Tomáš Bacigál

Multivariate smooth transition AR model with aggregation operators and application to exchange rates

Kybernetika, Vol. 43 (2007), No. 2, 245--254

Persistent URL: http://dml.cz/dmlcz/135770

KYBERNETIKA — VOLUME 43 (2007), NUMBER 2, PAGES 245 – 254

MULTIVARIATE SMOOTH TRANSITION AR MODEL WITH AGGREGATION OPERATORS AND APPLICATION TO EXCHANGE RATES

Tomáš Bacigál

An overview of multivariate modelling based on logistic and exponential smooth transition models with transition variable generated by aggregation operators and orders of auto and exogenous regression selected by information criterion separately for each regime is given. Model specification procedure is demonstrated on trivariate exchange rates time series. The application results show satisfactory improvement in fit when particular aggregation operators are used. Source code in the form of Mathematica package is provided.

Keywords: multivariate STAR, aggregation operator, information criterion, exchange rates

AMS Subject Classification: 37M10

1. INTRODUCTION

Smooth transition autoregression (STAR) as a part of regime-switching family of non-linear models uses piecewise linear structure which is easy to interpret as well as to forecast from. Transition between regimes is controlled by comparing realizations of the so-called transition variable to a threshold. This threshold and the parameter of transition function that describes how smooth/abrupt the change happens, are subject to estimation together with parameters of linear VAR model in each regime. The 2-regime multivariate STAR model can be written in matrix notation as

$$\boldsymbol{y}_t = \boldsymbol{\Phi}_1 \boldsymbol{X}_t (1 - G(z_t, \gamma, r)) + \boldsymbol{\Phi}_2 \boldsymbol{X}_t G(z_t, \gamma, r) + \boldsymbol{\varepsilon}_t, \tag{1}$$

which contains k-dimensional vector of modelled variables $\boldsymbol{y}_t$, $k \times K$ parameter matrix $\boldsymbol{\Phi}_j$ corresponding to jth regime, $K \times 1$ regressor $\boldsymbol{X}_t$, white noise series $\boldsymbol{\varepsilon}_t$ and transition function G with transition variable z_t, smooth parameter γ and threshold r as its parameters. Dimension K is number of all (current and/or lagged) variables included in regression. In the following, we use two different transition functions. The first-order logistic function

$$G(z_t, \gamma, r) = \frac{1}{1 + e^{-\gamma(z_t - r)}}, \qquad \gamma > 0, \tag{2}$$

causes the swap of regimes permanent as transition variable passes the threshold, whereas exponential function

$$G(z_t, \gamma, r) = 1 - e^{-\gamma(z_t - r)^2}, \qquad \gamma > 0, \tag{3}$$

allows the second regime to arise only for z_t close to threshold.

2. AGGREGATION OPERATOR

Besides the optionality in transition function, a considerable freedom is left with the transition variable as well. Whilst usually just a lagged value of one of the modelled variables is employed, $y_{i,t-d}$ for $i \in \{1, \dots k\}$ and some integer $d > 0$, we use here a more general way of generating z_t, in which last d values of some observed time series $\{\xi_i, i = 1, \dots t\}$ is taken and so called aggregation operator $\mathcal{A}$ is applied to get one-dimensional quantity, such as

$$z_t = \mathcal{A}(\xi_{t-1}, \dots, \xi_{t-d}). \tag{4}$$

Typical continuous aggregation operator (see [1]) on the real line ($R^d \to R$) is weighted mean. The diversity is achieved by setting weights $w_i \in [0, 1]$, $i \leq d$ (with $\sum_{i=1}^{d} w_i = 1$), for instance as constant ($w_i = 1/d$, arithmetic mean $\mathcal{M}$) or decreasing sequence based on convex bijection φ of $[0, 1]$, where the weights $w_i = \varphi(\frac{d-i+1}{d}) - \varphi(\frac{d-i}{d})$ for $i = 1, \dots d$, and we denote the aggregation operator as $\mathcal{W}_2$ if $\varphi(x) = x^2$ and $\mathcal{W}_3$ if $\varphi(x) = x^3$. In the experiment bellow there is used also weighted mean with Sierpinski carpet ($w_i = p(1-p)^{i-1}$ for $i < d$, $w_d = (1-p)^{d-1}$, with probability $p \in [0, 1]$) and Fibonacci weights ($w_i = \frac{f_{d-i+1}}{f_{d+2}-1}$, $i = 1, \dots d$, where f_i is the Fibonacci number satisfying the recurrence relation $f_i = f_{i-1} + f_{i-2}$ with $f_1 = f_2 = 1$).

Other type of weighted mean is achieved by applying weights to input sequence ordered by increasing values instead of time. It is called ordered weighted average (OWA) operator and covers also the maximum ($\mathcal{MAX}$) and minimum ($\mathcal{MIN}$).

3. TESTING FOR LINEARITY

There are two tests that should precede STAR model specification procedure. Both are simple and perform well in finite samples, yet they do not encounter the problem of unidentified nuisance parameters under null hypothesis.

The first one, proposed by Tsay in [7], put linearity against threshold non-linearity (i. e. abrupt transition between regimes) using the linear regression $\boldsymbol{y}_t = \boldsymbol{\Phi}\boldsymbol{X}_t + \boldsymbol{\varepsilon}_t$ rearranged according to the increasing order of transition variable z_t so that $\boldsymbol{y}_{t(i)} = \boldsymbol{\Phi}\boldsymbol{X}_{t(i)} + \boldsymbol{\varepsilon}_{t(i)}$. Here $t(i)$ is the time index of $z_{(i)}$ which denotes the ith smallest element of z_t for $i = 1, \dots n$, shortly $z_{t(i)} = z_{(i)}$. Thus the arranged regression transforms a threshold model into changepoint problem. Detection of such changepoint can be done via predictive residuals $\boldsymbol{\varepsilon}_{t(i+1)}$. If $\boldsymbol{y}_t$ is linear, the recursive least square

estimator of the arranged regression is consistent so that the predictive residuals approach white noise and, obviously, they are uncorrelated with regressor $\boldsymbol{X}_{t(i+1)}$.

Let $\hat{\boldsymbol{\Phi}}_{\tilde{n}}$ be a least squares estimate of $\boldsymbol{\Phi}$ from data points associated with $\tilde{n}$ smallest values of z_t and

$$\hat{\boldsymbol{\eta}}_{t(\tilde{n}+1)} = \frac{\boldsymbol{y}_{t(\tilde{n}+1)} - \hat{\boldsymbol{\Phi}}_{\tilde{n}} \boldsymbol{X}_{t(\tilde{n}+1)}}{\left[1 + \{\boldsymbol{X}'_{t(\tilde{n}+1)} \boldsymbol{V}_{\tilde{n}} \boldsymbol{X}_{t(\tilde{n}+1)}\right]^{1/2}}$$

be the standardized residuals of the 1-step-ahead prediction in the rearranged regression with $\boldsymbol{V}_{\tilde{n}} = \left[\sum_{i=1}^{\tilde{n}} \boldsymbol{X}_{t(i)} \boldsymbol{X}'_{t(i)}\right]^{-1}$. Then from the regression $\hat{\boldsymbol{\eta}}_{t(j)} = \boldsymbol{\Psi} \boldsymbol{X}_{t(j)} + \boldsymbol{w}_{t(j)}$, $j = \tilde{n}_0 + 1, \ldots n$, where $\tilde{n}_0$ denotes starting point of the recursive estimation (usually $\tilde{n}_0 \approx 3\sqrt{n}$), we test hypothesis $H_0 : \ \boldsymbol{\Psi} = \boldsymbol{0}$ versus $H_1 : \ \boldsymbol{\Psi} \neq \boldsymbol{0}$ with the test-statistic

$$C = [n - \tilde{n}_0 - K] \times (\ln |\boldsymbol{S}_0| - \ln |\boldsymbol{S}_1|), \tag{5}$$

where

$$\boldsymbol{S}_0 = \frac{1}{n - \tilde{n}_0} \sum_{j=\tilde{n}_0+1}^{n} \hat{\boldsymbol{\eta}}_{t(j)} \hat{\boldsymbol{\eta}}'_{t(j)}, \qquad \boldsymbol{S}_1 = \frac{1}{n - \tilde{n}_0} \sum_{j=\tilde{n}_0+1}^{n} \hat{\boldsymbol{w}}_{t(j)} \hat{\boldsymbol{w}}'_{t(j)}.$$

Under the null that $\boldsymbol{y}_t$ is linear, C is asymptotically a χ^2 random variable with kK degrees of freedom.

The second test, originally developed in [6] for univariate case, assumes STAR model framework in general and expresses the null hypothesis as equality $\Phi_1 = \Phi_2$ or $\gamma = 0$ against the alternative hypothesis of inequality. The problem with the presence of unidentified nuisance parameters under the null is resolved via approximating the transition function $G(z_t, \gamma, r)$ by suitable Taylor series around $\gamma = 0$. In case of LSTAR the 3rd-order Taylor approximation appears suitable (cf. [6]), whilst 2nd order is recommended for ESTAR (see [3]). This replacement yields the auxiliary regression

$$y_t = \beta_{0,0} + \boldsymbol{\beta}'_0 \boldsymbol{X}_t + \boldsymbol{\beta}'_1 \boldsymbol{X}_t z_t + \boldsymbol{\beta}'_2 \boldsymbol{X}_t z_t^2 + \boldsymbol{\beta}'_3 \boldsymbol{X}_t z_t^3 + e_t, \tag{6}$$

for logistic transition and (6) with extra term $\boldsymbol{\beta}'_4 \boldsymbol{X}_t z_t^4$ for exponential transition function, where $\boldsymbol{\beta}_i = (\beta_{i,0}, \beta_{i,1}, \ldots \beta_{i,K-1})'$, $i = 0, 1, 2, 3, 4$, are functions of the parameters $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$, γ and r. The null hypothesis $\gamma = 0$ now corresponds to $H'_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \boldsymbol{\beta}_3 (= \boldsymbol{\beta}_4) = \boldsymbol{0}$ (also $\boldsymbol{e}_t = \boldsymbol{\varepsilon}_t$) and can be tested by a standard LM-type test. Note that if z_t is one of the variables included in $\boldsymbol{X}_t$, the terms $\beta_{i,0} z_t^i$, $i = 1, 2, 3, 4$, should be dropped from the auxiliary regression to avoid perfect multicollinearity.

Next, the test statistic is computed as

$$LM = n(\ln |\boldsymbol{\Sigma}_0| - \ln |\boldsymbol{\Sigma}_1|) \tag{7}$$

where

$$\boldsymbol{\Sigma}_0 = n^{-1} \sum_t \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}'_t, \qquad \boldsymbol{\Sigma}_1 = n^{-1} \sum_t \hat{\boldsymbol{e}}_t \hat{\boldsymbol{e}}'_t,$$

with $\hat{\boldsymbol{\varepsilon}}$ being residuals estimated from null hypothesis regression $\boldsymbol{y}_t = \boldsymbol{\Phi}\boldsymbol{X}_t + \boldsymbol{\varepsilon}_t$ and $\hat{\boldsymbol{e}}$ from regressing $\hat{\boldsymbol{\varepsilon}}$ on $\boldsymbol{X}_t$ and $\boldsymbol{X}_t z_t^i$, $i = 0, 1, 2, 3, (4)$.

Under the null hypothesis of linearity, LM has an asymptotic χ^2 distribution with $3kK$ degrees of freedom in case of logistic transition and $4kK$ if exponential STAR is considered as alternative model.

4. MODEL SPECIFICATION STRATEGY

Testing for linearity, where the linear model under null hypothesis is usually selected in phase of preliminary specification by information criteria, can be used to choose correct transition variable. It is the choice of variable ξ_t and delay d in (4) that, if correct, minimize the p-value of Tsay or LM test. Model specification then continues by joined procedure of parameters estimation and selection of appropriate order in both autoregression and regression on exogenous variables (that means the maximum lag used in regressor $\boldsymbol{X}_t$). Then we do diagnostic tests on residuals and respecification if necessary. Finally, the models that perform best in the in-sample fit are examined for their out-of-sample performance by comparison of forecast errors.

4.1. Model selection and estimation

Orders of endogenous and exogenous regression selected for linear model and used in linearity tests can easily be inappropriate for STAR model building. Sometimes it is desirable to distinguish between regimes in terms of lags, and if so, it is fair to penalize the inclusion of the additional parameters in $\boldsymbol{\Phi}_1$, $\boldsymbol{\Phi}_2$ (associated with additional lags) not for the whole sample size but only for the number of regime-corresponding observations n_j. In [7] such information criteria are designed for threshold VAR model (TAR), here we provide generalization to suit m-regimes smooth transition VAR (STAR) model.

Let us define regressor in jth regime,

$$\boldsymbol{X}_{j,t} = (1, \boldsymbol{y}'_{t-1}, \dots \boldsymbol{y}'_{t-p_j}, \boldsymbol{x}'_{t-1}, \dots \boldsymbol{x}'_{t-q_j})', \qquad j = 1, 2, \dots m, \tag{8}$$

where $\boldsymbol{x}_t$ stands for l-dimensional exogenous variable, p_j is the order of endogenous and q_j of exogenous regression, then $K_j = kp_j + lq_j + 1$ is dimension of regressor $\boldsymbol{X}_{j,t}$. As for the number of observations, n_j is not necessarily an integer, but rather a weight ($\sum_{j=1}^{m} n_j = n$) defined as $n_j = \sum_t \Delta G_{j,t}$ with $\Delta G_{j,t} = G_{j-1,t} - G_{j,t}$, where $G_{j,t} = G_j(z_t, \gamma_j, r_j)$ is the transition function corresponding to jth regime, $G_{0,t} = 1$ and $G_{m,t} = 0$. Then BIC information criterion based on multivariate m-regimes STAR model

$$\boldsymbol{y}_t = \sum_{j=1}^{m} \boldsymbol{\Phi}_j \boldsymbol{X}_{j,t} \Delta G_{j,t} + \boldsymbol{\varepsilon}_t \tag{9}$$

can be expressed as

$$BIC(\boldsymbol{p}, \boldsymbol{q}) = \sum_{j=1}^{m} \left(n_j \ln |\hat{\boldsymbol{\Sigma}}_j| + \ln(n_j) k K_j \right), \tag{10}$$

where $\boldsymbol{p} = (p_1, \dots p_m)$, $\boldsymbol{q} = (q_1, \dots q_m)$, and

$$\hat{\boldsymbol{\Sigma}}_j = \frac{1}{n_j} \sum_t (\boldsymbol{y}_t - \hat{\boldsymbol{\Phi}}_j \boldsymbol{X}_{j,t})(\boldsymbol{y}_t - \hat{\boldsymbol{\Phi}}_j \boldsymbol{X}_{j,t})' \Delta G_{j,t}$$

is estimated covariance matrix. Other information criteria, such as AIC or HQIC, are computed in a similar way.

Estimation of the model parameters $\boldsymbol{\theta} = (\boldsymbol{\Phi}, \boldsymbol{\gamma}, \boldsymbol{r})$, where $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_1, \dots \boldsymbol{\Phi}_m)$, $\boldsymbol{\gamma} = (\gamma_1, \dots \gamma_{m-1})$, $\boldsymbol{r} = (r_1, \dots r_{m-1})$ is the problem of minimizing the trace of $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \sum_t (\boldsymbol{y}_t - \boldsymbol{\Phi} \boldsymbol{X}_t)(\boldsymbol{y}_t - \boldsymbol{\Phi} \boldsymbol{X}_t)'$, where $\boldsymbol{X}_t(\boldsymbol{\gamma}, \boldsymbol{r}) = (\boldsymbol{X}'_{1,t} \Delta G_1, \dots \boldsymbol{X}'_{m,t} \Delta G_m)'$ and ΔG_j again is function of (γ_j, r_j). This minimizing can be performed directly by nonlinear least squares (NLS) routine

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \operatorname{Tr} \left(\boldsymbol{\Sigma}(\boldsymbol{\Phi}, \boldsymbol{\gamma}, \boldsymbol{r}) \right),$$

for which several iterative optimization algorithms are available in statistical software. Alternatively, for fixed values of $\boldsymbol{\gamma}$ and $\boldsymbol{r}$ the model is linear in the parameters $\boldsymbol{\Phi}$, so that these can be (conditionally upon $\boldsymbol{\gamma}, \boldsymbol{r}$) estimated by ordinary least squares (OLS) as

$$\hat{\boldsymbol{\Phi}}(\boldsymbol{\gamma}, \boldsymbol{r}) = \left(\sum_t \boldsymbol{X}_t(\boldsymbol{\gamma}, \boldsymbol{r}) \boldsymbol{X}_t(\boldsymbol{\gamma}, \boldsymbol{r})' \right)^{-1} \left(\sum_t \boldsymbol{X}_t(\boldsymbol{\gamma}, \boldsymbol{r}) \boldsymbol{y}'_t \right) \tag{11}$$

and

$$(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{r}}) = \underset{(\boldsymbol{\gamma}, \boldsymbol{r})}{\operatorname{argmin}} \operatorname{Tr} \left(\boldsymbol{\Sigma}(\hat{\boldsymbol{\Phi}}(\boldsymbol{\gamma}, \boldsymbol{r}), \boldsymbol{\gamma}, \boldsymbol{r}) \right). \tag{12}$$

As the NLS need not always result in global minimum immediately, the conditional OLS grid search can help to define starting values for NLS. However, there is still a notorious problem with parameter $\boldsymbol{\gamma}$ that converges too slowly so that its estimate is rather imprecise (thus may appear insignificant) unless a large amount of observations (z_t) is available in the neighbourhood of the threshold $\boldsymbol{r}$. Especially when $\boldsymbol{\gamma}$ is large, rescaling it becomes important (see [4], p. 123). Also, for ensuring reliable estimates of $\boldsymbol{\Phi}$, each regime should contain at least about 15 % of observations, which limits the choice of $\boldsymbol{r}$.

4.2. Diagnostic on residuals

One of the basic diagnostics to detect model misspecification is the multivariate Ljung–Box (also known as portmanteau) test for residual autocorrelations. Let $\{\hat{\boldsymbol{\varepsilon}}_t\}$, $t = 1, \dots n$, be the estimated residuals from the best-fitting model and $\hat{\boldsymbol{\Gamma}}_i = n^{-1} \sum_t (\hat{\boldsymbol{\varepsilon}}_t - \hat{\bar{\boldsymbol{\mu}}})(\hat{\boldsymbol{\varepsilon}}_t - \hat{\bar{\boldsymbol{\mu}}})'$, $i = 0, 1, \dots D$ be autocovariance matrix where $\hat{\bar{\boldsymbol{\mu}}} = n^{-1} \sum_t \hat{\boldsymbol{\varepsilon}}_t$ and D is some maximum lag. Then under null hypothesis of no serial dependence in residuals the test statistic

$$Q = D^2 \sum_{i=1}^{D} \frac{1}{n-i} \operatorname{Tr}(\hat{\boldsymbol{\Gamma}}'_i \hat{\boldsymbol{\Gamma}}_0^{-1} \hat{\boldsymbol{\Gamma}}_i \hat{\boldsymbol{\Gamma}}_0^{-1})$$

has asymptotic χ^2 distribution with $k^2 D$ degrees of freedom.

Another diagnostics include tests for normality, parameter constancy and residual linearity (see [2]).

4.3. Forecasts comparison

Consider $\boldsymbol{y}_t$ being described by a general STAR model

$$\boldsymbol{y}_t = F(\boldsymbol{X}_t, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}_t$$

for some nonlinear function F. The optimal h-step-ahead forecast of $\boldsymbol{y}_{t+h}$ at time t is given by $\hat{\boldsymbol{y}}_{t+h|t} = \mathrm{E}[y_{t+h}|\Omega_t]$, where Ω_t denotes history of the time series up to and including the observation at time t. One possible approach to approximate the conditional expectation is using Monte Carlo method which gives

$$\hat{\boldsymbol{y}}_{t+h|t} = \frac{1}{s} \sum_{i=1}^{s} F(\hat{\boldsymbol{y}}_{t+h-1|t} + \boldsymbol{\varepsilon}_i, \boldsymbol{\theta}), \tag{13}$$

where s is some sufficiently large number and the $\boldsymbol{\varepsilon}_i$ are drawn from the presumed distribution of $\boldsymbol{\varepsilon}_{t+h-1}$.

The forecast (prediction) error is defined as $\boldsymbol{e}_{t+h|t} = \boldsymbol{y}_{t+h} - \hat{\boldsymbol{y}}_{t+h|t}$.

Given P additional observations $\{\boldsymbol{y}_{n+i}\}$, $i = 1, \dots P$, the out-of-sample performance of two rival models, say A and B, can effectively be compared either directly by their mean squared prediction error, $MSPE = P^{-1} \sum_{i=1}^{P} \boldsymbol{e}'_{n+i|n+i-h} \boldsymbol{e}_{n+i|n+i-h}$, or by the modified Diebold–Mariano test.

Let $d_i = g(\boldsymbol{e}_{n+i|n+i-h,A}) - g(\boldsymbol{e}_{n+i|n+i-h,B})$, $j = 1, \dots P$, be a loss differential with a loss function g (e. g., $g(\boldsymbol{a}) = \boldsymbol{a}'\boldsymbol{a}$), and $\boldsymbol{e}_{\_,M}$ denotes forecast error generated from a model M. Null hypothesis that there is no qualitative difference between the forecasts from model A and B is expressed as $H_0 : \bar{d} = 0$, where $\bar{d} = P^{-1} \sum_{i=1}^{P} d_i$ is average loss differential. The modified Diebold–Mariano test statistic (see [5]) is given by

$$MDM = \sqrt{P + 1 - 2h + h(h-1)/P}\, \frac{\bar{d}}{\sqrt{\omega}}, \tag{14}$$

where ω is proportional to asymptotic variance of $\bar{d}$ and suggested to be estimated as

$$\hat{\omega} = \hat{\gamma}_0 + 2 \sum_{i=1}^{h-1} \hat{\gamma}_i \qquad \text{with} \quad \hat{\gamma}_i = \frac{1}{P} \sum_{j=i+1}^{P} (d_j - \bar{d})(d_{j-i} - \bar{d}),$$

assuming that h-step-ahead forecast errors are serially correlated up to the order $h - 1$.

Under the null hypothesis the MDM test statistic has Student's t-distribution with $(P - 1)$ degrees of freedom.

5. APPLICATION

The data we put under investigation are monthly exchange rates of Czech, Slovak and Polish currency to Euro in the period of 5.25 years. As seen in Figure 1, these time series are non-stationary so their first differences are used for modelling instead. Since Polish Zloty dominates the currencies of V4 countries, it is chosen to be the transition variable. We model the CZK, SKK and PLN time series as 3-dimensional (endogenous) variable $\boldsymbol{y}_t$ without considering any exogenous ones. Mean of the data is $(-0.042, -0.038, 0.011)$ and determinant of covariance matrix 0.000516. Initially a linear VAR is fitted to data with order found by minimizing information criteria (Figure 1). The determinant of covariance matrix of residuals from the VAR(1) is 0.000234, which represents 55 % coverage in fit by the model.

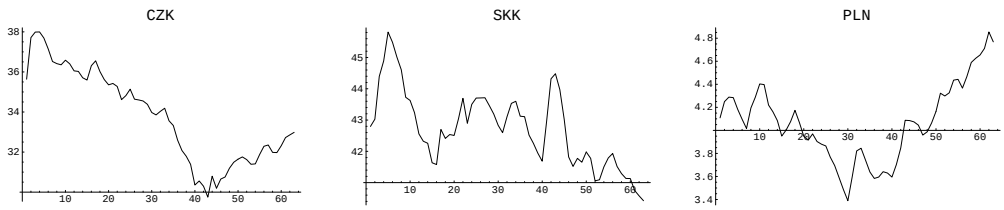

Fig. 1. Exchange rates of three currencies to Euro observed monthly for 5.25 years.

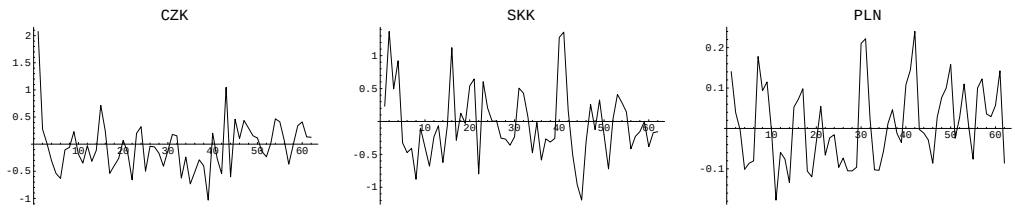

Fig. 2. First differences of the exchange rates.

The order of linear VAR model is used in linearity testing via Tsay test (against TAR) or LM-type test (against STAR) while letting the maximum delay d vary from 1 to 12. With some aggregation operators, setting d too high causes singularity in computation of test statistics. Parameters estimation and model selection is performed in one step by computing BIC conditionally over parameters grid $r \times \gamma$, $r = \{-0.1, -0.09, \ldots 0.12\}$ and $\gamma = \{0.1, 0.5, 1, 2, 3, 4, 5, 10, 20\}$. Estimated residuals are subjected to test for serial independence and if they pass, 1-step-ahead forecasts for next 3 years are drawn from the model. To compare models, there can be considered either in-sample (determinant of estimated covariance matrix, $|\boldsymbol{\Sigma}|$) or out-of-sample (mean square prediction error, MSPE) fit, however, the final "hit parade" is based on ranking from modified Diebold–Mariano test, where the models are being compared to each other in pairs and the instant evaluation points ($-1, 0$ or 1) are summed to constitute a score for each model. The higher the score is, the

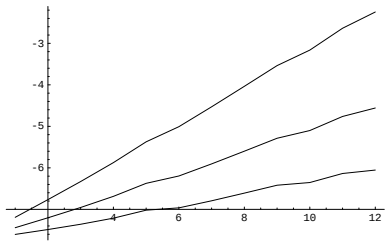

Fig. 3. AIC, HQIC and BIC (from above) versus order of VAR.

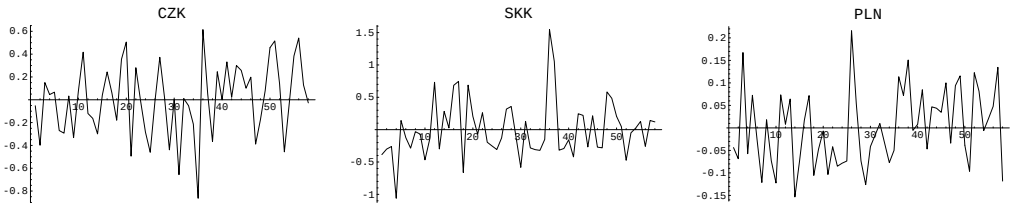

Fig. 4. Residuals from the winning multivariate 2-regime ESTAR model.

better position model occupies on the list. Finally, the forecast errors from LSTAR and ESTAR compete for overall placing.

From results in table, it is important to note that improvement in fit caused by inclusion of aggregation operators is indeed significant. The first operator in the list ($\mathcal{LAST}$) corresponds to the usual setup $z_t = \xi_{t-d}$ and obviously it does not belong to the most accurate models. According to predictive accuracy, the best model among LSTARs is the one with $\mathcal{MEAN}$ aggregation operator applied over last four months of PLN exchange rates, while the best ESTAR (considering in-sample fit as supplementary criterion) uses $\mathcal{MIN}$ over last 3 months.

All of the nonlinear models achieve better performance within sample than linear VAR model. The fit ratio of winning ESTAR model with $|\mathbf{\Sigma}| = 0.000152$ is over 70 % . However, the forecast performance displays smaller difference and VAR(1) with MSPE= 0.5328 is serious competitor to the nonlinear models placing itself to 3rd position overall.

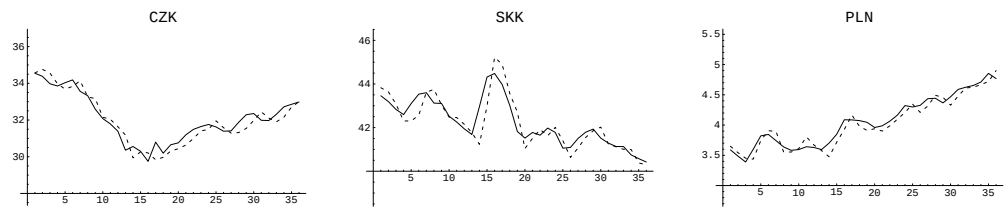

Fig. 5. Original data and forecasts from the best multivariate 2-regime ESTAR model.

Table. Results of linearity testing and model specification.

2-regimes multivariate LSTAR												overall
Aggreg. operator	Tsay test		LM test		parameters estimate				LB test	Forecast accuracy	ranking from DM	ranking from DM
	d	p-value	d	p-value	d	r	γ	$\lvert\boldsymbol{\Sigma}\rvert$	p-value	MSPE	test	test
$\mathcal{LAST}$	3	0.0064	9	0.00001	9	-0.20	20	0.000180	0.5052	0.6995	3	6
$\mathcal{MAX}$	4	0.0349	7	$< 10^{-6}$	7	0.04	20	0.000140	0.3902	0.7287	2	4
$\mathcal{MIN}$	3	0.0173	3	0.00016	3	-0.10	20	0.000130	0.9711	0.5006	2	2
$\mathcal{MEAN}$	4	0.0251	4	0.00012	4	-0.10	20	0.000103	0.9936	0.5084	1	**1**
$\mathcal{W}_2$	6	0.0158	5	0.00003	5	0.12	20	0.000104	0.8911	0.5227	1	2
$\mathcal{W}_3$	10	0.0083	6	0.00002	6	0.12	20	0.000102	0.8804	0.5406	2	3
$\mathcal{W}_F$	8	0.0139	5	$< 10^{-6}$	5	-0.10	20	0.000106	0.9195	0.5502	2	2
$\mathcal{W}_{S(0.8)}$	10	0.0102	4	0.00001	4	-0.10	0.1	0.000107	0.8664	0.5139	1	2
2-regimes multivariate ESTAR												
$\mathcal{LAST}$	3	0.0064	2	0.00002	2	0.12	0.2	0.000164	0.6133	0.5654	2	2
$\mathcal{MAX}$	4	0.0349	7	$< 10^{-6}$	7	0.80	20	0.000130	0.3087	0.7601	3	5
$\mathcal{MIN}$	3	0.0173	3	$< 10^{-6}$	3	0.04	0.2	0.000152	0.7487	0.4913	2	**1**
$\mathcal{MEAN}$	4	0.0251	4	$< 10^{-6}$	4	0.12	0.2	0.000149	0.2431	0.5153	2	2
$\mathcal{W}_2$	6	0.0158	4	$< 10^{-6}$	4	-0.07	5	0.000150	0.2607	0.5418	2	3
$\mathcal{W}_3$	10	0.0083	8	$< 10^{-6}$	8	-0.07	10	0.000158	0.0530	0.5170	2	3
$\mathcal{W}_F$	8	0.0139	3	$< 10^{-6}$	3	-0.06	2	0.000155	0.3354	0.5008	1	**1**
$\mathcal{W}_{S(0.8)}$	10	0.0102	8	0.00001	8	-0.08	10	0.000156	0.0597	0.5393	2	3

Also note that judging entirely from a single measure, either $|\boldsymbol{\Sigma}|$ or MSPE, need not be the only nor the best way to choose the model. To be correct, besides the Diebold–Mariano test of predictive accuracy it could be desirable also to perform additional diagnostic tests on residuals to check the model adequacy. However, at this stage the above investigation is sufficient enough to highlight the methods described in this article.

6. CONCLUSION

In this paper we have given an overview of multivariate modelling based on logistic and exponential smooth transition models with some useful generalizations, mainly generating transition variable by aggregation operators and computing information criteria with orders specific for each of m regimes. The model specification procedure has been demonstrated on trivariate exchange rates time series. The application results show satisfactory improvement in fit when particular aggregation operators are used. The routines were programmed in system Mathematica 5.2, and the source code with all utility functions available as Mathematica package is provided on the author's webpage: www.math.sk/bacigal/homepage/research_en.html.

ACKNOWLEDGEMENT

This work was partially supported by grants APVT-20-003204 and VEGA 1/3006/06.

(Received January 10, 2007.)

REFERENCES

[1] T. Calvo, A. Kolesárová, M. Komorníková, and R. Mesiar: Aggregation operators: Properties, classes and construction methods. In: Aggregation Operators (T. Calvo, G. Mayor, and R. Mesiar, eds.), Physica–Verlag, New York 2002, pp. 3–140.

[2] D. van Dijk, T. Teräsvirta, and P. H. Franses: Smooth Transition Autoregressive Models – A Survey of Recent Developments. Econometric Institute Research Report EI2000-23, 2000.

[3] A. Escribano and O. Jordá: Improved testing and specification of smooth transition regression models. In: Nonlinear Time Series Analysis of Economic and Financial Data (P. Rothman, ed.), Kluwer, Boston 1999, pp. 289–319.

[4] C. W. J. Granger and T. Teräsvirta: Modelling Nonlinear Economic Relationships. Oxford University Press, Oxford 1993.

[5] D. I. Harvey, S. J. Leybourne, and P. Newbold: Testing the equality of prediction mean squared errors. Internat. J. Forecasting *13* (1997), 281–291.

[6] R. Luukkonen, P. Saikkonen, and T. Teräsvirta: Testing linearity against smooth transition autoregressive models. Biometrika *75* (1988), 491–499.

[7] R. S. Tsay: Testing and modeling multivariate threshold models. J. Amer. Statist. Assoc. *93* (1998), 1188–1202.

Tomáš Bacigál, Faculty of Civil Engineering, Slovak University of Technology, Bratislava. Slovak Republic.
e-mail: bacigal@math.sk