The Multi-State Markov Switching Model

Edoardo Otranto*

Dipartimento di Economia, Impresa e Regolamentazione (DEIR),

Sassari, Italy

Abstract

In many real phenomena the behaviour of a certain variable, subjected to different regimes, depends on the state of other variables or the same variable observed in other subjects, so the knowledge of the state of the latter could be important to forecast the state of the former. In this paper a particular multivariate Markov Switching model is developed to represent this case. The transition probabilities of this model are characterized by the dependence on the regime of the other variables. The estimation of the

*Correspondence to: Edoardo Otranto, Dipartimento di Economia, Impresa e Regolamentazione (DEIR), Università di Sassari, Via Sardegna 59, I-07100 Sassari, Italy, e-mail: eotranto@uniss.it

transition probabilities provides useful informations for the researcher to forecast the regime of the variables analyzed. Theoretical background and an application are shown.

KEY WORDS regime-switching; multivariate time series; transition probabilities

1. Introduction

The Markov Switching (MS hereafter) model, introduced in econometrics by Hamilton (1989), has been largely used for the study of phenomena with two or more unobservable regimes, represented by states of a Markov chain. For example, it was used for the representation of business cycle (Hamilton, 1989), for the study of segmented trends in exchange rates (Engel and Hamilton, 1990, Engel, 1994), for the evaluation of monetary effects on output (Garcia and Schaller, 2002). Generally, univariate approaches are used for the applications in these fields and the extensions to multivariate cases look for common states for the same variable observed in different subjects (countries, markets, firms, etc.) or for several different variables. For example, Kim and Nelson (1998) extract a common component subjected to two different regimes from 4 economic variables and interpret it as the common cycle; Krolzig (1997) extends the MS model to the VAR case, considering a common state for all the variables considered.

The multivariate MS model has not received large attention, essentially because it does not provide easily interpretable results in terms of inference on the regimes (see, for example, the results in Engel and Hamilton, 1990). Maybe this fact is due to the common practice of econometricians to attribute to each regime an economic interpretation; for example, in the analysis of the business cycle, two states are considered representing respectively the recession and the growth regimes. Anyway the correspondence between regimes and economic periods is subjective (see, for example, the considerations in Anas and Ferrara, 2002), and the inference on the regime is important only for particular phenomena, as the business cycle, in which the purpose is to subdivide the time interval in spans with a certain interpretation. When the interest is devoted to the forecast of the regime, the focus is on the interpretations of the probabilities deriving by the Markovian structure imposed in the model. The model proposed in this paper deals with the latter kind of interest.

In this paper, we deal with n variables depending on a certain number of regimes; they could be different variables or the same variable observed on n subjects. The starting point runs from the considerations that it is well difficult to suppose that the same regime structure is followed by all the n variables; this could be a convenient hypothesis for the business cycle analysis again, to extract a common component representing the comovements among variables. But in other applications, such as disease diffusion, financial markets, economic development, the regimes follow different dynamics in different variables. On the other hand, some links among the dynamics of the unobserved regimes probably exist and the regime of a variable could lead the regime of another variable; for example, the economic growth of a "leader" country would influence the economy of other countries in the next time and the knowledge of a regime for a variable could help to forecast the regime of another variable.

In this paper we propose an MS model in a multivariate framework, allowing different variables representing the regimes. The dynamics of the regime of a variable depends on the dynamics of the regime in the other variables. We call this model the Multi-State Markov Switching (MSMS) model. The aim of this paper is to formalize the MSMS model and to show its usefulness.

In the next section the new model will be illustrated and in section 3 an application will analyze the relationship existing between the turnover and new orders series to show as the probabilities change when the state of the leading variable changes; the results will be compared, in terms of forecasting performance, with the multivariate MS model and two univariate MS models. Final remarks follow.

2. The Model Specification

Let $\mathbf{y}_t$ an nx1 multivariate time series observed for n different variables. We suppose that it follows a VAR process:

$$\mathbf{y}_t = \mathbf{B}\boldsymbol{\mu}_{\mathbf{S}_t} + \sum_{h=1}^{p} \boldsymbol{\Phi}_h \mathbf{y}_{t-h} + \boldsymbol{\varepsilon}_t, \qquad t = 1, ..., T \tag{2.1}$$

The vector $\boldsymbol{\mu}_{\mathbf{s}_t}$ contains n constants and it is indicized by a multiple random variable $\mathbf{S}_t = (s_{1,t}, s_{2,t}, ..., s_{n,t})$. We call this variable the multiple regime (or multiple state) at time t. The univariate random variable $s_{j,t}$ $(j = 1, ..., n)$ represents the state at time t for the variable j and can assume k possible values; for the sake of simplicity we suppose $k = 2$ and denote these states with 0 and 1 (the extension to a generic k is simple but notationally cumbersome). The difference with respect to the multivariate MS model is that the realization at time t of the dichotomous variable s_{jt} can be different from the realization of s_{it} $(i \neq j)$. The nxn matrix $\mathbf{B} = \{b_{ij}\}$ $(i, j = 1, ..., n)$ is a matrix of coefficients with elements on the diagonal equal one. In practice, the intercept of the variable y_{jt}, contained in the vector $\mathbf{y}_t$, is equal to

$$\mu_{s_{j,t}} + \beta_{j1}\mu_{s_{1,t}} + ... + \beta_{jn}\mu_{s_{n,t}}$$

in which it is possible to separate an "autonomous" intercept $\mu_{s_{j,t}}$ from the effect due to the other variables. $\mathbf{\Phi}_h$ $(h = 1, ..., p)$ is a matrix of unknown coefficients.

The n-vector of disturbances $\boldsymbol{\varepsilon}_t$ is uncorrelated along the time but correlated across the variables, to take into account their mutual influence. We suppose that $\boldsymbol{\varepsilon}_t$ is normally distributed with mean $\mathbf{0}$ and covariance matrix:

$$\mathbf{\Sigma}_{\mathbf{S}_t} = \begin{bmatrix} \sigma^2_{1s_{1,t}} & \rho_{12}\sigma_{1s_{1,t}}\sigma_{2s_{2,t}} & \cdots & \rho_{1n}\sigma_{1s_{1,t}}\sigma_{ns_{n,t}} \\ \rho_{12}\sigma_{1s_{1,t}}\sigma_{2s_{2,t}} & \sigma^2_{2s_{2,t}} & \cdots & \rho_{2n}\sigma_{2s_{2,t}}\sigma_{ns_{n,t}} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_{1s_{1,t}}\sigma_{ns_{n,t}} & \rho_{2n}\sigma_{2s_{2,t}}\sigma_{ns_{n,t}} & \cdots & \sigma^2_{ns_{n,t}} \end{bmatrix}$$

where ρ_{ij} represents the correlation between couples of disturbances $(i, j = 1, ..., n)$. Using this notation, we mean that also the variance and the covariances change with the multiple regime $\mathbf{S}_t$. Clearly the number of coefficients to be estimated increases with n, so, in some cases, it will be convenient to adopt some reparameterization for the coefficients ρ_{ij}; for example, if the n variables represent spatial units, we can use:

$$\rho_{ij} = \rho^{d_{ij}}$$

where d_{ij} is a known distance measure between the units i and j. In the example of section 3, we use only 2 variables, so reparameterization does not need.

Of course (2.1) can be modified considering an MA structure or constraining some coefficients to be zero. If the multiple regime $\mathbf{S}_t$ is a single variable, the previous model is a simple MS model, with the same regime for all the variables.

The multiple regime $\mathbf{S}_t$ is not observable and its distribution is unknown; we suppose that it follows an ergodic Markov chain, in which the probability of a particular realization of $\mathbf{S}_t$ depends only on the realization of $\mathbf{S}_{t-1}$. The possible realizations of $\mathbf{S}_t$ are 2^n, so the Markovian transition probability matrix contains 2^nx2^n elements; the generic element of this matrix will be:

$$\Pr\left[\mathbf{S}_t|\mathbf{S}_{t-1}\right] = \Pr\left[s_{1,t}, s_{2,t}, ..., s_{n,t}|s_{1,t-1}, s_{2,t-1}, ..., s_{n,t-1}\right].$$

The states $s_{1,t}$, $s_{2,t}$, ..., $s_{n,t}$, conditional on $\mathbf{S}_{t-1} = (s_{1,t-1}, s_{2,t-1}, ..., s_{n,t-1})$, are mutually independent so that:

$$\Pr\left[\mathbf{S}_t|\mathbf{S}_{t-1}\right] = \Pr\left[s_{1,t}|\mathbf{S}_{t-1}\right] \Pr\left[s_{2,t}|\mathbf{S}_{t-1}\right] ... \Pr\left[s_{n,t}|\mathbf{S}_{t-1}\right]. \tag{2.2}$$

The probabilities $\Pr\left[s_{j,t}|\mathbf{S}_{t-1}\right]$ can be parameterized with logistic functions[1] to

[1] Filardo (1994) and Diebold *et al.* (1994) use the logistic function to parameterize the time-varying transition probabilities of a univariate MS model.

take into account the influence of the regimes relative to the other variables:

$$\begin{aligned}
\Pr\left[s_{j,t}=0|s_{j,t-1}=0,\mathbf{S}_{-j,t-1}\right] &= \frac{\exp\left(\vartheta_{j,00}+\boldsymbol{\theta}'_{j,0}\mathbf{S}_{-j,t-1}\right)}{1+\exp\left(\vartheta_{j,00}+\boldsymbol{\theta}'_{j,0}\mathbf{S}_{-j,t-1}\right)} \\
\Pr\left[s_{j,t}=1|s_{j,t-1}=0,\mathbf{S}_{-j,t-1}\right] &= 1-\frac{\exp\left(\vartheta_{j,00}+\boldsymbol{\theta}'_{j,0}\mathbf{S}_{-j,t-1}\right)}{1+\exp\left(\vartheta_{j,00}+\boldsymbol{\theta}'_{j,0}\mathbf{S}_{-j,t-1}\right)} \\
\Pr\left[s_{j,t}=1|s_{j,t-1}=1,\mathbf{S}_{-j,t-1}\right] &= \frac{\exp\left(\vartheta_{j,11}+\boldsymbol{\theta}'_{j,1}\mathbf{S}_{-j,t-1}\right)}{1+\exp\left(\vartheta_{j,11}+\boldsymbol{\theta}'_{j,1}\mathbf{S}_{-j,t-1}\right)} \\
\Pr\left[s_{j,t}=0|s_{j,t-1}=1,\mathbf{S}_{-j,t-1}\right] &= 1-\frac{\exp\left(\vartheta_{j,11}+\boldsymbol{\theta}'_{j,1}\mathbf{S}_{-j,t-1}\right)}{1+\exp\left(\vartheta_{j,11}+\boldsymbol{\theta}'_{j,1}\mathbf{S}_{-j,t-1}\right)}
\end{aligned} \tag{2.3}$$

In this case $\boldsymbol{\theta}_{j,\iota}=\left(\vartheta_{j,\iota 1},\ldots,\vartheta_{j,\iota(j-1)},\vartheta_{j,\iota(j+1)},\ldots\vartheta_{j,\iota n}\right)'$ $(\iota=0,1)$ is a $(n-1)$-vector of unknown parameters and $\mathbf{S}_{-j,t-1}$ is equal to the multiple regime vector $\mathbf{S}_{t-1}$, excluding the $j-th$ element $s_{j,t-1}$.

The estimation of the vectors $\boldsymbol{\theta}_{j,\iota}$ $(\iota=0,1)$ provides the estimation of the conditional probabilities for each regime and the estimation of the full matrix with elements expressed by (2.2). Of course, the number of parameters to be estimated could result large and could affect the goodness of the estimation procedure; anyway, generally the priors of the researcher would constraint some parameters in $\boldsymbol{\theta}_{j,\iota}$ equal zero, as we will see in the successive section.

The estimation of the transition probabilities matrix can be utilized to perform the well-known procedure of filtering proposed by Hamilton (1990) and Kim (1994) for simple MS models and to provide the 2^n filtered probabilities $\Pr[\mathbf{S}_t|\mathbf{\Psi}_t]$, where $\mathbf{\Psi}_t$ represents the information available at time t. In practice, this procedure is valid for a generic $\varsigma-$states MS model; the same procedure can be utilized for the MSMS model, putting $\varsigma = 2^n$ and imposing (2.2) and the parameterization (2.3) to obtain the transition probabilities matrix. The filtering procedure is presented in the final appendix.

Resuming, the main results of the estimation procedure are:

1. the estimates of the probabilities $\Pr[s_{j,t}|\mathbf{S}_{t-1}]$: they show how the probability to stay in the same regime or to change it for the variable j changes when the regimes relative to the other variables change;

2. the estimate of the transition probabilities matrix $\mathbf{P} = \{\Pr[\mathbf{S}_t|\mathbf{S}_{t-1}]\}$: putting the possible states at time $t-1$ in correspondence to the rows of $\mathbf{P}$ and the possible states at time t in correspondence to the columns of $\mathbf{P}$, the rows will indicate what is the most probable scenario (a particular combination of regimes for the various variables) at time t given a certain scenario at time $t-1$.

These interpretations would be clarified in the next sections, applying the (2.1)-(2.2) model to real series.

3. An Example: Forecasting the Italian Turnover Index

In this section we show a simple example in which the prior informations about the phenomenon can help to specify the dependencies among the states, favoring the specification of a particular MSMS model.

It is a common practice to consider the industrial new orders as a leading indicator of the industrial turnover. We consider the monthly trend-cycle series of indices of the Italian new orders and turnover from January 1990 to June 2002 (obtained using the TRAMO-SEATS routine, Gómez and Maravall, 1996). To specify the MSMS model we can use this prior information, in the sense that we can use a model in which the new orders can determine the level and the regime of the turnover, but not vice-versa. We use a model like (2.1) with $p = 1$; in addition, in the matrix $\mathbf{\Phi}_1$ only the diagonal elements were significant. Labeling with 1 the turnover and 2 the new orders, the model for the first differences of the logs of the variables is:

$$\begin{aligned} \mathbf{y}_t &= \begin{bmatrix} 1 & \beta \\ 0 & 1 \end{bmatrix} \boldsymbol{\mu}_{\mathbf{S}_t} + \begin{bmatrix} \boldsymbol{\phi}_{1,1} & 0 \\ 0 & \boldsymbol{\phi}_{2,1} \end{bmatrix} \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t \\ \boldsymbol{\varepsilon}_t &\sim IIN\left(\mathbf{0}, \begin{bmatrix} \sigma^2_{1,s_{1,t}} & \rho\sigma_{1,s_{1,t}}\sigma_{2,s_{2,t}} \\ \rho\sigma_{1,s_{1,t}}\sigma_{2,s_{2,t}} & \sigma^2_{2,s_{2,t}} \end{bmatrix} \right) \end{aligned}$$

For the estimation of the transition probabilities matrix we use the parameterization (2.3); using our priors on the dependence relationships, they are parameterized as:

$$
\begin{aligned}
\Pr\left[s_{1,t}=0|s_{1,t-1}=0,s_{2,t-1}\right] &= \frac{\exp\left(\vartheta_{1,00}+\vartheta_{1,02}s_{2,t-1}\right)}{1+\exp\left(\vartheta_{1,00}+\vartheta_{1,02}s_{2,t-1}\right)} \\
\Pr\left[s_{2,t}=0|s_{2,t-1}=0,s_{1,t-1}\right] &= \frac{\exp\left(\vartheta_{2,00}\right)}{1+\exp\left(\vartheta_{2,00}\right)} \\
\Pr\left[s_{1,t}=1|s_{1,t-1}=1,s_{2,t-1}\right] &= \frac{\exp\left(\vartheta_{1,10}+\vartheta_{1,12}s_{2,t-1}\right)}{1+\exp\left(\vartheta_{1,10}+\vartheta_{1,12}s_{2,t-1}\right)} \\
\Pr\left[s_{2,t}=1|s_{2,t-1}=1,s_{1,t-1}\right] &= \frac{\exp\left(\vartheta_{2,10}\right)}{1+\exp\left(\vartheta_{2,10}\right)}
\end{aligned}
$$

The estimates of the coefficients are shown in Table 1. Note that the intercept of the turnover index ($\mu_{1h}+\beta\mu_{2r},\ h,r=0,1$) changes with the regimes s_{1t} and s_{2t}, adapting to the various scenarios; in particular it will be -0.33, -0.01, 0.24, 0.56.

Of more interest is the estimation of probabilities in (2.3), because we have a clear interpretation of the possibility to change the regime for a single series, given a certain scenario in the previous month. Let:

$$
p_{i,hr}^{(j)}=\Pr\left[s_{j,t}=i|s_{1,t-1}=h,s_{2,t-1}=r\right] \qquad j=1,2 \quad i,h,r=0,1
$$

These probabilities are shown in Table 2.

The influence of s_{2t} on the changes in regime for the turnover are evident; in fact, the probability of variable 1 to stay in regime 0, when the variable 2 was in regime 0 in the previous month, is 0.95, but decreases to 0.48 when variable 2 was in regime 1. In a similar way, the turnover remains in regime 1 with a 86% of chances, when the new orders was in the same regime, but switches to regime 0 with a probability equal to 47% ($p^{(1)}_{0,10} = 1 - p^{(1)}_{1,10}$) when the variable 2 was in regime 0.

Let us denote:

$$p_{ih,rv} = \Pr\left[s_{1,t} = i, s_{2,t} = h | s_{1,t-1} = r, s_{2,t-1} = v\right], \; i, h, r, v = 0, 1;$$

we can obtain by (2.2) the transition probability matrix shown in Table 3.

The rows of this matrix sum up to one. In this case we can deduce the most probable scenario at time t, given the scenario at the previous month. When both the variables are in the same state at time $t-1$, it is probable to maintain this scenario in the next month. When the new orders index is in a different state with respect to the turnover at time $t-1$, the most probable scenario at time t will be the persistence of this scenario or the alignment of turnover to the same scenario of new orders.

The performance of the MSMS specification with respect to the classical multivariate MS model and the univariate MS model can be evaluated in terms of forecasts of the turnover index. We estimate a multivariate MS (MMS) model, with a similar specification of the previous MSMS:

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\mu}_{\mathbf{s}_t} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \\ \boldsymbol{\varepsilon}_t &\sim IIN\left(\mathbf{0}, \begin{bmatrix} \sigma^2_{1,s_{1,t}} & \rho\sigma_{1,s_{1,t}}\sigma_{2,s_{2,t}} \\ \rho\sigma_{1,s_{1,t}}\sigma_{2,s_{2,t}} & \sigma^2_{2,s_{2,t}} \end{bmatrix}\right) \end{aligned}$$

where s_t is the common state variable and $\boldsymbol{\Phi}_1 = \begin{bmatrix} \boldsymbol{\phi}_{1,1} & \boldsymbol{\phi}_{1,2} \\ 0 & \boldsymbol{\phi}_{2,2} \end{bmatrix}$ (the coefficient $\boldsymbol{\phi}_{1,2}$ in this case has resulted significant). The inadequacy of a multivariate model with a common state is confirmed by the fact that, for both the variables, the intercepts in the two states were not significantly different. In practice, for MMS model, only the variances are switching.

Some improvement is obtained estimating univariate MS models only for the turnover. Two models are adopted for this case: a purely autoregressive model (UMS) and a model with new orders as explicative variable (UXMS); the latter is:

$$\begin{aligned} y_{1,t} &= \mu_{\mathbf{s}_t} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \varepsilon_{1,t}, \\ \varepsilon_{1,t} &\sim IIN\left(0, \sigma^2_{1,s_t}\right). \end{aligned}$$

Of course, $y_{1,t}$ indicate the turnover index whereas $\mathbf{\Phi}_1 = \left[\begin{array}{cc} \boldsymbol{\phi}_{1,1} & \boldsymbol{\phi}_{1,2} \end{array} \right]$, and UMS can be obtained by UXMS constraining $\boldsymbol{\phi}_{1,2} = 0$.[2]

The forecasting performance is evaluated with four typical indices (Root Mean Squared Error-$RMSE$, Mean Absolute Error-MAE, Theil U statistic-U, Theil U statistic measured on percentage changes-U_Δ; see, for example, Greene, 2000, section 7.11.2). For the MSMS model, the forecast $\widehat{\mathbf{y}}_t$, conditional on the information at time t, is obtained by (the hat indicates the estimates):

$$\widehat{\mathbf{y}}_t = \sum_{i=1}^{2} \sum_{h=1}^{2} \left(\widehat{\mathbf{B}} \widehat{\boldsymbol{\mu}}_{\mathbf{S}_t} + \widehat{\mathbf{\Phi}}_1 \mathbf{y}_{t-1} \right) \Pr\left[s_{1t} = i, s_{2t} = h | \mathbf{\Psi}_t \right],$$

[2]Another univariate model, the time varying transition probability one (Filardo, 1994, Diebold *et al.*, 1994), was estimated to forecast the turnover, using the lagged new orders as exogenous variable in the transition probabilities time varying formula, but it has not produced good results, so it is not furthermore described.

whereas for the MMS, UMS, UXMS models:

$$\widehat{\mathbf{y}}_t = \sum_{i=1}^{2} \left(\widehat{\boldsymbol{\mu}}_{\mathbf{s}_t} + \widehat{\boldsymbol{\Phi}}_1 \mathbf{y}_{t-1} \right) \Pr\left[s_t = i | \boldsymbol{\Psi}_t \right].$$

In Table 4 the results are showed. For all the loss functions, the best performance is shown by the MSMS model, which strongly improves the MMS and the UMS approaches. The good results in terms of Theil statistics, with respect to the other models, reflects the ability of MSMS model to track the turning points in the series.

The presence of the new orders as exogenous variable in UXMS helps to improve the univariate model, obtaining results slightly worse than MSMS model; anyway, the MSMS model provides various additional information about the comovements between the variables and the probable scenarios.

Table 3.1: Estimates of the MSMS model and relative standard errors

μ_{10}	μ_{11}	μ_{20}	μ_{21}	β	$\phi_{1,1}$	$\phi_{2,1}$
-0.06 (0.05)	0.51 (0.05)	-0.04 (0.04)	0.23 (0.06)	-1.16 (0.40)	0.75 (0.02)	0.77 (0.03)
σ_{10}	σ_{11}	σ_{20}	σ_{21}	ρ		
0.14 (0.02)	0.11 (0.01)	0.31 (0.02)	0.43 (0.03)	0.66 (0.04)		
$\vartheta_{1,00}$	$\vartheta_{1,02}$	$\vartheta_{2,00}$	$\vartheta_{1,10}$	$\vartheta_{1,12}$	$\vartheta_{2,10}$	
2.89 (0.84)	-2.95 (1.11)	1.45 (0.49)	0.12 (0.63)	1.71 (1.02)	1.16 (0.35)	

Table 3.2: Conditional probabilities for the regimes of each series

$p^{(1)}_{0,00}$	$p^{(1)}_{0,01}$	$p^{(1)}_{1,10}$
0.947	0.485	0.531
$p^{(1)}_{1,11}$	$p^{(2)}_{0,00} = p^{(2)}_{0,10}$	$p^{(2)}_{1,01} = p^{(2)}_{1,11}$
0.862	0.809	0.762

Table 3.3: Estimated transition probabilities matrix

	$p_{00,..}$	$p_{01,..}$	$p_{10,..}$	$p_{11,..}$
$p_{..,00}$	0.767	0.181	0.042	0.010
$p_{..,01}$	0.115	0.369	0.123	0.393
$p_{..,10}$	0.380	0.089	0.430	0.101
$p_{..,11}$	0.033	0.105	0.205	0.657

Table 3.4: Loss functions to evaluate the forecasts with several models.

	$RMSE$	MAE	$U(10^2)$	U_Δ
MSMS	0.2744	0.2124	0.2795	0.4036
MMS	0.2878	0.2228	0.2932	0.4223
UMS	0.2874	0.2253	0.2928	0.4201
UXMS	0.2769	0.2169	0.2822	0.4059

4. Concluding Remarks

In this paper a new approach to analyze multivariate series subject to changes in regimes was proposed and we name it MSMS model. It consists in an extension of the classical MS model to the multivariate case, but allowing separate not independent state variables for each series. The main novelty of this model is that the state of each variable does not depend only on its state at the previous time, but also on the states of the other variables. In this way it is possible to capture the comovements among variables and their dependencies. In particular, the MSMS model provides useful interpretation in the dynamics of the states, pointing out the most probable scenarios for the next time, given the actual states, and the effects of changes in regimes of the other variables.

The structure of the MSMS model seems to be proper for the analysis of the same variable observed in different countries (space-time series) and variables linked by leading-lagging relationships. Our example deals with economic variables, but this model could be useful in biometrics (for example for the study of disease diffusion among countries) or in social sciences.

In this paper we have stressed the specification aspects and the methodological background, whereas the hypothesis test theory was not treated. It will be interesting to analyze if the methodologies developed for the classical MS models to detect the number of regimes (Hansen, 1992 and 1996, Garcia, 1995, Otranto and Gallo 2002) could be extended to the MSMS models.

The extension to $k > 2$ states is conceptually simple,but the dimension of the transition probabilities matrix will increase to k^nxk^n; the estimators could be inefficient and their convergence would not be guaranteed. Computational techniques developed for MS models (for example, MCMC algorithms described in Kim and Nelson, 1999) could help.

Other possible extensions can consider as switching also the AR coefficients in (2.1) or insert exogenous observed variables in (2.3), to obtain the time-varying transition probabilities (Filardo, 1994, Diebold *et al.*, 1994).

Appendix

In this appendix the steps to filter the series and obtain the various probabilities explained in the paper are detailed. The filter is a simple extension of the well-known filter of Hamilton (1990) and Kim (1994); the reader is referred to the chapter 22 of Hamilton (1994) for the proofs of the following relationships, whereas here the only algorithms of filtering and smoothing are described.

Let $\boldsymbol{\xi}_{t|t}$ the (2^nx1) vector containing the probabilities of each possible multiple states, conditional on the information at time t; in other terms, the generic element of $\xi_{t|t}$ is $\Pr\left[s_{1,t}=i, s_{2,t}=j, \ldots s_{n,t}=k|\mathbf{\Psi}_t\right]$. Furthermore, let $\boldsymbol{\eta}_t$ the (2^nx1) vector containing the densities of $\mathbf{y}_t$ corresponding to the 2^n possible multiple states; for the normality of the n-vector $\boldsymbol{\varepsilon}_t$ in (2.1), the generic element of $\boldsymbol{\eta}_t$ is the density of $\mathbf{y}_t$ conditional on a particular multiple state $\mathbf{S}_t$:

$$(2\pi)^{-n/2}\left|\mathbf{\Sigma}_{\mathbf{s}_t}\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\mathbf{y}_t-\mathbf{B}\boldsymbol{\mu}_{\mathbf{S}_t}-\sum_{h=1}^{p}\mathbf{\Phi}_h\mathbf{y}_{t-h}\right)'\mathbf{\Sigma}_{\mathbf{s}_t}^{-1}\left(\mathbf{y}_t-\mathbf{B}\boldsymbol{\mu}_{\mathbf{S}_t}-\sum_{h=1}^{p}\mathbf{\Phi}_h\mathbf{y}_{t-h}\right)\right\}.$$

Denoting with $\odot$ the element-by-element product and with $\mathbf{1}$ a (2^nx1) vector with all the elements equal 1, the filter is composed by the following two steps:

1. $\boldsymbol{\xi}_{t|t}=\frac{\boldsymbol{\xi}_{t|t-1}\odot\boldsymbol{\eta}_t}{\mathbf{1}'\left(\boldsymbol{\xi}_{t|t-1}\odot\boldsymbol{\eta}_t\right)}$

2. $\boldsymbol{\xi}_{t+1|t}=\mathbf{P}'\boldsymbol{\xi}_{t|t}$

Note that the denominator of the point 1) is the density of a mixture of 2^n Normal distributions and represents the density function $f\left(\mathbf{y}_t|\mathbf{\Psi}_t\right)$, not conditional on the multiple state. Iterating the two steps for $t=1,...,T$, the log-likelihood $\sum_{t=1}^{T}\log f\left(\mathbf{y}_t|\mathbf{\Psi}_t\right)$ will be obtained.

To start the algorithm, a starting value $\boldsymbol{\xi}_{1|0}$ has to be chosen. Using the properties of the ergodic Markov chains, an appropriate choice are the unconditional probabilities:

$$\boldsymbol{\xi}_t = \Pr\left[s_{1,t} = i, s_{2,t} = j, \ldots s_{n,t} = k\right];$$

they can be obtained by (see Hamilton, 1994):

$$\boldsymbol{\xi}_t = \left(\mathbf{D}'\mathbf{D}\right)^{-1}\mathbf{D}'\mathbf{i}_{2^n+1},$$

where $\mathbf{i}_{2^n+1}$ is the (2^n+1)-th column of a (2^n+1)x(2^n+1) identity matrix and:

$$\mathbf{D} = \begin{bmatrix} \mathbf{I}_{2^n} - \mathbf{P}' \\ \\ \mathbf{1}' \end{bmatrix}$$

where $\mathbf{I}_{2^n}$ is the $(2^n$x$2^n)$ identity matrix.

To obtain the smoothing probabilities for each multiple state, it is sufficient to iterate for $t = T - 1$, ..., 1,the following formula:

$$\boldsymbol{\xi}_{t|T} = \boldsymbol{\xi}_{t|t} \odot \left\{\mathbf{P}\left[\boldsymbol{\xi}_{t+1|T}\,(\div)\,\boldsymbol{\xi}_{t+1|t}\right]\right\}$$

where $(\div)$ denotes the element-by-element division. In this case, the starting values $\boldsymbol{\xi}_{T|T}$, $\boldsymbol{\xi}_{t|t}$ and $\boldsymbol{\xi}_{t+1|t}$ are provided by the previous filtering. If we are interested in the inference on the regime, $\Pr\left[s_{j,t} = i|\boldsymbol{\Psi}_T\right]$ $(j = 1, ..., n,\ i = 0, 1)$ is obtained summing the 2^{n-1} elements of $\boldsymbol{\xi}_{t|T}$ with $s_{j,t} = i$.

References

[1] Anas J, Ferrara L, 2002. A Comparative Assessment of Parametric and Non Parametric Turning Points detection Methods: the Case of the Euro-zone Economy. manuscript, presented to *Colloquium on Modern Tools for Business Cycle Analysis*, Luxembourg, 28-29 November 2002.

[2] Diebold FX, Lee JH, Weinbach GC, 1994. Regime Switching with Time-Varying Transition Probabilities. in *Nonstationary Time Series Analysis and Cointegration,* Hargreaves CP (ed); Oxford University Press 283-302.

[3] Engel C, 1994. Can the Markov Switching Model Forecast Exchange Rates?. *Journal of International Economics* **36**: 151-165.

[4] Engel C, Hamilton JD, 1990. Long Swings in the Dollar: Are they in the Data and Do Markets Know It?. *American Economic Review* **80**: 689-713

[5] Filardo AJ, 1994. Business-Cycle Phases and Their Transitional Dynamics. *Journal of Business and Economic Statistics* **12**: 299-308.

[6] Garcia R, 1998. Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models. *International Economic Review* **39**: 763-88.

[7] Garcia R, Schaller H, 2002. Are the Effects of Monetary Policy Asymmetric? *Economic Inquiry* **40**: 102-119.

[8] Gómez V, Maravall A, 1996. Programs SEATS and TRAMO: instructions for the user. *Working Paper* No. 9628, Bank of Spain.

[9] Greene WH, 2000. *Econometric Analysis*, 4nd edn. Prentice-Hall: Upper Saddle River.

[10] Hamilton JD, 1989. A New Approach to the Economic Analysis of Time Series in the Business Cycle. *Econometrica* **57**: 357-384.

[11] Hamilton JD, 1990. Analysis of Time Series Subject to Changes in Regime. *Journal of Econometrics* **45**: 39-70.

[12] Hamilton JD, 1994. *Time Series Analysis*, Princeton University Press: Princeton.

[13] Hansen BE, 1992. The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP. *Journal of Applied Econometrics* **7**, S61:S82.

[14] Hansen BE, 1996. Erratum: the Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP. *Journal of Applied Econometrics* **11** 195-198.

[15] Kim CJ, 1994. Dynamic Linear Models with Markov-Switching. *Journal of Econometrics* **60**: 1-22.

[16] Kim CJ, Nelson CR 1998. Business Cycle Turning Points, A New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime-Switching. *Review of Economics and Economic Statistics* **80**: 188-201.

[17] Kim CJ, Nelson CR, 1999. *State-Space Models with Regime Switching.* MIT Press: Cambridge.

[18] Krolzig HM, 1997. *Markov Switching Vector Autoregressions. Modelling, Statistical Inference and Application to Business Cycle Analysis.* Springer: Berlin.

[19] Otranto E, Gallo GM, 2002. A Nonparametric Bayesian Approach to Detect the Number of Regimes in Markov Switching Models. *Econometric Reviews* **21**: 477-496.

Author's biography:

Edoardo Otranto is researcher in Statistics at University of Sassari (Italy). He holds a PhD in Statistics from University of Bari. His main scientific interests are time series analysis, signal extraction, nonparametric statistics.