C O N T R I B U T I   D I   R I C E R C A   C R E N O S

ADDING FLEXIBILITY TO MARKOV SWITCHING MODELS

Edoardo Otranto

WORKING PAPERS

2015/09

CUEC

Centro Ricerche Economiche Nord Sud
(CRENoS)
Università di Cagliari
Università di Sassari

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.it
info@crenos.it

CRENoS – Cagliari
Via San Giorgio 12, I-09100 Cagliari, Italia
tel. +39-070-6756406; fax +39-070- 6756402

CRENoS - Sassari
Via Torre Tonda 34, I-07100 Sassari, Italia
tel. +39-079-213536; fax +39-079-213002

Title: ADDING FLEXIBILITY TO MARKOV SWITCHING MODELS

ISBN: 978 88 8467 939 0

First Edition: August 2015

Adding Flexibility to Markov Switching Models

Edoardo Otranto

Department of Cognitive Sciences, Educational and Cultural Studies and CRENoS,

Università di Messina, Via Concezione, 6 - 98121 Messina, Italy (*eotranto@unime.it*)

Abstract

Very often time series are subject to abrupt changes in the level, which are generally represented by Markov Switching (MS) models, hypothesizing that the level is constant within a certain state (regime). This is not a realistic framework because in the same regime the level could change with minor jumps with respect to a change of state; this is a typical situation in many economic time series, such as the Gross Domestic Product or the volatility of financial markets. We propose to make the state *flexible*, introducing a very general model which provides oscillations of the level of the time series within each state of the MS model; these movements are driven by a forcing variable. The flexibility of the model allows for consideration of extreme jumps in a parsimonious way (also in the simplest 2-state case), without the adoption of a larger number of regimes; moreover this model increases the interpretability and fitting of the data with respect to the analogous MS model. This approach can be applied in several fields, also using unobservable data. We show its advantages in three distinct applications, involving macroeconomic variables, volatilities of financial markets and conditional correlations.

Keywords: abrupt changes; goodness of fit; Hamilton filter; smoothed changes; time–varying parameters

JEL Classification: C22, C32, C5, C58

1 Introduction

The nonlinear behavior of many economic time series characterized by abrupt changes in the level has been object of several studies in the last decades, favouring the development of switching regime models; see Hamilton (2015) for a review of this kind of models in a macroeconomic framework, Franses and Van Dijk (2000) in the case of financial time series. In particular, the change in the level of the time series happens in unknown points along the time; this characteristic has favored the large success of models providing an estimation of the change-point and an inference on the regime. Among them we recall the Markov Switching Autoregressive (MS–AR) model of Hamilton (1989 and 1990),[1] the

[1]The AR–MS model was originally proposed by Lindgren (1978), but its diffusion and success in the economic framework is due to the works of Hamilton.

multiple change-point model of Chib (1998), the smooth transition model of Teräsvirta (2004) and their extensions.

The simple form of the MS model, its clear interpretation and the possibility to infer on the regime using the Hamilton (1990) filter, justify its enormous success in the statistical and econometric literature and its application in several fields of economics. In general the MS model provides a good fitting, capturing the presence of abrupt changes, but the consideration of fixed coefficients in correspondence of the different states could be a rigid constraint. This problem could be solved with a larger number of states, which involves computational problems (high computational time, convergence, identification, Markov chains with zero elements). Also, the identification of the number of states k is an open problem; very often a 2-state MS model is adopted to avoid the loss of efficiency in estimation due to the small possible number of observations falling in a certain state. Similarly, a higher number of states could imply the absence of transitions from a state to another one (there are not cases in which the state changes from regime i to regime j), with the corresponding coefficients of the transition probability matrix identically equal to zero (see, for example, Hamilton and Susmel, 1994).

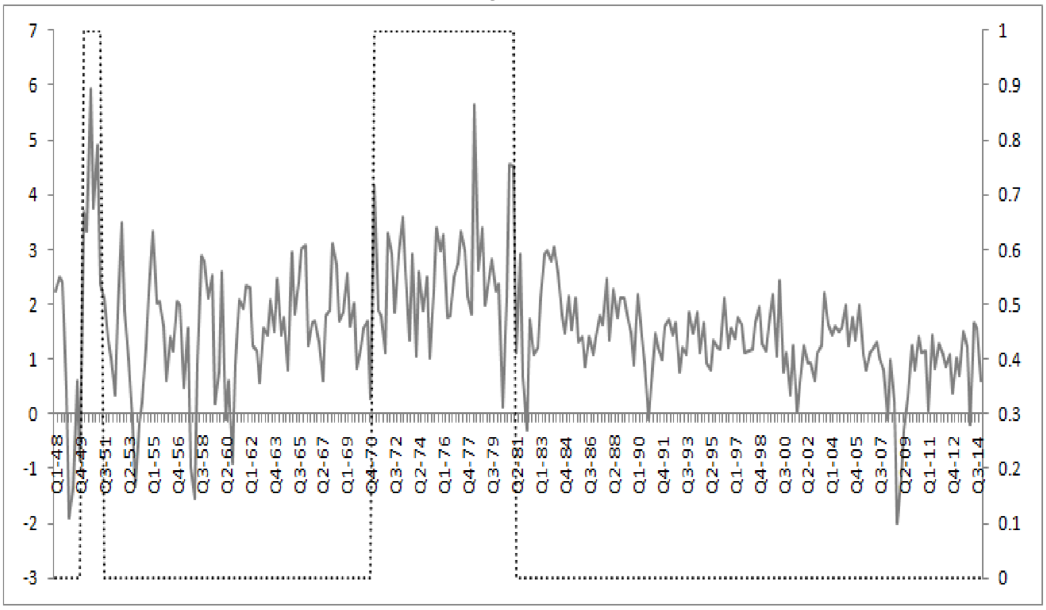
A graphical example would help to understand the main motivations of this work. Let us observe the first differences of the log quarterly U.S. Gross Domestic Product (GDP) series in Figure 1. Representing this series with a MS model with two states[2] we obtain the inference on the regime (denoted with the labels 0 and 1) illustrated by the dotted line; in practice two periods are identified in state 1, corresponding to the highest peaks in the time span considered: the boom in 1950 after the second World War and the decade 1971-1980. Notice the different behavior of these intervals; in the first one the levels of growth are high in the full period, whereas in the second interval there are frequent oscillations. In the last case the assignment to regime 1 is due to the high peaks (similar to those of 1950) approximately at the extremes of the interval; in particular, in correspondence of Q1 71, Q2 78, Q4 80 and Q1 81, the level of the series is more than 4, whereas in the other dates it is similar to the average level of the series in state 0.

It is likely that a better fitting could be obtained providing the possibility to change the parameters within the regimes, adding a certain flexibility to the MS model. In practice in Figure 1 a certain gain in fitting could be obtained if we allow a change in level within the regimes, with the possibility to isolate in a single state the highest peaks of the series. For this purpose we propose a new model, called Flexible State Markov Switching (FSMS) model, considering time varying coefficients within each state. The within–state dynamics of the coefficients is driven by forcing variables, which can be also non observable. We develop a very general framework, extending different MS models and providing examples of their applications.

The paper is structured in the following way: in the next section we will describe the new model proposed, underlying how the estimation procedure can follow the steps proposed in Hamilton (1990). Section 3 is devoted to three examples of application of the FSMS model, dealing with the U.S. GDP, the volatility of the Nasdaq 100 index and the conditional correlations of the components of the Dow Jones Industrial Average index, comparing our model with the corresponding *classical* MS model. Some final remarks

[2]We will illustrate it more in detail in section 3.1.

Figure 1: First differences of the logarithms of Quarterly U.S. GDP series (gray continuous line, left axis) from QI:1948 to QIV:2014 and inference on the regime using an AR(2) MS model with two states (dotted line, right axis).



will conclude the paper.

2 The Flexible State Markov Switching Model

Let y_t the variable object of study. It can be a single variable or a vector of variables; moreover it is not necessary that y_t is observed. In a MS framework, we consider different models in correspondence of the different states; these models have the same structural form and differ only for the value of some coefficients (called the switching coefficients). We call $m_{i,t}$ the model representing y_t in state i. Formally, in a MS model with two states we have:

$$y_t = (1-s_t)m_{0,t} + s_t m_{1,t}$$
$$\boldsymbol{P} = \begin{bmatrix} p_{00} & p_{01} \\ \\ p_{10} & p_{11} \end{bmatrix} \tag{2.1}$$

where $s_t = 0, 1$ represents the state at time t and $\boldsymbol{P}$ is the transition probability matrix relative to an ergodic Markov Chain with elements $p_{ji} = Pr(s_t = i|s_{t-1} = j)$ and, when $i \neq j$, $p_{ji} = 1 - p_{ii}$.

We consider the possibility that, within each regime, the model can vary in a range of models which differ for the values of the switching parameters (representing the level of the series y_t) in the following way:

$$y_t = (1-s_t)\left[m^*_{0,t} f_{0,t} + m^*_{1,t}(1-f_{0,t})\right] + s_t\left[m^*_{1,t} f_{1,t} + m^*_{2,t}(1-f_{1,t})\right] \tag{2.2}$$

In practice in state 0 the model is a weighted mean of $m^*_{0,t}$ and $m^*_{1,t}$, with weights derived from the function $f_{0,t}$ ranging in $[0, 1]$; in state 1 the model is a weighted mean of $m^*_{1,t}$

and $m^*_{2,t}$, with weights derived from the function $f_{1,t}$ (again ranging in $[0,1]$). The level increases when the model changes from $m^*_{0,t}$ to $m^*_{1,t}$ and from $m^*_{1,t}$ to $m^*_{2,t}$, so we need adequate reparameterizations to obtain this constraint. Each $m^*_{i,t}$ ($i = 0,1,2$) depends on a set of common coefficients $\boldsymbol{\varphi}$ and a set of specific coefficients $\boldsymbol{\vartheta}_i$ (the switching coefficients). The functions $f_{0,t}$ and $f_{1,t}$ depend on a variable x_t (the *forcing* variable), not necessarily observed, and a different set of coefficients $\boldsymbol{\zeta}_h$ ($h = 0,1$). They can assume different specifications; in our experience good representations can be given by smooth transition functions and logistic functions. We call them *within state dynamic* (wsd) functions.

The estimation of (2.2) does not involve particular problems, including $\boldsymbol{\zeta}_0$ and $\boldsymbol{\zeta}_1$ among the set of unknown coefficients. In practice the density of y_t is a mixture of four densities $f(y_{s_t,s_{t-1},t}|\Im_{t-1},\boldsymbol{\pi})$ obtained from the four possible combinations of s_t and s_{t-1}, where $\Im_t$ represents the amount of information available at time t and $\boldsymbol{\pi} = (\boldsymbol{\varphi}', \boldsymbol{\vartheta}'_0, \boldsymbol{\vartheta}'_1, \boldsymbol{\vartheta}'_2, \boldsymbol{\zeta}'_0, \boldsymbol{\zeta}'_1)'$. The weights of each component of the mixture are given by the probabilities $Pr(s_t, s_{t-1}|\Im_{t-1})$, derived from the Hamilton (1990) filter. More in detail, the iterative steps (from $t = 1$ until $t = T$) to obtain these probabilities are:

1. $Pr(s_t = i, s_{t-1} = j|\Im_{t-1}) = p_{ji} Pr(s_{t-1} = j|\Im_{t-1})$;
2. $f(y_t|\Im_{t-1},\boldsymbol{\pi}) = \sum_{i=1}^{2}\sum_{j=1}^{2} Pr(s_t = i, s_{t-1} = j|\Im_{t-1}) f(y_{i,j,t}|\Im_{t-1},\boldsymbol{\pi})$;
3. $Pr(s_t = i, s_{t-1} = j|\Im_t) = \frac{Pr(s_t=i,s_{t-1}=j|\Im_{t-1}) f(y_t|s_t=i,s_{t-1}=j,\Im_{t-1},\boldsymbol{\pi})}{f(y_t|\Im_{t-1},\boldsymbol{\pi})}$;
4. $Pr(s_t = i|\Im_t) = \sum_{j=1}^{2} Pr(s_t = i, s_{t-1} = j|\Im_t)$

The starting probability $Pr(s_0 = j)$ to be used in the first step at the first iteration can be given by the ergodic probability $\boldsymbol{p}_e$ such that $\boldsymbol{P}'\boldsymbol{p}_e = \boldsymbol{p}_e$. The previous scheme is the basic Hamilton filter to obtain the filtered probabilities at each time; it is possible to obtain an inference on the regimes using the full available information $\Im_T$ from the Kim (1994) algorithm, starting from the results of the Hamilton filter. It consists in iterating the following two steps (from $T-1$ to 1):

1. $Pr(s_t = j, s_{t+1} = i|\Im_T) = \frac{Pr(s_{t+1}=i|\Im_T) Pr(s_t=j|\Im_t) p_{ji}}{Pr(s_{t+1}=i|\Im_t)}$
2. $Pr(s_t = j|\Im_T) = \sum_{i=1}^{2} Pr(s_t = j, s_{t+1} = i|\Im_T)$

The starting probability $Pr(s_T = i|\Im_T)$ is obtained at the final iteration of the Hamilton filter; from the same Hamilton filter we derive the probabilities $Pr(s_t = j|\Im_t)$ and $Pr(s_{t+1} = i|\Im_t) = \sum_{j=1}^{2} Pr(s_{t+1} = i, s_t = j|\Im_t)$ used in the step 1 of the Kim algorithm. The inference on the regime of Figure 1 is obtained from the smoothed probabilities of the MS(2)–AR(2) model, described in subsection 3.1, assigning the observations with $Pr(s_t = 0|\Im_T) > 0.5$ to regime 0 (regime 1 otherwise).

In the following section we will propose three different specifications of (2.2), which extend three well known MS models to include the flexible structure within the regimes.

3 Examples of Applications

As said, model (2.2) is expressed in a very general form and can include the extension of all MS models. In this section we propose the extension in three fields in which MS models are largely used. In the first case we extend the analysis of the U.S. GDP by an MS AR model, similar to the one adopted by Hamilton (1990); in the second case we consider the analysis of the volatility of the Nasdaq 100 U.S. index extending the MS–GARCH model (see, for example, Dueker, 1997); finally we extend the analysis of the correlations of the 30 assets compounding the Dow Jones Industrial Average index (DJ) extending the Regime Switching Dynamic Correlation (RSDC) model proposed by Pelletier (2006). Notice that in the first case the variable analyzed (GDP) is observed, whereas in the last two applications they are estimated by the model (the conditional variance in the second case and the conditional correlations in the third one). Also, to show the several potential approaches to estimate the FSMS model, we will adopt as forcing variables three different quantities: the estimated lagged level of GDP in the first application, the filtered probabilities of state 0 derived from the Hamilton filter in the second application, the forecasts of the exogenous VXD series[3] in the third application. Notice again that in the first and second case the forcing variable is not observed and it is obtained as a sub-product of the Maximum Likelihood estimation procedure of the model; in the third case it is obtained from an independent estimation procedure. Finally, we will adopt both the smooth transition function and the logistic function as swd functions; the differences are not relevant, so we will describe only one case for each example.

In all the examples the FSMS model is compared with the corresponding MS model and the corresponding model without regimes; the comparison is conducted in terms of log–likelihood functions and specific loss functions generally used in these three different frameworks.

3.1 FSMS–AR Model

Let us suppose that the variable of interest y_t is the U.S. GDP. In Figure 1 we have shown the behavior of the first differences of the log of GDP from the first quarter of 1948 to the fourth quarter of 2014 (source Federal Reserve Economic Data of St. Louis Bank). The MS–AR model adopted is the following:

$$\begin{aligned} m_{s_t,t} &= \mu_{s_t} + \phi_1(y_{t-1} - \mu_{s_{t-1}}) + \phi_2(y_{t-2} - \mu_{s_{t-2}}) + \varepsilon_t \quad s_t = 0,1 \\ \varepsilon_t &\sim IIN(0,\sigma^2) \end{aligned} \tag{3.1}$$

In the case of FSMS model the first equation in (3.1) is substituted by:

$$m^*_{s_t,t} = \mu_{s_t} + \phi_1(y_{t-1} - \mu^*_{t-1,s_{t-1}}) + \phi_2(y_{t-2} - \mu^*_{t-2,s_{t-2}}) + \varepsilon_t \quad s_t = 0,1,2 \tag{3.2}$$

where:

$$\mu^*_{t,s_t} = \begin{cases} \mu_0 f_{0,t} + \mu_1(1 - f_{0,t}) & if \quad s_t = 0 \\ \mu_1 f_{1,t} + \mu_2(1 - f_{1,t}) & if \quad s_t = 1 \end{cases}$$

[3]The VXD index is a proxy of the volatility of the DJ index, constructed in the same way as the well known VIX index, which is referred to the Standard & Poor 500 index.

Table 1: Parameter estimates (standard errors in parentheses) and evaluation criteria of AR(2), MS–AR(2) and FSMS–AR(2) models for the US GDP data set

	μ_0	μ_1	μ_2	ϕ_1	ϕ_2	σ	p_{00}	p_{11}
AR	1.722			0.053	-0.001	0.895		
	(22.842)			(0.094)	(0.002)	(0.082)		
MS	1.307	2.815		0.398	-0.040	0.892	0.985	0.935
	(0.096)	(0.354)		(0.072)	(0.014)	(0.048)	(0.008)	(0.026)
FSMS	0.069	2.402	4.760	0.452	0.071	0.806	0.984	0.626
	(1.354)	(6.717)	(1.109)	(0.317)	(0.241)	(0.061)	(0.007)	(0.565)
	pv_1	pv_5	pv_{10}	$Log-Lik$	AIC		MSE	U
AR	0.000	0.000	0.000	-453.99	3.405		1.193	1.556
MS	0.189	0.000	0.000	-366.95	2.780		0.921	0.551
FSMS	0.565	0.276	0.042	-349.38	2.687		0.873	0.550

Note: the symbol pv_i indicates the p-value of the Ljung-Box statistic to verify the presence of autocorrelation at lag i.

The only switching coefficient is the mean of the process μ_{s_t}; to ensure the increase in the level when changing the state we impose that $\mu_0 \leq \mu_1 \leq \mu_2$. Moreover the wsd function chosen is the smooth transition function:

$$f_{i,t} = [1 + \exp(-\gamma_i(\mu^*_{t-1} - c_i))]^{-1}, \text{with} \quad \gamma_i > 0 \quad i = 0, 1. \tag{3.3}$$

where

$$\mu^*_t = Pr(s_t = 0|\Im_t)\mu^*_{t,0} + Pr(s_t = 1|\Im_t)\mu^*_{t,1}$$

is the estimated mean at time t; we simply are hypothesizing that the most recent (estimated) value of the unconditional level of GDP can drive the movements in level within the states.

Table 1 shows the estimation of the parameters of the two alternative MS models, jointly with the linear AR(2) model, and some statistics. The three models have different estimates of the common coefficients; in particular the unconditional mean in regime 0 estimated with FSMS includes levels until 2.40, whereas in regime 1 it varies between 2.40 and 4.76; the constant unconditional mean obtained from the linear AR(2) model is equal to 1.72. In practice model FSMS considers in regime 1 only the highest peaks of the series, whereas, in the same regime, model MS considers the levels of the series around a mean equal to 2.81 with a larger standard deviation than the FSMS case. Moreover, the duration of regime 1 (obtained by $1/(1-p_{11})$) is in average 16.7 quarters with model MS and 2.7 quarters with model FSMS; this is a consequence of the interpretation of regime 1 of FSMS, which contains only the highest peaks. Of course, given this characteristic, only a few observations fall in state 1, explaining the large standard errors in some coefficients of model FSMS.

In Figure 2 we show the two estimated wsd functions; notice as both the functions are approximately constant around 0.43 and 0 respectively with abrupt jumps, reaching the maximum value 1 in correspondence of the highest peaks of GDP; this result implies that, when the state is 0, the estimated mean is around 1.3, as in the MS model, whereas near to μ_1 (if the state is 0) and μ_2 (if the state is 1) in correspondence of the highest peaks. This

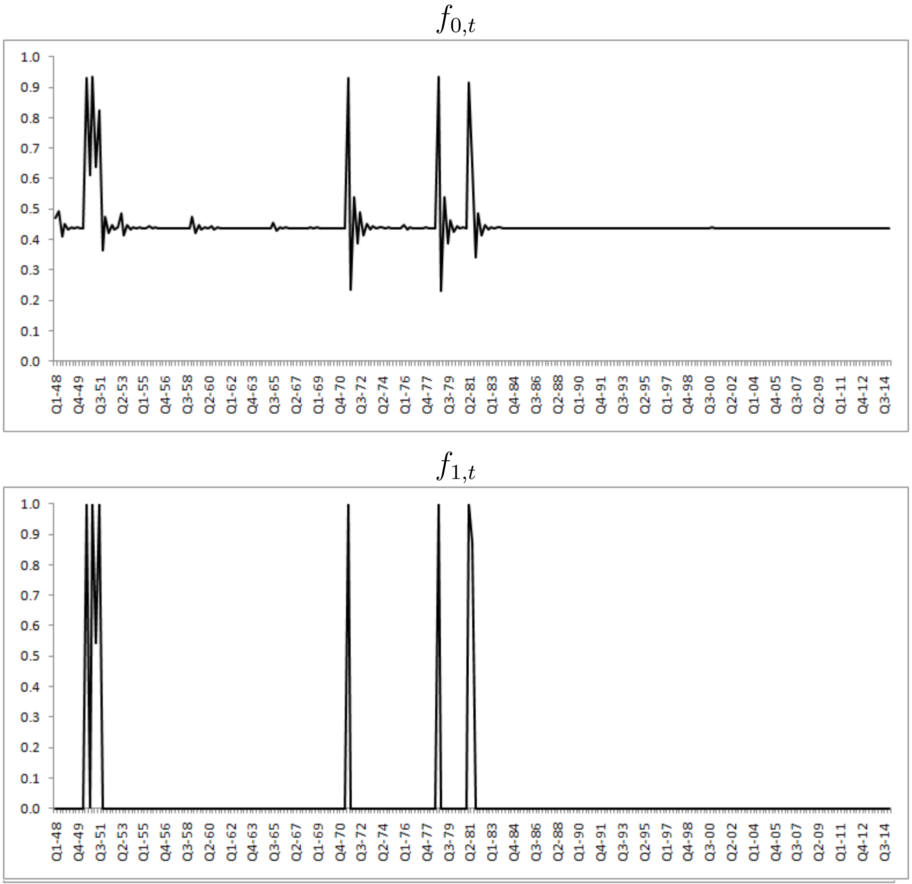


Figure 2: Smooth transition swd functions for the US GDP data set.

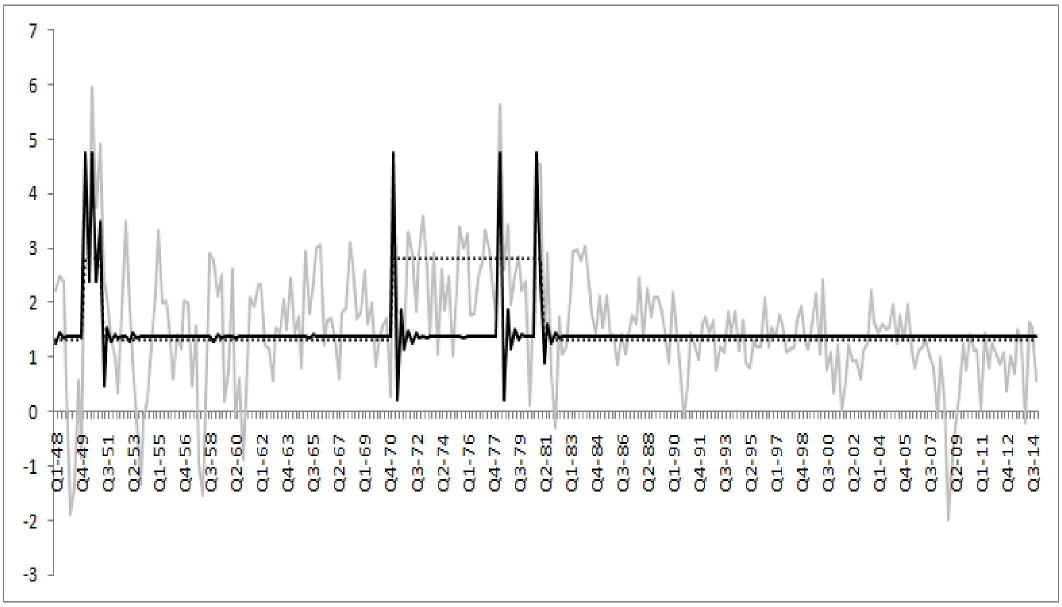
would be clearer observing Figure 3, where the GDP is plotted with the estimated means obtained by the two alternative models. Notice as the unconditional means are practically the same in both models, excluding the two sub-periods underlined in the Introduction. In 1950 the FSMS model shows some movements in the mean following the dynamics of the GDP series; in the period 1971-1980 only the peaks higher than 4 are assigned to state 1, whereas the rest, more coherently, belongs to state 0.

This major flexibility is confirmed in terms of fitting. In the bottom part of Table 1 we show some criteria to evaluate the model fitting. The FSMS model outperforms the others in terms of AIC; also the Mean Squared Error (MSE) decreases sensibly in the FSMS case with respect to the linear and the MS models. A specific loss function for the analysis of the GDP is expressed by the Theil U (see, for example, Greene, 2008), given by:

$$U = \sqrt{\frac{\sum_t (\Delta y_t - \Delta \hat{y}_t)^2}{\sum_t (\Delta y_t)^2}}$$

where $\hat{y}_t$ is the estimated y_t, $\Delta y_t = (y_t - y_{t-1})/y_{t-1}$ and $\Delta \hat{y}_t = (\hat{y}_t - y_{t-1})/y_{t-1}$. Theil U is particularly useful to detect the capability of the model to identify the turning points; better performances are expressed by lower values (as MSE). It is evident the better pre-

Figure 3: First differences of the logarithms of Quarterly U.S. GDP series (gray continuous line) from QI:1948 to QIV:2014 and estimated unconditional mean using a MS–AR(2) model with two states (dotted line) and a FSMS–AR(2) model with two states (black continuous line).



dictive accuracy, in terms of turning points, of the MS models respect to the linear one, with a small preference for the FSMS model respect to the MS one.

In the same Table we show the p-values of the Ljung-Box statistic for residuals in correspondence of lags 1, 5 and 10; we can notice as the corresponding test rejects the null hypothesis of autocorrelation for the linear model, after lag 1 for the MS model, whereas the FSMS seems not affected by residual autocorrelation.

In practice, the comparison of the three models for the US GDP data set seems to favor the FSMS model not only in terms of interpretability of the regimes but also in terms of fitting and diagnostic checking.

3.2 FMSM–GARCH Model

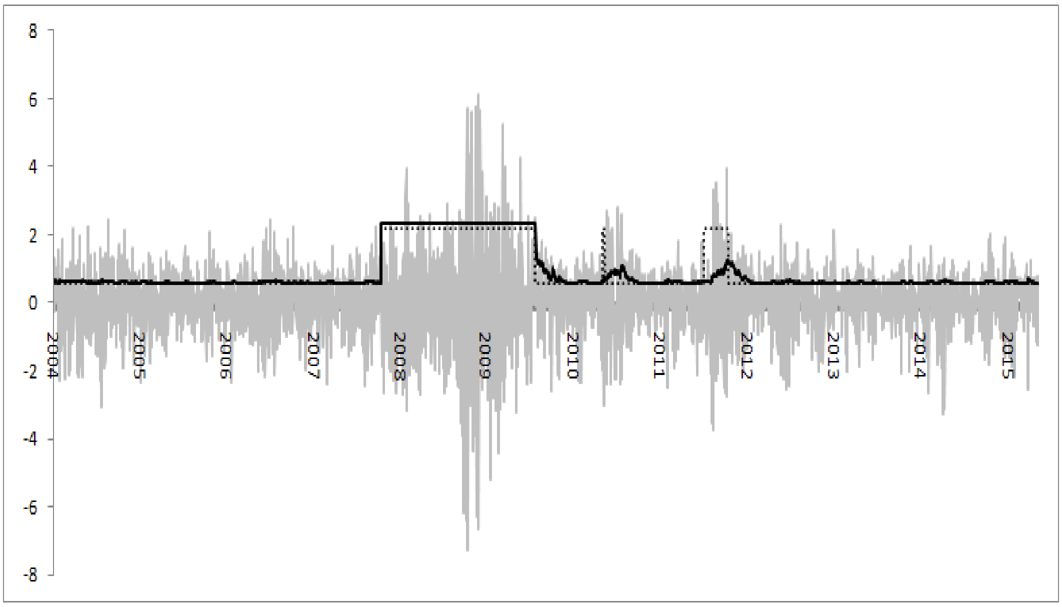
A recent large use of MS models concerns the analysis of the volatility of financial time series, in particular inserting a Markovian dynamics in the coefficients of the GARCH model (Bollerslev, 1986).

Let us consider the series of the returns[4] of the Nasdaq index from 2 January 2004 to 5 May 2015 (2840 daily observations; source Oxford-Man Institute realised library); the gray line of Figure 4 illustrates the dynamics of this series, where it is evident the high volatility (i.e. large conditional variance) period between October 2007 and July 2009, corresponding to the world financial crisis.[5] The model we adopt is a MS–GARCH(0,1)

[4] In financial analysis the return at time t is the logarithm of the ratio of the price of the asset at time t and the price of the asset at time $t-1$, multiplied by 100.

[5] The dates on the x axis of Figure 4 are put in correspondence of the beginning of each year.

Figure 4: Returns of the Nasdaq 100 index (gray continuous line) from 2 Jan 2004 to 5 May 2015 and estimated unconditional variance using a MS Asymmetric GARCH(0,1) model with two states (dotted line) and a FSMS Asymmetric GARCH(0,1) model with two states (black continuous line).



with asymmetric effects;[6] calling r_t the series of the returns, we hypothesize that the conditional distribution of r_t is given by:

$$r_t|\mathfrak{I}_{t-1} \sim N(\mu, y_t)$$

and the conditional variance y_t follows the structure (2.1) in the case of MS model and (2.2) in the case of FSMS model; the components of the FSMS models are given by:

$$\begin{aligned} m^*_{s_t,t} &= \omega_{s_t} + \beta y_{s_{t-1},t-1} + \gamma I(r_{t-1} < 0)\varepsilon^2_{t-1} \quad & s_t = 0, 1, 2 \\ f_{j,t} &= \frac{exp(a_j + b_j Pr(s_{t-1} = 0)|\mathfrak{I}_{t-1})}{1 + exp(a_j + b_j Pr(s_{t-1} = 0)|\mathfrak{I}_{t-1})} \quad & j = 0, 1 \end{aligned} \tag{3.4}$$

where $I(\cdot)$ is the indicator function which provides the asymmetric effects due to the sign of the return at time $t-1$.

In the MS case $m_{s_t,t}$ is the same of $m^*_{s_t,t}$, but $s_t = 0, 1$. In model (3.4) the wsd function is a logistic driven by the filtered probabilities of state 0 obtained, step by step, from the Hamilton filter during the estimation procedure.

In MS–GARCH models a path dependence problem arises, due to the unobservability of the volatility and the dependence on all the past values of the state s_t: at the end of the t-th step of the Hamilton filter, it would be necessary to keep track of all possible paths

[6]We have estimated, both in MS and FSMS cases, the usual GARCH(1,1) model, but the coefficients corresponding to the lagged effect of the returns were estimated equal to zero. The asymmetric effects are referred to the fact that the level of volatility in general increases in correspondence of the most recent negative returns (see Glosten et al., 1993).

Table 2: Parameter estimates (standard errors in parentheses) and evaluation criteria of Asymmetric GARCH(0,1) and MS and FSMS Asymmetric GARCH(0,1) models for the Nasdaq 100 data set

	μ	ω_0	ω_1	ω_2	β	γ	p_{00}	p_{11}
GARCH	0.024	0.035			0.893	0.107		
	(0.015)	(0.002)			(0.005)	(0.005)		
MS	0.014	0.037	0.138		0.873	0.127	0.999	0.994
	(0.002)	(0.003)	(0.008)		(0.007)	(0.007)	(0.000)	(0.000)
FSMS	0.013	0.023	0.152	0.152	0.870	0.130	0.999	0.997
	(0.013)	(0.001)	(0.009)		(0.007)	(0.007)	(0.000)	(0.000)
	pv_1	pv_5	pv_{10}	$Log-Lik$	AIC		MSE	ρ
GARCH	0.522	0.530	0.932	-3766.68	2.656		0.414	0.793
MS	0.541	0.505	0.890	-3743.43	2.642		0.403	0.798
FSMS	0.525	0.502	0.883	-3742.09	2.644		0.401	0.799

Note: the symbol pv_i indicates the p-value of the Ljung-Box statistic to verify the presence of autocorrelation at lag i. MSE compares the estimated conditional variance with the realized kernel variance (both standardized), whereas ρ is relative to their correlation.

taken by the regime until time t, making the model intractable. Several solutions were proposed in literature; see, for example, Gray (1996), Klaassen (2002), Haas et al. (2004). The proposal we adopt is similar to the one of Dueker (1997), who extends the method of Kim (1994), developed in a state-space framework with a simple approximation. The Kim approximation consists in collapsing at each step of the Hamilton filter the 4 possible values of the estimated volatility $y_{s_t,s_{t-1},t}$ into 2 values, by:

$$y_{s_t,t} = \frac{\sum_{j=1}^{2} Pr(s_t = i, s_{t-1} = j|\mathfrak{I}_t) y_{i,j,t}}{Pr(s_t = i|\mathfrak{I}_t)} \tag{3.5}$$

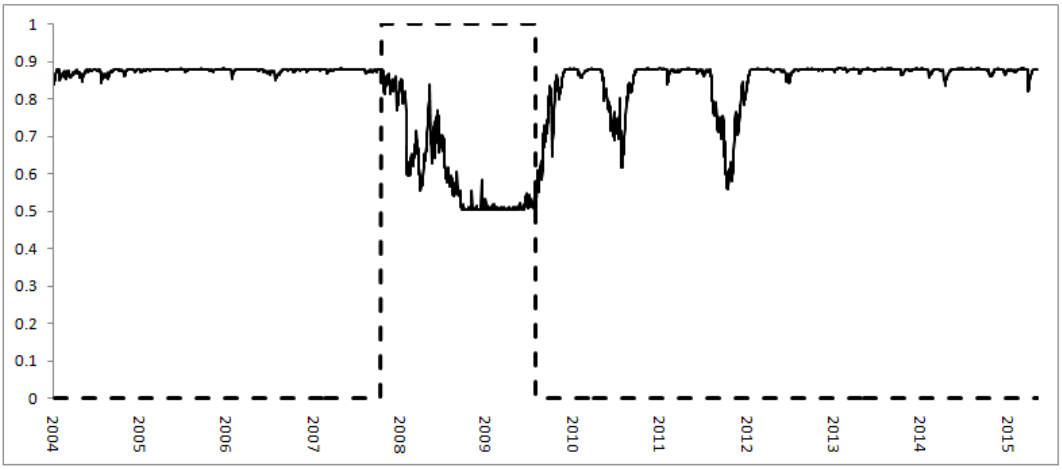
Kim (1994), using real data, and Gallo and Otranto (2015), via simulations in an asymmetric multiplicative error model framework, show that the Kim approximation provides good results, with decreasing bias when the sample size increases. This approximation can be used also for the FSMS model.

In Table 2 we show the estimates of the common coefficients, jointly with the results for the simple GARCH case; for the FSMS case we obtain that $\omega_1 = \omega_2$, so, to avoid overparameterization, we have fixed the wsd function $f_{1t} = 1$ for each t; in practice we impose that state 1 has a constant unconditional level of the volatility, whereas it can change within state 0.

Given the impossibility to observe the variance, the MSE is calculated comparing the estimated unconditional variance y_t with the realized kernel variance (Barndorff-Nielsen et al., 2008), a realized volatility estimator with the property of robustness to market microstructure noise;[7] both the variables are standardized to avoid scale problems. Moreover the correlation between the two variables could be an alternative indicator which bypasses the problems linked to the scale.

[7]The realized volatility is considered an unbiased estimator of the variance of financial markets with several desirable properties; see Andersen et al. (2000) and (2003).

Figure 5: Logistic swd function for state 0 of the Nasdaq 100 data set (continuous line) and inference on the state with FSMS–GARCH(0,1) model with two states (dotted line).



The MS and FSMS models outperform the simple GARCH model, but, differently from the GDP example, the estimations of the two alternative Markovian models provide very similar results, with uncorrelated residuals, similar behavior with respect to the realized kernel variance and levels of the unconditional volatility, which, for state i, is given by:

$$uv_i = \frac{\omega_i}{1 - \beta - \gamma/2} \tag{3.6}$$

In Figure 5 we show the behavior of the swd function for state 0 with the inference on the regime obtained from the FSMS model; it is almost constant around 0.87 for a large part of the span considered, implying an unconditional variance uv_0 around 0.61 (the linear combination of ω_0 and ω_1 with weights equal to f_{0t} and $(1 - f_{0t})$ respectively). The exceptions are the period of crisis October 2007-July 2009, which belongs to state 1 and has fixed unconditional variance $uv_1 = 2.33$, and two brief successive periods, the end of July 2010 and the end of 2011, in which there is an increase in volatility, with lower level with respect to the long previous crisis. We can observe in Figure 4 the unconditional volatilities (3.6) estimated with the two models; both the models assign the first long high volatility period to state 1 with similar estimated level, but MS assigns also the two brief successive periods to state 1, with the same level of unconditional volatility of October 2007–July 2009. More coherently, the FSMS model assigns these periods to state 0 but the levels have a clear increase with respect to the rest of state 0 with a smooth return to the previous level. Similarly, the switch to state 0 after the long crisis of October 2007-July 2009 is not abrupt, but gradual.

3.3 FSMS RSDC Model

A recent strand of the literature is devoted to the analysis of the time-varying conditional correlation of a set of financial time series; this is due to the empirical evidence that financial time series are subject to comovements and this relationship is more evident in the periods of high volatility (see, for example, Longin and Solnik, 1995). Several models have been developed to represent the dynamic conditional correlation, adopting some reparameterization to avoid the so-called *curse of dimensionality*: generally a large set of assets (or financial indices) are considered and the number of parameters is explosive. Feasible models were proposed by Engle (2002), Tse and Tsui (2002), Silvennoinen and Teräsvirta (2015), to cite just a few. The increase in correlation in correspondence of high volatility regimes can be developed inserting a MS dynamics in the conditional correlation models; this idea was developed, for example, by Billio and Caporin (2005), Pelletier (2006), Otranto (2010). In particular, Pelletier (2006) proposes the simple RSDC model, in which the full correlation matrix can switch from a regime to another one, without path dependence problems; referring to the notation in (2.1), y_t represents the conditional correlation matrix $\mathbf{R}_t$ and m_i is the conditional correlation matrix $\boldsymbol{R}_i$ in state i. Calling n the number of time series considered, each conditional correlation matrix will contain $n(n-1)/2$ coefficients, requiring a reparameterization to reduce the number of unknown parameters. We follow the proposal of Bauwens and Otranto (2015), adopting:

$$\begin{array}{l} \boldsymbol{R}_i = \bar{\boldsymbol{R}}\lambda_i + \boldsymbol{I}_n(1-\lambda_i) \quad i = 0, 1 \\ \lambda_0 \in [0, 1],\ \lambda_1 \in [1, 1/\bar{r}_{max}], \end{array} \tag{3.7}$$

where $\boldsymbol{I}_n$ is the $n \times n$ identity matrix, $\bar{r}_{max}(> 0)$ is the maximum correlation coefficient in the sample correlation matrix $\bar{\boldsymbol{R}}$. We extend the RSDC model to the FSMS case; we adopt a reparameterization as (3.7) with $i = 0, 1, 2$ with the constraints $0 \leq \lambda_0 \leq \lambda_1 \leq 1$ and $\lambda_2 \in [1, 1/\bar{r}_{max}]$. The wsd functions are represented by smooth transition functions as (3.3), using the expected value of volatility at time t as forcing variable.

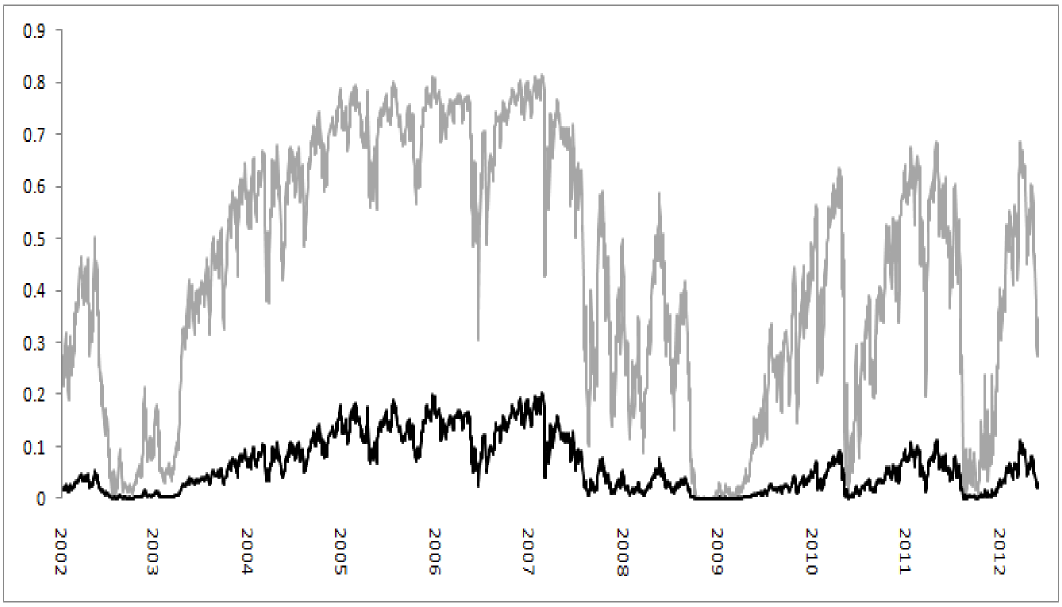
We apply the RSDC and the FSMS RSDC models to the same data set used by Bauwens and Otranto (2015), developing their Volatility Dependent Conditional Correlation (VDCC) models, relative to the 30 assets composing the DJ index from January 2, 2002 to May 23, 2012 (2617 daily returns for each series; source: Yahoo Finance). Also we consider, as a benchmark, the Constant Conditional Correlation (CCC) model of Bollerslev (1990), in which the correlation matrix is constant in the full span considered ($\boldsymbol{R}_t = \bar{\boldsymbol{R}}$). In their work Bauwens and Otranto (2015) hypothesize that the one-step ahead forecasts of the VXD index drive the transition probabilities of the RSDC model; in a similar way we use these forecasts, obtained by a 2–state MS AR(4) model, as forcing variable for the swd functions. We adopt the two-step estimation procedure of Engle (2002): in the first step we estimate 30 univariate asymmetric GARCH(1,1) models to represent the conditional variances of each asset, and in the second step we estimate the correlation parts on the standardized residuals derived from the first step.

In Table 3 we show the estimation results relative to the λ coefficients and the transition probabilities; the inference on the regime is very similar, given the similar estimation of the transition probabilities. Moreover the fitting, evaluated in terms of AIC, favors the FSMS model, which shows an increase in the likelihood function of 67 points respect to

Table 3: Parameter estimates (standard errors in parentheses) and evaluation criteria of MS–RSDC and FSMS–RSDC models for the DJ data set

	λ_0	λ_1	λ_2	p_{00}	p_{11}
MS	0.323	1.092		0.377	0.880
	(0.039)	(0.004)		(0.070)	(0.009)
FSMS	0.003	0.661	1.118	0.380	0.874
	(0.041)	(0.081)	(0.008)	(0.056)	(0.010)
		$Log-Lik$	AIC		VP
CCC		-18794.02	14.407		4.790
MS		-17283.43	13.252		4.788
FSMS		-17216.45	13.205		4.697

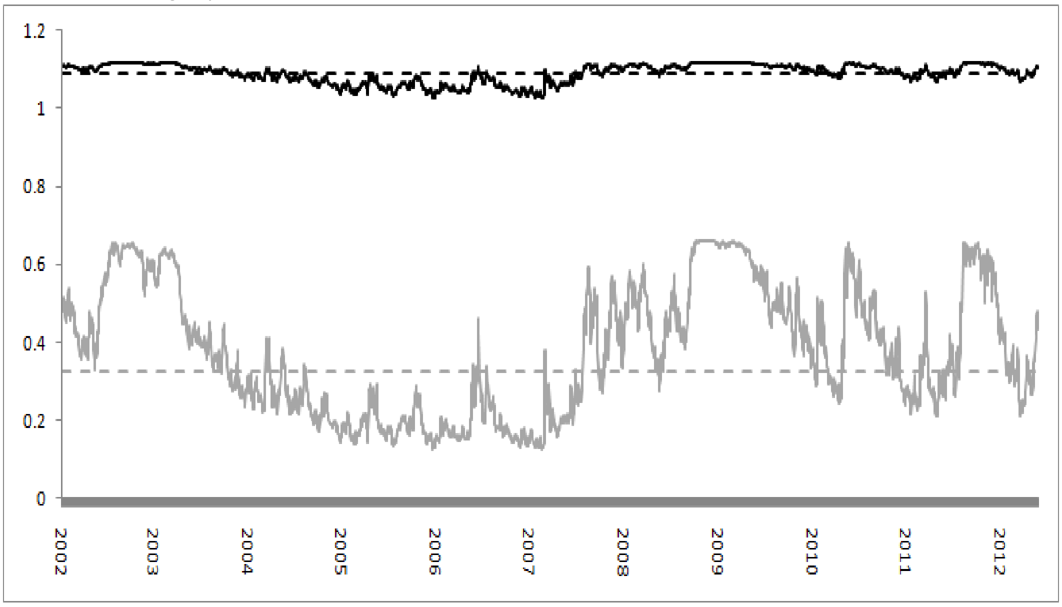
Figure 6: Smooth transition swd function for state 0 (gray line) and state 1 (black line) of the DJ data set.



the MS model and more than 1500 respect to the CCC model. In Figure 6 the behaviors of the two swd functions is illustrated; it is clear that regime 0 is subject to larger oscillations, as better illustrated in Figure 7, where the values of λ_0 and λ_1 of the MS model are compared with the FSMS time-varying values of $\lambda_0 f_{0,t} + \lambda_1(1 - f_{0,t})$ and $\lambda_1 f_{1,t} + \lambda_2(1 - f_{1,t})$ respectively.

Of course in this case it does not make sense to evaluate the performance of the models by MSE because the true correlation is unknown and there are not logical proxies, as in the volatility case. When dealing with financial markets, it is frequent to compare the performance of alternative models in terms of minimum variance portfolio rather than statistical criteria. Following Engle and Colacito (2006), we consider a set of 30 expected returns, each one equal to $1/\sqrt(30)$, and we construct two alternative portfolios with the 30 assets with weights minimizing the portfolio variance (see Markowitz, 1959); using the same expected returns and the same conditional variance for both the portfolios, the

Figure 7: Time-varying λ coefficients for state 0 (gray line) and state 1 (black line) of the DJ data set, derived from the FSMS–RSDC model; the dotted lines are in correspondence of the fixed λ_0 (gray) and λ_1 (black) derived from the MS–RSDC model.



differences will be due only to the conditional correlation matrices. The best model will be the one with minimum sample portfolio variance. The VP index in Table 3 represents the portfolio variance obtained from the two models; notice as the FSMS model performs better than MS and CCC (the last one performs better than the MS case). It is also interesting to note that FSMS portfolio performs better than the alternative RSDC models with transition probabilities proposed by Bauwens and Otranto (2015); in particular the so–called Time-Varying Transition Probability RSDC model had a VP equal to 4.789, whereas the Double–Chain RSDC model equal to 4.790 (see Bauwens and Otranto, 2015, for details about these models).

4 Final Remarks

We have proposed an extension of the MS model to allow the switching coefficients to vary within the states, increasing the flexibility of the model; as a consequence we are able to gain in interpretability of the dynamics of the time series and in terms of fitting. The framework adopted is very general, in the sense that we have illustrated our theory using a very general model, whereas the examples of Section 3 want to show as this general framework can be adapted to specific MS models and applications. In our applications we have considered as switching the coefficients that affect the levels of the time series (in particular the mean in the GDP example, the unconditional volatility in the Nasdaq example and the scale coefficients of the conditional correlation matrix in the DJ data set). This is made because the analysis of the levels can be made in graphical terms and provides a better appreciation of the interpretability of the new model; anyway it is possible to extend the analysis considering as switching and flexible also the other coefficients of the models, or a subset of flexible switching parameters and a subset of fixed switching parameters; in fact these cases are consistent with both equations (2.1) and (2.2), representing the general MS and FSMS models respectively.

We have considered only 2-state FSMS models; the formal extension to a larger number of states is not difficult, but, as in the MS case, it involves some computational problems, in particular when a few observations belong to a single state. Anyway, the flexibility added to the MS model changing the switching coefficients within the regimes provides the possibility to capture extreme jumps that, in a classical MS model, would belong to a further regime; in other words a FSMS model with two states would be sufficient to represent several real cases of series subject to change in regime and abrupt jumps.

The comparison with the classical MS models was performed in in–sample terms, given the characteristics that the FSMS models would capture (interpretability and fitting); an out–of–sample analysis is beyond the scope of this work, but it could be interesting to perform it, although there are no reasons to think that FSMS has a better out–of–sample performance than MS.

In the examples we have considered also the corresponding models without switching coefficients (AR, GARCH and CCC models) to verify the increase in fitting including the presence of regimes in the analysis; for the first two examples we have tried also to estimate Smooth Transition models (see Teräsvirta, 1994, for the AR case, and Teräsvirta, 2009, for the GARCH case). They do not show relevant movements in the parameters and, in practice, are equivalent to the corresponding AR and GARCH models illustrated in this work.

The changes within the states in FSMS models are driven by selected forcing variables; a nice characteristics is that these variables are not necessarily observed, but they could be derived also during the estimation process. For example, in the GARCH case we have used the filtered probabilities, derived from the Hamilton filter, to drive the within state dynamics of the unconditional volatility. This idea is particularly appealing when observed forcing variables are not available.

References

[1] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2000). Great Realizations. *Risk* 13, 105–108.

[2] Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica* 71, 579–625.

[3] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N. (2008). Designing Realised Kernels to Measure the Ex–post Variation of Equity Prices in the Presence of Noise. *Econometrica* 76, 1481–1536.

[4] Bauwens, L., Otranto, E. (2015). Modeling the Dependence of Conditional Correlations on Market Volatility. *Journal of Business and Economic Statistics*, forthcoming, DOI:10.1080/07350015.2015.1037882.

[5] Billio, M., Caporin, M. (2005). Multivariate Markov Switching Dynamic Conditional Correlation GARCH Representations for Contagion Analysis, *Statistical Methods and Applications* 14, 145–161.

[6] Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31, 307–327.

[7] Bollerslev, T. (1990). Modelling the Coherence in Short–run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *Review of Economics and Statistics* 72, 498–505.

[8] Chib, S. (1998). Estimation and Comparison of Multiple Change-Point Models, *Journal of Econometrics* 86, 221–241.

[9] Dueker, M.J. (1997). Markov Switching in GARCH Processes and Mean-reverting Stock-market Volatility. *Journal of Business and Economic Statistics* 15, 26–34.

[10] Engle, R., Colacito, R. (2006). Testing and Evaluating Dynamic Correlations for Asset Allocation, *Journal of Business and Economic Statistics* 22, 367–381.

[11] Engle, R. F. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics*, 20, 339–350.

[12] Franses, P.H. and D. van Dijk (2000). *Non-Linear Time Series Models in Empirical Finance*. Cambridge University Press.

[13] Gallo, G.M., Otranto, E. (2015). Forecasting Realized Volatility with Changing Average Volatility Levels. *International Journal of Forecasting* 31, 620–634.

[14] Glosten, L., Jagannathan, R., Runkle, D. (1993). On the Relation between Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance* 48, 1779–1801.

[15] Gray, S.F. (1996). Modeling the Conditional Distribution of Interest Rates as a Regime-switching Process. *Journal of Financial Economics*, 42, 27–62.

[16] Greene, W.H. (2008). *Econometric Analysis*. Prentice Hall, Englewood Cliffs.

[17] Haas, M., Mittnik, S., Paolella, M.S. (2004). A New Approach to Markov–Switching GARCH Models. *Journal of Financial Econometrics* 2, 493–530.

[18] Hamilton, J.D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 57, 357–384.

[19] Hamilton, J.D. (1990). Analysis of Time Series Subject to Changes in Regime, *Journal of Econometrics* 45, 39–70.

[20] Hamilton, J.D. (2015). Macroeconomic Regimes and Regime Shifts. Forthcoming in Taylor,J. and Uhlig, H. (Eds.) *Handbook of Macroeconomics*, Vol. 2.

[21] Hamilton, J.D., Susmel, R. (1994). Autoregressive Conditional Heteroskedasticity and Changes in Regime, *Journal of Econometrics* 64, 307–333.

[22] Kim, C. (1994). Dynamic Linear Models with Markov Switching. *Journal of Econometrics* 60, 1–22.

[23] Klaassen, F. (2002). Improving GARCH Volatility Forecasts with Regime-Switching GARCH. *Empirical Economics*, 27, 363–394.

[24] Longin, F., Solnik, B. (1995). Is the Correlation in International Equity Returns Constant: 1960–1990?. *Journal of International Money and Finance* 14, 3–26.

[25] Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Wiley, New York.

[26] Otranto, E. (2010). Asset Allocation Using Flexible Dynamic Correlation Models with Regime Switching. *Quantitative Finance* 10, 325–338.

[27] Pelletier, D. (2006). Regime Switching for Dynamic Correlations, *Journal of Econometrics*, 131, 445–473.

[28] Silvennoinen, A., Teräsvirta, T. (2015). Modelling Conditional Correlations of Asset Returns: A Smooth Transition Approach, *Econometric Reviews* 34, 174197.

[29] Teräsvirta, T. (2004). Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models, *Journal of the American Statistical Association* 89, 208–218.

[30] Teräsvirta, T. (2009). Recent Developments in GARCH Modeling. In Andersen, T.G., Davis, R.A., Kreiß, J.-P., Mikosch, T. (Eds.) *Handbook of Financial Time Series* 17–42. Springer-Verlag, Berlin Heidelberg.

[31] Tse, Y.K., Tsui, A.K.C. (2002). A Multivariate GARCH Model with Time–varying Correlations, *Journal of Business and Economic Statistics* 20, 351–362. White,

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.it

- **15/08** *Paul Belleflamme, Dimitri Paolini,* "Strategic Promotion and Release Decisions for Cultural Goods"
- **15/07** *Gianpiero Meloni, Dimitri Paolini, Manuela Pulina,* "The Great Beauty: Public Subsidies in the Italian Movie Industry"
- **15/06** *Elias Carroni, Berardino Cesi, Dimitri Paolini,* "Peer Group, Distance and tuition fees: when widening university participation is still better"
- **15/05** *Bianca Biagi, Maria Gabriela Ladu,* "Productivity and employment dynamics: new evidence from Italian regions"
- **15/04** *Luca De Benedictis, Anna Maria Pinna,* "Islands as 'bad geography'. Insularity, connectedness, trade costs and trade"
- **15/03** *Massimo Del Gatto, Carlo S. Mastinu,* "Geography, Cultural Remoteness and Economic Development: A Regional Analysis of the Economic Consequences of Insularity"
- **15/02** *Malika Hamadi, Andréas Heinen,* "Firm Performance when Ownership is very Concentrated: Evidence from a Semiparametric Panel"
- **15/01** *Gerardo Marletto, Francesca Mameli, Eleonora Pieralice,* "Top-down *and* Bottom-up. Testing a mixed approach to the generation of priorities for sustainable urban mobility"
- **14/14** *Fabio Cerina, Luca G. Deidda,* "Reward from public office and selection of politicians by parties"
- **14/13** *Roberta Melis, Alessandro Trudda,* "Mixed pension systems sustainability"
- **14/12** *Gerardo Marletto,* "Socio-technical dynamics and political institutions: A multilevel Darwinian framework of sustainability transitions"
- **14/11** *Andrea Pinna,* "Shall We Keep Early Diers Alive?"
- **14/10** *Gianpiero Meloni, Dimintri Paolini, Juan de Dios Tena,* "American Beauty: an analysis of U.S. movies revenues in the global market"
- **14/09** *Silvia Balia, Rinaldo Brau, Emanuela Marrocu,* "Free patient mobility is not a free lunch. Lessons from a decentralised NHS"
- **14/08** *Gerardo Marletto,* "A deliberative-participative procedure for sustainable urban mobility – Findings from a test in Bari (Italy)"
- **14/07** *Manuela Deidda,* "Insularity and economic development: a survey"
- **14/06** *Edoardo Otranto, Massimo Mucciardi, Pietro Bertuccelli,* "Spatial Effects in Dynamic Conditional Correlations"
- **14/05** *Francesco Quatraro, Stefano Usai,* "Are knowledge flows all alike? Evidence from European regions"
- **14/04** *Angelo Antoci, Fabio Sabatini, Mauro Sodini,* "Online and offline social participation and social poverty traps. Can social networks save human relations?"
- **14/03** *Anna Bussu, Claudio Detotto,* "The bi-directional relationship between gambling and addictive substances"
- **14/02** *Alessandro Fiori, Tadas Gudaitis,* "Optimal Individual Choice of Contribution to Second Pillar Pension System in Lithuania"

Finito di stampare nel mese di Dicembre 2015
Presso Centro Stampa dell'Università degli Studi di Cagliari
Via Università 40
09125 Cagliari

www.crenos.it

9 788884 679390