Stochastic models for forecasting inflation rate. Empirical evidence from Romania

Ionut-Cristian Baciu

*Corresponding author.
E-mail address: ionutb86@yahoo.com

1. INTRODUCTION

Although being slightly perceptible, inflation is one of the most difficult to define and bound complex phenomena. With all the differences between the various currents of economic thinking on the nature and effects of inflation, we can speak of a consensus on the need to control the inflation process.
An important effect of inflation affecting consumers and companies is the uncertainty generated when price changes vary significantly over time. The volatile inflation can be caused by major fluctuations in the demand and/or supply in the economy. The volatile and unpredictable inflation rates make difficult the long-term planning and thus the investments and savings are discouraged.

The uncertainty about inflation causes two types of economic effects. The first affects companies and consumers that must take economic decisions taking into account the future inflation forecasted. (‘ex ante’ effect). The uncertainty of the inflation affects the financial markets by raising interest rates in the long term, leading to uncertainty regarding other variables that are important in economic decisions and encourages companies to direct their resources in order to avoid certain risks related to inflation (Devereaux, 1989).

Inflation discourages saving, people prefer present satisfactions instead of the future ones. Short term investments are searched against the costly ones but intended for modern construction and reconstruction of economic sectors where the prospect of profit is in more far away horizon. Investors should adjust for inflation, but a greater increase of uncertainty will force them to claim higher efficiencies to invest, leading to increased financing costs.

The second is the ‘ex post’ effect when the inflation rate differs significantly from its projection. An inflation that is better than expected leads to a transfer of wealth from the creditor to the debtor. The latter is favored because repayment is done with money whose purchasing power is “penalized” by the inflation rate.

When inflation is high, anti-inflationary measures are adopted. These measures act to lower the inflation but increase the variability of the phenomenon. In addition, it is created a state of uncertainty on the forecast of inflation in the future as the impact of policies against inflation does not occur immediately but after a certain period of time.

The purpose of this paper is to identify the econometric model that provides the most accurate predictions for the inflationary phenomenon, studied for the period January 1997 – August 2013 in Romania.

The paper is structured as follows: Section 2 presents a brief review of the literature on the use of autoregressive models. Section 3 presents aspects of the methodology used to estimate an econometric model for the time series and making a forecast for September 2013. In Section 4, using data series with monthly inflation rate for the period studied, three models were estimated AR (autoregressive model), MA (moving average model) and ARMA (autoregressive-moving-average model) to describe the evolution of the monthly inflation rate. Based on specific criteria, it is chosen the model that will help to forecast the rate of inflation for September 2013. The key findings of the study can be found in Section 5.

2.2. Literature review

Based on studies by Box and Jenkins, the autoregressive moving average (ARMA) models have become the main tools for modeling and forecasting time series. The theoretical justification is based on the Wold’s decomposition theorem for stationary parametric processes.

The estimation of an ARMA model includes three main steps: identifying the type of model, parameter estimation and forecasting. The most important step is the one involving the determination of AR and MA parties and of the model and it is performed using the information provided by the autocorrelation and partial autocorrelation functions. One major requirement of the ARMA model is that the time series must be linear and stationary (Wu and Chan, 2011).

Autoregressive integrated moving average (ARIMA) modelling is a specific subset of univariate modelling, in which a time series is expressed in terms of past values of itself (the autoregressive component) and current and lagged values of a “white noise” error term (the moving average component).

Box and Jenkins found that a high number of non-stationary time series can be modeled by using ARIMA integrated and moving average models (Turtlean, 2007; Jemna, 2012). The authors accept the idea of a non-stationary stochastic process but provide a method to eliminate its influence by differentiation.

In the literature, there are numerous studies using Box-Jenkins methodology for modeling inflation phenomenon. Fritzler et all.(2002) find that for austrian inflation, univariate models outperform multivariate models at short horizon. Meyler et all.(1998) used two different approaches – the Box-Jenkins methodology and the objective penalty function methods for identifying appropriate ARIMA models, for Irish inflation.

The studies conducted by Junttila (2001), Pufnik and Kunovac (2006), Suleman and Sarpong (2012), confirmed that the ARIMA models tends to perform better in terms of forecasting compared to other time series models.
3. 3. Research methodology

The main purpose of modeling the time series using stochastic processes is to explain the manner of the phenomenon’s evolution and to make predictions based on the model estimated.

Depending on the type of time series, several categories of stochastic processes as models for time series are used: autoregressive processes, moving average processes and composite models based on them.

In the development of such models there are the following considerations:

- the evolution of economic phenomena is under the existing resources, the capacities already created. The variables in the economy have an inertial nature, a strong autoregressive component being present.
- the moving average component is the effect of some unpredictable events on variables.

3.1. Autoregressive process (AR)

A stationary series, Y_t, follows a process AR (p) if the condition if fullfilled:

\[ Y_t = \beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p} + \epsilon_t, \]

where \( \epsilon_t \sim N(0, \sigma^2) \) stationary time series, \( E(\epsilon_t) = 0, E(\epsilon_t^2) = \sigma^2, E(\epsilon_t, \epsilon_s) = 0, \) if \( t \neq s; \) \( \beta_0, \beta_1, \ldots, \beta_p - \) parameters

The autoregressive models are characterized by the fact that the value of variable Y at time t depends on the previous values of the variable. One of AR models used to explain the unpredictability nature of financial asset price evolution is Random Walk model, where \( p=1, \beta_0 = 0, \beta_1 = 1. \)

The representation of the model is \( Y_t = Y_{t-1} + \epsilon_t. \)

As a result, the value of a series in a given period depends on the value of the series in the previous period and a random term whose value is expected to be zero.

3.2. Moving average process (MA)

The \( Y_t \) process follows a moving average process of q order, if it is defined by the equality:

\[ Y_t = \mu + \alpha_0 \epsilon_t + \alpha_1 \epsilon_{t-1} + \ldots + \alpha_q \epsilon_{t-q}, \]

where \( \epsilon_t \sim N(0, \sigma^2) \) stationary time series, \( E(\epsilon_t) = 0, E(\epsilon_t^2) = \sigma^2, E(\epsilon_t, \epsilon_s) = 0, \) if \( t \neq s; \) \( \alpha_0, \alpha_1, \ldots, \alpha_q - \) parameters

3.3. Autoregressive moving average process (ARMA)

A model of ARMA autoregressive moving average type \( (p, q) \) contains both a component of autoregressive type and a component of moving average type:

\[ Y_t = \beta_0 + \beta_1 Y_{t-1} + \ldots + \beta_p Y_{t-p} + \epsilon_t - \alpha_1 \epsilon_{t-1} - \ldots - \alpha_q \epsilon_{t-q}, \]

where \( p \) is the ordinal of autoregressive part, \( q \) the ordinal of moving average and \( \epsilon_t \) is a process of white noise type.

Models AR (p) and MA (p) can be regarded as a particular form of model ARMA (p, q):

- AR(p) = ARMA(p,0)
- MA(q) = ARMA(0,q)

Autocorrelation function measures the correlation between series values at different time distances. Usually, it is used the autocorrelation function for selection that is defined as follows:

\[ P_y(k) = \frac{\sum (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum (Y_t - \bar{Y})^2} = \frac{\text{cov}(Y_t, Y_{t-k})}{\sigma^2_y} \]
where $Y_t$ is initial series, $\sigma^2$ is variance.

The graphic of autocorrelation function for various k lags is called correlogram.

The partial autocorrelation function is determined as a partial correlation coefficient between the variable at $t$ time and that at $t-k$ time. The partial correlation coefficient measures the intensity of the connection between the two variables, $Y_t$ and $Y_{t-k}$ controlling the influence of variables for a different gap.

3.4. Autoregressive Integrated Moving Average (ARIMA) process

Autoregressive Integrated Moving Average processes are built as a generalization of ARMA processes that do not meet the condition of stationarity. A time series may be stationed by applying the difference operator. Each integration order corresponds to the level of differentiation of the series analyzed.

Using Box-Jenkins method the necessary steps modeling a time series using an ARIMA process are determined. The estimation stages are:

- the analysis of the series stationarity
- identification of the model’s type
- parameters estimation of the econometric models
- validation and making predictions

Testing the stationarity of the series is achieved by applying the unit root tests (Dickey-Fuller, ADF and Phillips-Perron tests). If the series is not stationary, data correction is performed using the difference operator. The number of differentiation of the series to obtain a stationary series is equal with the order of integration, that is equal to $d$.

Identification of the model’s type

The identification of the process involves determining the $p$, $d$, $q$ parameters, allowing the choice of ARIMA. The identification is performed using autocorrelation and partial autocorrelation functions after having checked the stationarity of the series.

For a stationary series, the two correlograms are analyzed to identify the process that is most suitable to the data series. The following possibilities can be distinguished:

- if there is an $h$ value equal to $q$ starting from the value of the autocorrelation function drops suddenly to zero, then it is used a MA($q$) process for the processing of the series.
- where the value of the partial autocorrelation function decreases to zero, starting with a gap value equal to $p$, then the time series is processed through an AR($p$) process
- ARMA model is used when a type of process is not obvious or when the correlograms indicate a process involving both components.

The next stage is to estimate the parameters of the model. This stage provides information about how appropriate is the model chosen.

The estimated model is subjected to testing for validation. At this stage the model’s parameters are tested using t-Student test. The hypotheses for the error variable from the model are checked: the average of the errors to be zero, the errors to be normally distributed, to not admit the phenomenon of autocorrelation. If errors do not follow the conventional hypotheses, it returns to the stage of model identification.

To choose among several models the appropriate one, there are used a series of indicators applied to the series of residuals estimated using the autoregressive model ($R^2$, $F$) or indicators such as Akaike (AIC) or Schwartz criterion.

Based on the selected model, various analyzes and forecasts are made. To get the forecast, all the changes that have been applied to time series in the modeling process are taken into account.
4. 4. Empirical Study

4.1. Data used

The data used are represented by the monthly data series from January 1997 to August 2013 for the inflation rate in Romania. The data source is represented by the annual reports of the National Bank of Romania. The empirical analysis is provided using the statistical software EViews.

We applied on the series a logarithmic transformation to remove the negative effect induced by non-stationarity in variance. After this transformation, we will get a series with approximately the same trend but with milder amplitudes of the variance.

For this series applies the Box – Jenkins’s methodology to estimate a model for the time series and to realise a prediction for September 2013.

In order to estimate such a model we need to ensure that the series is stationary.

4.2. The analysis of the series stationarity

Hypothesis: H₀: the time series has a unit root
H₁: the time series is stationary

According to Augmented Dickey-Fuller test, for a significance apron of 0.05, the test’s value is -1.72 and is higher than the theoretical critique value of the table, which is equal to -2.87. Using the p value, the null hypothesis is accepted for a certain level of relevance, whenever the p probability is higher than that level of relevance. In conclusion we accept the null hypothesis according to which the series is non-stationary(table A1 – appendix).

To determine the order of integration of the series is tested the stationarity of the first difference series. As the test value is lower than the critical value for any level of relevance, we can say that at 5% level of relevance, the null hypothesis is rejected.

Applying the unitary root test for the differentiation series, according to the table below (table 1), shows that it is stationary, so the basis series is integrated by the first order.

<table>
<thead>
<tr>
<th>Table 1. Augmented Dickey-Fuller test</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-statistic</td>
</tr>
<tr>
<td>ADF test statistic</td>
</tr>
<tr>
<td>Test critical value</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

4.3. Identification of the model’s type

Using the data series with a monthly frequency of the inflation rate for the period January 1997- August 2013 there were estimated three models AR, MA and ARMA to describe the inflation rate.

Based on the autocorrelation coefficients and on the partial correlation coefficients are determined the start autoregressive models for the analysis of the data series (figure A2 – appendix).

At this stage we will determine the p and q parameters from the graph of the autocorrelation and partial autocorrelation functions of the stationary series, considering that:

- the autocorrelation function decreases sharply to zero after the first two terms, so it is anticipated a MA process (2)
- the value of the partial autocorrelation function decreases to zero starting with the second value, so it is recommended an AR process (1).

For a time series to be integrated by the order 1, the autocorrelation coefficients should be close to 1, and the autocorrelation coefficients for the first difference should be (statistically significant) lower than 1.
4.4. Parameters estimation of the econometric models

Since the constant from the model is not significant for a risk of 5%, this can be removed from the model. In this case, the estimated coefficient of the autoregressive component becomes 0.99.

The estimated model is:

$$ \Delta Y_t = 0.9938 \Delta Y_{t-1} + \epsilon_t $$

Estimation of the MA(2) model:

$$ \Delta Y_t = 2.589 + \epsilon_t - 1.676 \epsilon_{t-1} - 0.955 \epsilon_{t-2} $$

Estimation of the ARMA(1,2) model:

$$ \Delta Y_t = 0.9931 \Delta Y_{t-1} + \epsilon_t - 0.3129 \epsilon_{t-1} - 0.2374 \epsilon_{t-2} $$

The statistical indicators for the three types of models are presented in the following table:

<table>
<thead>
<tr>
<th></th>
<th>AR(1)</th>
<th>MA(2)</th>
<th>ARMA(1,2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted R-squared</td>
<td>0.989593</td>
<td>0.905757</td>
<td>0.993166</td>
</tr>
<tr>
<td>Akaike info criterion</td>
<td>-1.519265</td>
<td>0.711269</td>
<td>-1.657553</td>
</tr>
<tr>
<td>Schwarz criterion</td>
<td>-1.502716</td>
<td>0.760744</td>
<td>-1.607905</td>
</tr>
</tbody>
</table>

According to all three criteria (the highest value of $R^2$ and, respectively, the lowest values recorded by the informational criteria) and taking into account the differentiation of the series from the previous stage, for modelling the initial time series, it is chosen the ARIMA specification (1,1,2).

4.5. Testing the validity of the ARIMA model (1,1,2)

The model coefficients are significantly different from zero, according to table 3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>L_RINFL</td>
<td>0.9931</td>
<td>244.63</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(1)</td>
<td>0.3129</td>
<td>4.55</td>
<td>0.0000</td>
</tr>
<tr>
<td>MA(2)</td>
<td>0.2374</td>
<td>3.44</td>
<td>0.0007</td>
</tr>
</tbody>
</table>

According to the statistical results, the module of the characteristic polynomial is lower than 1, and therefore the equation is stable (table A2 – appendix).

4.5.1. Testing the hypothesis that the error average is zero

After comparing the calculated value with the theoretical value of the test, it has observed that the null hypothesis is accepted with a probability of 0.95.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 4. Testing the hypothesis according to which the errors average is null</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5. Heteroskedasticity Test: Breusch-Pagan-Godfrey

<table>
<thead>
<tr>
<th>Method</th>
<th>Value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-statistic</td>
<td>0.1396</td>
<td>0.8891</td>
</tr>
</tbody>
</table>

4.5.2. Testing the hypothesis of homoscedasticity

When the Chi-Square significance is greater than 0.05, we are taking the decision to accept the hypothesis of homoscedasticity with a probability of 0.95.

4.5.3. Testing the hypothesis of errors normality

The distribution of the monthly evolutions of the inflation rate has the average close to zero, presents a negative asymmetry and the kurtosis has a value which means that this distribution is leptokurtic (Figure A1 - appendix).

In this case, the probability associated to the Jarque Bera test is zero and thus is rejected the null hypothesis (the series is normally distributed).

4.5.4. Testing the autocorrelation errors

Q-statistic and its associated probability represents a statistical test which has as null hypothesis that there is no autocorrelation up to lag k. The probability associated to the Q-statistic test is superior to the level of relevance, the null hypothesis is accepted (nonexistent autocorrelation errors). According to the errors correlogram, there are no serial autocorrelation of the errors, up to lag 8 (see figure A3 – appendix).

4.6. Inflation forecast

The estimation of the ARIMA model (1,1,2) is valid, in conclusion the time series can be represented by an ARMA (1,2) process.

The econometric model estimated for inflation rate is an ARIMA(1,1,2):

\[ \Delta Y_t = 0.9931 \Delta Y_{t-1} + \epsilon_{t-1} - 0.3129 \epsilon_{t-1} - 0.2374 \epsilon_{t-2} \]

Based on the estimated model is obtained the projection value of the inflation rate for September 2013. To achieve the predictions, it takes into account all the changes that the time series has incurred in the modelling process.

Thus, the foretokened inflation rate for September 2013 is 3.01%.

In September 2013, the inflation measured over the last 12 months decreased very strongly from 3.67% to just 1.88%, below the 2% threshold recommended by the European Central Bank.

The impact of reducing the cost of bread was very important due to the large relative weight of the bakery group in the consumer’s basket (6.73%). Comparing the projected rate of inflation with the actual one from September 2013, that was 1.88%, results that there is a significant difference between projection and actual value.

5. Conclusions

The inflationary phenomenon is considered a sign of an imbalance in the economy, which is why it is necessary to control the inflation. The inflationary expectations have an important role in the conduct of the monetary policy and this is why the monetary authorities use the inflation forecasts so that monetary policy to be effective.

In this paper we applied the Box-Jenkins methodology in order to estimate a model for time series consists of records monthly inflation rate between January 1997 and August 2013. Based on the best econometric model after comparing various criteria, we achieved prediction for September 2013.
The study results that there is a significant difference between the prediction made on the selected model and the actual inflation rate. Estimation of ARIMA(p,d,q) models have certain limitations: are not suitable for medium and long-term forecasts, after q terms projected, value remains the same and parameters can be very different from one estimation to another.

The long-term analysis of the evolution of economic variables often reveals that their version is not constant over time. Thus, there are introduced Autoregressive Conditional Heteroskedastic (ARCH) models that are taking into account the information contained in the conditional variance of the process.

References

Appendix A

Table A1. Augmented Dickey-Fuller test

<table>
<thead>
<tr>
<th></th>
<th>t-statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADF test statistic</td>
<td>-1.7252</td>
<td>0.4169</td>
</tr>
</tbody>
</table>

Test critical value

<table>
<thead>
<tr>
<th></th>
<th>1% level</th>
<th>5% level</th>
<th>10% level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-3.4653</td>
<td>-2.8768</td>
<td>-2.5750</td>
</tr>
</tbody>
</table>

Table A2. Inverse roots of ARMA polynomial

<table>
<thead>
<tr>
<th>MA Root(s)</th>
<th>Modulus</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.1564 ±0.4614i</td>
<td>0.4872</td>
<td>3.3108</td>
</tr>
</tbody>
</table>

Figure A1. Histogram errors

Figure A2. Autocorrelation and partial autocorrelation functions of inflation rate

Figure A3. Correlogram of the errors