UDC 330:43 / 51:77

Econometric modelling of nonstationary processes

M. Lytsenko, T.O. Marynych
Sumy State University, Ukraine

Econometric research of nonstationary time series on causality, cointegration relation and adequate simulation methods was conducted. VAR and VEC models were found to be the most appropriate ways to make reliable prediction and scenario analysis of macro financial data under unstable economic conditions. These econometric techniques were approbated on the financial indicators of Ukrainian economy.

Key words: *nonstationary time series, causality, cointegration, vector autoregression, vector error correction model, scenario analysis.*

1. Introduction

The choice of methods and models to obtain reliable forecasts, as well as to establish relationships between the study variables is of particular relevance in terms of macroeconomic and financial instability. Adequate prediction and scenario analysis techniques become an essential part of decision-making activity of the government and business authorities thus improving their effectiveness.

The vector autoregression approach (*VAR*), proposed by C. Sims in 1980 [1] as an alternative to the theory–based structural modeling, is a common way to estimate nonstationary time series. This method eliminates the problem of determining endogenous and exogenous variables, considers econometric equations with lagged variables. J. Hamilton (1994) continued *VAR* research examining on the problem of reaching stationarity of the incoming data [2]. Campbell (1997) and Tsay (2001) focused on the *VAR* analysis of the financial data and financial markets' time seires. Engle and Granger (1987) introduced vector error correction models (*VECM*) which resolve the drawback of the *VAR* modeling considering both short-term and long-term aspects of the dynamics of the observed variables [3]. This effect was reached through introduction of the cointegration phenomenon that explores the stationary linear combinations of the time series. Currently *VAR* and *VEC* models are widely used for scenario analysis which considers alternative possible dynamics of endogenous variables under various paths of exogenous factors.

2. Formulation of the problem

Time series representing Ukrainian economy and financial sector are highly volatile and limited in the number of observations. Thus the task is to observe the features of the time series under unstable economic conditions, identify the problems of their modeling using *VAR/VEC* techniques and evaluate their prognostic qualities.

For this purpose the monthly macroeconomic statistics of the National bank of Ukraine [4] and State statistics department [5] within 2007-2014 years was used. It includes relative time series representing: share of loans and deposits in foreign currency (*Credfor; Depfor*); the ratio of total deposits to the monetary aggregate M2 (*DepM2*); the ratio of monetary base to the international foreign reserves; index of consumer prices (*Consprice*); index of industrial production prices (Indprice); industrial production index (*Indprod*); export and import growth rate (*Exp; Imp*); index of real effective exchange rate (*Reer*); interbank exchange rate (*MBrate*).

3. Solution of the problem and results

Econometric research was conducted in EViews software and involved the following steps:

1) graphical and descriptive statistics analysis of the time series;
2) seasonal and cyclical adjustment, logarithm transformation, sample division;
3) time series causality research using Granger test;
4) testing of the time series' stationarity;
5) checking time series with the same integration order for cointegration using Johansen test;
6) *VAR* and *VEC* models estimation and verification;
7) prediction using impulse response functions, variance decomposition and scenario analysis.

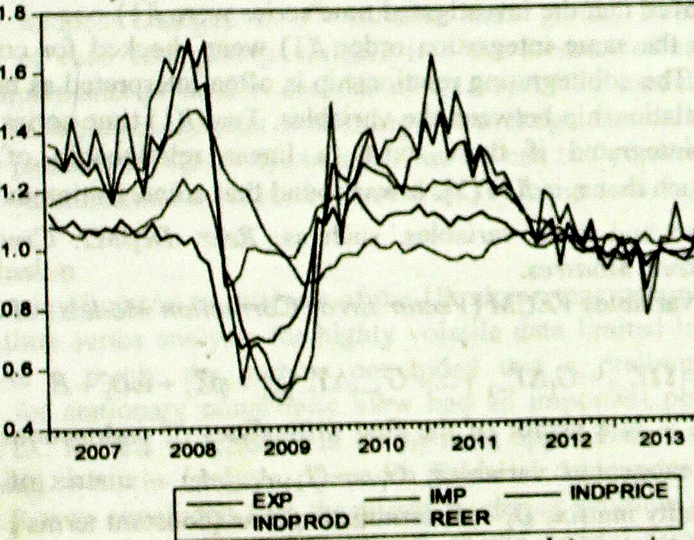

Fig. 1. Dynamics of the investigated time series

Figure 1 shows that investigated time series are nonstationary and include trend, cyclical and seasonal components. Thus preliminary preparation was done to get appropriate data for modeling. *Imp* and *Exp* time series were seasonally adjusted. All time series were divided into two periods to remove abnormal data of the 2008-2009 crisis period. Logarithm transformation was eliminated due to the relative nature of the time series. Trend was removed using first differences approach [2].

Granger test for causality was the next step in our research. It involved the construction of two-dimensional auto regression models:

$$y_t = a_0 + a_1 y_{t-1} + ... + a_p y_{t-p} + b_1 x_t + ... + b_p x_{t-p}$$
$$x_t = a_0 + a_1 x_{t-1} + ... + a_p x_{t-p} + b_1 y_t + ... + b_p y_{t-p} \quad (1),$$

where p – autoregression order for time series x_t, y_t.

Variable x is considered causal to variable y ($x \to y$), if, other things being equal, x contributes to the forecast of y and y doesn't significantly influence the forecast of x. If there is an interdependence between x and y, that each of the variables makes a significant contribution to the forecast of another one ($x \leftrightarrow y$), there would likely be a third variable z, which affects both variables.

Granger test showed the following results: *Reer* → *Indprice* on 1-6 lags; *Reer* → *Indprod* on 4-6 lags; *Reer* ↔ *Credfor* and *Depfor* on 2-6 lags; *Indprice* ↔ *Indprod; Indprice* → *Credfor* and *Indprice* → *DepM2* on 1-6 lags; *Exp, Imp* → *Indprice* on 1-4 lags; *Exp* ↔ *Imp* on 1-6 lags; *Indprod* → *Credfor* and *Indprod* ↔ *Exp* on lags 1-6; *MbasRes* → *Reer* on 1-6 lags; *MbasRes* → *Depfor* and *Credfor* on lags 3-6.

To check time series' stationarity we used the formal econometric methods:

1) construction of autocorrelation and partial correlation function (correlogram) which show the degree of closeness of the statistical relationship between t observations of the time series;

2) Augmented Dickey-Fuller Unit Root test (*ADF*). Its null hypothesis states that time series has a unit root and is nonstationary – $I(1)$. Existence of the unit root implies that its first differences make stationary time series $I(0)$ [2].

Both methods proved that the investigated time series were $I(1)$.

Time series with the same integration order $I(1)$ were checked for cointegration using Johansen test. The cointegrating relationship is often interpreted as being a long run or equilibrium relationship between the variables. Two $I(1)$ time series $y_{1,t}$ and $y_{2,t}$ are said to be cointegrated if there exists a linear relationship of the form $z_t = \beta_1 y_{1,t} + \beta_2 y_{2,t}$ such that z_t is $I(0)$ [3]. It was found that cointegration exists in time series vectors of two and three variables, such as: *Reer, DepM2, Credfor; Reer, Indprod, IndPrice; Reer, Mbasres*.

For these sets of variables *VECM* (*Vector Error Correction Models*) of the general form (2) were built:

$$G_0 \Delta Y_t = \Omega Y_{t-1} + G_1 \Delta Y_{t-1} + ... + G_{p-1} \Delta Y_{t-p+1} + \varphi X_t + \omega D_t + E_t \qquad (2),$$

where $\Delta Y_t = Y_t - Y_{t-1}$ – is a vector of the first differences of endogenous variables; X_t – is a vector of exogenous variables; $G_i = -(I_k, A_1, ... A_i)$ – matrix of short-term parameters; I_k – identity matrix; D_t – deterministic terms (constant terms , linear time trends etc.); E_t – vector of stochastic random errors; $\Omega = \alpha\beta^T$ – matrix of long-term parameters, where β – cointegration mattix, α – adjustment coefficients, that show the speed of turning back the long-term equilibrium trajectory.

According to the results of the Granger causality test *VAR* models of the general form (3) were built:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-1} ... + A_p y_{t-p} + \varepsilon_t \qquad (3),$$

where y_t – k-dimensional vector of endogenous variables; c – constant terms vector; A_p – coefficient matrix of ($k \times k$); p – lags' number; ε_t – error terms vector (E(ε_t)=0). If the stationarity condition is fulfilled and random disturbances ε_t is a "white noise", then equations of *VAR(p)* models can be estimated by the ordinary least squares method (OLS).

The practical issues of the *VAR* analysis concerned selection of the appropriate models and interpretation of the results. Estimation of the lag structure in *VAR* modeling was an important part of the first part of research. The number of lags for each model was determined on the basis of the minimum lag structure criteria: Likelihood Ratio tests, FPE (final prediction error), the AIC (Akaike information), the SC (Schwarz), the HQ (Hannan-Quinn) criteria and residual correlograms of appropriate *VAR*s. It was investigated that the average model's lag structure consists of 2-3 lags. Models were checked for adequacy using R^2 coefficient, t-statistics for coefficients of endogenous lagged variables and residuals statistics (residuals correlation matrix; autocorrelation, normality and heteroskedasticity testing).

Interpretation of *VAR* models is different from classical regression analysis: explanation of autoregression coefficients is replaced by impulse functions and variance decomposition analysis. Impulse function is a partial derivative $Y_{j,t+k}$ with time horizon k for a separate shock or impulse change of one variable at a time moment t in relation to other variables. Impulse functions associate the current value of the error with future values of Y_t. Thus, dynamic simulation of an external shock (one standard deviation increase) for each of the endogenous variables is performed, and then the system response to this impulse is examined on the forecasted time horizon. Variance decomposition shows how the error in j^{th} equation is important to explain the sudden changes in i^{th} variable thus giving information about the contribution of each endogenous variable into the forecast of each other. Scenario analysis complements prediction analysis of *VAR* and *VEC* models by setting one of the models' variables exogenous or entering a new exogenous variable and specifying the possible paths of its future dynamics. Thus we can test predicted dynamics of the endogenous variables under alternative scenarios for exogenous factors.

4. Conclusion

On the basis of empirical analysis of the Ukrainian macroeconomical statistics the features of time series analysis for highly volatile data limited in observations were observed. As a result, the authors concluded that a preliminary stage of data preparation for stationary congeneric view had an important place in the research. VAR and VEC models were found to be the most appropriate ways to make reliable prediction and scenario analysis of macro financial data under unstable economic conditions. It was concluded that the major modeling problem concerned residuals diagnostics and methods for bringing them to the normal view. Consideration of relative seasonally adjusted indicators on separated time intervals which show relatively stable dynamics of the variables or introduction of dummy variables displaying crisis events to the models improve their adequacy. Models interpretation was completed by scenario analysis which appeared to be a useful tool for stress-testing under adverse and favorable economic conditions.

LITERATURE

1. C. Sims. Macroeconomics and reality // Econometrica. – 1980. – Series 48 (1). – P. 1-48.
2. J. Hamilton. Time series analysis // Princeton University Press. – 1994.
3. R. Engle, C. Granger. Cointegration and Error Correction: representation, estimation and testing // Econometrica – 1987. – Series 55 (2). – P. 251-276.
4. Statistics of the National bank of Ukraine [Electronic resource]. – Mode access: www.bank.gov.ua.

   Government department statistics [Electronic resource]. – Mode access: www.ukrstat.gov.ua.

03.05.2015.