Dipartimento di Statistica
"Giuseppe Parenti"

Dipartimento di Statistica "G. Parenti" – Viale Morgagni 59 – 50134 Firenze - www.ds.unifi.it

WORKING PAPER 2012/02

Realized Volatility and Change of Regimes

Giampiero Gallo,
Edoardo Otranto

Università degli Studi
di Firenze

Realized Volatility and Change of Regimes*

Giampiero M. Gallo

Dipartimento di Statistica "G. Parenti", Università di Firenze

Via G. B. Morgagni, 59 - 50134 Firenze, Italy (*gallog@ds.unifi.it*)

Edoardo Otranto

Dipartimento di Scienze Cognitive e della Formazione and CRENoS,

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

This version: June 2012

Abstract

Persistence and occasional abrupt changes in the average level characterize the dynamics of high frequency based measures of volatility. Since the beginning of the 2000s, this pattern can be attributed to the dot com bubble, the quiet period of expansion of credit between 2003 and 2006 and then the harsh times after the burst of the subprime mortgage crisis. We conjecture that the inadequacy of many econometric volatility models (a very high level of estimated persistence, serially correlated residuals) can be solved with an adequate representation of such a pattern. We insert a Markovian dynamics in a Multiplicative Error Model to represent the conditional expectation of the realized volatility, allowing us to address the issues of a slow moving average level of volatility and of a different dynamics across regime. We apply the model to realized volatility of the S&P500 index and we gauge the usefulness of such an approach by a more interpretable persistence, better residual properties, and an increased goodness of fit.

Keywords: MEM models, regime switching, realized volatility, volatility persistence

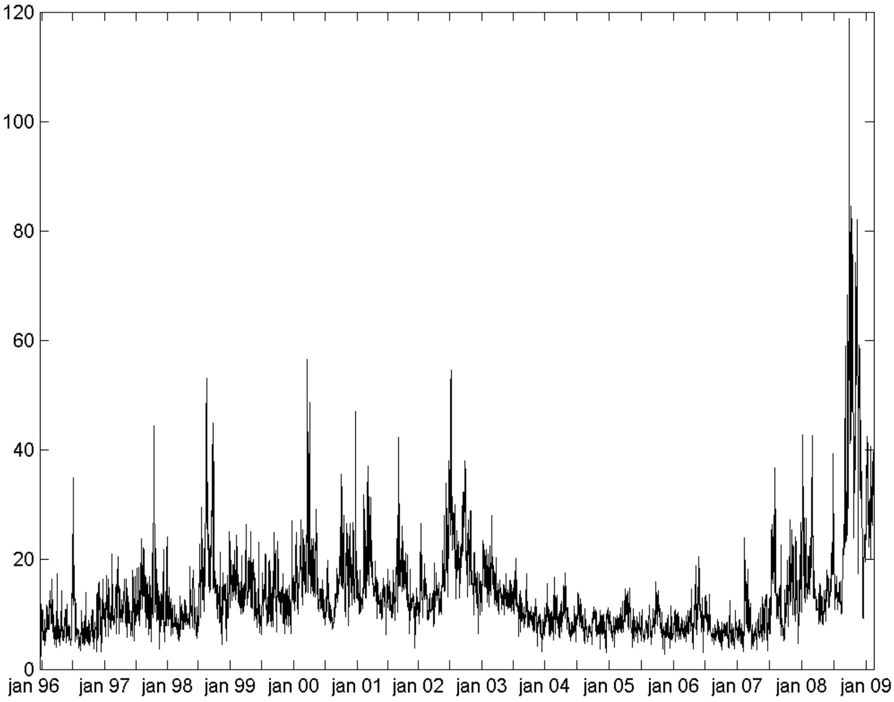
1 Introduction

A consolidated literature in financial econometrics is devoted to measuring asset volatility exploiting the information contained in asset price data collected at a very high frequency. The volatility estimators, known as realized volatility measures, have allowed for deeper insights into the dynamics of volatility, traditionally analyzed in a modeling and forecasting framework within the GARCH paradigm (Engle, 1982; Bollerslev, 1986; and further extensions – for a review, see Teräsvirta, 2009). Starting from the plain realized volatility, largely studied in Andersen et al. (2000, 2003), other measures have been introduced to take into consideration the presence of jumps, and other market microstructure issues (for a review, see Andersen et al., 2010). The most recent addition in the family of volatility estimators is the realized kernel volatility (developed by Barndoff-Nielsen et al., 2008), designed to possess robustness to market microstructure noise.

While volatility measurement from an end-of-day perspective has reached a mature stage, the question of volatility forecasting is still open, since there is still a wide debate as of which model is

*Thanks are due to participants in the following conferences: ERCIM 2010 (London, 10-12 December 2010), ECTS 2011 (Monte Porzio Catone, 13-14 June 2011), SCO 2011 (Padova, 19-21 September 2011). Financial support from Italian MIUR under Grant 20087Z4BMK_002 is gratefully acknowledged.

Figure 1: Realized kernel volatility of S&P500 index.



appropriate to characterize volatility dynamics, based on the current information available (cf. among others, Brownlees and Gallo, 2010; Shephard and Sheppard, 2010; Hansen et al., 2011; Cipollini et al., 2012). As a matter of fact, the visual inspection of the time series profile of high frequency based measures of volatility reveals a high degree of persistence: this is the case, for example, of the behavior of the kernel volatility of the S&P500 index, relative to the 500 most important US companies, shown in Figure 1 on a period between January 3, 1996 and February 27, 2009.[1] One can conjecture the presence of changing levels of the prevailing average volatility by subperiods: the series shows in fact alternating regimes which visually involve changes in the level but may also correspond to differences in the dependence in the series. This is particularly clear in the past ten years, with the turbulence leading to the burst of the tech bubble, the 2001 recession, the low level of volatility in mid decade and then the explosion of uncertainty following the subprime mortgage crisis.

The modeling effort which we will present in this paper aims at addressing the empirical regularities that ensue from the analysis of kernel volatility, keeping model diagnostics as a guideline for correct specification. The volatility series can be modeled as the product of a time-varying scale factor evolving autoregressively and a random disturbance (Multiplicative Error Models – MEM – developed by Engle, 2002, and expanded by Engle and Gallo, 2006); as it is applied to non–negative values, a MEM has the advantage of capturing dynamics without resorting to logs thus producing forecasts of volatility (and not of log–volatility).

For series such as the one at hand, the original MEM with an asymmetric response of volatility to the sign of lagged returns still suffers from residual autocorrelation. The question as of which features in the series need to be accommodated is addressed here by explicitly considering the existence of volatility regimes. With the introduction of a new member in the MEM family possessing Markov Switching properties (Hamilton, 1989, 1990), we want to represent the existence of different phases of volatility (quiet periods, turmoil periods and brief abnormal peaks), through changes in regimes

[1]Data are expressed as percentage annualized volatility, i.e. the square root of the realized variance series taken from the *Oxford-Man Institute's realised library* version 0.1 (Heber et al., 2009), and multiplied by $\sqrt{252} * 100$.

in the MEM data generating process. Our proposal is a MEM where Markov Switching parameters capture the presence of regimes with different dynamics. We try several specifications, starting from the most general and then imposing some constraints on the coefficients. Finally, further exploiting the information deriving from the sign of the lagged returns, we also propose a new time-varying probability Markov Switching model. Both sets of models provide several interesting interpretations deriving from the parameter estimation and a correct identification of the volatility related events, favoring a better fit relative to the base MEM.

Our work has several points of contact with various contributions in the literature: Markov Switching models have received an increasing attention in financial econometrics, starting from the SWARCH model, proposed by Hamilton and Susmel (1994), the MS GARCH model of Dueker (1997) and Klaassen (2002), or the recent multivariate extensions (Edward and Susmel, 2003, Higgs and Worthington, 2004, Gallo and Otranto, 2007, 2008). The issue of time varying underlying level of volatility is addressed by Engle and Rangel (2008), who adapt a spline function to capture a low frequency component of volatility (which they connect to macroeconomic factors). A different type of spline function, the P-Spline, is used by Brownlees and Gallo (2010) in a MEM context with the same motivation.

The paper is organized as follows: in the next section we introduce the Markov Switching extensions within the MEM framework. After discussing the data, in section 3 we show in detail the empirical results using the S&P500 stock index, illustrating the different performance with respect to the classical MEM model. Some final remarks will conclude the paper.

2 The Asymmetric Multiplicative Error Model with Markov Switching

The basic MEM idea is introduced in Engle (2002) and successively developed in Engle and Gallo (2006); Cipollini et al. (2012) suggest an extension to the multivariate case. The volatility x_t of a certain financial time series is modeled as the product of a time varying scale factor μ_t, representing the conditional mean of x_t, which follows a GARCH-type dynamics, and a positive valued error ε_t:

$$\begin{aligned} & x_t = \mu_t \varepsilon_t, \qquad \varepsilon_t | \Psi_{t-1} \sim Gamma(a, 1/a) \text{ for each } t \\ & \mu_t = \omega + \alpha x_{t-1} + \beta \mu_{t-1} + \gamma D_{t-1} x_{t-1} \\ & D_t = \begin{cases} 1 & \text{if } r_t < 0 \\ 0 & \text{if } r_t \geq 0 \end{cases} \end{aligned} \tag{2.1}$$

where Ψ_t represents the information available at time t. This base specification takes the presence of asymmetric responses of volatility to the sign of the returns (Engle and Gallo, 2006), similarly to the TGARCH model (Zakoïan, 1994) and the coefficient γ captures a stronger response to negative returns. We call this model *Asymmetric MEM* (AMEM); setting γ to zero gives us the simple MEM. Constraints can be imposed to ensure the positiveness of μ_t ($\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, $\gamma \geq 0$) and the stationarity of the process (persistence $(\alpha + \beta + \gamma/2)$ less than 1 under zero median of the returns). The Gamma distribution depends only on a single parameter a, providing a mean and a variance of the conditional error equal to 1 and $1/a$ respectively. Correspondingly, the conditional mean and variance of x_t are μ_t and μ_t^2/a respectively. Further lags can be added to the specification.

In this context, the unconditional mean of the volatility across the entire period is given by:

$$\mu = \frac{\omega}{1 - \alpha - \beta - \gamma/2}, \tag{2.2}$$

As noted in the introduction, this feature may be too restrictive for series such as our leading example, especially if we want to allow for sudden and persistent changes in the level of the series. In order to extend the capabilities of the model to capture extreme events which may have an impact on market dynamics, we introduce switching parameters that follow a discrete Markov chain. We define the Markov–Switching AMEM (MS–AMEM) as:

$$\begin{aligned} x_t = \mu_{t,s_t}\varepsilon_t, \qquad & \varepsilon_t|\Psi_{t-1} \sim Gamma(a_{s_t}, 1/a_{s_t}) \text{ for each } t \\ \mu_{t,s_t} = \omega + \textstyle\sum_{i=1}^{n} k_i I_{s_t} & + \alpha_{s_t} x_{t-1} + \beta_{s_t}\mu_{t-1,s_{t-1}} + \gamma_{s_t} D_{t-1} x_{t-1} \end{aligned} \tag{2.3}$$

where s_t is a discrete latent variable which ranges in $[1,\dots,n]$, representing the regime at time t. I_{s_t} is an indicator equal to 1 when $s_t \leq i$ and 0 otherwise; $k_i \geq 0$ and $k_1 = 0$. In other terms, the constant in regime j is given by $(\omega + \sum_{i=1}^{j} k_i)$. The changes in regime are driven by a Markov chain, such that:

$$Pr(s_t = j|s_{t-1} = i, s_{t-2}, \dots) = Pr(s_t = j|s_{t-1} = i) = p_{ij} \tag{2.4}$$

The positiveness and stationary constraints given for (2.1) hold within each regime in (2.3) as well.

It is useful to impose a particular reparameterization for β_{s_t} to guarantee a certain coherence between the regime and the level of volatility, ensuring that states of volatility (from low to high) increase with the regime identifier. For example, let us assume two regimes corresponding to low and high volatility, respectively, and that at time t an abrupt jump occurs in the volatility level; this would force a sizeable increase in the value of x_t and s_t will switch to 2. Let us now think that at time $t+1$ the level will stay in the same regime with $x_{t+1} \simeq x_t$; this event would not be captured by model (2.3), because the high value of x_{t+1} will correspond to a small value of the intercept $\omega + k_{s_{t+1}}$, pushing s_{t+1} to revert to 1. To allow for a correct identification of the state (cf. in Hamilton, 1990 in the context of business cycle analysis), it is more appropriate to subtract the mean of the expected volatility from each x_t in the second equation of (2.3):

$$\mu_{t,s_t} = \omega + \sum_{i=1}^{n} k_i I_{s_t} + \alpha^*_{s_t}(x_{t-1} - \mu_{t-1,s_{t-1}}) + \beta^*_{s_t}\mu_{t-1,s_{t-1}} + \gamma^*_{s_t} D_{t-1}(x_{t-1} - \mu_{t-1,s_{t-1}})$$

This model is equivalent to model (2.3), with unconditional expected value within state j, $j = 1,\dots,n$, equal to:

$$\mu_j = \frac{\omega + \sum_{i=1}^{j} k_i}{1 - \alpha^*_j - \gamma^*_j/2 - (\beta^*_j - \alpha^*_j - \gamma^*_j/2)}$$

so that, from (2.2), $\alpha_j = \alpha^*_j$, $\gamma_j = \gamma^*_j$, $\beta_j = \beta^*_j - \alpha_j - \gamma_j/2$. It is convenient to estimate β^*_j, with the constraint $(\alpha_j + \gamma_j/2) \leq \beta^*_j < 1$, and then obtain β_j from it. This constraint, together with the particular reparameterization of the constant in (2.3), achieves the desired property that the constant term itself increases with the volatility level.

The estimation of model (2.3) does not present particular problems because it can be performed along the lines of the Hamilton filter and smoother (Hamilton, 1994, ch.22). The only computational problem is due to the dependence of μ_{t,s_t} on s_{t-1}, its regime at time $t-1$. This means that, evaluating the likelihood recursively, we need to keep track of all possible paths taken by the regime between $t = 1$ and $t = T$, involving a non tractable model. A typical solution adopted in this case is the one proposed by Kim (1994), dealing with a similar problem with a state-space MS model (see the many examples shown in Kim and Nelson, 1999). After each step of the Hamilton filter, at time t we collapse the n^2 possible values of μ_t into n values, by an average over the probabilities at time $t-1$:

$$\hat{\mu}_{t,s_t} = \frac{\sum_{i=1}^{n} Pr[s_{t-1} = i, s_t = j|\Psi_t]\hat{\mu}_{t,s_{t-1},s_t}}{Pr[s_t = j|\Psi_t]} \tag{2.5}$$

where a hat indicates the estimate of the unknown variable and the probabilities in (2.5) are obtained by the Hamilton filter.

Finally, the asymmetry deriving from the sign of the returns may affect not only the average level within a certain regime, but also the transition probabilities. For this purpose, we suggest an extension to the model, called *Asymmetry in Probability MS–AMEM* (AsyP–MS–AMEM), with the same expression (2.3) for the level of volatility, whereas the transition probabilities are given by:

$$Pr(s_t|s_{t-1}, r_{t-1}) = p_{ij,t} = \begin{cases} p_{ij}^{-} & \text{if } r_{t-1} < 0 \\ p_{ij}^{+} & \text{if } r_{t-1} \geq 0 \end{cases} \tag{2.6}$$

A possible reparameterization of the transition probabilities could be made using a multinomial logit:

$$\begin{aligned} p_{ij,t} &= \frac{\exp(\phi_{ij}+\vartheta_{ij}D_{t-1})}{1+\sum_{h=1}^{n-1}\exp(\phi_{ih}+\vartheta_{ih}D_{t-1})} \\ p_{in,t} &= 1-\sum_{j=1}^{n-1} p_{ij,t} \\ i &= 1,...,n; \quad j = 1,...n-1 \end{aligned} \tag{2.7}$$

with $p_{ij,t} = p_{ij}^{-}$ when $D_{t-1} = 1$ and $p_{ij,t} = p_{ij}^{+}$ when $D_{t-1} = 0$. Expression (2.7) can be considered a particular time-varying transition probability MS model, proposed in the econometric literature by Filardo (1994) and Diebold et al. (1994). The advantage of this reparameterization is that the MS–AMEM can be nested into the AsyP–MS–AMEM, by constraining all the coefficients ϑ_{ij} to be 0, and a comparison can be made via a likelihood ratio test. Tests based on the likelihood function cannot be used to compare the MEM and AMEM with respect to the corresponding MS models because of the presence of nuisance parameters present only under the alternative hypothesis; in this case, with the proper caution, a classical BIC and AIC could provide some information (see Psaradakis and Spagnolo, 2003); in particular the AIC seems to choose the correct state dimension more successfully than the BIC, provided that the parameter changes are not too small and the hidden Markov chain is fairly persistent.

3 Regimes in the Volatility of the S&P500 Index

As a complement to the profile of the S&P500 volatility series shown in Figure 1 before, the descriptive statistics in Table 1 confirm the compatibility with the presence of regimes, especially a very large range (touching almost 120% at the maximum relative to a mean of about 13%) with a thick right tail (high kurtosis). Time dependence is reflected in the autocorrelation function, reported at lag 1 (one day), 5(one week), 22 (one month), which is characterized by slowly declining high values, a fact typically seen as evidence of the presence of regimes (as an alternative to a long memory explanation, not discussed here).

3.1 Model Specification

We start from estimating the benchmark models MEM and AMEM, which will be used as guidelines in what follows. The estimation results are reported in Table 2. We pinpoint to a small but statistically significant threshold coefficient γ. We calculated the information criteria and some loss functions of interest, namely, the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE) (a comparison based on the levels of volatility) and Theil's U (the latter calculated using the first differences of observed and forecasted data to detect the capability of the model to capture the turning points). On all accounts the AMEM is favored, revealing that the presence of asymmetric effects must be taken into account. The estimated persistence, measured as $(\alpha + \beta)$ for the MEM, and $(\alpha + \beta + \gamma/2)$ for the AMEM, is very high and around 0.94 for both models.

Table 1: Descriptive statistics for S&P500 realized kernel volatility (in annualized percentage terms). Sample: January 3, 1996 to February 27, 2009.

Mean	13.363
Median	11.392
Min	2.387
Max	118.75
St.Dev.	8.611
Skewness	3.415
Kurtosis	20.132
ACF(1)	0.821
ACF(5)	0.698
ACF(22)	0.525

Table 2: S&P500 realized kernel volatility (in annualized percentage terms). Coefficient estimates for the two benchmark models MEM and AMEM (standard errors in parentheses), together with likelihood-based criteria and loss functions. Sample: January 3, 1996 to February 27, 2009.

	ω	α	β	γ	a	
MEM	0.721	0.412	0.532		13.444	
	(0.068)	(0.028)	(0.032)		(0.427)	
AMEM	0.690	0.312	0.591	0.083	13.842	
	(0.078)	(0.031)	(0.035)	(0.008)	(0.452)	
	log-lik	AIC	BIC	RMSE	MAE	Theil-U
MEM	-8438.48	5.175	5.182	4.591	2.873	0.390
AMEM	-8389.66	5.145	5.155	4.490	2.811	0.381

As with other MS models, it is crucial to avoid over–parameterization in order to achieve a parsimonious representation with reliable parameter estimates: we start from the most general specification (2.3)-(2.6) with 3 regimes, involving 27 coefficients.[2] In Table 3 we show the estimation results for the parameters present in equation (2.3), whereas in Table 4 we report the estimated transition probabilities leaving to Section 3.2 some comments about the dynamics and the interpretation of the regimes. The differences of the transition probabilities p_{ij}^{+} and p_{ij}^{-} are not so large so as to suggest a formal Wald test on the equality restrictions. The results are shown at the bottom of Table 4 in the form of p–values associated with the null hypothesis reported. The only hypothesis that is rejected is $p_{21}^{+} = p_{21}^{-}$ and $p_{22}^{+} = p_{22}^{-}$ at the significance level of 5%. With this in mind, the evidence (cf. also how similar estimated parameters of MS–AMEM and the AsyP–MS–AMEM are from Table 3) points to no major asymmetry in the transition probabilities and thus favors the MS–AMEM. With an eye to a simplification of the specification, we can calculate the Wald test statistics for the joint hypotheses in the MS–AMEM:

$$\alpha_i = \alpha_j$$
$$\beta_i = \beta_j$$

for each $i, j = 1, 2, 3$ and $i \neq j$. The corresponding p-values are 0.09 for $(i, j) = (1, 2)$ and 0.00 for $(i, j) = (1, 3)$ and $(2, 3)$. On this basis we suggest a new model constrained to have the same MEM dynamics in regimes 1 and 2; the estimation results are shown in the last column of Tables 3 and 4. We label this model as MS–AMEM(c), with c as a reference to it being *constrained*. LR test, AIC and BIC favor this last model relative to the previous two models, with very similar values of the loss functions. As a further check, we have also estimated a model with the same MEM dynamics

[2]For the sake of completeness, we have also considered the case of 2 regimes; the outcome is in favor of the models with three regimes, based on the AIC and BIC.

Table 3: S&P500 realized kernel volatility (in annualized percentage terms). Coefficient estimates for Markov Switching MEM specifications with three regimes and Gamma innovations: i. AsyP–MS–AMEM, ii. MS–AMEM and iii. MS–AMEM with parameters on the volatility dynamics constrained to be the same across regimes 1 and 2 (standard errors in parentheses). Likelihood-based criteria and in sample forecasting performance are reported in the lower portion. Sample: January 3, 1996 to February 27, 2009.

	AsyP–MS–AMEM	MS–AMEM	MS–AMEM(c)
ω	1.858	1.872	1.731
	(0.163)	(0.229)	(0.121)
k_2	0.645	0.685	1.040
	(1.128)	(0.228)	(0.038)
k_3	5.103	5.188	5.005
	(1.880)	(1.326)	(1.276)
α_1	0.200	0.199	0.180
	(0.020)	(0.028)	(0.015)
α_2	0.158	0.161	
	(0.079)	(0.029)	
α_3	0.259	0.257	0.259
	(0.058)	(0.058)	(0.051)
β_1	0.524	0.525	0.561
	(0.040)	(0.056)	(0.028)
β_2	0.602	0.594	
	(0.155)	(0.058)	
β_3	0.345	0.343	0.339
	(0.111)	(0.108)	(0.093)
γ_1	0.075	0.076	0.080
	(0.002)	(0.007)	(0.010)
γ_2	0.082	0.083	
	(0.018)	(0.014)	
γ_3	0.137	0.143	0.143
	0.019	0.018	0.018
a_1	15.919	15.808	15.818
	(0.379)	(0.632)	(0.623)
a_2	19.006	18.742	18.613
	(2.616)	(1.452)	(1.248)
a_3	10.831	10.946	11.013
	(0.814)	(0.751)	(0.753)
log-lik	-8326.58	-8328.77	-8329.12
AIC	5.120	5.118	5.116
BIC	5.171	5.157	5.150
RMSE	4.452	4.428	4.433
MAE	2.635	2.632	2.635
Theil-U	0.367	0.367	0.368

Table 4: S&P500 realized kernel volatility (in annualized percentage terms). Markov Switching MEM specifications with three regimes and Gamma innovations: estimates of transition probabilities (standard errors in parentheses; p_{i3} ($i = 1, 2, 3$) is computed as $1 - p_{i1} - p_{i2}$). P-values associated with Wald test statistics for the equality of the rows of the transition probability matrix in the AsyP–MS–AMEM.

	AsyP–MS–AMEM		MS–AMEM	MS–AMEM(c)
	p_{ij}^{+}	p_{ij}^{-}		
p_{11}	0.980	0.995	0.989	0.989
	(0.006)	(0.003)	(0.002)	(0.001)
p_{12}	0.017	0.000	0.007	0.007
	(0.008)	(0.000)	(0.001)	(0.001)
p_{13}	0.003	0.005	0.004	0.004
p_{21}	0.000	0.018	0.007	0.007
	(0.000)	(0.007)	(0.001)	(0.001)
p_{22}	0.994	0.956	0.977	0.977
	(0.017)	(0.016)	(0.002)	(0.002)
p_{23}	0.006	0.026	0.026	0.026
p_{31}	0.012	0.000	0.006	0.004
	(0.008)	(0.000)	(0.002)	(0.001)
p_{32}	0.054	0.026	0.042	0.041
	(0.023)	(0.012)	(0.001)	(0.006)
p_{33}	0.934	0.974	0.952	0.955

AsyP–MS–AMEM	
Null Hypothesis:	p-value
$p_{11}^{+} = p_{11}^{-}$ and $p_{12}^{+} = p_{12}^{-}$	0.069
$p_{21}^{+} = p_{21}^{-}$ and $p_{22}^{+} = p_{22}^{-}$	0.017
$p_{31}^{+} = p_{31}^{-}$ and $p_{32}^{+} = p_{32}^{-}$	0.359

Table 5: Diebold-Mariano statistics for comparison of in sample forecast errors.

	MEM	AMEM	MS–AMEM(c)
		Squared Errors	
MEM		3.192	1.908
AMEM	-3.192		0.609
MS–AMEM(c)	-1.908	-0.609	
		Absolute Errors	
MEM		5.070	7.880
AMEM	-5.070		6.130
MS–AMEM(c)	-7.880	-6.130	
		Squared Errors in first differences	
MEM		5.139	6.017
AMEM	-5.139		5.521
MS–AMEM(c)	-6.017	-5.521	

in all the regimes, with the constant term and the coefficient of the Gamma distribution as the only switching parameters, but the LR and Wald tests again favor the MS–AMEM(c) which will be kept as the MS model of reference to be compared with the MEM and the AMEM.

In terms of goodness of fit, the MS–AMEM(c) seems to be clearly superior relative to the MEM and the AMEM, with lowest AIC and BIC, in spite of the larger number of estimated coefficients, and the lowest RMSE, MAE and Theil's U. To verify whether the differences among the models synthesized by loss functions measures are significant, we have applied a Diebold-Mariano test (Diebold and Mariano, 1995), with the correction proposed by Harvey et al. (1997), to test the null hypothesis of no difference in the accuracy of two competing forecasts. In practice we verify if the mean of the differences of the squared forecast errors, absolute forecast errors and squared forecast errors in first differences of each pair of models is zero. In Table 5 we show the Diebold-Mariano statistics: a positive sign of the statistic in position (i, j) indicates that the model in the $j - th$ column has, on average, better performance than the model in the $i - th$ row. Whether that performance is significant is gauged by the statistic's value relative to a Student's t distribution.[3]

The fact that the MS–AMEM(c) is significantly better than the MEM and the AMEM in terms of absolute errors (see MAE values in Table 3) and squared errors in the first differences (see Theil's U values) is confirmed. Since its corresponding column in Table 5 is always positive, the results of the Diebold Mariano test show that MS–AMEM(c) is the best model. The analysis of squared errors is a puzzling case, because there is a significant better performance of AMEM with respect to MEM, whereas MS–AMEM(c) is not significantly better than the other two models. Maybe this is due to isolated large errors in the MS model, which become even larger when squared.

One of the motivations to adopt a MS model to analyze volatility is the presence of autocorrelated residuals in the MEMs with constant parameters. In Table 6 we show the Ljung-Box statistics (Q) and the corresponding p-value (p) for the three models to check how uncorrelated the residuals are up to lag 20.[4] What we observe is that the model with three regimes are able to capture the strong residual dependence structure still present in the MEM and the AMEM.

[3]Often the Diebold–Mariano test, or, more generally, the predictive ability tests, are applied with out-of-sample forecasts, but Inoue and Kilian (2004) show, both analytically and via Monte Carlo simulations, that out-of-sample tests have lower power than their in-sample counterparts. In our case, the out-of-sample behavior could be misleading for the high volatility regime that characterizes the latest period of our time series; in fact, the out-of-sample forecasts would evaluate just the behavior of the models in a particular phase of volatility and not the global forecasting capability of the models across different regimes.

[4]For MS–AMEM(c) we have used the generalized residuals, introduced by Gourieroux et al. (1987) for latent variable models, defined as $E(\hat{\varepsilon}_t|\Psi_{t-1}) = \sum_{i=1}^{3} \hat{\varepsilon}_{s_t,t} t Pr(s_t = i|\Psi_{t-1})$, where $\hat{\varepsilon}_{s_t,t}$ are the residuals at time t derived from the parameters of the model in state s_t.

Table 6: Ljung-Box Q statistics and corresponding p-value.

	MEM		AMEM		MS–AMEM(c)	
lag	Q	p-value	Q	p-value	Q	p-value
1	16.22	0.00	18.55	0.00	2.27	0.13
2	20.39	0.00	22.43	0.00	2.90	0.23
3	35.40	0.00	35.26	0.00	5.45	0.14
4	35.89	0.00	35.54	0.00	9.18	0.06
5	36.71	0.00	36.02	0.00	12.82	0.03
6	36.90	0.00	36.23	0.00	12.99	0.04
7	36.92	0.00	36.23	0.00	13.08	0.07
8	36.98	0.00	36.24	0.00	13.15	0.11
9	39.72	0.00	38.30	0.00	15.13	0.09
10	41.04	0.00	39.17	0.00	16.32	0.09
11	41.55	0.00	39.23	0.00	16.53	0.12
12	43.91	0.00	40.81	0.00	17.03	0.15
13	46.26	0.00	42.70	0.00	18.45	0.14
14	48.23	0.00	44.26	0.00	19.16	0.16
15	53.19	0.00	49.69	0.00	22.49	0.10
16	55.80	0.00	51.43	0.00	23.07	0.11
17	56.21	0.00	52.47	0.00	23.69	0.13
18	57.21	0.00	53.22	0.00	24.01	0.15
19	63.01	0.00	59.06	0.00	27.20	0.10
20	69.37	0.00	66.43	0.00	33.06	0.03

3.2 Inference on the regimes

Going back to Table 4 we are now in a position to comment on the transition probabilities in the light of the discussion about the model specification above. The probabilities to stay in the same regime allow for estimating the duration in a certain regime i as $\frac{1}{1-p_{ii}}$. While such probabilities are high across all MS models, in reference to our favorite MS–AMEM(c) model, the estimates imply, on average, a 91 days permanence in the state of low volatility, which decreases to 43 days for the intermediate volatility state, and to 22 days for the high volatility state. This result is consistent with the empirical evidence that the turmoil periods have a lower duration with respect to quieter spells.

Some further insights are gained by looking at the off-diagonal elements of the transition probability matrix. Being in regime 1 there is a very low probability to switch to either of the other two regimes. From the regime of intermediate volatility there is a higher probability to move to the high volatility regime than to revert to a low volatility regime. By the same token, note that the downward transition from the high volatility states occurs preferably with a move to the intermediate state: joint with the considerations above, there seems to be a strong interaction between regimes two and three while the period of low volatility is a sort of self standing regime.

These comments are further validated by calculating the smoothed probabilities $Pr(s_t|\Psi_T)$ to obtain an inference on the regimes. Each period is assigned to a regime based on the mode of the probabilities of being in one of the three states for each t; to be true, these probabilities are, in general, near zero or one, denoting a clear–cut association of each period with either regime.

Each state corresponds to a different average level of volatility, namely

$$\mu_j = \frac{\omega + \sum_{i=1}^{j} k_i}{1 - \alpha_j - \beta_j - \gamma_j/2} \quad j = 1, 2, 3. \tag{3.1}$$

From our estimates of the MS–AMEM(c), we derive the corresponding levels of volatility by regime which are 7.89% in regime 1, 12.62% in regime 2 and 23.53% in regime 3. We can superimpose such average volatility levels to the observed series as done in the Panel A. of Figure 2. Bursts of volatilities, as well as sudden reductions in their values correspond to a discrete change in the average value around which volatility follows its dynamics. Moreover, the estimated persistence is clearly different than the one in the MEM and AMEM (0.94 for both models, cf. Table 2) and also within each regime: based on the values in Table 3, we estimate it to be equal to 0.78 during the low and medium volatility, and to a smaller 0.67 in the high volatility regime. These results are consistent with the intuitive notion that turmoil periods have a lower persistence when compared to the quiet periods. Moreover, the regime 3 of high volatility depends mainly on the most recent observation and on the sign of the returns (the values of α_3 and γ_3 are higher and the value of β_3 is lower, relative to corresponding coefficients of regime 1 and 2). It is interesting to note also that $\gamma_3 \simeq 2\gamma_1$, showing that bad news has a bursting effect in turmoil periods. Finally, values of the shape parameter a_{s_t} across regimes are also of interest. Larger values of a correspond to a density approaching the Normal distribution, whereas smaller values polarize the random variable to take on either small or large values. In our case, the estimated values are the largest for the state of medium volatility and the smallest for the state of high volatility, in keeping with the interpretation of the regimes.

Further insights on the value added in considering our Markov Switching AMEM can be gained by dividing the original data by the regime specific average volatility as done in Panel B. of Figure 2. It is apparent that the model manages to remove the underlying slow–moving trend in the evolution of volatility: now the series oscillates around 1, with the exception of the episode started in September 2008 with the collapse of Lehman Brothers which lasted for a relatively short period.

From the Panel A. of Figure 2 it seems that, after a first year (1996), in which the volatility is low with an abrupt, but brief, peak in July, we have a period of medium volatility until the middle of 1998, which includes the Asian crisis; from July 1998 until the beginning of 2003 there are many peaks of high volatility, consistent with the Russian crisis of August 1998, the dot-com bubble (which has its peak on March 2000), the 2001 recession. After this long period, there is a quiet period between October 2003 and July 2007, with only one sudden passage to the regime of high volatility on July 24, 2007. On such a day, the value of the index dropped by 2% and started a short–lived bearish rally around the first measures taken by the Fed in response to the first signs of the liquidity crisis later exploded in August 2007.

Low volatility is fairly persistent and experienced at the beginning of the sample period, a brief bout in 1998, and then from 2003 to 2007. The worsening of the financial crisis is marked by a permanent change to regime 3 until the end of the sample period.

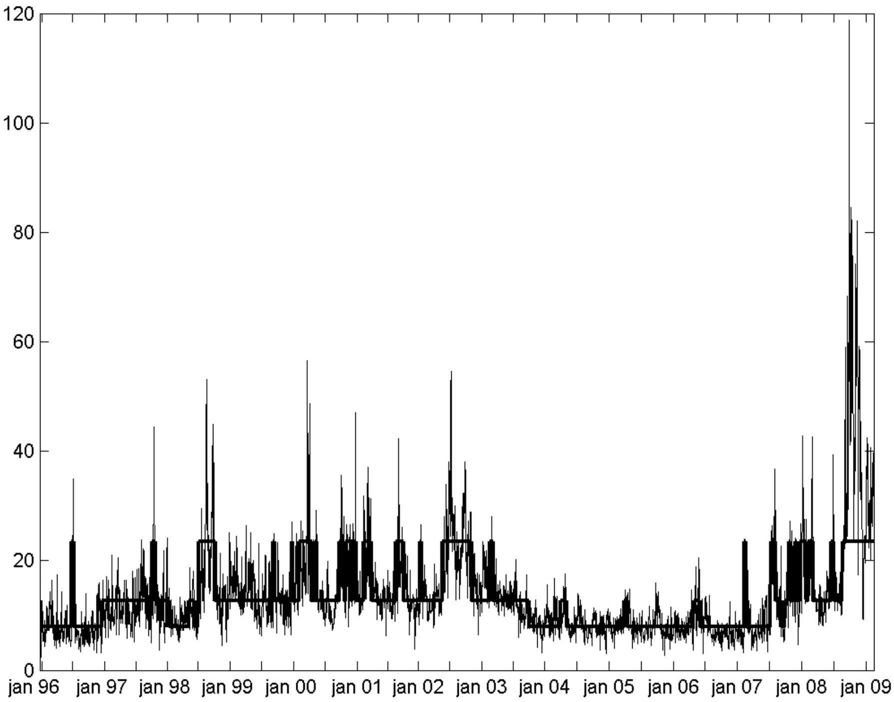
The Panel B. of Figure 2 shows a fairly regular behavior of the transformed series with some signs of autocorrelation (captured by the multiplicative structure of the model) with much less pronounced peaks relative to the original series. The fact that the observed series is well above the average level in the last portion of the sample may be an indication that the height of the crisis may be marked by an even higher average level of volatility, and a fourth (very short lived) regime would in principle accommodate that. In results not reported here this is indeed the case, but it implies a worsening of the inference in the other regimes. We opt not to fit an extra regime given the exceptional nature of the market volatility dynamics following the demise of Lehman Brothers. By the same token, the Ljung–Box autocorrelation statistics of Table 6 are fairly reassuring as of the little impact that the burst of volatility at the end of 2008 has on the overall picture.

3.3 Estimation accuracy

In order to evaluate the properties of our volatility estimates, we resort to the calculation of 95% confidence intervals for the expected volatility values, using the three models. Given that $E(x_t|\Psi_{t-1}) = \mu_t$ and $Var(x_t|\Psi_{t-1}) = \hat{\mu}_t^2/\hat{a}$, we compute the lower and upper values of the intervals, respectively, as the 2.5 and the 97.5 percentiles for the Gamma distribution with parameters $\hat{a}$ and $\hat{\mu}_t/\hat{a}$. For the MS–AMEM(c), paralleling the way $\hat{\mu}_t$ is computed, we derive the intervals' bounds as the weighted

Figure 2: S&P500 realized kernel volatility series. Features of the regimes identified by the MS–AMEM(c).

A. Original series with regime–specific average volatilities in bold line.



B. Volatility divided by the regime–specific average volatilities.

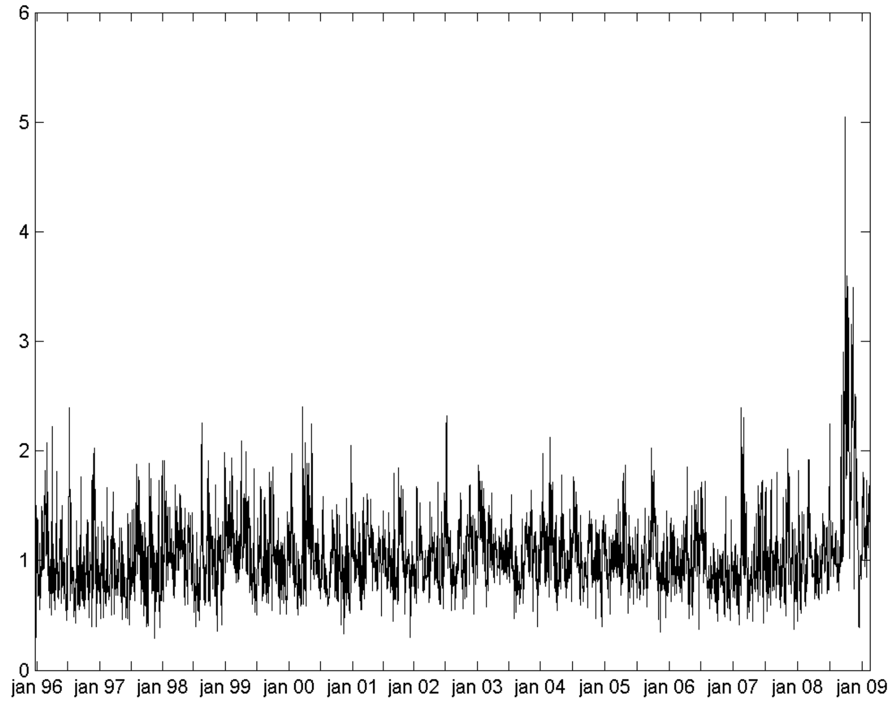


Table 7: Average widths and standard deviations of the 95% confidence intervals of the conditional expected volatility for the whole period and within each regime identified by MS–AMEM(c).

	Whole		Regime 1		Regime 2		Regime 3	
	Mean	st.dev.	Mean	st.dev.	Mean	st.dev.	Mean	st.dev.
MEM	14.09	7.38	9.22	1.86	14.05	3.02	24.68	10.79
AMEM	13.89	7.35	9.09	1.94	13.84	3.08	24.38	10.77
MS–AMEM(c)	13.32	7.99	7.94	1.41	12.26	2.59	27.62	8.50

Table 8: Percentage of errors, distinguished by overprediction and underprediction. Percentage shares of the overprediction and underprediction errors by regimes.

	Overall			Overprediction $\%$ by Regime			Underprediction $\%$ by Regime		
	total	over	under	1	2	3	1	2	3
MEM	5.5	1.9	3.6	44.4	38.1	17.5	17.8	33.0	49.2
AMEM	5.4	2.1	3.3	48.5	32.4	19.1	16.7	36.1	47.2
MS–AMEM(c)	4.7	2.1	2.6	44.1	47.1	8.8	34.9	40.7	24.4

average of the relevant percentiles from a distribution $Gamma(\hat{a}_{s_t}, \hat{\mu}_{s_t}/\hat{a}_{s_t})$ in each regime, with weights equal to the corresponding smoothed probabilities.

In Table 7 we show the mean width of the intervals (together with their standard deviation) as values computed across the T intervals and then the corresponding values by each regime identified by the MS–AMEM(c). For the full period, we can note that the MS–AMEM(c) provides on average smaller intervals (more precise), with a larger standard deviation which we interpret as a capability of being more flexible in adapting to the volatility of volatility. In turn, this feature is consistent with the mean values by regime: in regime 1 and 2 the average width of intervals calculated under MS–AMEM(c) is substantially low relative to the others, whereas in the third regime of very high volatility we get a larger average width with an overall smaller dispersion (lower standard deviations).

Note that conditional expected volatilities obtained by the MS–AMEM(c) explicitly consider a different underlying regime–specific average volatility. Relative to other models, this should imply fewer observed volatilities falling outside the interval bounds. For a nominal confidence interval of 95% and for each model, we show the percentage of out–of–bound observations in the left part of Table 8, distinguishing conditional overprediction (i.e. the observed data falls below the lower bound of the interval) from underprediction (the observed data falls over the upper bound of the interval). The relative incidence of such occurrences is lower for the MS–AMEM(c) (4.7% against more than 5.4% for the other models) and it is equally spread between overprediction and underprediction. By the same token, the MEM and the AMEM underpredict more frequently. We find it instructive to further decompose the incidence of overprediction, respectively, underprediction obtaining the share in each regime (central and right part of the Table 8). It is clear that the MEM and the AMEM, in general, have a decreasing pattern of overprediction as the volatility regime increases while they have higher shares of underprediction with higher volatility regimes. By contrast, the MS–AMEM(c) has patterns balanced between over– and underprediction, concentrating the out–of–bound observations in the first and (more so) in the second regime.

4 Concluding Remarks

Judging upon the results of the analysis of the leading case of the S&P500 index, the introduction of regimes and the inclusion of asymmetric effects in a MEM does improve the fitting performance of the model and increases the possibility of interpreting and capturing crucial events in the volatility dynamics. The presence of residual autocorrelation which is fairly strong in the MEM and AMEM is substantially reduced when adopting a regime switching specification. This seems to be reassuring for the series at hand which presents enormous peaks very often, especially in the last part of the series analyzed.

The model proposed presents several possible specifications: with or without threshold parameters, with or without a threshold change in the transition probabilities, with all the parameters switching or just a part of them. The advantage is that we can nest the resulting models, so it is possible to choose the best model simply using Wald tests or information criteria. Further extensions such as the introduction of a fourth regime are encumbered by the difficulty of reaching convergence: results not reported here show that while a better fit is obtained on the more extreme episodes, there is a general worsening of the fit and of the prediction capabilities for the other regimes.

Briefly put, this good performance can be replicated also for other time series; we have experimented this new model for other American stock indices, differing by degrees of capitalization: the Dow Jones industrial index, which is based on the main 30 stocks exchanged on Wall Street; the S&P400 Midcap index, relative to 400 companies of the mid-cap equities sector; Russell 3000, which deals with the first 3000 companies with larger capitalization; Russell 1000 (the first 1000 companies considered in Russell 3000); Russell 2000 (the last 2000 companies considered in Russell 3000). For the Dow Jones index we have obtained a performance similar to the one of the S&P500. For the S&P400 Midcap index and the Russell indices we obtain a better performance in terms of fitting and in sample forecasting performance with respect to the MEM and AMEM, but not of the residual uncorrelatedness. An interesting fact is that, for these four series, we can capture large part of residual autocorrelation providing a time-varying smooth transition variance, not dependent on regimes, but with a proper dynamics. The idea is that the dynamics of the coefficient of the Gamma distribution is driven by the following model (see Teräsvirta, 2009):

$$a_t = a_0 + a_1(1 + exp(-\delta[r_{t-1} - c]))^{-1}$$

where $a_0 > 0$, $a_1 \geq 0$, $\delta > 0$ and c are unknown parameters. In terms of inference on the regime, the results show a behavior of the Russell 1000 and the Russell 3000 very similar to the one of the S&P500 and the Dow Jones, and a separate behavior of S&P400 Midcap and Russell 2000, similar to one another. In practice the subdivision in regimes seems linked to the degree of liquidity behind the indices.

A detailed analysis and the comparison of several indices is contained in Gallo and Otranto (2012). The lesson we learn is that the increasing flexibility of the distribution hypothesized for the error terms favors the possibility to respect the statistical hypothesis of the model; in practice the correct hypothesis about the error distribution seems to be a crucial task to increase the quality of the model estimation.

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] Andersen, T.G., Bollerslev, T.,Diebold, F.X. (2010). Parametric and nonparametric volatility measurement, in Aït-Sahalia, Y., Hansen, L.P. (eds.): *Handbook of Financial Econometrics*, pages 67-138, North-Holland. Amsterdam.

[4] 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.

[5] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31, 307-321.

[6] Bollerslev, T., Wooldridge, J.M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances, *Econometric Reviews*, 11, 143-172.

[7] Brownlees, C.T., Gallo, G.M. (2010). Comparison of volatility measures: a risk management perspective, *Journal of Financial Econometrics*, 8, 29-56.

[8] Cipollini, F., Engle, R.F., Gallo, G.M. (2012). Semiparametric vector MEM, forthcoming, *Journal of Applied Econometrics*.

[9] Diebold F.X., Lee, J.H., Weinbach, G.C. (1994). Regime switching with time-varying transition probabilities, in Hargreaves, P. (ed.): *Nonstationary Time Series Analysis and Cointegration*. Oxford: Oxford University Press, 283-302.

[10] Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.

[11] Dueker, M.J. (1997). Markov switching in GARCH processes and mean-reverting stock-market volatility, *Journal of Business and Economic Statistics*, 15, 26-34.

[12] Edwards, S., Susmel, R. (2003). Interest-rate volatility in emerging markets, *The Review of Economics and Statistics*, 85, 328-348.

[13] Engle, R.F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica* 50, 987-1008.

[14] Engle, R.F. (2002). New frontiers for ARCH models, *Journal of Applied Econometrics* 17, 425-446.

[15] Engle, R.F, Rangel, J.G. (2008). The Spline-GARCH model for low-frequency volatility and its global macroeconomic causes, *Review of Financial Studies*, 21, 1187-1222.

[16] Filardo, A. J. (1994). Business-cycle phases and their transitional dynamics, *Journal of Business and Economic Statistics*, 12, 299-308.

[17] Engle, R.F., Gallo, G.M. (2006). A multiple indicators model for volatility using intra-daily data, *Journal of Econometrics* 131, 3-27.

[18] Gallo, G.M., Otranto, E. (2007): Volatility transmission across markets: a Multichain Markov Switching model, *Applied Financial Economics*, 17, 659-670.

[19] Gallo, G.M., Otranto, E. (2008): Volatility spillovers, interdependence and comovements: A Markov Switching approach, *Computational Statistics and Data Analysis* 52, 3011-3026.

[20] Gallo, G.M., Otranto, E. (2012): Volatility Swings in the US Financial Markets, forthcoming, in Grigoletto, M., Lisi, F., Petrone, S. (eds.): *Complex Models and Computational Methods in Statistics*, Springer. .

[21] Gourieroux, C., Monfort, A. and Trognon, E. R. A. (1987): Generalized residuals. *Journal of Econometrics* 34, 532.

[22] Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57 357-384.

[23] Hamilton, J.D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, 45, 39-70.

[24] Hamilton, J.D. (1994). *Time series Analysis*. Princeton University Press: Princeton.

[25] Hamilton, J.D., Susmel, R. (1994). Autoregressive conditional hetero-skedasticity and changes in regime. *Journal of Econometrics*, 64, 307-333.

[26] Hansen, R.P., Huang, Z., Shek, H.H. (2011). Realized GARCH: a joint model of returns and realized measures of volatility. *Journal of Applied Econometrics*, in press, DOI: 10.1002/jae.1234

[27] Harvey, D., Leybourne, S., Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, 13, 281-291.

[28] Heber, G., Lunde, A., Shephard, N., Sheppard, K. (2009). *OMI's realised library*, version 0.1. Oxford-Man Institute. University of Oxford.

[29] Higgs, H., Worthington, A.C. (2004). Transmission of returns and volatility in art markets: a multivariate GARCH analysis, *Applied Economics Letters*, 11, 217-222.

[30] Inoue, A., Kilian, L. (2004). In-sample or out-of-sample tests of predictability? Which one should we use? *Econometric Reviews*, 23, 371-402.

[31] Kim, C.J. (1994). Dynamic linear models with Markov-switching, *Journal of Econometrics*, 60, 1-22.

[32] Kim, C.J., Nelson, C.R. (1999). *State-Space Models with Regime Switching*. MIT Press, Cambridge, Massachusetts.

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

[34] Psaradakis, Z., Spagnolo, F. (2003). On the determination of the number of regimes in Markov-Switching autoregressive models, *Journal of Time Series Analysis*, 24, 237-252.

[35] Shephard, N., Sheppard, K. (2010). Realising the future: forecasting with high-frequency-based volatility (HEAVY) models, *Journal of Applied Econometrics* 25, 197-231.

[36] Teräsvirta, T. (2009). An Introduction to univariate GARCH models, in Andersen, T.G., Davis, R.A., Kreiss, J.-P., Mikosch, T. (eds.): *Handbook of Financial Time Series*, pages 17-42, Springer. Berlin-Heidelberg.

[37] Zakoïan, J.M. (1994). Threshold Heteroskedastic Models, *Journal of Economic Dynamic and Control*, 18, 931-955.