ECONOMICS
& Sociology

Mihaela Simionescu, The Use of VARMA Models in Forecasting Macroeconomic Indicators, *Economics & Sociology*, Vol. 6, No 2, 2013, pp. 94-102. **DOI: 10.14254/2071-789X.2013/6-2/9**

Mihaela Simionescu
PhD
Academy of Economic Studies
Bucharest, Romania
E-mail: mihaela_mb1@yahoo.com

THE USE OF VARMA MODELS IN FORECASTING MACROECONOMIC INDICATORS

Received: July, 2013
1st Revision: September, 2013
Accepted: October, 2013

DOI: 10.14254/2071-789X.2013/6-2/9

ABSTRACT. Although the scalar components methodology used to build VARMA models is rather difficult, the VAR models application being easier in practice, the forecasts based on the first models have a higher degree of accuracy. This statement is demonstrated for variables like the 3-month Treasury bill rate and the spread between the 10 year government bond yield, where the quarterly data are from the U.S. economy (horizon: first quarter of 2001 – second quarter of 2013). It was used a better measure of accuracy than those used in literature till now, the generalized forecast error of second moment, that was adapted to measure relative accuracy.

JEL Classification: C11, C13, C51

Keywords: macroeconomic forecasts, VARMA models, accuracy, scalar components methodology, full information maximum likelihood, canonical correlation.

Introduction

The *aim of this paper* is to provide forecasts based on a Vector autoregressive moving average model for the U.S. economy, forecasts that are better than the predictions based on vector auto-regressive model. Moreover, the *novelty* of this research is given by the proposal of a new accuracy measure for multivariate forecasts like those based on VARMA model: the generalized forecast error of second moment. Usually, in literature the trace and the determinant of the mean square errors matrix is used as an accuracy indicator, but the measure proposed in this paper is better, being invariant to elementary operations with variables. The empirical study is based on quarterly data of four macroeconomic variables: (inflation rate (π_t) the GDP growth rate (g_t), the 3-month Treasury bill rate (r_t the spread between the 10 year government bond yield and the 3-month Treasury bill rate (s_t). More alternative econometric models were proposed (VARMA, VAR and AR (auto-regressive) models), but the superiority of VARMA model is demonstrated in terms of forecasts' accuracy.

1. VARMA models in literature

Vector autoregressive moving average models (VÄRMA) and the VAR ones are used in econometrics, particularly in time series analysis to reveal the cross-correlations between series, exceeding the isolated analysis of the data series.

Amid the successes of univariate ARMA models in developing forecasting, it was made the passing to VARMA models for multivariate context. The purpose of their introduction is consistent with the Granger definition of causality and it is related to the improvement of forecast accuracy by using a model with interrelated variables. For the first time these models were used by Quenouille (1957). Since then, they have been the subject of researches made by Tiao and Box (1981), Tiao and Tsay (1983, 1989), Tsay (1989), Wallis (1977), Zellner and Palm (1974). In all these cases, the number of variables was small, no more than 3 variables. Another problem was the inability to identify VARMA representations. These problems were analyzed by Hannan (1970, 1976, 1979, 1981), Dunsmuir and Hannan (1976) and Akaike (1974). Hannan and Deistler (1988) are the first that provide a theoretical presentation of VARMA models representation. Lautkepohl (1991, 2002) and Reinsel (1993) analyze the forecasts based on these models. ARIMA models, although widely used, fail to describe the dynamics of all relationships between selected variables.

The VARMA models VARMA models (vector autoregressive moving average) are the result of Wold decomposition theorem for multivariate stationary series, as shown by Athanasopoulos and Vahid (2007). Kascha and Trenkler (2011) state that there are very few studies on the performance evaluation of forecasts based on VARMA or cointegrated VARMA models. Poskitt (2003), Athanasopoulos and Vahid (2008) and Kascha and Trenkler (2011) evaluated the accuracy of forecasts made using VARMA models and they obtain a good performance, exceeding the one of VAR models.

VARMA models have been used by many researchers as Quenouille (1957), Hannan (1969), Tunnicliffe-Wilson (1973), Hillmer and Tiao (1979), Tiao and Box (1981), Tiao and Tsay (1989), Tsay (1991), Poskitt (1992), Lütkepohl (1993), Lütkepohl and Poskitt (1996), Reinsel (1997), Tiao (2001), G. Athanasopoulos and Vahid F. (2004, 2005, 2006, 2007, 2008). However, finite order VAR models are preferred to the VARMA ones, since in literature there was no question about their alternative use and the identification of VAR is easier, many software allowing the development of these models. Economic theory is not in accordance with the process modeling using VAR, the moving average terms couldn't be excluded. Cooley and Dwyer (1998) argue that macroeconomic time series modeling using VAR models is not consistent with the economic theory. But the difficulty of the VARMA methodology imposed the selection of VAR models, whose results are quite good. Likelihood function is based on the normality assumption and it can be recursively determined, the first p observations being set, and the next being zero. Starting from the state space form of the model this likelihood function can be exactly calculated. The determination of VARMA(p, q) model orders is quite difficult, given the fact that the parameters must follow certain restrictions. Kascha and Mertens (2009) realized a comparative analysis of the identification of structural form for VAR and VARMA models and for the representation in state space form.

Feunou (2009) used a VARMA model to represent the yield curve, eliminating the restrictions on co-integration. Dufour and Pelletier (2005) have proposed a modified information criterion to determine the VARMA orders, this being only a generalization of the Hannan and Rissanen (1982) criterion. Mainassara(2010) brought a change in AIC criterion used in VARMA models selection by Tsay and Hurvich (1989), resulting the AICc criterion, which is an almost unbiased estimator of the Kullback-Leibler divergence. If d data sets are analyzed, the number of parameters to be estimated is: $(p+q+3)d^2$. Choosing a too small VARMA (p, q) order implies inconsistent estimators and a too large order bring a decrease in forecast accuracy, as showed Mainassara (2010).

Procedures for specifying and estimating co-integrated VARMA models have been developed by Yap and Reinsel (1995), Lutkepohl and Claessen (1997), Poskitt (2003, 2006, 2009), all these procedures being based on the "echelon form". This form consists in a set of restrictions for parameters to ensure that the rest of the parameters are obtained using likelihood function. Kascha and Trenkler (2011) extend the representations of Dufour and Pelletier (2008) that were valid for the non-stationary series.

Kascha and Trenkler (2011) started from the last significant results related to the VARMA models, proposing a strategy for specification and estimation for co-integrated series. The authors made predictions based on these models for the U.S. interest rate.

Athanasopoulos and Vahid (2007) showed that the forecasts based on VARMA models are better than the ones based on VAR.

In literature there are several methods to identify the VARMA. Athanasopoulos and Vahid (2008a) identified two methodologies that can be applied to obtain a unique identification of VARMA. The authors made comparisons of the performance of forecasts made on VARMA models. The first methodology is an extension of Tiao's and Tsay's one (1989). The second methodology, the echelon form one, involves the estimation of the Kronecker indices, calculated as the maximum rank of each row from each equation of the model, and the specification the canonical echelon form. Kronecker indices are estimated using the least squares method applied to regressions. The innovations estimates with lag are derived from the first stage of a VAR presented by Hannan and Rissanen (1982). Kronecker indices are determined in the second stage using a model selection criterion, as shown by Hannan and Diestler (1988) and Lutkepohl and Poskitt (1996). This methodology is very simple, being used by Akaike (1974, 1976), Kailath (1980) and Kavalieris Hannan (1984), Solo (1986), Hannan and Deistler (1988), Tsay (1991), LÄutkepohl (1993), Nsiri and Roy (1992.1996), Poskitt (1992), LÄutkepohl and Poskitt (1996). Using Monte Carlo simulations, Athanasopoulos and Vahid (2006) evaluated the ability of the two methodologies to identify the VARMA models. Based on the real data, the authors compared the performance of VARMA models that used the two methodologies.

Tiao and Say (1989) proposed a first method, fairly criticized, but then it was improved by Athanasopoulos and Vahid (2006). Their methodology consists of three stages:

- Identification of the scalar components of the model (SCM) by applying canonical correlation tests between different sets of variables;
- Identification of the structural form of the model using the same tests and some certain logical deductions;
- Estimation of model using as method the full information maximum likelihood (FIML).

Tiao's and Say's (1989) methodology estimates the parameters in two stages, but in the first stage a deviation of standard errors results, this error being corrected later in the three stages version proposed of Athanasopoulos and Vahid (2006).

In 1989, George Tiao and Ruey Tsay presented their SCM methodology to the Royal Society of Statistics. The critiques related to their methodology were formulated by Chatfield, Hannan, Reinsel and Tunnicliffe-Wilson, but excluding Tsay's intervention in 1991, the methodology was developed later only by two authors. Athanasopoulos and Vahid are those who have extended the methodology, the results of its application being periodically published, especially in the last 10 years.

Tiao's and Say's methodology critiques are related to the determination of transformation matrix can be summarized as follows:

- the use of transformation matrix does not lead to the most efficient estimators for the parameters;
- the standard errors can't be calculated for estimated parameters in matrix A;

- the total number of estimated parameters in A should be included among the model parameters to reduce the number of degrees of freedom;
- the identification of VARMA (p, q) is based on transformed variables and not on the original ones.

All these problems are solved by Athanasopoulos and Vahid (2008b), which provide a formula for determining the number of redundant parameters of matrix A. They describe the procedure by which certain parameters are normalized to 1, but they are different from those that should be set to zero. The authors give up to the estimated canonical covariates by choosing an estimate of full information maximum likelihood parameters. However, they keep the way of determining the order of the K scalar components.

2. The VARMA modelling methodology based on scalar components

In order to identify the VARMA model I examined the presence of simple structures in the process. The scalar component methodology of Tiao and Say (1989) considers a K-dimensional VARMA (p, q) model ($x_t = \Phi_1 x_{t-1} + ... + \Phi_p x_{t-p} + \eta_t - \Theta_1 \eta_{t-1} - ... - \Theta_q \eta_{t-q}$), where is a non-zero linear combination: $z_t = \alpha' x_t$ follows a SCM (p, q) process if α satisfies the following properties:

$$\alpha'\Phi_{p_1} \neq 0^T, 0 \leq p_1 \leq p$$
$$\alpha'\Phi_l = 0^T, l = p_1 + 1, ..., p$$
$$\alpha'\Theta_{q_1} \neq 0^T, 0 \leq q_1 \leq q$$
$$\alpha'\Theta_{q_l} = 0^T, l = q_1 + 1, ..., q$$

Scalar random variable admits an ARMA representation with orders that vary from p1 to q1, depending on lags from 1 to p1 and its innovations depend on lags from 1 to q1. The identification starts with the SCM (0,0), which is actually a white noise. The basic idea is to find out K linearly independent vectors K that achieve a rotation operation of VARMA(p, q) process in a new process with a dynamic structure, but with fewer parameters. The linearly independent vectors form a matrix $A = (\alpha_1, ..., \alpha_k)'$. Then, $z_t = Ax_t$.

The VARMA process with transformed variables keeps all the rows of zero restrictions from AR component, the parameters matrix from MA having the form:

$$z_t = \Phi_1^* z_{t-1} + ... + \Phi_p^* z_{t-p} + \varepsilon_t - \Theta_1^* \varepsilon_{t-1} - ... - \Theta_q^* \varepsilon_{t-q} \ (1),$$

where $\Phi_i^* = A\Phi_i A^{-1}, \varepsilon_t = A\eta_t, \Theta_i^* = A\Theta_i A^{-1}$.

If we have identified two scalar components $z_{r,t} = SCM(p_r, q_r)$ *and* $z_{s,t} = SCM(p_s, q_s)$, *where* $p_r > p_s, q_r > q_s$, the lags of $z_{s,t}$ from the right part of the dynamic equation for $z_{r,t}$ can be expressed in terms of variables from right side of $z_{r,t}$ .can take values between 1 and the minimum of $\{p_r - p_s, q_r - q_s\}$. The parameters on the right side of the dynamic equation for $z_{r,t}$ can be determined only if the maximum lag order is set to zero.

Using canonical correlation tests, the form of models with scalar components embedded is identified. The SCM (0,0) combination is a linear one, the canonical correlations between past and present being identified by a simple generalization made by Hotelling

(1935) for time series. The squares of these canonical correlations will be noted with $\hat{\lambda}_1 < \hat{\lambda}_2 < ... < \hat{\lambda}_K$. The likelihood ratio test is applied when the null hypothesis is that there are at least s scalar components and the alternative one refers to the existence of less than s unpredictable components. The test statistic is: $C(s) = -(n-h)\sum_{k=1}^{s}\ln(1-\hat{\lambda}_k) \sim \chi^2_{s\cdot[(h-1)K+s]}$.

Consistent estimates of scalar components are given by the canonical co-variances corresponding to the insignificant canonical correlations. Generalized method of moments based on the test with the same hypothesis as the above one has the statistics, as shown in Anderson and Vahid (1998): $(n-h)\sum_{k=1}^{s}\hat{\lambda}_k$.

Using the squared canonical correlations between $x_{p,t} \equiv (x'_t,...,x'_{t-p})'$ and $x_{h,t-1} \equiv (x'_{t-1},...,x'_{t-1-h})', h \geq p$ and a similar test a SCM(p,0) is determined.

SCM (p, j) are linear combinations of $x_{p,t}$, for which linear predictions can't be made in history before t-j moment. By a structure of weighted matrix obtained applying the generalized method of moments is determined a linear combination, which is a moving average of order j. In this context a test of over-identifying the restrictions is applied. Tiao and Say (1989) proposed a statistic: $C(s) = -(n-h-j)\sum_{k=1}^{s}\ln(1-\frac{\hat{\lambda}_k}{d_k}) \sim \chi^2_{s\cdot[(h-p)K+s]}$. d_k is a correction factor that arises because the canonical variations may be moving average processes of order j. Thus, $d_k = 1 + 2\sum_{v=1}^{j}\hat{\rho}_v(\hat{r}'_k x_{p,t})\hat{\rho}_v(\hat{g}'_k x_{h,t-1-j})$, $\hat{\rho}_v(.)$ being the autocorrelation of order v corresponding to the argument and the terms in brackets are the canonical variances of the k canonical correlation. Higher value orders are identified below by testing the orders. Tiao and Say (1989) pointed out the results of the tests in a table that represents the rules for identifying the orders of SCMs. The two authors have obtained a consistent estimator of the transformation matrix A, starting from the estimated canonical coefficients. They identify the appropriate null eigenvectors the applying the statistical tests.

Athanasopoulos and Vahid (2008a) present the following rules to ensure unique identification of the system:

- The model structure does not change if each row of the matrix A is multiplied by a constant. This allows the normalization of the parameters on each row by one. Using tests of predictability for subsets of variables, we verify that a parameter set to zero is not normalized by one.
- Any linear combination of SCM(p_1,q_1) and SCM(p_2,q_2) is a SCM(max(p_1,p_2),(q_1,q_2)). When there are only two scalar components in the SCM(p_1,q_1), random multiples could be included without changing the structure. Because in this case the line of matrix A corresponding to SCM(p_1,q_1) is not identified, the parameter in column k corresponding to the line from A is normalized by 1. The parameter from line k that corresponds to SCM(p_2,q_2) is restricted to zero.
- If $p_1 = p_2$ and $q_1 = q_2$ sub-matrix identity is formed and the previous rule is applied twice. If there is only one SCM with an AR / MA of minimal, the corresponded row from A is uniquely identified.

In the original methodology an estimator for A ($\hat{A}$) was obtained and $z_t = \hat{A}\cdot x_t$ was determined, then $\Phi^*_1,...,\Phi^*_p$ *and* $\Theta^*_1,...,\Theta^*_q$ which have many null restrictions. The improved methodology of Athanasopoulos and Vahid (2008a) rewrite the original system variables and by identification of A restrictions obtain estimates using as method the full information maximum likelihood.

Setting at zero the MA coefficients is equivalent to replacing the MA process variables on the right side of the equation ($z_{t-1},...,z_{t-p}$) with variables $x_{t-1},...,x_{t-p}$, maintaining the system structure. Taking into account the replacement of $z_{t-1},...,z_{t-p}$ with $A\cdot x_{t-1},...,A\cdot x_{t-p}$ and the obtain of the system $z_t = \psi_1 z_{t-1} + ... + \psi_p z_{t-p} + \varepsilon_t - \Theta_1^* \varepsilon_{t-1} - ... - \Theta_q^* \varepsilon_{t-q}$ (2), Athanasopoulos and Vahid (2008b) shows that $\psi_1,...,\psi_p$ have the same zero restrictions as $\Phi_1^*,...,\Phi_p^*$. This lemma leads to: $A\cdot x_t = \psi_1 x_{t-1} + ... + \psi_p x_{t-p} + \varepsilon_t - \Theta_1^* \varepsilon_{t-1} - ... - \Theta_q^* \varepsilon_{t-q}$ (3), in which the parameters satisfy the same restrictions as matrix parameters from the right part of the equation (1). Since not all matrix parameters are free, the system is still unidentified. This situation restricts the matrix A to have a uniquely determined system.

The matrix A is identified if and only if the single matrix H so that $HAx_t = H\psi_1 x_{t-1} + ... + H\psi_p x_{t-p} + H\varepsilon_t - H\Theta_1^* \varepsilon_{t-1} - ... - H\Theta_q^* \varepsilon_{t-q}$ has the same restrictions as (3) and it is the unit matrix of order K.

It is also assumed that the k row from the system is SCM(p_i, q_i). Null restrictions on the right side of the system show that the row k of matrix H may differ from that of an identity matrix, if there are other SCM (p_i, q_i) models. The row of rank k from matrix is transformed into a row of identity matrix by normalization to 1 of an item in this row.

A canonical representation SCM VARMA has the following characteristics:

i. The orders for SCM are as small as possible;
ii. In order to obtain a unique identification, the all redundant parameters from transformed matrix A are restricted;
iii. Zero restriction used to determine the number of redundant parameters is set for the corresponding coefficients of MA process.

3. An empirical example of VARMA modeling methodology

For four macroeconomic variables (inflation rate (π_t) the GDP growth rate (g_t) the 3-month Treasury bill rate (r_t the spread between the 10 year government bond yield and the 3-month Treasury bill rate (s_t) we have formed series of quarterly data from the U.S. economy. These variables are used in recent researches in models with observable factors or in the new Keynesian DSGE ones. Federal Reserve Economics Database (FRED) was used to create the data series. Quarterly interest rate data were obtained from monthly data as average. Inflation rate used in VARMA model is different from the one published by FRED, being calculated by multiplying by 400 the difference between the logarithm of consumer price index in the last month of the current quarter and the last month of the previous one. Levin (1999) recommends the first order differentiation of interest rate in monetary policy models. His indication is argued by the fact that it can develop rules for monetary policy less affected by model uncertainty. One of the reasons for which variables as interest rate in first difference are used in this study and the GDP growth rate is related to the need of applying canonical correlation tests with a chi-square asymptotic distribution for stationary series.

I used quarterly data for the period first quarter 1955-fourth quarter 2000 to build VAR and VARMA models and make predictions based on these models for the horizon: first quarter of 2001 – second quarter of 2013. The identified models are: VARMA (2,1), VAR with 3 lags when the selection criteria is AIC and VAR model with two lags when the selection criteria is BIC, univariate AR (AR(1) for GDP growth rate, AR(3) for inflation rate and AR(2) for the other variables.

The procedure applied by Athanasopoulos and Vahid (2007) to ensure that the element that must be set to zero is not normalized to one starts with a SCM of minimum order. One of the variables is excluded and the test of predictability is applied for the rest of the variables. If the test is rejected, then the eliminated variable coefficient is set to 1 and the rest of the coefficients are zero. If the test is accepted (an SCM is formed with the rest of variables) the coefficient corresponding to the eliminated variable is set to zero and the test continues after the elimination of another variable. Tests applied in this case are GMM tests, which are tests of generalized method of moments proposed by Hansen (1982).

1. The identification of scalar components

Tiao's and Say's (1989) methodology for this stage consists in two steps, at which Athanasopoulos and Vahid (2008b) add a rule of elimination.

The two steps are:

a. The determination of the overall order

All null canonical correlations between $x_{m,t}$ and $x_{m,t-1-j}$ are determined, beginning with m = 0 and j = 0. A table with two parts is built.

Determine all canonical correlations between the null şi and , începând cu m=0 şi j=0. Since m = 0 and j = 0. It is a table composed of two parts. We start from the top left corner and we look the first occurrence of zero eigenvalues s + K, where s is the number of null eigenvalues in position (p-1, q-1) of the table. It is considered that (p, q) is the general order of the system. In case we identify several orders of this form, we will select only one using an information criterion.

Table 1. Criterion table

	j				
m	0	1	2	3	4
0	140,11	90	3,47	2,57	2,55
1	3,2	0,9	0,98	0,94	0,94
2	1,23	1,05	1,23	0,9	1,03
3	1,05	1,08	0,92	1,05	1
4	0,89	0,92	1,93	0,9	0,98

Table 2. Root table

	j				
m	0	1	2	3	4
0	1	1	1	1	1
1	1	4	4	4	4
2	1	5	9	9	9
3	3	9	11	15	16
4	4	10	17	19	24

The VARMA model is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -0{,}82 & -0{,}03 & 0{,}15 & 1 \end{pmatrix} \cdot \begin{pmatrix} g_t \\ \pi_t \\ s_t \\ \Delta r_t \end{pmatrix} = \begin{pmatrix} 0{,}72 \\ 0{,}5 \\ 0{,}23 \\ -0{,}64 \end{pmatrix} + \begin{pmatrix} 0{,}13 & -0{,}05 & 0{,}41 & 0{,}55 \\ -0{,}38 & 0{,}89 & 0{,}3 & 0{,}24 \\ -0{,}04 & 0{,}04 & 1 & -0{,}11 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} g_{t-1} \\ \pi_{t-1} \\ s_{t-1} \\ \Delta r_{t-1} \end{pmatrix} +$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -0{,}02 & -0{,}07 & -0{,}15 & 0{,}05 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} g_{t-2} \\ \pi_{t-2} \\ s_{t-2} \\ \Delta r_{t-2} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0{,}62 & -0{,}57 & -1{,}23 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} e_{1,t-1} \\ e_{2,t-1} \\ e_{3,t-1} \\ e_{4,t-1} \end{pmatrix} + \begin{pmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \\ e_{4,t} \end{pmatrix}$$

The determination of scalars components orders

We tested the null canonical correlations between $x_{m,t}$ and $x_{m+(q-j),t-1-j}$, where m = 0,..., p and j = 0 ,..., q. The SCM (m, j) includes all the scalar components of smaller order, in (m, j) position there are s scalar components will be and scaling components, where s = min {m-p1 +1, j-1} for each SCM (p1, q1).

2. The place of restrictions of identification

Identification rules are applied to determine the structure of the matrix A.

The parameters estimation method is the full information maximum likelihood presented by Durbin (1963). This method provides estimates for both parameters, as well as standard errors of parameters, including the free ones.

The trace and the determinant of the mean square errors matrix are classical measures of forecast accuracy, used by Athanasopoulos and Vahid (2007). We used the generalized forecast error second moment as measure of accuracy. This is calculated according to Clements and Hendry (1993) as a determinant of the expected value of the vector forecast errors for future times on the horizon of interest. If we study a number until h quarters, this indicator is calculated as:

$$GFESM = \left| E \left| \begin{pmatrix} e_{t+1} \\ e_{t+2} \\ \dots \\ e_{t+h} \end{pmatrix} \right| \cdot \left| \begin{pmatrix} e_{t+1} \\ e_{t+2} \\ \dots \\ e_{t+h} \end{pmatrix}^T \right| \right|$$

e_{t+h} –dimensional forecast error of the model of n variables on the forecasting horizon h.

It is considered that GFESM is a better measure of accuracy because it is invariant to elementary operations with variables, unlike the trace of MSFE, and also it is a measure invariant to elementary operations for the same variables on different forecasting horizons, unlike the trace or the determinant of MSFE.

We propose as a measure to compare the accuracy of forecasts based on these models a new indicator: the ratio GFESM relative to the VARMA model.

We calculated the ratio GFESM relative to the VARMA model one quarter ahead for the two models (h = 1). For multivariate models we also calculated GFESM separately.

Table 3. Ratio of the GFESM relative to the VARMA model for one step ahead forecasts

Type of model	GDP growth rate	Inflation	s_t	Δr_t	
VAR (Akaike information criterion) model	0.93	1.01	1.22**	1.59*	1.69
VAR (Schwartz information criterion) model	0.91	1.06	1.24*	2.06*	2.11
AR (auto-regressive) model	0.82	1.03	1.09	1.22*	-
Naïve model	0.81	1.35	6.82*	1.45*	-

** For a significance level of 5% the ratio differs significantly from 1
** For a significance level of 10% the ratio differs significantly from 1

Analyzing the table above, the VARMA model provides forecasts with a higher degree of accuracy than VAR models for variables s_t and Δr_t. Knowing that these variables are not affected by structural shocks it is likely that forecasts based on VARMA models are better than those based on VAR models.

Conclusions

Scalar components methodology used in building VARMA models is quite difficult to apply in practice, but on small time horizons, the forecasts based on these models are better than others for variables unaffected by structural shocks. This conclusion has been reached by other researchers, Athanasopoulos and Vahid, but indicators used to measure the accuracy were the classical ones: the trace and the determinant of MSFE. In this study, the accuracy is evaluated using as indicator the generalized forecast error second moment. We introduced a new measure for evaluating the relative accuracy in order to make comparisons between forecasts: the ratio of the GFESMs relative to the VARMA model.

In a future research it would be interesting to estimate the VARMA model using the state space form for which the methods based on Kalman filter are used for optimization.

References

Athanasopoulos, G., Poskitt, D., S. and Vahid, F. (2007), Two canonical VARMA forms: Scalar component models vis-´a-vis the Echelon form, *Working paper* 10/07, Department of Econometrics and Business Statistics, Monash University.

Athanasopoulos, G. and Vahid, F. (2008a), A complete VARMA modelling methodology based on Scalar Components, *Journal of Time Series Analysis* 29(3), pp. 533–554.

Athanasopoulos, G. and Vahid, F. (2008b), VARMA versus VAR for macroeconomic forecasting, *Journal of Business and Economic Statistics* 26(2), pp. 237–252.

Clements, M., P., Hendry, D., F. (1993), On the limitations of comparing mean squared forecast errors (with discussions), *Journal of Forecasting* 12, pp. 617-637.

Kantsukov, M., Linnas, A. (2013), Risk propensity of corporate financial executives: the comparison of 3 countries, *Actual Problems of Economics*, 148(10), pp. 310-318.

Mainassara, Y., B. (2010), Selection of weak VARMA models by Akaike's information criteria, *MPRA Paper*, No. 23412, http://mpra.ub.uni-muenchen.de/23412/

Tiao, G., Tsay, R. (1989), Model specification in multivariate time series (with discussions), *Journal of the Royal Statistical Society* B 51, pp. 157-213.

http://research.stlouisfed.org/fred2/ (Federal Reserve Economic Database).