National Seminar Series

November 26, 2020
Maciej Augustyniak

Multifractal Discrete Stochastic Volatility
Multifractal discrete stochastic volatility

Maciej Augustyniak$^1$
Arnaud Dufays$^2$
Kassimou Abdoul Haki Maoude$^1$

$^1$Université de Montréal
$^2$Université Laval & Essec Business School

CANSSI National Seminar Series
Nov. 26, 2020
What is this talk about?

- Financial modelling with hidden Markov models (HMMs)
- A little bit of history
- Classic use of HMMs in financial modelling
- Advantages / disadvantages of the classic approach
- Markov switching multifractal model
- Multifractal discrete stochastic volatility
HMMs in financial modelling
Introduction

Financial time series occasionally exhibit dramatic breaks

- Financial crises
- Changes in government policy

The econometrician James D. Hamilton popularized the use of HMMs in economics and finance

- Regime switching models, Markov switching models
Introduction

A potential model to describe the consequences of a dramatic change in the behaviour of a single variable $y_t$ is:

$$y_t = \begin{cases} 
\omega_1 + \phi y_{t-1} + \epsilon_t, & \text{if state 1,} \\
\omega_2 + \phi y_{t-1} + \epsilon_t, & \text{if state 2.}
\end{cases}$$

In the paper by Hamilton (1989), the variable $y_t$ was 100 times the quarterly change in the log of U.S. real GNP over 1951–1984:

“An empirical application of this technique to postwar U.S. real GNP suggests that the periodic shift from a positive growth rate to a negative growth rate is a recurrent feature of the U.S. business cycle, and indeed could be used as an objective criterion for defining and measuring economic recessions. The estimated parameter values suggest that a typical economic recession is associated with a 3% permanent drop in the level of GNP.”
Financial time series

The previous time series (U.S. real GNP) was a quarterly time series. Whenever available, it is typical to model financial data over a daily frequency.

- **Returns** on a financial asset with time-\( t \) price \( P_t \):

\[
r_t = \log \left( \frac{P_t}{P_{t-1}} \right) \approx \frac{P_t - P_{t-1}}{P_{t-1}}
\]

- **Realized variances** constructed from high-frequency intraday returns (5-minute intervals):

\[
RV_t = \sum_{i=1}^{n} r_{t,i}^2
\]
Why would we want to model returns on financial assets?

- Economic scenario generator
- Risk assessment
- Investment decision / asset allocation / diversification
- Forecasting
Stylized facts of financial time series
Basic modelling framework
Basic modelling framework

- Returns (demeaned)
  \[ r_t = \sqrt{V_t \epsilon_t}, \quad \epsilon_t \sim \text{i.i.d. } N(0, 1) \]

- Realized variances (recall that \( RV_t = \sum_{i=1}^{n} r_{t,i}^2 \))
  \[ RV_t = V_t \eta_t, \quad \eta_t \sim \text{i.i.d. Gamma}(\alpha, \theta = 1/\alpha) \]

- We are interested in modelling the (latent) return variance \( V_t \)

- Remark: \( r_t^2 \) and \( RV_t \) are noisy observations of \( V_t \) since:
  \[ r_t^2 = V_t \epsilon_t^2, \quad \epsilon_t \sim \text{i.i.d. } \chi^2(1) \]
GARCH approach

- $V_t$ is $\mathcal{F}_{t-1}$-measurable

- In the context of returns:

  $$V_t = \omega + \alpha r_{t-1}^2 + \beta V_{t-1}$$

- In the context of realized variances:

  $$V_t = \omega + \alpha R V_{t-1} + \beta V_{t-1}$$

- Leverage effect: add term

  $$\gamma r_{t-1}^2 \mathbb{1}_{\{r_{t-1} < 0\}} \quad \text{or} \quad \gamma R V_{t-1} \mathbb{1}_{\{r_{t-1} < 0\}}$$
HMM approach

\[ r_t = \mu_{s_t} + \sqrt{V_{s_t}} \epsilon_t, \quad \epsilon_t \sim \text{i.i.d. } N(0, 1) \]

\( s_t \in \{1, 2, 3, 4\} \) a Markov chain with t.p.m. \( P \)

\[
\begin{array}{cccc}
  j & \mu_j & \sqrt{252} V_j \\
  1 & 0.1\% & 8\% \\
  2 & 0.0\% & 15\% \\
  3 & -0.1\% & 25\% \\
  4 & -0.3\% & 50\% \\
\end{array}
\]

\[
P = \begin{pmatrix}
0.97 & 0.03 & 0.00 & 0.00 \\
0.04 & 0.95 & 0.00 & 0.01 \\
0.00 & 0.01 & 0.98 & 0.01 \\
0.00 & 0.00 & 0.04 & 0.96
\end{pmatrix}
\]
HMM approach

Rydén, Teräsvirta and Åsbrink (1998) showed that HMMs can reproduce reasonably well most of the stylized facts

- Volatility clustering
- Heavy tails
- Bonus: nice interpretation

However:

- They argue that the model is “doomed from the start” for replicating the high degree of persistence in volatility that is empirically observed
- Can only generate an autocorrelation function that eventually decays exponentially
HMM approach

- If $V_t$ is a Markov chain on $N$ states, then

$$\text{Corr}[V_t, V_{t+k}] = \sum_{i=2}^{N} w_i \lambda_i^k$$

where $(\lambda_2, \ldots, \lambda_N)$ are the eigenvalues of the transition probability matrix (other than 1)

- When $N = 2$, we have

$$\text{Corr}[V_t, V_{t+k}] = (p_{11} + p_{22} - 1)^k$$

- A Markov chain on 2 states is like an AR(1) process
HMM approach

Quadratic increase in the number of parameters

- $N(N - 1)$ parameters for transition matrix
- $N$ parameters for volatility states
- $N^2$ parameters in total

- $N = 4 \implies 16$ parameters
- $N = 10 \implies 100$ parameters
Towards
multifractal discrete stochastic volatility
Modelling objective

Construct a HMM that addresses the weaknesses of standard HMMs

1. Too few support points
2. Autocorrelations decay too fast
3. The number of parameters increases quadratically with the number of states

Factorial hidden Markov (FHM) model construction

Factorial hidden Markov (FHM) model

- FHM processes (Ghahramani and Jordan, 1997, *Machine Learning*) include multiple hidden Markov chains that evolve independently of each other and that are combined to produce the final state.

- **Remark:** FHM models are a particular case of hierarchical hidden Markov models (Fine et al., 1998, *Machine Learning*), which consists in layers of hidden Markov chains (the chains are no more assumed independent).
Example of a FHM model with two Markov chains

\[ V_t = C_t^{(1)} \cdot C_t^{(2)} \]

where

\[ C_t^{(1)} \in \{A, B\} \quad \text{and} \quad C_t^{(2)} \in \{a, b\} \]
Example of a FHM model with two Markov chains

\( V_t = C_t^{(1)} \cdot C_t^{(2)} \) is as a Markov chain on 4 states:

\[ V_t \in \{Aa, Ab, Ba, Bb\} \]

\[
P_V = \begin{pmatrix}
    p_A & 1 - p_A \\
    1 - p_B & p_B \\
\end{pmatrix} \otimes \begin{pmatrix}
    p_a & 1 - p_a \\
    1 - p_b & p_b \\
\end{pmatrix} = 
\]

\[
\begin{pmatrix}
    p_A p_a & p_A(1 - p_a) & (1 - p_A)p_a & (1 - p_A)(1 - p_a) \\
    p_A(1 - p_b) & p_A p_b & (1 - p_A)(1 - p_b) & p_A(1 - p_a) \\
    (1 - p_B)p_a & (1 - p_B)(1 - p_a) & p_B p_a & p_B(1 - p_a) \\
    (1 - p_B)(1 - p_b) & (1 - p_B)p_b & p_B(1 - p_b) & p_B p_b \\
\end{pmatrix}
\]

- Nb of par: 8 instead of 16 for unrestricted HMM
- In general: \( 2^N \) states \( \Rightarrow \) \( 4N \) par instead of \( 2^{2N} \)
Markov switching multifractal (MSM) model
Markov switching multifractal (MSM) model

The MSM process is constructed from the product of $N$ independent two-state Markov chains, denoted by $V_{t}^{(i)}$, $i = 1 \ldots, N$:

$$V_{t}^{(i)} = \begin{cases} V_{t-1}, & \text{with probability } \phi_{i}, \\ \nu_{0}, & \text{with probability } (1 - \phi_{i})/2, \\ 2 - \nu_{0}, & \text{with probability } (1 - \phi_{i})/2. \end{cases}$$

- Common state space $\forall i$: $V_{t}^{(i)} \in \{\nu_{0}, 2 - \nu_{0}\}$, where $\nu_{0} \in (0, 1)$.
- Transition probability matrix:

$$P^{(i)} = \begin{pmatrix} (1 + \phi_{i})/2 & (1 - \phi_{i})/2 \\ (1 - \phi_{i})/2 & (1 + \phi_{i})/2 \end{pmatrix},$$

where $\phi_{i} = a^{b^{i-1}}$, $i = 1, \ldots, N$, with $a \in (0, 1)$ and $b \in (1, \infty)$. 
Markov switching multifractal (MSM) model

The MSM process models volatility by the product of these chains:

\[ V_t = \sigma^2 \frac{\prod_{i=1}^{N} V_t^{(i)}}{\prod_{i=1}^{N} \mathbb{E} \left[ V_t^{(i)} \right]} \]

Interpretation of the MSM process:

- Financial volatility is impacted by the arrival of news in the market whose impact persist for varying periods of time
Markov switching multifractal (MSM) model

- $V_t$ is a Markov chain on $2^N$ states

- But it can only take $N + 1$ distinct values in the set:

$$
\sigma^2 \nu_0^N, \sigma^2 \nu_0^{N-1}(2 - \nu_0), \sigma^2 \nu_0^{N-2}(2 - \nu_0)^2, \ldots, \sigma^2(2 - \nu_0)^N
$$

- Transition matrix: $P_V = \bigotimes_{i=1}^{N} P^{(i)}$

- Stationary distribution: $2^{-N} \mathbf{1}_{2^N}$

- 4 parameters in total (independent of $N$):

$$
\sigma^2 \in (0, \infty), \quad \nu_0 \in (0, 1), \quad a \in (0, 1), \quad b \in (1, \infty)
$$
Multifractal discrete stochastic volatility (MDSV)
Similarly to the MSM, the MDSV model that we propose is constructed from the product of $N$ independent Markov chains:

$$V_t = \sigma^2 \frac{\prod_{i=1}^{N} V_t^{(i)}}{\prod_{i=1}^{N} \mathbb{E}[V_t^{(i)}]}.$$

**However:**

- The Markov chains $V_t^{(i)}$ are no longer restricted to two states
- The MDSV process generalizes the MSM to dimensions larger than two
Multifractal discrete stochastic volatility (MDSV)

Each Markov chain $V_t^{(i)}$ has:

- $D$-dimensional state space of positive values $\nu^{(i)}$
- Stationary distribution $\pi^{(i)}$
- Transition matrix:

$$P^{(i)} = \phi_i \mathbf{1}_D + (1 - \phi_i) \mathbf{1}_D (\pi^{(i)})'$$
The specific structure imposed on the transition matrix implies that:

$$\log V_t^{(i)} = \varphi_i + \phi_i \log V_{t-1}^{(i)} + u_t^{(i)},$$

where $\varphi_i = (1 - \phi_i)(\pi^{(i)})' \log \nu^{(i)}$ and $u_t^{(i)}$ is a martingale difference sequence, and hence a white noise.

The MDSV process thus corresponds to an aggregation of AR(1) processes:

$$\log V_t = \text{constant} + \sum_{i=1}^{N} \log V_t^{(i)}.$$

The MDSV process can be interpreted as a multicomponent stochastic volatility model.
Link with multicomponent stochastic volatility

Andersen and Bollerslev (1997)\textsuperscript{1}

- They model log-volatility as an aggregation of AR(1) processes
- They argue that (asymptotically) this structure can induce long-memory or long-range dependence (hyperbolic versus exponential decay of the autocorrelation function)

---

Multifractal discrete stochastic volatility (MDSV)

Assumptions to reduce the amount of parameters:

\[ \nu^{(i)} = \nu = (\nu_1 \ \nu_2 \ \ldots \ \nu_D)', \ \forall i = 1, \ldots, N, \]
\[ \nu_j = \nu_0^{2-j}(2 - \nu_0)^{j-1}, \ \ j = 1, \ldots, D, \]
\[ \pi^{(i)} = \pi = (\pi_1 \ \pi_2 \ \ldots \ \pi_D)', \ \forall i = 1, \ldots, N, \]
\[ \pi_j = \binom{D-1}{j-1} \omega^{j-1}(1 - \omega)^{D-j}, \ \ j = 1, \ldots, D. \]
\[ \phi_i = a^{b^{i^{-1}}}, \ \ i = 1, \ldots, N, \]

The MDSV process depends on five parameters (independent of \( N \)):

\[ \sigma^2 \in (0, \infty), \ \nu_0 \in (0, 1), \ \ a \in (0, 1), \ \ b \in [1, \infty), \ \ \omega \in (0, 1) \]
Multifractal discrete stochastic volatility (MDSV)

We write MDSV($N, D$) to designate the MDSV process $V_t$ constructed from the product of $N$ Markov chains of dimension $D$.

Recall that:

$$V_t = \sigma^2 \frac{\prod_{i=1}^{N} V_t^{(i)}}{\prod_{i=1}^{N} \mathbb{E} \left[ V_t^{(i)} \right]}$$

$V_t$ is itself a Markov chain of dimension $D^N$ with state space $\nu$, transition matrix $P_V$, and stationary distribution $\pi_V$, where:

$$\nu = \sigma^2 \frac{\nu \otimes^N}{(\pi' \nu)^N}, \quad P_V = \otimes_{i=1}^{N} P^{(i)}, \quad \pi_V = \pi \otimes^N.$$

**Remark**: Although the state space is of dimension $D^N$, it contains $N(D - 1) + 1$ distinct values.
Multifractal discrete stochastic volatility (MDSV)

The MDSV process unifies under a single framework many regime switching models proposed in the econometric literature:

- The **MSM** process of Calvet and Fisher (2004) is a MDSV($N, D = 2$) with $\omega = 1/2$.

- The component-driven regime switching process (CDRS) of Fleming and Kirby (2013) is a MDSV($N, D = 2$) which does not assume a specific functional form for $\phi_i, i = 1, \ldots, N$.

- The discrete stochastic autoregressive volatility (DSARV) of Cordis and Kirby (2014) is a MDSV($N = 1, D$).

- The factorial hidden Markov volatility (FHMV) of Augustyniak, Bauwens and Dufays (2019) is a MDSV($N, D = 2$) with $b = 1$ and $\omega = 1/2$, but which assumes different state spaces for the Markov chains $V_t^{(i)}, i = 1, \ldots, N$. 

Theorem (Properties of the MDSV process)

(i) Autocovariance structure:

\[
\text{Cov}[V_{t+k}, V_t] = \pi' \Upsilon \left( P^k_V - \Pi_V \right) \upsilon \\
\quad = \sigma^4 \left( \prod_{i=1}^{N} \left( 1 + \left( \psi^{D-1} - 1 \right) \phi_i^k \right) - 1 \right), \quad k \geq 1.
\]

(ii) Asymptotic rate of convergence of \(P_V\) as \(k \to \infty\):

\[
P^k_V - \Pi_V = O \left( k^{D-2} a^k \right).
\]

(iii) The distribution of the time that the Markov chain spends in one of the (non-extremal) \(N(D - 1) + 1\) distinct values of the state space is not geometric (semi-Markov property).

Here, \(\Upsilon\) denotes the \(D_N \times D_N\) diagonal matrix with the elements of \(\upsilon\) on its diagonal, \(\Pi_V = 1_{D_N} \pi'\), and \(\psi = \frac{\nu_0^2 + 2\omega(2-\nu_0)}{(\nu_0 + 2\omega(1-\nu_0))^2} > 1\).
Application to data
Application to data

We assume that the MDSV process is related to returns $r_t$ and realized variances $RV_t$ in the following way:

$$r_t = \sqrt{V_t} \epsilon_t,$$

$$\log RV_t = \xi + \varphi \log V_t + \delta_1 \epsilon_t + \delta_2 (\epsilon_t^2 - 1) + \gamma \epsilon_t,$$

where $\xi \in \mathbb{R}$, $\varphi \in \mathbb{R}$, $\delta_1 \in \mathbb{R}$, $\delta_2 \in \mathbb{R}$ and $\gamma \in (0, \infty)$ are parameters, and $\epsilon_t$ and $\varepsilon_t$ are mutually and serially independent innovation processes with mean 0 and variance 1.

Time period for the data: 2000–2019.
Comparison of the joint fit to returns and realized variances

<table>
<thead>
<tr>
<th>Models</th>
<th>Real EGARCH</th>
<th>Real MS</th>
<th>.</th>
<th>MDSV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K</td>
<td>K</td>
<td>(N, D)</td>
<td>(N, D)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>(1, 1024)</td>
<td>(2, 32)</td>
</tr>
<tr>
<td>Np</td>
<td>10</td>
<td>11</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S&amp;P500 $(n = 4934)$</td>
</tr>
<tr>
<td>log-lik</td>
<td>-6785.6</td>
<td>-7747.1</td>
<td>-7141.6</td>
<td>-6765.6</td>
</tr>
<tr>
<td>AIC</td>
<td>-6795.6</td>
<td>-7758.1</td>
<td>-7164.6</td>
<td>-6776.6</td>
</tr>
<tr>
<td>BIC</td>
<td>-6828.1</td>
<td>-7793.8</td>
<td>-7239.4</td>
<td>-6812.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NASDAQ $(n = 4928)$</td>
</tr>
<tr>
<td>log-lik</td>
<td>-8607.8</td>
<td>-9719.4</td>
<td>-8999.1</td>
<td>-8636.4</td>
</tr>
<tr>
<td>AIC</td>
<td>-8617.8</td>
<td>-9730.4</td>
<td>-9022.1</td>
<td>-8647.4</td>
</tr>
<tr>
<td>BIC</td>
<td>-8650.3</td>
<td>-9766.2</td>
<td>-9096.9</td>
<td>-8683.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FTSE $(n = 4957)$</td>
</tr>
<tr>
<td>log-lik</td>
<td>-7730.8</td>
<td>-8311.5</td>
<td>-7916.2</td>
<td>-7816.4</td>
</tr>
<tr>
<td>AIC</td>
<td>-7740.8</td>
<td>-8322.5</td>
<td>-7939.1</td>
<td>-7827.4</td>
</tr>
<tr>
<td>BIC</td>
<td>-7773.3</td>
<td>-8358.3</td>
<td>-8014.0</td>
<td>-7863.2</td>
</tr>
</tbody>
</table>
Comparison of the forecasting performance of $RV_t$

\[
RMSFE(h) = \sqrt{\frac{1}{756 - h + 1} \sum_{t=1}^{756-h} \left( \frac{1}{h} \sum_{i=1}^{h} \hat{RV}_{t+h} - \frac{1}{h} \sum_{i=1}^{h} RV_{t+h} \right)^2}
\]

<table>
<thead>
<tr>
<th>horizon(h)</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>25</th>
<th>50</th>
<th>75</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDSV(3,10)</td>
<td>0.49</td>
<td>0.44</td>
<td>0.43</td>
<td>0.40</td>
<td>0.37</td>
<td>0.38</td>
<td>0.36</td>
</tr>
<tr>
<td>MDSV(10,2)</td>
<td><strong>0.44</strong></td>
<td><strong>0.41</strong></td>
<td><strong>0.40</strong></td>
<td><strong>0.39</strong></td>
<td>0.37</td>
<td><strong>0.37</strong></td>
<td><strong>0.35</strong></td>
</tr>
<tr>
<td>Real-EGARCH</td>
<td>0.45</td>
<td>0.43</td>
<td>0.44</td>
<td>0.43</td>
<td>0.42</td>
<td>0.45</td>
<td>0.46</td>
</tr>
<tr>
<td>Real-MS(2)</td>
<td>0.52</td>
<td>0.54</td>
<td>0.55</td>
<td>0.54</td>
<td>0.56</td>
<td>0.60</td>
<td>0.61</td>
</tr>
<tr>
<td>Real-MS(4)</td>
<td>0.48</td>
<td>0.53</td>
<td>0.55</td>
<td>0.66</td>
<td>0.88</td>
<td>1.02</td>
<td>1.11</td>
</tr>
<tr>
<td><strong>S&amp;P500</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDSV(3,10)</td>
<td>0.49</td>
<td>0.44</td>
<td>0.43</td>
<td>0.40</td>
<td>0.37</td>
<td>0.38</td>
<td>0.36</td>
</tr>
<tr>
<td>MDSV(10,2)</td>
<td><strong>0.44</strong></td>
<td><strong>0.41</strong></td>
<td><strong>0.40</strong></td>
<td><strong>0.39</strong></td>
<td>0.37</td>
<td><strong>0.37</strong></td>
<td><strong>0.35</strong></td>
</tr>
<tr>
<td>Real-EGARCH</td>
<td>0.45</td>
<td>0.43</td>
<td>0.44</td>
<td>0.43</td>
<td>0.42</td>
<td>0.45</td>
<td>0.46</td>
</tr>
<tr>
<td>Real-MS(2)</td>
<td>0.52</td>
<td>0.54</td>
<td>0.55</td>
<td>0.54</td>
<td>0.56</td>
<td>0.60</td>
<td>0.61</td>
</tr>
<tr>
<td>Real-MS(4)</td>
<td>0.48</td>
<td>0.53</td>
<td>0.55</td>
<td>0.66</td>
<td>0.88</td>
<td>1.02</td>
<td>1.11</td>
</tr>
</tbody>
</table>
Conclusion

- The MDSV process is proposed as a new HMM approach to model financial volatility.
- Unifies under a single framework many regime switching models proposed in the econometric literature, in particular it generalizes the MSM model to dimensions larger than two.
- Relationship to multicomponent stochastic volatility model.
- Large state space.
- Ability to generate a high degree of volatility persistence.
- Semi-Markov property on the distinct values of the state space.
- Parsimonious (5 parameters).
- Economic interpretation: arrival of news whose impact persist for varying periods of time.
Get your questions ready!
Q & A will start in 1 minute.

Please post your questions in the Q & A part, not the chat. You can upvote questions that are similar to your own, rather than typing a duplicate question. Use the thumbs up!