A Markov-Switching Multi-Fractal Inter-Trade Duration Model, with Application to U.S. Equities

Fei Chen (HUST)
Francis X. Diebold (UPenn)
Frank Schorfheide (UPenn)

June 26, 2012
Why More Focus on Durations Now?

- The duration point process is the ultimate process of interest. It determines everything else, yet it remains incompletely understood.
Why More Focus on Durations Now?

- The duration point process is the ultimate process of interest. It determines everything else, yet it remains incompletely understood.

- We need better understanding of information arrival, trade arrival, liquidity, volume, market participant interactions, links to volatility, etc.
Why More Focus on Durations Now?

- The duration point process is the ultimate process of interest. It determines everything else, yet it remains incompletely understood.

- We need better understanding of information arrival, trade arrival, liquidity, volume, market participant interactions, links to volatility, etc.
  - Financial market roots of the Great Recession
  - Purely financial-market events like the Flash Crash

- Duration literature lags the vol literature in an important way. Long memory is clearly present in calendar-time volatility and is presumably inherited from conditional intensity of arrivals in transactions time, yet there is little long-memory duration literature.
For What is “Big Data” Informative?

For trends? No

Andersen, Bollerslev, Christoffersen and Diebold (2012)

For volatilities? Yes: realized volatility.

Hautsch (2012)

– Trivially: trade-by-trade data needed for inter-trade durations

– Subtly: time deformation links volatilities to durations.

So big data informs us about vols which inform us about durations.
For What is “Big Data” Informative?

For trends? No

For volatilities? Yes: realized volatility.
Andersen, Bollerslev, Christoffersen and Diebold (2012)
For What is “Big Data” Informative?

For trends? No

For volatilities? Yes: realized volatility.
Andersen, Bollerslev, Christoffersen and Diebold (2012)

For durations? Yes: both trivially and subtly.
Hautsch (2012)

– Trivially: trade-by-trade data needed for inter-trade durations
– Subtly: time deformation links volatilities to durations.
So big data informs us about vols which inform us about durations.
Stochastic Volatility Model

\[ r_t = \sigma \sqrt{e^{h_t}} \cdot \varepsilon_t \]

\[ h_t = \rho h_{t-1} + \eta_t \]

\[ \varepsilon_t \sim iid\mathcal{N}(0,1) \]

\[ \eta_t \sim iid\mathcal{N}(0,\sigma^2_\eta) \]

\[ \varepsilon_t \perp \eta_t \]

Equivalently,

\[ r_t|\Omega_{t-1} \sim N(0,\sigma^2 e^{h_t}) \]
From Where Does Stochastic Volatility Come?

Time-deformation model of calendar time (e.g., “daily”) returns:

\[ r_t = \sum_{i=1}^{e^{ht}} r_i \]

\[ h_t = \rho h_{t-1} + \eta_t \]

(trade-by-trade returns \( r_i \sim iidN(0, \sigma^2) \), daily volume \( e^{ht} \))

\[ \implies r_t|\Omega_{t-1} \sim N(0, \sigma^2 e^{ht}) \]

- Volume/duration dynamics produce volatility dynamics
- Volatility properties inherited from duration properties
What Are the Key Properties of Volatility?

In general:

- Volatility dynamics fatten unconditional distributional tails
e.g., $r_t | \Omega_{t-1} \sim N(0, \sigma^2 e^{h_t}) \implies r_t \sim \text{“fat-tailed”}$
What Are the Key Properties of Volatility?

In general:

- Volatility dynamics fatten unconditional distributional tails
  
e.g., \( r_t \mid \Omega_{t-1} \sim N(0, \sigma^2 e^{ht}) \implies r_t \sim \text{“fat-tailed”} \)

- Volatility dynamics are persistent

Calvet and Fischer (2008), Multifractal Volatility: Theory, Forecasting, and Pricing, Elsevier
What Are the Key Properties of Volatility?

In general:

- Volatility dynamics fatten unconditional distributional tails
  e.g., $r_t|\Omega_{t-1} \sim N(0, \sigma^2 e^{ht}) \implies r_t \sim \text{“fat-tailed”}$

- Volatility dynamics are persistent

- Volatility dynamics are long memory
What Are the Key Properties of Volatility?

In general:

- Volatility dynamics fatten unconditional distributional tails. e.g., \( r_t | \Omega_{t-1} \sim N(0, \sigma^2 \epsilon^h_t) \implies r_t \sim \text{“fat-tailed”} \)

- Volatility dynamics are persistent

- Volatility dynamics are long memory

Elegant modeling framework that captures all properties:

Calvet and Fischer (2008), 
*Multifractal Volatility: Theory, Forecasting, and Pricing*, Elsevier
Roadmap

- Empirical regularities in durations
- The MSMD model
- Empirics
Twenty-Five U.S. Firms Selected Randomly from S&P 100

- Consolidated trade data extracted from the TAQ database
- 20 days, 2/1/1993 - 2/26/1993, 10:00 - 16:00
- 09:30-10:00 excluded to eliminate opening effects

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Company Name</th>
<th>Symbol</th>
<th>Company Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>ALCOA</td>
<td>ABT</td>
<td>Abbott Laboratories</td>
</tr>
<tr>
<td>AXP</td>
<td>American Express</td>
<td>BA</td>
<td>Boeing</td>
</tr>
<tr>
<td>BAC</td>
<td>Bank of America</td>
<td>C</td>
<td>Citigroup</td>
</tr>
<tr>
<td>CSCO</td>
<td>Cisco Systems</td>
<td>DELL</td>
<td>Dell</td>
</tr>
<tr>
<td>DOW</td>
<td>Dow Chemical</td>
<td>F</td>
<td>Ford Motor</td>
</tr>
<tr>
<td>GE</td>
<td>General Electric</td>
<td>HD</td>
<td>Home Depot</td>
</tr>
<tr>
<td>IBM</td>
<td>IBM</td>
<td>INTC</td>
<td>Intel</td>
</tr>
<tr>
<td>JNJ</td>
<td>Johnson &amp; Johnson</td>
<td>KO</td>
<td>Coca-Cola</td>
</tr>
<tr>
<td>MCD</td>
<td>McDonald’s</td>
<td>MRK</td>
<td>Merck</td>
</tr>
<tr>
<td>MSFT</td>
<td>Microsoft</td>
<td>QCOM</td>
<td>Qualcomm</td>
</tr>
<tr>
<td>T</td>
<td>AT&amp;T</td>
<td>TXN</td>
<td>Texas Instruments</td>
</tr>
<tr>
<td>WFC</td>
<td>Wells Fargo</td>
<td>WMT</td>
<td>Wal-Mart</td>
</tr>
<tr>
<td>XRX</td>
<td>Xerox</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: Stock ticker symbols and company names.
Overdispersion

Figure: **Citigroup Duration Distribution.** We show an exponential QQ plot for Citigroup inter-trade durations between 10:00am and 4:00pm during February 1993, adjusted for calendar effects.
Figure: **Citigroup Duration Time Series.** We show a time-series plot of inter-trade durations between 10:00am and 4:00pm during February 1993, measured in minutes and adjusted for calendar effects.
Figure: **Citigroup Duration Autocorrelations.** We show the sample autocorrelation function of Citigroup inter-trade durations between 10:00am and 4:00pm during February 1993, adjusted for calendar effects.
Roadmap

- Empirical regularities in inter-trade durations ✔
- The MSMD model
- Empirics
A Dynamic Duration Model

\[ d_i \sim \frac{\epsilon_i}{\lambda_i} \]

\[ \epsilon_i \sim iidExp(1) \]

How to parameterize the conditional intensity \( \lambda_i \)?
Markov Switching Multifractal Durations (MSMD)

\[ \lambda_i = \lambda \prod_{k=1}^{\bar{k}} M_{k,i} \]

\[ \lambda > 0, \ M_{k,i} > 0, \ \forall k \]
Markov Switching Multifractal Durations (MSMD)

\[ \lambda_i = \lambda \prod_{k=1}^{\bar{k}} M_{k,i} \]

\[ \lambda > 0, \ M_{k,i} > 0, \ \forall k \]

Independent intensity components \( M_{k,i} \) are Markov renewal processes:

\[ M_{k,i} = \begin{cases} 
\text{draw from } f(M) & \text{w.p. } \gamma_k \\
M_{k,i-1} & \text{w.p. } 1 - \gamma_k
\end{cases} \]
Markov Switching Multifractal Durations (MSMD)

\[
\lambda_i = \lambda \prod_{k=1}^{\bar{k}} M_{k,i}
\]

\[
\lambda > 0, \; M_{k,i} > 0, \; \forall k
\]

Independent intensity components \( M_{k,i} \)
are Markov renewal processes:

\[
M_{k,i} = \begin{cases} 
\text{draw from } f(M) & \text{w.p. } \gamma_k \\
M_{k,i-1} & \text{w.p. } 1 - \gamma_k
\end{cases}
\]

\( f(M) \) is identical \( \forall k \), with \( M > 0 \) and \( E(M) = 1 \)
Modeling Choices

Simple binomial renewal distribution $f(M)$:

$$M = \begin{cases} 
  m_0 & \text{w.p. } 1/2 \\
  2 - m_0 & \text{w.p. } 1/2,
\end{cases}$$

where $m_0 \in (0, 2]$
Modeling Choices

Simple binomial renewal distribution $f(M)$:

$$M = \begin{cases} 
m_0 & \text{w.p. } 1/2 \\
2 - m_0 & \text{w.p. } 1/2,
\end{cases}$$

where $m_0 \in (0, 2]$

Simple renewal probability $\gamma_k$:

$$\gamma_k = 1 - (1 - \gamma_{\bar{k}})^{b^{k-\bar{k}}}$$

$\gamma_{\bar{k}} \in (0, 1)$ and $b \in (1, \infty)$
Modeling Choices

Simple binomial renewal distribution $f(M)$:

\[ M = \begin{cases} 
  m_0 & \text{w.p. } 1/2 \\
  2 - m_0 & \text{w.p. } 1/2,
\end{cases} \]

where $m_0 \in (0, 2]$

Simple renewal probability $\gamma_k$:

\[ \gamma_k = 1 - (1 - \gamma_{\bar{k}})^{b^{k-\bar{k}}} \]

$\gamma_{\bar{k}} \in (0, 1)$ and $b \in (1, \infty)$

(The tightly-parametric approach is intentional!)
Renewal Probabilities

Figure: MSMD Intensity Component Renewal Probabilities. We show the renewal probabilities \( \gamma_k = 1 - (1 - \gamma_{\bar{k}})^{b^k - \bar{k}} \) associated with the latent intensity components \( M_k \), \( k = 3, \ldots, 7 \). We calibrate the MSMD model with \( \bar{k} = 7 \), and with remaining parameters that match our subsequently-reported estimates for Citigroup.
All Together Now

\[ d_i = \frac{\epsilon_i}{\lambda_i} \]

\[ \lambda_i = \lambda \prod_{k=1}^{\bar{k}} M_{k,i} \]

\[ M_{k,i} = \begin{cases} M & \text{w.p. } 1 - (1 - \gamma_k)^{b^{k-\bar{k}}} \\ M_{k,i-1} & \text{w.p. } (1 - \gamma_k)^{b^{k-\bar{k}}} \end{cases} \]

\[ M = \begin{cases} m_0 & \text{w.p. } 1/2 \\ 2 - m_0 & \text{w.p. } 1/2 \end{cases} \]

\[ \epsilon_i \sim iid \ exp(1), \; \bar{k} \in \mathbb{N}, \; \lambda > 0, \; \gamma_k \in (0, 1), \; b \in (1, \infty), \; m_0 \in (0, 2] \]

parameters \( \theta_{\bar{k}} = (\lambda, \gamma_{\bar{k}}, b, m_0)' \)

\( \bar{k} \)-dimensional state, \( M_i = (M_{1,i}, M_{2,i}, \ldots M_{\bar{k},i}) \)

\( 2^{\bar{k}} \) states
Figure: **QQ Plot, Simulated Durations.** We show a QQ plot for a simulated duration sample path for an MSMD model with sample size \(N = 22,988\) and parameters calibrated to match our subsequently-reported estimates for Citigroup.
Figure: **Time-series plots of simulated** $M_{1,i}$, ..., $M_{7,i}$, $\lambda_i$, and $d_i$.

We show simulated sample paths for an MSMD model with sample size ($N = 22,988$) and parameters calibrated to match Citigroup estimates.
Figure: Sample Autocorrelation Function, Simulated Durations. We show the sample autocorrelation function for a simulated sample path from an MSMD model with sample size ($N = 22,988$) and parameters calibrated to match our subsequently-reported estimates for Citigroup.
The MSMD autocorrelation function satisfies

$$\sup_{\tau \in \ell_{\bar{k}}} \left| \frac{\ln \rho(\tau)}{\ln \tau^{-\delta}} - 1 \right| \to 0 \quad \text{as} \quad \bar{k} \to \infty$$

$$\delta = \log_b E(\tilde{M}^2) - \log_b \{[E(\tilde{M})]^2\}$$

$$\tilde{M} = \begin{cases} \frac{2m_0^{-1}}{m_0^{-1} + (2-m_0)^{-1}} & \text{w.p. } \frac{1}{2} \\ \frac{2(2-m_0)^{-1}}{m_0^{-1} + (2-m_0)^{-1}} & \text{w.p. } \frac{1}{2} \end{cases}$$
Literature I:
Mean Duration vs. Mean Intensity

Mean Duration:

\[ d_i = \varphi_i \epsilon_i, \quad \epsilon_i \sim iid(1, \sigma^2) \]

– ACD: Engle and Russell (1998), ...
  – MEM: Engle (2002), ...

Mean Intensity:

\[ d_i \sim \frac{\epsilon_i}{\lambda_i}, \quad \epsilon_i \sim iidExp(1) \]

– MSMD
– Bauwens and Hautsch (2006)
– Bowsher (2006)
Literature II: Observation- vs. Parameter-Driven Models

Observation-Driven:

\[ \Omega_{t-1} \text{ observed (like GARCH)} \]

- ACD
- MEM as typically implemented
  - GAS

Parameter-Driven:

\[ \Omega_{t-1} \text{ latent (like SV)} \]

- MSMD
Literature III: Short Memory vs. Long Memory

Short Memory:

Quick (exponential) duration autocorrelation decay

- ACD as typically implemented
- MEM as typically implemented

Long-Memory:

Slow (hyperbolic) duration autocorrelation decay

- MSMD
  - FI-ACD: Jasiak (1999)
  - FI-SCD: Deo, Hsieh and Hurvich (2010)
Literature IV: Reduced-Form vs. Structural Long Memory

Reduced Form:

\[(1 - L)^d \lambda_t = \nu_t, \quad \nu_t \sim \text{short memory}\]

- FI-ACD
- FI-SCD

Structural:

\[\lambda_t = \nu_{1t} + \ldots + \nu_{Nt}, \quad \nu_{it} \sim \text{short memory}\]

- MSMD
Roadmap

- Empirical regularities in inter-trade durations ✅
- The MSMD model ✅
- Empirics
Twenty-Five U.S. Firms Selected Randomly from S&P 100

- Consolidated trade data extracted from the TAQ database
- 20 days, 2/1/1993 - 2/26/1993, 10:00 - 16:00
- 09:30-10:00 excluded to eliminate opening effects

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Company Name</th>
<th>Symbol</th>
<th>Company Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>ALCOA</td>
<td>ABT</td>
<td>Abbott Laboratories</td>
</tr>
<tr>
<td>AXP</td>
<td>American Express</td>
<td>BA</td>
<td>Boeing</td>
</tr>
<tr>
<td>BAC</td>
<td>Bank of America</td>
<td>C</td>
<td>Citigroup</td>
</tr>
<tr>
<td>CSCO</td>
<td>Cisco Systems</td>
<td>DELL</td>
<td>Dell</td>
</tr>
<tr>
<td>DOW</td>
<td>Dow Chemical</td>
<td>F</td>
<td>Ford Motor</td>
</tr>
<tr>
<td>GE</td>
<td>General Electric</td>
<td>HD</td>
<td>Home Depot</td>
</tr>
<tr>
<td>IBM</td>
<td>IBM</td>
<td>INTC</td>
<td>Intel</td>
</tr>
<tr>
<td>JNJ</td>
<td>Johnson &amp; Johnson</td>
<td>KO</td>
<td>Coca-Cola</td>
</tr>
<tr>
<td>MCD</td>
<td>McDonald’s</td>
<td>MRK</td>
<td>Merck</td>
</tr>
<tr>
<td>MSFT</td>
<td>Microsoft</td>
<td>QCOM</td>
<td>Qualcomm</td>
</tr>
<tr>
<td>T</td>
<td>AT&amp;T</td>
<td>TXN</td>
<td>Texas Instruments</td>
</tr>
<tr>
<td>WFC</td>
<td>Wells Fargo</td>
<td>WMT</td>
<td>Wal-Mart</td>
</tr>
<tr>
<td>XRX</td>
<td>Xerox</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table: Stock ticker symbols and company names.
Figure: Distribution of Duration Coefficients of Variation Across Firms. We show a histogram of coefficients of variation (the standard deviation relative to the mean), as a measure of overdispersion relative to the exponential. For reference we indicate Citigroup.
Figure: **Duration Autocorrelation Function Profile Bundle.** For each firm, we show the sample autocorrelation function of inter-trade durations between 10:00am and 4:00pm during February 1993, adjusted for calendar effects. For reference we show Citigroup in bold.
MSMD Likelihood Evaluation (Using $\bar{k} = 2$ for Illustration)

Each $M_k$, $k = 1, 2$ is a two-state Markov switching process:

$$\mathcal{P}(\gamma_k) = \begin{bmatrix} 1 - \gamma_k/2 & \gamma_k/2 \\ \gamma_k/2 & 1 - \gamma_k/2 \end{bmatrix}$$

Hence $\lambda_i$ is a four-state Markov-switching process:

$$\lambda_i \in \{\lambda s_1 s_1, \lambda s_1 s_2, \lambda s_2 s_1, \lambda s_2 s_2\}$$

$$\mathcal{P}_\lambda = \mathcal{P}(\gamma_1) \otimes \mathcal{P}(\gamma_2) \text{ (by independence of the } M_{k,i})$$

Likelihood function:

$$p(d_{1:n}|\theta_{\bar{k}}) = p(d_1|\theta_{\bar{k}}) \prod_{i=2}^{n} p(d_i|d_{1:i-1}, \theta_{\bar{k}})$$

Conditional on $\lambda_i$, the duration $d_i$ is $\text{Exp}(\lambda_i)$:

$$p(d_i|\lambda_i) = \lambda_i e^{-\lambda_i d_i}$$

Weight by state probabilities obtained by the Hamilton filter.
Figure: Maximized Log Likelihood Profile Bundle. We show likelihood profiles for all firms as a function of $\bar{k}$, in deviations from the value for $\bar{k} = 7$, which is therefore identically equal to 0. For reference we show Citigroup in bold.
Figure: Distributions of MSMD Parameter Estimates Across Firms, $\bar{k} = 7$. We show histograms of maximum likelihood parameter estimates across firms, obtained using $\bar{k} = 7$. For reference we indicate Citigroup.
Figure: Estimated Intensity Component Renewal Probability Profile Bundle, \( \tilde{k} = 7 \). For reference we show Citigroup in bold.
Figure: Empirical CDF of White Statistic $p$-Value, $\bar{k} = 7$. For reference we indicate Citigroup.
Figure: Distribution of Differences in BIC Values Across Firms. We compute differences as $D = \text{MSMD}(7) - \text{ACD}(1,1)$. We show a histogram. For reference we indicate Citigroup.
Roadmap

- Empirical regularities in inter-trade durations ✓
- The MSMD model ✓
- Empirics ✓
Future Directions

- More recent data
- Recently-introduced ultra-accurate transactions time stamps
- Panel of trading months. Structural change?