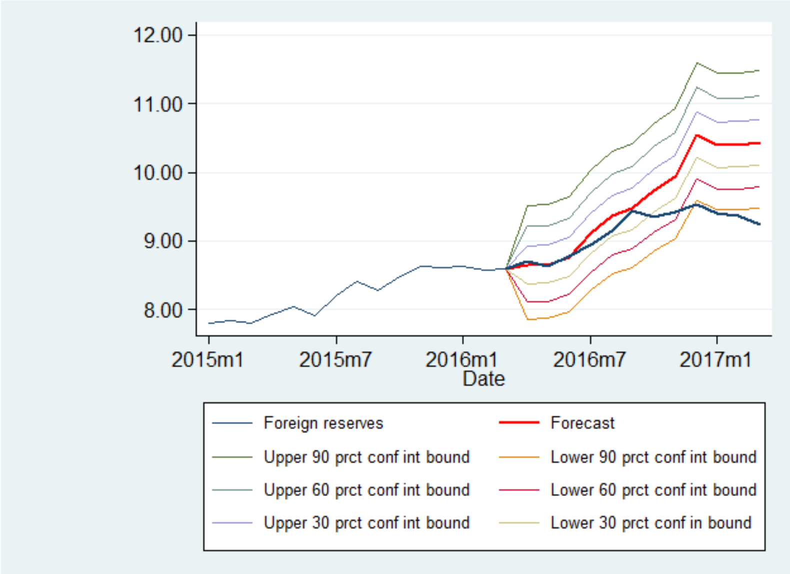
Table 9: MSE Decomposition

	Period 1	**Period 2**	**Period 3**	**Period 4**
EXPANDING WINDOW				
BP_ar2ma1_12	0.080	0.300	0.302	0.304
VP_ar2ma1_12	0.299	0.640	0.635	0.631
CP_ar2ma1_12	0.640	0.074	0.077	0.079
BP_ma3_12	0.067	0.301	0.303	0.306
VP_ma3_12	0.149	0.629	0.624	0.622
CP_ma3_12	0.803	0.084	0.086	0.086
BP_ma3_12seas	0.052	0.203	0.206	0.207
VP_ma3_12seas	0.032	0.042	0.042	0.041
CP_ma3_12seas	0.935	0.771	0.768	0.767
BP ma3mma12seas	0.057	0.200	0.202	0.203
VP_ma3mma12seas	0.039	0.037	0.038	0.037
CP_ma3mma12seas	0.923	0.779	0.777	0.776
CONSTANT WINDOW				
BP_ar2ma1_12	0.021	0.084	0.084	0.085
VP_ar2ma1_12	0.350	0.522	0.522	0.521
CP_ar2ma1_12	0.650	0.412	0.412	0.411
BP_ma3_12	0.012	0.080	0.080	0.079
VP_ma3_12	0.174	0.534	0.532	0.529
CP_ma3_12	0.835	0.406	0.407	0.411
BP_ma3_12seas	0.006	0.050	0.052	0.046
VP_ma3_12seas	0.006	0.008	0.006	0.006
CP_ma3_12seas	1.008	0.961	0.962	0.967
BP_ma3mma12seas	0.014	0.057	0.059	0.059
VP_ma3mma12seas	0.004	0.007	0.007	0.007
CP_ma3mma12seas	1.002	0.955	0.954	0.953

competing models, the specification SARIMA(0,1,1/3)(0,0,1) + seasonal dummies is considered to be the most appropriate model for forecasting the foreign reserves in Bosnia and Herzegovina. The significance of the seasonal dummy variables in the seasonality test suggests that deterministic seasonality is present in the series. Even though many of the test statistics considered show mixed results in terms of the forecasting performances of the alternative models, the mean squared error decomposition and Thail U2 test suggest that this model outperforms the others.

The fact that the third order moving average model has been chosen as being representative of the data generating process for the foreign reserves has the following interpretation. The moving-average model specifies that the foreign reserves depend linearly on the current and past values of a stochastic, imperfectly predictable error term. Shocks are propagated to future values of the series directly and are relatively short lived—they affect the series only for the current period and three periods into the future. Additionally, the multiplicative seasonality specification suggests that there are more distant regular seasonal shocks that affect the series.

Figure 7: Foreign reserves: forecasted and actual series (in KM billions)



Note: Coefficient uncertainty is not accounted for when calculationg confidence interval bounds.

The figure 8 shows the forecasted values of the original foreign reserves series based on the selected model for 12 months into the future in the form of a fan chart. One can see that the forecast traces the series fairly well in the first several months, but in the later periods it drifts apart. This is natural bearing in mind that the univariate time series models are generally intended for obtaining very short term forecasts and that the moving average specification has short memory, up to the specified number of lags.

5 Conclusion

This paper shows the testing and selection procedure for an appropriate univariate time series model for forecasting the foreign reserves in Bosnia and Herzegovina. The foreign reserves represent one of the most important macroeconomic indicators in the economy, especially in the context of functioning of the automatic adjustment mechanism of the currency board. Readily available forecasts of the foreign reserves are important from the perspective of maintaining macroeconomic and financial stability, foreign reserves management, drafting budgeting plans in the central bank, etc.

Visual inspection and formal analysis of the logarithm of the foreign reserves in the spirit of Box-Jenkins suggests that the series contains the unit root. Graphical analysis in the form of Franses graph and the seasonality test are suggestive of the presence of deterministic seasonality, the former less conclusively though.

A standard battery of the formal unit root tests at zero frequency indicates the presence of the unit root in the series. The tests with endogenous structural breaks substantiate this finding. The paper illustrates procedures for the standard HEGY seasonal unit root test and the tests with endogenous consideration of structural breaks. Towards that end, the critical values have been tabulated for the both types of tests based on Monte Carlo simulations. The HEGY test for monthly data confirms the presence of the unit root at zero frequency, but gives no indication of unit roots at other frequencies. Additionally, the HEGY tests with endogenous treatment of structural breaks, based on the innovational outlier model, generally provides evidence for this conclusion, even though there are some conflicting results of the seasonal unit root test when the break date is determined by maximizing the significance of the seasonal break dummies.

Finally, after running a horse race between competing ARIMA model specifications, the forecasting performance tests suggest that the third order moving average specification with multiplicative stochastic stationary seasonality and with a standard set of seasonal dummy variables included outperforms the alternative models. This leads us to the conclusion that shocks to the foreign reserves propagate directly to the future values of the series and that they are relatively short lived.

Future work can extend the HEGY test procedure with endogenous structural breaks to additive outlier model case, assuming the breaks have immediate effect on the series. Furthermore, the assumption of a normal distribution of breaks may be substituted for the case of large structural breaks. In both cases the test results should be compared to the appropriate simulated critical values—based on the testing procedure for the additive outlier model in the first case and on the generated series when the large breaks are introduced in the DGP in the second.

The univariate time series models for forecasting the foreign reserves should be supplemented with time series models with a richer structure, such as vector autoregressive (VAR) models or vector error correction models in the case of a cointegrating relationship between variables. Finally, more complex structural models for medium to long term forecasting could be developed.

References

Banerjee, A., R. L. Lumsdaine and J. H. Stock, "Recursive and sequential tests of the unit root and trend-break hypothesis: theory and international evidence", *Journal of Business and Economic Statistics*, Vol. 10, No. 3, pp. 271-287.

Beaulieu, J. and J. A. Miron (1993), "Seasonal unit roots in aggregate U.S. data", *Journal of Econometrics*, 55, pp. 305-328.

Box, G.E.P, G. Jenkins (1970), Time series analysis, forecasting and control, Holden-Day, San Francisco Clemente, J., A. Montanes and M. Reyes (1998), "Testing for a unit root in variables with a double change in the mean", *Economics Letters* 59, pp. 175-182

Dickey, D. A. and W.A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association* 74, pp. 427-431

Dickey, D. A. and W. A. Fuller (1981), "Likelihood ratio for autoregressive time series with a unit root", *]Econometrica* 49, pp. 1057-1072.

Diebold, F. X. and R. Mariano (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics* 13, pp. 253-265.

Elliott, G., T. J. Rothenborg and J. H. Stock (1996), Efficient tests for an autoregressive unit root,*Econometrica*, Vol. 64, pp. 813-836.

Enders, W. (2010), *Applied Econometric Time Series*, (3rd Edition ed.), Hoboken, NJ, USA: John Wiley & Sons

Frances, P. H. (1991), "Seasonality, non-stationarity and the forecasting of monthly time series", *International Journal of Forecasting* 7, pp. 199-208.

Franses, P. H. and B. Hobijn (1997), "Critical values for unit root tests in seasonal time series", *Journal of Forecasting*, 24(1), pp. 25-48.

Giacomimi, R. and H. White (2006), "Tests of conditional predictive ability", *Econometrica*, Econometric society, vol 74(6), pp. 1545-1578

Gujarati, D. N. (2004), *Basic Econometrics*, 4th Edition, McGrow-Hill Companies, New York

Hasler, U. and P. M. M. Rodrigues (2002),"Inference on the cointegration rank in fractionally integrated processes", *Journal of Econometrics* 100, pp. 167-185.

Hylleberg, S., Engle, R. F., Granger, C. W. and Yoo, B. S. (1990),"Seasonal Integration and Cointegration", *Journal of Econometrics* 44, pp. 215-238.

Kwiatkowski, D., P. C. B. Phillips, P. J. Schmidt and Y. Shin (1992), "Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root", *Journal of econometrics* 54, pp. 159-178.

IMF (2011), "Assessing Reserves Adequacy", Monetary and Capital Markets, Research and Strategy, Policy and Review Departments, IMF Washington D.C.

IMF (2013), *Sixth Edition of the IMF's Balance of Payments and International Investment Position Manual*, International Monetary Fund Publication Services, Washington, D.C.

Leybourne, S. J., T. C. Mills and P. Newbold (1998), "Spurious rejections by Dickey-Fuller test in the presence of a break under the null", *Journal of Econometrics*, Vol. 87, pp. 191-203.

Lumsdaine, R. L. and Papell, D. H. (1997), "Multiple trend breaks and the unit root hypothesis", *Review of Economics and Statistics* 79 (2), pp. 212-218.

Mendez, P. M. (2015), "Seasonal unit roots and structural breaks in agricultural time series: monthly exports and domestic supply in Argentina", *MPRA* Paper No. 63831

Nelson, C. R. and C.I. Plosser (1982), "Trends and random walks in macroeconomic time series", *Journal of Monetary Economics* 10, pp. 139-162.

Perron, P. (1989), "The Great Crash, the oil price shock and the unit root hypothesis", *Econometrica* 57, pp. 1361-1401.

Perron, P. (1997), "Further Evidence on breaking trend functions in macroeconomic variables", *Economic Research Program Memorandum* No. 350, Princeton University

Perron, P. and T. Vogelsang (1992), "Nonstationarity and level shifts with an application to purchasing power parity", *Journal of Business and Economic Statistics* 10(3), pp. 301-320.

Phillips, P. and P. Perron (1988), "Testing for a unit root in time series regression", *Biometrica* 75, pp. 335-346. Sid, S. E. and D. A. Dickey (1984), Testing for unit roots in autoregressive moving average models of unknown order, Biometrika, Vol. 71, pp. 599-607.

Zivot, E. and D.W.K. Andrews (1992), "Further evidence on the Great Crash, the oil price shock and the unit root hypothesis", *Cowles Foundation Discussion Paper* No. 944, Yale University

A Appendix

HEGY test auxiliary equations

$$s_{8,t} = (1 - L^{12})s_t \tag{7}$$

$$s_{1,t} = (1 + L)(1 + L^2)(1 + L^4 + L^8)s_t \tag{8}$$

$$s_{2,t} = -(1 - L)(1 + L^2)(1 + L^4 + L^8)s_t \tag{9}$$

$$s_{3,t} = -(1 - L^2)(1 + L^4 + L^8)s_t \tag{10}$$

$$s_{4,t} = -(1 - L^4)(1 - \sqrt{3}L + L^2)(1 + L^2 + L^4)s_t t \tag{11}$$

$$s_{5,t} = -(1 - L^4)(1 + \sqrt{3}L + L^2)(1 + L^2 + L^4)s_t \tag{12}$$

$$s_{6,t} = -(1 - L^4)(1 - L^2 + L^4)(1 - L + L^2)s_t \tag{13}$$

$$s_{7,t} = -(1 - L^4)(1 - L^2 + L^4)(1 + L + L^2)s_t \tag{14}$$

Table 10: HEGY test for monthly data - critical values

	Constant, dummies and no trend						**Constant, dummies and trend**					
t-statistics	**0.01**	**0.05**	**0.10**				**0.01**	**0.05**	**0.10**			
$\boldsymbol{\pi 1}$	−3.32	−2.76	−2.47				−3.84	−3.30	−3.00			
$\boldsymbol{\pi 2}$	−3.31	−2.76	−2.47				−3.26	−2.72	−2.43			
t-statistics	**0.005**	**0.025**	**0.05**	**0.95**	**0.975**	**0.995**	**0.005**	**0.025**	**0.05**	**0.95**	**0.975**	**0.995**
$\boldsymbol{\pi 3}$	−2.85	−2.15	−1.86	1.87	2.20	2.80	−2.85	−2.13	−1.82	1.82	2.13	2.81
$\boldsymbol{\pi 4}$	−3.95	−3.52	−3.25	−0.50	−0.23	0.39	−4.04	−3.45	−3.21	−0.46	−0.19	0.33
$\boldsymbol{\pi 5}$	−3.92	−3.39	−3.11	−0.08	0.21	0.79	−3.82	−3.34	−3.06	−0.09	0.20	0.80
$\boldsymbol{\pi 6}$	−4.00	−3.47	−3.21	−0.48	−0.19	0.41	−3.91	−3.42	−3.17	−0.45	−0.16	0.41
$\boldsymbol{\pi 7}$	−0.86	−0.25	0.04	3.14	3.41	3.93	−0.77	−0.19	0.14	3.08	3.33	3.89
$\boldsymbol{\pi 8}$	−4.01	−3.48	−3.22	−0.43	−0.15	0.43	−3.96	−3.49	−3.22	−0.43	−0.14	0.43
$\boldsymbol{\pi 9}$	−3.56	−3.02	−2.70	0.83	1.21	1.86	−3.50	−2.93	−2.61	0.79	1.11	1.67
$\boldsymbol{\pi 10}$	−4.02	−3.50	−3.24	−0.45	−0.18	0.43	−3.92	−3.42	−3.18	−0.46	−0.14	0.44
$\boldsymbol{\pi 11}$	−1.80	−1.13	−0.79	2.71	3.05	3.65	−1.70	−1.09	−0.75	2.66	2.95	3.50
$\boldsymbol{\pi 12}$	−4.06	−3.51	−3.24	−0.48	−0.17	0.47	−4.04	−3.49	−3.21	−0.48	−0.21	0.28
F-statistics				**0.90**	**0.95**	**0.99**				**0.90**	**0.95**	**0.99**
$\boldsymbol{\pi 3, \pi 4}$				5.24	6.25	8.24				5.13	6.06	8.35
$\boldsymbol{\pi 5, \pi 6}$				5.22	6.17	8.20				5.10	6.09	7.95
$\boldsymbol{\pi 7, \pi 8}$				5.27	6.23	8.39				5.14	6.19	8.19
$\boldsymbol{\pi 9, \pi 10}$				5.30	6.25	8.30				5.07	5.93	7.86
$\boldsymbol{\pi 11, \pi 12}$				5.30	6.35	8.46				5.09	6.11	8.31
$\boldsymbol{\pi 3, ..., \pi 12}$				3.91	4.36	5.17				3.87	4.34	5.17

Note: The critical values are obtained by 10,000 Monte Carlo simulations. DGP: $y = y(-12) + \epsilon, \epsilon \sim N(0,1)$. Number of observations equals 202.

Table 11: Critical values of HEGY test for monthly data with endogenous breaks

	Specification without trend			**Specification with trend**		
t-statistics	**0.99**	**0.95**	**0.90**	**0.99**	**0.95**	**0.90**
$\boldsymbol{\pi 1}$	-4.86	-4.28	-3.96	-5.33	-4.75	-4.45
$\boldsymbol{\pi 2}$	-4.81	-4.23	-3.93	-4.84	-4.24	-3.96
F-statistics	**0.90**	**0.95**	**0.99**	**0.90**	**0.95**	**0.99**
$\boldsymbol{\pi 3, \pi 4}$	11.34	12.75	16.24	11.33	12.78	15.82
$\boldsymbol{\pi 5, \pi 6}$	11.55	12.87	15.90	11.34	12.78	15.90
$\boldsymbol{\pi 7, \pi 8}$	11.54	12.89	16.13	11.45	12.88	16.04
$\boldsymbol{\pi 9, \pi 10}$	11.40	12.79	16.12	11.35	12.87	16.24
$\boldsymbol{\pi 11, \pi 12}$	5.31	6.31	8.93	5.33	6.39	8.91

Note: The critical values are obtained by 5000 Monte Carlo simulations.The test is based on the Innovational Outlier model. Brake dates are determined on the basis of the minimum t, maximum F statistics criteria.

Table 12: Critical values of HEGY test for monthly data with endogenous breaks

	Specification without trend			**Specification with trend**		
t-statistics	**0.99**	**0.95**	**0.90**	**0.99**	**0.95**	**0.90**
$\pi 1$	-3.83	-3.14	-2.81	-4.24	-3.62	-3.31
$\pi 2$	-3.77	-3.14	-2.80	-3.80	-3.12	-2.80
F-statistics	**0.90**	**0.95**	**0.99**	**0.90**	**0.95**	**0.99**
$\pi 3, \pi 4$	6.64	8.01	10.94	6.67	7.98	11.01
$\pi 5, \pi 6$	6.85	8.06	10.82	6.74	8.10	11.01
$\pi 7, \pi 8$	6.66	8.10	10.97	6.80	8.18	11.47
$\pi 9, \pi 10$	6.71	7.94	10.64	6.75	7.98	10.79
$\pi 11, \pi 12$	6.84	8.04	11.11	6.83	8.31	11.13

Note: The critical values are obtained by 5000 Monte Carlo simulations.The test is based on the Innovational Outlier model. Brake dates are determined on the basis of the maximum value of the t statistics on a dummy shift variable.

Table 13: One step ahead forecast evaluation statistics

	BIAS	**SE**	**RMSE**	**MAE**	**MAPE**
Expanding window					
ar2ma1_12	-0.006	0.019	0.020	0.016	2.421
ma3_12	-0.006	0.021	0.021	0.017	2.757
ma3_12seas	-0.005	0.023	0.023	0.019	4.170
ma3mma12seas	-0.006	0.023	0.024	0.019	4.151
Constant window					
ar2ma1_12	-0.003	0.018	0.018	0.015	2.182
ma3_12	-0.002	0.019	0.019	0.016	2.347
ma3_12seas	-0.002	0.019	0.019	0.016	2.672
ma3mma12seas	-0.002	0.019	0.019	0.016	2.695

Table 14: Two step ahead forecast evaluation statistics

	BIAS	**SE**	**RMSE**	**MAE**	**MAPE**
Expanding window					
ar2ma1_12	-0.013	0.019	0.023	0.019	3.864
ma3_12	-0.013	0.020	0.023	0.019	3.880
ma3_12seas	-0.012	0.023	0.026	0.019	4.490
ma3mma12seas	-0.011	0.023	0.025	0.019	4.422
Constant window					
ar2ma1_12	-0.006	0.020	0.021	0.017	2.940
ma3_12	-0.006	0.020	0.021	0.017	2.883
ma3_12seas	-0.004	0.019	0.020	0.016	2.869
ma3mma12seas	-0.005	0.019	0.020	0.016	3.088

Table 15: Three step ahead forecast evaluation statistics

	BIAS	SE	RMSE	MAE	MAPE
Expanding window					
ar2ma1_12	-0.013	0.019	0.023	0.019	3.883
ma3_12	-0.013	0.020	0.023	0.019	3.901
ma3_12seas	-0.012	0.023	0.026	0.019	4.470
ma3mma12seas	-0.011	0.023	0.025	0.019	4.401
Constant window					
ar2ma1_12	-0.006	0.020	0.021	0.017	2.931
ma3_12	-0.006	0.020	0.021	0.017	2.866
ma3_12seas	-0.004	0.019	0.019	0.016	2.854
ma3mma12seas	-0.005	0.019	0.020	0.016	3.043

Table 16: Four step ahead forecast evaluation statistics

	BIAS	SE	RMSE	MAE	MAPE
Expanding window					
ar2ma1_12	-0.013	0.019	0.023	0.019	3.905
ma3_12	-0.013	0.019	0.023	0.019	3.914
ma3_12seas	-0.012	0.023	0.026	0.019	4.490
ma3mma12seas	-0.011	0.023	0.025	0.019	4.432
Constant window					
ar2ma1_12	-0.006	0.020	0.021	0.017	2.959
ma3_12	-0.006	0.020	0.021	0.017	2.882
ma3_12seas	-0.004	0.019	0.019	0.016	2.859
ma3mma12seas	-0.005	0.019	0.020	0.016	3.099

Table 17: Diebold Mariano test between alternative specifications

	Period 1			**Period 2**			**Period 3**			**Period 4**		
	Constant window			**Constant window**			**Constant window**			**Constant window**		
	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE	MSE	MAE	MAPE
ar2ma1_12	0.00	0.01	2.18	0.00	0.02	2.94	0.00	0.02	2.93	0.00	0.02	2.96
ma3_12	0.00	0.02	2.35	0.00	0.02	2.88	0.00	0.02	2.87	0.00	0.02	2.88
Difference	−0.00	−0.00	−0.17	0.00	0.00	0.06	0.00	0.00	0.06	0.00	0.00	0.08
p value	0.18	0.12	0.52	0.29	0.40	0.13	0.05	0.35	0.17	0.28	0.23	0.13
ar2ma1_12	0.00	0.01	2.18	0.00	0.02	2.94	0.00	0.02	2.93	0.00	0.02	2.96
ma3_12seas	0.00	0.02	2.67	0.00	0.02	2.87	0.00	0.02	2.85	0.00	0.02	2.86
Difference	−0.00	−0.00	−0.49	0.00	0.00	0.07	0.00	0.00	0.08	0.00	0.00	0.10
p value	0.62	0.28	0.12	0.66	0.70	0.88	0.63	0.71	0.87	0.63	0.66	0.84
ar2ma1_12	0.00	0.01	2.18	0.00	0.02	2.94	0.00	0.02	2.93	0.00	0.02	2.96
ma3mma12seas	0.00	0.02	2.70	0.00	0.02	3.09	0.00	0.02	3.04	0.00	0.02	3.10
Difference	−0.00	−0.00	−0.51	0.00	0.00	−0.15	0.00	0.00	−0.11	0.00	0.00	−0.14
p value	0.80	0.56	0.28	0.68	0.82	0.68	0.70	0.84	0.76	0.71	0.87	0.71
ma3_12	0.00	0.02	2.35	0.00	0.02	2.88	0.00	0.02	2.87	0.00	0.02	2.88
ma3_12seas	0.00	0.02	2.67	0.00	0.02	2.87	0.00	0.02	2.85	0.00	0.02	2.86
Difference	0.00	−0.00	−0.33	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.02
p value	0.97	0.84	0.47	0.68	0.73	0.97	0.64	0.73	0.98	0.64	0.70	0.96
ma3_12	0.00	0.02	2.35	0.00	0.02	2.88	0.00	0.02	2.87	0.00	0.02	2.88
ma3mma12seas	0.00	0.02	2.70	0.00	0.02	3.09	0.00	0.02	3.04	0.00	0.02	3.10
Difference	0.00	0.00	−0.35	0.00	0.00	−0.20	0.00	0.00	−0.18	0.00	0.00	−0.22
p value	0.84	0.97	0.54	0.70	0.85	0.52	0.71	0.87	0.58	0.73	0.91	0.50
ma3_12seas	0.00	0.02	2.67	0.00	0.02	2.87	0.00	0.02	2.85	0.00	0.02	2.86
ma3mma12seas	0.00	0.02	2.70	0.00	0.02	3.09	0.00	0.02	3.04	0.00	0.02	3.10
Difference	0.00	0.00	−0.02	−0.00	−0.00	−0.22	−0.00	−0.00	−0.19	−0.00	−0.00	−0.24
p value	0.00	0.00	0.87	0.55	0.00	0.01	0.01	0.03	0.01	0.03	0.04	0.01

Note:
Numbers in the table represent:
- value of a specific criteria for particular ARIMA specification,
- difference (the first minus the second value in each block),
- and statistical significance of the H0 hypothesis that forecast values are equal.