Forecasting Realized Volatility with Changing Average Levels

Giampiero M. Gallo* Edoardo Otranto**

* Dipartimento di Statistica, Informatica, Applicazioni (DiSIA) G. Parenti – Università di Firenze
** Dipartimento di Scienze Cognitive e della Formazione and CRENOS – Università di Messina
Outline

Introduction

On Today’s Menu

MEM and MS–MEM

Regimes in the Volatility of S&P500
  Modeling
  Inference on Regimes
  In– and Out–of–sample Forecasting

Extensions

Conclusions
Persistence as a Challenge to Volatility Modeling

Background

- Direct measurement of volatility: ideal properties of the estimators *ex post*
- Modeling for forecasting purposes: dynamic models
- Clustering features are present
- Choice of modeling volatility or log–volatility
- Residual diagnostics as a guideline to specification
- Possible recalcitrant residual correlation as a guide to misspecification
Persistence as a Challenge to Volatility Modeling

Background

- Direct measurement of volatility: ideal properties of the estimators \textit{ex post}
- \textbf{Modeling for forecasting purposes: dynamic models}
  - Clustering features are present
  - Choice of modeling volatility or log–volatility
  - Residual diagnostics as a guideline to specification
  - Possible recalcitrant residual correlation as a guide to misspecification
Persistence as a Challenge to Volatility Modeling

Background

- Direct measurement of volatility: ideal properties of the estimators *ex post*
- Modeling for forecasting purposes: dynamic models
- Clustering features are present
  - Choice of modeling volatility or log–volatility
  - Residual diagnostics as a guideline to specification
  - Possible recalcitrant residual correlation as a guide to misspecification
Persistence as a Challenge to Volatility Modeling

Background

- Direct measurement of volatility: ideal properties of the estimators *ex post*
- Modeling for forecasting purposes: dynamic models
- Clustering features are present
- Choice of modeling volatility or log–volatility
  - Residual diagnostics as a guideline to specification
  - Possible recalcitrant residual correlation as a guide to misspecification
Persistence as a Challenge to Volatility Modeling

Background

- Direct measurement of volatility: ideal properties of the estimators \textit{ex post}
- Modeling for forecasting purposes: dynamic models
- Clustering features are present
- Choice of modeling volatility or log–volatility
- Residual diagnostics as a guideline to specification
- Possible recalcitrant residual correlation as a guide to misspecification
Persistence as a Challenge to Volatility Modeling

Background

- Direct measurement of volatility: ideal properties of the estimators ex post
- Modeling for forecasting purposes: dynamic models
- Clustering features are present
- Choice of modeling volatility or log–volatility
- Residual diagnostics as a guideline to specification
- Possible recalcitrant residual correlation as a guide to misspecification
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al., 2003)
  - Quasi long memory (HAR on log–vol; Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  - Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  - Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  - Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
- Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al. 2003)
  - Quasi long memory (HAR on log–vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  - Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  - Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  - Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
- Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al. 2003;)
  - Quasi long memory (HAR on log-vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  - Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  - Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  - Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
- Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al. 2003;)
  - Quasi long memory (HAR on log-vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  - Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  - Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  - Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
- Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al. 2003;)
  - Quasi long memory (HAR on log-vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  - Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  - Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  - Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
- Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al. 2003;)
  - Quasi long memory (HAR on log-vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
- Markov Switching linear model on vol (Maheu and McCurdy, 2002)
- Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
- Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
- Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

▶ Underlying movements at low frequency
▶ A variety of interpretations/choices
  ▶ Levels vs logs; Variance vs Volatility
  ▶ Long memory (ARFIMA on log–vol; Andersen et al. 2003)
  ▶ Quasi long memory (HAR on log-vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  ▶ Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  ▶ Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  ▶ Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  ▶ Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
  ▶ Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al. 2003;)
  - Quasi long memory (HAR on log-vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  - Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  - Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  - Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
  - Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al. 2003)
  - Quasi long memory (HAR on log-vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  - Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  - Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  - Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
- Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
Nonlinear Effects in Volatility

How to model such a series

- Underlying movements at low frequency
- A variety of interpretations/choices
  - Levels vs logs; Variance vs Volatility
  - Long memory (ARFIMA on log–vol; Andersen et al. 2003;)
  - Quasi long memory (HAR on log-vol: Corsi, 2009; extensions: McAleer and Medeiros, 2008)
  - Spline fitting (van Bellegem and von Sachs, 2004; Engle and Rangel, 2008; Brownlees and Gallo, 2010)
  - Markov Switching linear model on vol (Maheu and McCurdy, 2002)
  - Markov Switching and fractionally integrated dynamics (Bordignon and Raggi, 2010)
  - Smooth multiplicative component in a GARCH framework (Amado and Teräsvirta, 2012)
- Interpretation of the time-varying unconditional volatility (Engle and Rangel, 2008)
In This Paper

Two strong choices

- Concentrate on Multiplicative Error Models
  - (over GARCH) Use volatility rather than squared returns
  - (over linear models) Directly applied to volatility (and not logs)
  - QMLE interpretation ensures consistency
- Adopt a Markov Switching approach
  - Exploit well-established properties
  - Classification of periods by regime
In This Paper

Two strong choices

- Concentrate on Multiplicative Error Models
  - (over GARCH) Use volatility rather than squared returns
  - (over linear models) Directly applied to volatility (and not logs)
  - QMLE interpretation ensures consistency
- Adopt a Markov Switching approach
  - Exploit well established properties
  - Classification of periods by regime
In This Paper

Two strong choices

▶ Concentrate on Multiplicative Error Models
  ▶ (over GARCH) Use volatility rather than squared returns
  ▶ (over linear models) Directly applied to volatility (and not logs)
  ▶ QMLE interpretation ensures consistency

▶ Adopt a Markov Switching approach
  ▶ Exploit well established properties
  ▶ Classification of periods by regime
In This Paper

Two strong choices

- Concentrate on Multiplicative Error Models
  - (over GARCH) Use volatility rather than squared returns
  - (over linear models) Directly applied to volatility (and not logs)
  - QMLE interpretation ensures consistency

- Adopt a Markov Switching approach
  - Exploit well established properties
  - Classification of periods by regime
In This Paper

Two strong choices

- Concentrate on Multiplicative Error Models
  - (over GARCH) Use volatility rather than squared returns
  - (over linear models) Directly applied to volatility (and not logs)
  - QMLE interpretation ensures consistency

- Adopt a Markov Switching approach
  - Exploit well established properties
  - Classification of periods by regime
In This Paper

Two strong choices

▶ Concentrate on Multiplicative Error Models
  ▶ (over GARCH) Use volatility rather than squared returns
  ▶ (over linear models) Directly applied to volatility (and not logs)
  ▶ QMLE interpretation ensures consistency

▶ Adopt a Markov Switching approach
  ▶ Exploit well established properties
  ▶ Classification of periods by regime
In This Paper

Two strong choices

- Concentrate on Multiplicative Error Models
  - (over GARCH) Use volatility rather than squared returns
  - (over linear models) Directly applied to volatility (and not logs)
  - QMLE interpretation ensures consistency

- Adopt a Markov Switching approach
  - Exploit well established properties
  - Classification of periods by regime
Why a MEM?

Modeling non-negative time series

GARCH conditional variance is the expectation of squared returns but volatility is measured directly

- A lot of information available in financial markets is positive valued:
  - ultra-high frequency data (within a time interval: range, volume, number of trades, number of buys/sells, durations)
  - daily volatility estimators (realized volatility, daily range, absolute returns)

- Time series exhibit persistence which can be modeled à la GARCH
Why a MEM?

**Modeling non-negative time series**

GARCH conditional variance is the expectation of squared returns but volatility is measured directly

- A lot of information available in financial markets is positive valued:
  - ultra-high frequency data (within a time interval: range, volume, number of trades, number of buys/sells, durations)
  - daily volatility estimators (realized volatility, daily range, absolute returns)
- Time series exhibit persistence which can be modeled à la GARCH
Why a MEM?

Modeling non-negative time series

GARCH conditional variance is the expectation of squared returns but volatility is measured directly

- A lot of information available in financial markets is positive valued:
  - ultra-high frequency data (within a time interval: range, volume, number of trades, number of buys/sells, durations)
  - daily volatility estimators (realized volatility, daily range, absolute returns)
- Time series exhibit persistence which can be modeled à la GARCH
Why a MEM?

Modeling non-negative time series

GARCH conditional variance is the expectation of squared returns but volatility is measured directly

- A lot of information available in financial markets is positive valued:
  - ultra-high frequency data (within a time interval: range, volume, number of trades, number of buys/sells, durations)
  - daily volatility estimators (realized volatility, daily range, absolute returns)

- Time series exhibit persistence which can be modeled à la GARCH
Why a MEM?

Modeling non-negative time series
GARCH conditional variance is the expectation of squared returns but volatility is measured directly

- A lot of information available in financial markets is positive valued:
  - ultra-high frequency data (within a time interval: range, volume, number of trades, number of buys/sells, durations)
  - daily volatility estimators (realized volatility, daily range, absolute returns)

- Time series exhibit persistence which can be modeled à la GARCH
Multiplicative Error Models

- Extension of GARCH approach to modeling the expected value of processes with positive support (Engle, 2002; Engle and Gallo, 2006)

- Autoregressive Conditional Duration is a special case. Absolute returns, high-low, number of trades in a certain interval, volume, realized volatility can be modeled with the same structure.

- Rather than calling the models Autoregressive Conditional Volatility, Autoregressive Conditional Volume, etc. call them MEM.

- Ease of estimation.

- Possibility of expanding the information set (main interesting results).
Multiplicative Error Models

- Extension of GARCH approach to modeling the expected value of processes with positive support (Engle, 2002; Engle and Gallo, 2006)
- Autoregressive Conditional Duration is a special case. Absolute returns, high-low, number of trades in a certain interval, volume, realized volatility can be modeled with the same structure
- Rather than calling the models Autoregressive Conditional Volatility, Autoregressive Conditional Volume, etc. call them MEM
- Ease of estimation
- Possibility of expanding the information set (main interesting results)
Multiplicative Error Models

- Extension of GARCH approach to modeling the expected value of processes with positive support (Engle, 2002; Engle and Gallo, 2006)

- Autoregressive Conditional Duration is a special case. Absolute returns, high-low, number of trades in a certain interval, volume, realized volatility can be modeled with the same structure

- Rather than calling the models Autoregressive Conditional Volatility, Autoregressive Conditional Volume, etc. call them MEM

- Ease of estimation

- Possibility of expanding the information set (main interesting results)
Multiplicative Error Models

- Extension of GARCH approach to modeling the expected value of processes with positive support (Engle, 2002; Engle and Gallo, 2006)
- Autoregressive Conditional Duration is a special case. Absolute returns, high-low, number of trades in a certain interval, volume, realized volatility can be modeled with the same structure
- Rather than calling the models Autoregressive Conditional Volatility, Autoregressive Conditional Volume, etc. call them MEM
- Ease of estimation
- Possibility of expanding the information set (main interesting results)
Multiplicative Error Models

- Extension of GARCH approach to modeling the expected value of processes with positive support (Engle, 2002; Engle and Gallo, 2006)
- Autoregressive Conditional Duration is a special case. Absolute returns, high-low, number of trades in a certain interval, volume, realized volatility can be modeled with the same structure
- Rather than calling the models Autoregressive Conditional Volatility, Autoregressive Conditional Volume, etc. call them MEM
- Ease of estimation
- Possibility of expanding the information set (main interesting results)
The Asymmetric MEM

\[ X_t = \mu_t \varepsilon_t, \quad \varepsilon_t \sim \text{Gamma}(a, 1/a) \text{ for each } t \]

\[ \mu_t = \omega + \alpha x_{t-1} + \beta \mu_{t-1} + \gamma D_{t-1} x_{t-1} \]

\[ D_t = \begin{cases} 
1 & \text{if } r_t < 0 \\
0 & \text{if } r_t \geq 0
\end{cases} \]
The Markov Switching–AMEM

\[ X_t = \mu_{t,s_t} \varepsilon_t, \quad \varepsilon_t | s_t \sim \text{Gamma}(a_{s_t}, 1/a_{s_t}) \text{ for each } t \]

\[
\begin{align*}
\mu_{t,s_t} &= \omega + \sum_{i=1}^{n} k_i I_{s_{t}} + \alpha_{s_t} (x_{t-1} - \mu_{t-1,s_{t-1}}) + \\
&\quad + \beta^*_{s_t} \mu_{t-1,s_{t-1}} + \gamma_{s_t} D_{t-1} (x_{t-1} - \mu_{t-1,s_{t-1}})
\end{align*}
\]

\( s_t \) is a discrete latent variable ranging in \([1, \ldots, n]\) (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 \( k_1 = 0 \) (non decreasing levels of volatility passing to higher regimes)
MS–AMEM

\[ Pr(s_t = j \mid s_{t-1} = i, s_{t-2}, \ldots) = Pr(s_t = j \mid s_{t-1} = i) = p_{ij} \]

The unconditional expected value within state \( j, j = 1, \ldots, n \), is equal to:

\[ \mu_j = \frac{\omega + \sum_{i=1}^{j} k_i}{1 - \alpha_j - \beta_j - \gamma_j/2} \]

where \( \beta_j = \beta_j^* - \alpha_j - \gamma_j/2 \). Under this reparameterization, the mean equation of the MS-AMEM is:

\[ \mu_{t,s_t} = \omega + \sum_{i=1}^{n} k_i l_{s_t} + \alpha_{s_t} x_{t-1} + \beta_{s_t} \mu_{t-1,s_{t-1}} + \gamma_{s_t} D_{t-1} x_{t-1} \]
Estimation Issues

- Hamilton filter and smoother with Kim (1994) approximation to avoid path dependence

- After each step of the Hamilton filter, at time $t$ we collapse the $n^2$ possible values of $\mu_t$ into $n$ values, by an average over the probabilities at time $t-1$:

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

<table>
<thead>
<tr>
<th>statistic</th>
<th>full</th>
<th>in-sample</th>
<th>out-of-sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>15.54</td>
<td>15.60</td>
<td>15.00</td>
</tr>
<tr>
<td>median</td>
<td>13.10</td>
<td>13.10</td>
<td>12.95</td>
</tr>
<tr>
<td>min</td>
<td>3.49</td>
<td>3.49</td>
<td>4.50</td>
</tr>
<tr>
<td>max</td>
<td>153.19</td>
<td>153.19</td>
<td>62.43</td>
</tr>
<tr>
<td>st. dev.</td>
<td>10.02</td>
<td>10.16</td>
<td>8.76</td>
</tr>
<tr>
<td>skewness</td>
<td>3.26</td>
<td>3.34</td>
<td>2.06</td>
</tr>
<tr>
<td>kurtosis</td>
<td>20.67</td>
<td>21.43</td>
<td>5.73</td>
</tr>
<tr>
<td>$\rho(1)$</td>
<td>0.82</td>
<td>0.82</td>
<td>0.74</td>
</tr>
<tr>
<td>$\rho(5)$</td>
<td>0.70</td>
<td>0.71</td>
<td>0.57</td>
</tr>
<tr>
<td>$\rho(22)$</td>
<td>0.53</td>
<td>0.55</td>
<td>0.41</td>
</tr>
</tbody>
</table>
MEM/AMEM Estimation Results

<table>
<thead>
<tr>
<th></th>
<th>( \omega )</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
<th>( \alpha )</th>
<th>log-lik</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEM</td>
<td>1.05</td>
<td>0.40</td>
<td>0.53</td>
<td>13.44</td>
<td></td>
<td>-7834.23</td>
<td>5.47</td>
<td>5.48</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMEM</td>
<td>0.94</td>
<td>0.27</td>
<td>0.61</td>
<td>0.11</td>
<td>14.21</td>
<td>-7752.06</td>
<td>5.42</td>
<td>5.43</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.51)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Gamma density for the AMEM
## Residual Autocorrelation Results

<table>
<thead>
<tr>
<th>lag</th>
<th>MEM</th>
<th>AMEM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q</td>
<td>p-value</td>
</tr>
<tr>
<td>1</td>
<td>5.44</td>
<td>0.02</td>
</tr>
<tr>
<td>2</td>
<td>7.09</td>
<td>0.03</td>
</tr>
<tr>
<td>3</td>
<td>7.74</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>10.96</td>
<td>0.03</td>
</tr>
<tr>
<td>5</td>
<td>14.39</td>
<td>0.01</td>
</tr>
<tr>
<td>6</td>
<td>15.36</td>
<td>0.02</td>
</tr>
<tr>
<td>7</td>
<td>17.01</td>
<td>0.02</td>
</tr>
<tr>
<td>8</td>
<td>19.81</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Nonparametric Bayesian Regime Identification

Otranto and Gallo (2002) identify the number of regimes in Markov switching models, based on the detection of the empirical posterior distribution of the number of regimes, via Gibbs sampling, using the nonparametric Bayesian techniques derived from the Dirichlet process theory.

<table>
<thead>
<tr>
<th>$A$</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.93</td>
<td>0.07</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>0.21</td>
<td>0.50</td>
<td>0.39</td>
<td>0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>0.33</td>
<td>0.91</td>
<td>0.09</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Strong indication of the presence of regimes and three is the favored number of regimes.
**MS–AMEM Estimation Results**

<table>
<thead>
<tr>
<th></th>
<th>Mean Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \omega )</td>
</tr>
<tr>
<td></td>
<td>( \alpha_1 )</td>
</tr>
<tr>
<td>MS(3)–AMEM</td>
<td>1.92 (0.13)</td>
</tr>
<tr>
<td>MS(4)–AMEM</td>
<td>1.92 (0.18)</td>
</tr>
</tbody>
</table>

Common dynamics detected between regimes 1 and 3 in MS(3). Numerical difficulties encountered with MS(4)
Gamma density by regime for the MS(3)–AMEM

![Gamma density graph](image)

- Regime 1
- Regime 2
- Regime 3
Gamma density by regime for the MS(4)-AMEM
### Residual Autocorrelation Results

<table>
<thead>
<tr>
<th>lag</th>
<th>MEM Q</th>
<th>MEM p-value</th>
<th>AMEM Q</th>
<th>AMEM p-value</th>
<th>MS(3)–AMEM Q</th>
<th>MS(3)–AMEM p-value</th>
<th>MS(4)–AMEM Q</th>
<th>MS(4)–AMEM p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.44</td>
<td>0.02</td>
<td>5.12</td>
<td>0.02</td>
<td>0.84</td>
<td>0.36</td>
<td>15.93</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>7.09</td>
<td>0.03</td>
<td>7.38</td>
<td>0.02</td>
<td>1.81</td>
<td>0.41</td>
<td>22.19</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>7.74</td>
<td>0.05</td>
<td>7.77</td>
<td>0.05</td>
<td>2.34</td>
<td>0.51</td>
<td>22.22</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>10.96</td>
<td>0.03</td>
<td>9.72</td>
<td>0.05</td>
<td>2.63</td>
<td>0.62</td>
<td>25.15</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>14.39</td>
<td>0.01</td>
<td>12.31</td>
<td>0.03</td>
<td>3.58</td>
<td>0.61</td>
<td>31.52</td>
<td>0.00</td>
</tr>
<tr>
<td>6</td>
<td>15.36</td>
<td>0.02</td>
<td>13.15</td>
<td>0.04</td>
<td>3.58</td>
<td>0.73</td>
<td>37.58</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>17.01</td>
<td>0.02</td>
<td>14.22</td>
<td>0.05</td>
<td>3.78</td>
<td>0.81</td>
<td>41.57</td>
<td>0.00</td>
</tr>
<tr>
<td>8</td>
<td>19.81</td>
<td>0.01</td>
<td>16.81</td>
<td>0.03</td>
<td>3.83</td>
<td>0.87</td>
<td>48.17</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Discussion on Transition Probabilities

- Duration in a certain regime $i$: $\left(\frac{1}{1 - p_{ii}}\right)$
- MS(3)–AMEM: average 87 days in Regime 1; 28 days in Regime 2; 13 days in Regime 3. MS(4)–AMEM: expected duration equal to 77, 18, 15 and 5 days respectively.
- Off–diagonal elements: strong interaction between regimes 2 and 3 while the period of low volatility is a sort of self standing regime. From Regime 2 higher probability to move to Regime 3 (and vice versa) than to revert to Regime 1.
- MS(4)–AMEM. Period classification is similar but with frequent changes in regime
- many isolated dots in the figure (largest smoothed probability not always close to 1. Regime 3 and 4 in the MS(4)-AMEM capture most of Regime 3 in the MS(3)–AMEM, with a higher interaction between Regimes 1 and 2.
Inference on the Regimes: MS(3)-AMEM
Inference on the Regimes: MS(4)-AMEM

![Graph showing changes in average volatility over time](chart.png)
Unconditional Volatilities by Regime

<table>
<thead>
<tr>
<th></th>
<th>No regime</th>
<th>Regime 1</th>
<th>Regime 2</th>
<th>Regime 3</th>
<th>Regime 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEM</td>
<td>14.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMEM</td>
<td>14.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS(3)–AMEM</td>
<td></td>
<td>9.21</td>
<td>14.39</td>
<td>28.80</td>
<td></td>
</tr>
<tr>
<td>MS(4)–AMEM</td>
<td></td>
<td>9.28</td>
<td>13.74</td>
<td>26.76</td>
<td>56.97</td>
</tr>
<tr>
<td>MS(3)–AMEM(d)</td>
<td></td>
<td>9.20</td>
<td>14.35</td>
<td>28.37</td>
<td></td>
</tr>
</tbody>
</table>
Different Classification by Regime

Side-By-Side Chart
Volatility relative to regime–specific averages - MS(3)–AMEM

Gallo & Otranto
Changing Average Volatility
Paris, Mar 30, 2015
Volatility relative to regime–specific averages - MS(4)–AMEM

Gallo & Otranto
Changing Average Volatility
Paris, Mar 30, 2015
The behavior of the MS(4)–AMEM points to some instability in the estimation and in the interpretation of the results, with undesirable autocorrelation generated. Capture the burst with a dummy in the MS(3)–AMEM.

### Mean Coefficients

<table>
<thead>
<tr>
<th></th>
<th>$\omega$</th>
<th>$k_2$</th>
<th>$k_3$</th>
<th>$\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>1.91</td>
<td>0.47</td>
<td>3.51</td>
<td>71.75</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.31)</td>
<td>(0.65)</td>
<td>(14.20)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\gamma_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>0.14</td>
<td>0.13</td>
<td>0.14</td>
<td>0.58</td>
<td>0.67</td>
<td>0.58</td>
<td>0.13</td>
<td>0.08</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Gamma coefficients

<table>
<thead>
<tr>
<th></th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>17.71</td>
<td>23.62</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>(0.94)</td>
<td>(2.11)</td>
<td>(0.92)</td>
</tr>
</tbody>
</table>

### Likelihood-based criteria

<table>
<thead>
<tr>
<th></th>
<th>Log-Lik</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>-7594.89</td>
<td>5.32</td>
<td>5.36</td>
</tr>
</tbody>
</table>
Volatility relative to regime–specific averages - MS(3)–AMEM(d)

Changing Average Volatility

Paris, Mar 30, 2015 32 / 49
Residual Autocorrelation Results

<table>
<thead>
<tr>
<th>lag</th>
<th>MEM Q</th>
<th>p-value</th>
<th>AMEM Q</th>
<th>p-value</th>
<th>MS(3)–AMEM Q</th>
<th>p-value</th>
<th>MS(4)–AMEM Q</th>
<th>p-value</th>
<th>MS(3)–AMEM(d) Q</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.44</td>
<td>0.02</td>
<td>5.12</td>
<td>0.02</td>
<td>0.84</td>
<td>0.36</td>
<td>15.93</td>
<td>0.00</td>
<td>1.17</td>
<td>0.28</td>
</tr>
<tr>
<td>2</td>
<td>7.09</td>
<td>0.03</td>
<td>7.38</td>
<td>0.02</td>
<td>1.81</td>
<td>0.41</td>
<td>22.19</td>
<td>0.00</td>
<td>1.91</td>
<td>0.38</td>
</tr>
<tr>
<td>3</td>
<td>7.74</td>
<td>0.05</td>
<td>7.77</td>
<td>0.05</td>
<td>2.34</td>
<td>0.51</td>
<td>22.22</td>
<td>0.00</td>
<td>2.36</td>
<td>0.50</td>
</tr>
<tr>
<td>4</td>
<td>10.96</td>
<td>0.03</td>
<td>9.72</td>
<td>0.05</td>
<td>2.63</td>
<td>0.62</td>
<td>25.15</td>
<td>0.00</td>
<td>2.53</td>
<td>0.64</td>
</tr>
<tr>
<td>5</td>
<td>14.39</td>
<td>0.01</td>
<td>12.31</td>
<td>0.03</td>
<td>3.58</td>
<td>0.61</td>
<td>31.52</td>
<td>0.00</td>
<td>3.36</td>
<td>0.64</td>
</tr>
<tr>
<td>6</td>
<td>15.36</td>
<td>0.02</td>
<td>13.15</td>
<td>0.04</td>
<td>3.58</td>
<td>0.73</td>
<td>37.58</td>
<td>0.00</td>
<td>3.39</td>
<td>0.76</td>
</tr>
<tr>
<td>7</td>
<td>17.01</td>
<td>0.02</td>
<td>14.22</td>
<td>0.05</td>
<td>3.78</td>
<td>0.81</td>
<td>41.57</td>
<td>0.00</td>
<td>3.55</td>
<td>0.83</td>
</tr>
<tr>
<td>8</td>
<td>19.81</td>
<td>0.01</td>
<td>16.81</td>
<td>0.03</td>
<td>3.83</td>
<td>0.87</td>
<td>48.17</td>
<td>0.00</td>
<td>3.59</td>
<td>0.89</td>
</tr>
</tbody>
</table>
### MSE and MAE Results

<table>
<thead>
<tr>
<th>Method</th>
<th>In-sample</th>
<th></th>
<th>Out-of-sample</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>MAE</td>
<td>MSE</td>
<td>MAE</td>
<td>MSE</td>
</tr>
<tr>
<td>MEM</td>
<td>29.10</td>
<td>3.25</td>
<td>29.61</td>
<td>3.72</td>
<td>38.27</td>
<td>3.98</td>
</tr>
<tr>
<td>AMEM</td>
<td>27.06</td>
<td>3.16</td>
<td>27.65</td>
<td>3.62</td>
<td>35.16</td>
<td>3.82</td>
</tr>
<tr>
<td>MS(3)–AMEM</td>
<td>25.57</td>
<td>2.91</td>
<td>27.00</td>
<td>3.62</td>
<td>31.05</td>
<td>3.70</td>
</tr>
<tr>
<td>MS(4)–AMEM</td>
<td><strong>22.54</strong></td>
<td><strong>2.81</strong></td>
<td><strong>28.95</strong></td>
<td><strong>3.63</strong></td>
<td><strong>35.58</strong></td>
<td><strong>3.97</strong></td>
</tr>
<tr>
<td>MS(3)–AMEM(d)</td>
<td>23.01</td>
<td>2.89</td>
<td><strong>26.96</strong></td>
<td><strong>3.61</strong></td>
<td><strong>30.92</strong></td>
<td><strong>3.68</strong></td>
</tr>
</tbody>
</table>

*Note: The values are averages across different datasets and scenarios.*
## Diebold Mariano tests

### Absolute Errors

<table>
<thead>
<tr>
<th></th>
<th>MEM</th>
<th>AMEM</th>
<th>MS(3)–AMEM</th>
<th>MS(4)–AMEM</th>
<th>MS(3)–AMEM(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEM</td>
<td>*</td>
<td>♦</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>AMEM</td>
<td></td>
<td>♦</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS(3)–AMEM</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td>♦</td>
</tr>
<tr>
<td>MS(4)–AMEM</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS(3)–AMEM(d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>♦</td>
</tr>
</tbody>
</table>

### Squared Errors

<table>
<thead>
<tr>
<th></th>
<th>MEM</th>
<th>AMEM</th>
<th>MS(3)–AMEM</th>
<th>MS(4)–AMEM</th>
<th>MS(3)–AMEM(d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEM</td>
<td>*</td>
<td>♦</td>
<td>*</td>
<td>*</td>
<td>♦</td>
</tr>
<tr>
<td>AMEM</td>
<td></td>
<td>♦</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS(3)–AMEM</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>MS(4)–AMEM</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS(3)–AMEM(d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>♦</td>
</tr>
</tbody>
</table>

The model by column is significantly better in forecasting performance than the model by row (at a 5% confidence level) if a symbol is present: in-sample (⋆), 1-step (∗) and 2-steps (♦) ahead out-of-sample.
Results of the Model Confidence Set

Model Confidence Set procedure (Hansen, Lunde and Nason, 2011): set of models containing the best model at 95% confidence

<table>
<thead>
<tr>
<th>Model</th>
<th>Absolute errors</th>
<th>Squared errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEM</td>
<td>◊</td>
<td>◊</td>
</tr>
<tr>
<td>AMEM</td>
<td>◊</td>
<td>◊•</td>
</tr>
<tr>
<td>MS(3)–AMEM</td>
<td>⋆◊</td>
<td>⋆•</td>
</tr>
<tr>
<td>MS(4)–AMEM</td>
<td>⋆◊</td>
<td>⋆◊•</td>
</tr>
<tr>
<td>MS(3)–AMEM(d)</td>
<td>⋆◊•</td>
<td>⋆◊•</td>
</tr>
</tbody>
</table>

Prediction confidence intervals - March 2005

![Graph showing prediction confidence intervals for volatility, AMEM, and MS(3)-AMEM, MS(4)-AMEM, and MS(3)-AMEM(d) across 31 days.](image)
Prediction confidence intervals - October 2008
Prediction confidence intervals - March 2011
Extending the Information Set

Realized Kernel Volatility and VIX
Time–varying Transition Probability

TVTP as a function of the volatility index $v_t$.

**TVTP–MS(3)–AMEM**

\[ x_t = \mu_{t,s_t} \varepsilon_t, \quad \varepsilon_t | s_t \sim \text{Gamma}(a_{s_t}, 1/a_{s_t}) \text{ for each } t \]

\[
\mu_{t,s_t} = \omega + \sum_{i=1}^{n} k_i l_{s_t} + \alpha_{s_t}(x_{t-1} - \mu_{t-1,s_{t-1}}) + \beta^*_{s_t} \mu_{t-1,s_{t-1}} + \gamma_{s_t} D_{t-1}(x_{t-1} - \mu_{t-1,s_{t-1}})
\]

\[
D_t = \begin{cases} 
1 & \text{if } r_t < 0 \\
0 & \text{if } r_t \geq 0 
\end{cases}
\]

\[
p_{ij,t} = \frac{\exp(\theta_{ij} + \phi_{i,j} v_{t-1})}{1 + \exp(\theta_{i,j} + \phi_{i,j} v_{t-1})}
\]

(1)
Rough or Smooth?

Smooth transition driven by $v_t$ \textbf{ST–AMEM}

$$x_t = \mu_t \varepsilon_t, \quad \varepsilon_t \sim \text{Gamma}(a, 1/a) \text{ for each } t$$

$$\mu_t = \omega + (\alpha_0 + \alpha_1 F(v_{t-1}))x_{t-1} + (\beta_0 - \beta_1 F(v_{t-1}))\mu_{t-1} + \gamma D_{t-1}x_{t-1}$$

$$F(v_t) = (1 + \exp(-g(v_t - c)))^{-1}$$

(2)
Mixing Frequencies - Quasi Long Memory

Borrow Corsi’s (2009) HAR (heterogeneous dynamics)

**HMEM** Errors enter multiplicatively into **HAR**

\[ x_t = \mu_t \varepsilon_t, \quad \varepsilon_t \sim \text{Gamma}(a, 1/a) \]

\[ \mu_t = \omega + \alpha_D x_{t-1} + \alpha_W \bar{x}_{t-1}^{(5)} + \alpha_M \bar{x}_{t-1}^{(22)} \]

**MS(3)–HMEM** Insert regimes

\[ x_t = \mu_{t,s_t} \varepsilon_t, \quad \varepsilon_t | s_t \sim \text{Gamma}(a_{s_t}, 1/a_{s_t}) \]

\[ \mu_{t,s_t} = \omega + \sum_{i=1}^{n} k_i I_{s_t} + \alpha_{D,s_t} x_{t-1} + \alpha_{W,s_t} \bar{x}_{t-1}^{(5)} + \alpha_{M,s_t} \bar{x}_{t-1}^{(22)} \]
### Autocorrelation Properties

<table>
<thead>
<tr>
<th>lag</th>
<th>TVTP–MS(3)–AMEM</th>
<th>ST–AMEM</th>
<th>HMEM</th>
<th>MS(3)–HMEM</th>
<th>HAR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q</td>
<td>p-value</td>
<td>Q</td>
<td>p-value</td>
<td>Q</td>
</tr>
<tr>
<td>1</td>
<td>52.05</td>
<td>0.00</td>
<td>0.04</td>
<td>0.84</td>
<td>1.59</td>
</tr>
<tr>
<td>2</td>
<td>76.10</td>
<td>0.00</td>
<td>0.90</td>
<td>0.64</td>
<td>38.25</td>
</tr>
<tr>
<td>3</td>
<td>99.15</td>
<td>0.00</td>
<td>1.76</td>
<td>0.62</td>
<td>40.74</td>
</tr>
<tr>
<td>4</td>
<td>132.53</td>
<td>0.00</td>
<td>3.29</td>
<td>0.51</td>
<td>46.21</td>
</tr>
<tr>
<td>5</td>
<td>177.49</td>
<td>0.00</td>
<td>5.45</td>
<td>0.36</td>
<td>46.24</td>
</tr>
<tr>
<td>6</td>
<td>207.78</td>
<td>0.00</td>
<td>6.02</td>
<td>0.42</td>
<td>47.98</td>
</tr>
<tr>
<td>7</td>
<td>234.29</td>
<td>0.00</td>
<td>6.39</td>
<td>0.50</td>
<td>52.32</td>
</tr>
<tr>
<td>8</td>
<td>264.40</td>
<td>0.00</td>
<td>7.72</td>
<td>0.46</td>
<td>56.87</td>
</tr>
<tr>
<td>9</td>
<td>293.85</td>
<td>0.00</td>
<td>10.39</td>
<td>0.32</td>
<td>64.57</td>
</tr>
<tr>
<td>10</td>
<td>332.39</td>
<td>0.00</td>
<td>15.80</td>
<td>0.11</td>
<td>74.82</td>
</tr>
<tr>
<td>11</td>
<td>363.26</td>
<td>0.00</td>
<td>20.36</td>
<td>0.04</td>
<td>89.91</td>
</tr>
<tr>
<td>12</td>
<td>394.20</td>
<td>0.00</td>
<td>22.55</td>
<td>0.03</td>
<td>97.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Gallo & Otranto  
Changing Average Volatility  
Paris, Mar 30, 2015
## Estimated unconditional volatility

<table>
<thead>
<tr>
<th></th>
<th>No regime</th>
<th>Regime 1</th>
<th>Regime 2</th>
<th>Regime 3</th>
<th>Regime 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEM</td>
<td>14.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMEM</td>
<td>14.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS(3)–AMEM</td>
<td></td>
<td>9.21</td>
<td>14.39</td>
<td>28.80</td>
<td></td>
</tr>
<tr>
<td>MS(4)–AMEM</td>
<td></td>
<td>9.28</td>
<td>13.74</td>
<td>26.76</td>
<td>56.97</td>
</tr>
<tr>
<td>MS(3)–AMEM(d)</td>
<td></td>
<td>9.20</td>
<td>14.35</td>
<td>28.37</td>
<td></td>
</tr>
<tr>
<td>TVTP–MS(3)–AMEM</td>
<td></td>
<td>8.69</td>
<td>14.37</td>
<td>55.99</td>
<td></td>
</tr>
<tr>
<td>ST-AMEM</td>
<td></td>
<td>7.20</td>
<td>27.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMEM</td>
<td>14.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS(3)–HMEM</td>
<td></td>
<td>9.13</td>
<td>7.30</td>
<td>67.33</td>
<td></td>
</tr>
</tbody>
</table>
Some limitations of the ST–AMEM
## Best Forecasting Sets - MCS

<table>
<thead>
<tr>
<th>Model</th>
<th>Absolute errors</th>
<th>Squared errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEM</td>
<td>◇</td>
<td>◇</td>
</tr>
<tr>
<td>AMEM</td>
<td>◇ ●</td>
<td>◇ ●</td>
</tr>
<tr>
<td>MS(3)–AMEM</td>
<td>◇ ●</td>
<td>◇ ●</td>
</tr>
<tr>
<td>MS(4)–AMEM</td>
<td>◇</td>
<td>◇ ●</td>
</tr>
<tr>
<td>MS(3)–AMEM(d)</td>
<td>◇ ●</td>
<td>◇ ●</td>
</tr>
<tr>
<td>TVTP–MS(3)–AMEM</td>
<td>★ ◇ ●</td>
<td>★ ●</td>
</tr>
<tr>
<td>ST–AMEM</td>
<td>◇ ●</td>
<td>◇ ●</td>
</tr>
<tr>
<td>HMEM</td>
<td></td>
<td>●</td>
</tr>
<tr>
<td>MS(3)–HMEM</td>
<td>◇</td>
<td>◇</td>
</tr>
<tr>
<td>HAR</td>
<td></td>
<td>●</td>
</tr>
</tbody>
</table>
Conclusions

- Evidence of “excessive” persistence in realized volatility
- Issue of time–varying unconditional volatility
- Address nonlinearity versus long memory
  - statistical fit of a flexible function of time (spline)
  - smooth transition (multiplicative structure needed cf. Amado and Teräsvirta, 2012)
  - Markov Switching
- In Markov Switching abrupt changes in line with market episodes
- Transition probabilities could be made dependent on some interpretable forcing event (some results with lagged VIX).
Realized Traffic and Change of Regimes!