High Frequency Quoting: Short-Term Volatility in Bids and Offers

Joel Hasbrouck
Stern School, NYU

Financial Econometrics Conference
Toulouse School of Economics
Disclaimers

- I teach in an entry-level training program at a large financial firm that is generally thought to engage in high-frequency trading has been named as a defendant in an HFT lawsuit.
- I serve on a CFTC advisory committee that discusses issues related to high frequency trading.
- I accept honoraria for presentations at events sponsored by financial firms.
What does quote volatility look like?

- In US equity markets, a bid or offer can originate from any market participant.
  - “Traditional” dealers, retail and institutional investors.
- Bids and offers from all trading venues are consolidated and disseminated in real time.
  - The highest bid is the National Best Bid (NBB)
  - The lowest offer is the National Best Offer (NBO)
- Next slide: the NBBO for AEPI on April 29, 2011
Figure 1. AEPI bid and offer, April 29, 2011
Figure 1. AEPI bid and offer on April 29, 2011 (detail)
Features of the AEPI episodes

- Extremely rapid *oscillations* in the bid.
- Start and stop abruptly
- Mostly one-sided
  - activity on the ask side is much smaller
- Episodes don’t coincide with large long-term changes in the stock price.
Quote volatility: why worry?

- **Noise**
  - The quotes are price signals. Noise degrades the value of these signals.
- **Execution price risk (for marketable orders and dark trades)**
  - We don’t know and can’t time exactly when our order will reach the market.
  - Quote volatility links arrival uncertainty to execution price risk.
Quote volatility: the questions

- What is its economic meaning and importance?
- How should we measure it?
- Is it elevated? Relative to what?
- Has it increased along with wider adoption of high-speed trading technology?
Context and connections

- Analyses of high frequency trading (HTF)
- Traditional volatility modeling
- Methodology: time scale resolution and variance estimation
- Economic models of dynamic oligopolistic pricing.
Traditional volatility modeling

- Mainstream ARCH, GARCH, and similar models focus on fundamental/informational volatility.
  - Statistically: volatility in the unit-root component of prices.
  - Economically important for portfolio allocation, derivatives valuation and hedging.

- Quote volatility is non-informational
  - Statistically: short-term, stationary, transient volatility
  - Economically important for trading and market making.
Statistics are local variances about local means.
Connection to pre-averaging

- Local averaging of price levels is used to remove microstructure noise prior to modeling fundamental variances.
- The local volatility is generally not studied.
- Here, it is the focus.
Computational issues

- In computing a local average ...
  - How long should the averaging period be?
  - How should the averaging periods be aligned?

- Wavelet transformations simply provide computationally efficient techniques for
  - considering a range of averaging periods
  - obtaining alignment-invariant estimates.
The origins of high frequency quoting: Suggestions from economic theory

- Price volatility can result from randomized strategies.
  - Varian (1980)
- Edgeworth cycles
  - Progressive undercutting until all producers but one exit the market
  - The remaining producer raises his price to the monopoly level. Repeat.
  - Masking and Tirole (1988)
Descriptive statistics:
computation and interpretation
Local variances about local means

\( n = \) length of averaging interval. Depends on trader’s latency and order strategies: we want a range of \( n \)
To assess economic importance, I present the volatility estimates in three ways.

- In mils ($0.001) per share
- In basis points
- As a short-term/long-term variance ratio
The short/long variance ratio

- For a random walk with per period variance $\sigma^2$, the variance of the $n$-period difference is $n\sigma^2$.

- An conventional variance ratio might be
  
  \[ V = \frac{60 \times \text{one minute return variance}}{\text{one hour return variance}} \]

- For a random walk, $V = 1$.
  - Microstructure: we usually find $V > 1$.

- Extensively used in microstructure studies: Barnea (1974); Amihud and Mendelson (1987); etc.
The empirical analysis

CRSP Universe 2001-2011. (Share code = 10 or 11; average price $2 to $1,000; listing NYSE, Amex or NASDAQ)

In each year, chose 150 firms in a random sample stratified by dollar trading volume

2011 April TAQ with one-millisecond time stamps

High-resolution analysis

2001-2011 April TAQ data with one-second time stamps

Lower-resolution analysis
Figure 2. Wavelet variance ratios across time scale and dollar volume quintiles

**Chart Description:**
- **Y-axis:** Normalized quote variance
- **X-axis:** Time scale (milliseconds)
- **Time Scale Markings:**
  - 100 ms
  - 1 s
  - 10 s
  - 1 m
  - 20 m
  - 10 ms
  - 100 ms
  - 1,000 ms
  - 10.0 sec
  - 100.0 sec
  - 16.7 min
  - 166.7 min
- **Avg Dollar Volume Ranks:**
  - 1 (low)
  - 2
  - 3
  - 4
  - 5 (high)

The chart illustrates the wavelet variance ratios across different time scales and dollar volume quintiles.
Variance ratios: short term volatility is much higher than we’d expect relative to a random-walk.

In mils per share or basis points, average short term volatility is economically meaningful, but small.
Historical analysis

CRSP Universe 2001-2011. (Share code = 10 or 11; average price $2 to $1,000; listing NYSE, Amex or NASDAQ)

In each year, chose 150 firms in a random sample stratified by dollar trading volume

2011 April TAQ with one-millisecond time stamps

High-resolution analysis

2001-2011 April TAQ data with one-second time stamps

Lower-resolution analysis
High-resolution analysis ...
... with low resolution data

- TAQ with millisecond time stamps only available from 2006 onwards
- TAQ with one second time stamps available back to 1993.
- Can we draw inferences about subsecond variation from second-stamped data?
- Yes, if we are confident in the ordering of the data.
Recall the constant intensity Poisson process ...

- $N(t)$ = no. of events in an interval $(0, t)$
- $s_i$ = arrival time of event $i$
- If $N(t) = n$, then $s_1, s_2, ..., s_n$ have the same distribution as the order statistics in a sample of $n$ independent $U(0, t)$ random variables.
- This suggests that millisecond remainders can be easily simulated.
Table 5. Summary statistics, historical sample, 2001-2011 (only odd numbered years are shown)

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th>2003</th>
<th>2005</th>
<th>2007</th>
<th>2009</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. firms</td>
<td>137</td>
<td>141</td>
<td>144</td>
<td>150</td>
<td>145</td>
<td>149</td>
</tr>
<tr>
<td>NYSE</td>
<td>106</td>
<td>51</td>
<td>48</td>
<td>55</td>
<td>56</td>
<td>47</td>
</tr>
<tr>
<td>Amex</td>
<td>16</td>
<td>10</td>
<td>8</td>
<td>14</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>15</td>
<td>80</td>
<td>88</td>
<td>81</td>
<td>84</td>
<td>96</td>
</tr>
<tr>
<td>Avg. daily trades</td>
<td>167</td>
<td>231</td>
<td>448</td>
<td>970</td>
<td>1,993</td>
<td>1,346</td>
</tr>
<tr>
<td>Avg. daily quotes</td>
<td>1,525</td>
<td>1,470</td>
<td>6,004</td>
<td>12,521</td>
<td>41,571</td>
<td>24,053</td>
</tr>
<tr>
<td>Avg. daily NBB changes</td>
<td>128</td>
<td>210</td>
<td>611</td>
<td>772</td>
<td>1,787</td>
<td>1,225</td>
</tr>
<tr>
<td>Avg. daily NBO changes</td>
<td>127</td>
<td>226</td>
<td>729</td>
<td>789</td>
<td>1,789</td>
<td>1,146</td>
</tr>
<tr>
<td>Avg. price</td>
<td>$20.57</td>
<td>$14.41</td>
<td>$16.10</td>
<td>$15.81</td>
<td>$11.25</td>
<td>$15.77</td>
</tr>
<tr>
<td>Market equity cap</td>
<td>$976</td>
<td>$205</td>
<td>$348</td>
<td>$480</td>
<td>$382</td>
<td>$690</td>
</tr>
<tr>
<td>$ Million</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5. Summary statistics, historical sample, 2001-2011 (only odd numbered years are shown)

<table>
<thead>
<tr>
<th></th>
<th>2001</th>
<th>2003</th>
<th>2005</th>
<th>2007</th>
<th>2009</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. firms</td>
<td>137</td>
<td>141</td>
<td>144</td>
<td>150</td>
<td>145</td>
<td>149</td>
</tr>
<tr>
<td>NYSE</td>
<td>106</td>
<td>51</td>
<td>48</td>
<td>55</td>
<td>56</td>
<td>47</td>
</tr>
<tr>
<td>Amex</td>
<td>16</td>
<td>10</td>
<td>8</td>
<td>14</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>15</td>
<td>80</td>
<td>88</td>
<td>81</td>
<td>84</td>
<td>96</td>
</tr>
<tr>
<td>Avg. daily trades</td>
<td>167</td>
<td>231</td>
<td>1,993</td>
<td>1,346</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. daily quotes</td>
<td>1,525</td>
<td>1,470</td>
<td>41,571</td>
<td>24,053</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Avg. daily NBB changes  | 128  | 210  | 611  | 772  | 1,787| 1,225| 23% CAGR
| Avg. daily NBO changes  | 127  | 226  | 729  | 789  | 1,789| 1,146| 32% CAGR
| Avg. price              | $20.57| $14.41| $16.10| $15.81| $11.25| $15.77|
| Market equity cap $ Million | $976 | $205 | $348 | $480 | $382 | $690 |
What statistics to consider?

- Long-term volatilities changed dramatically over the sample period.
- Variance ratios (normalized to long-term volatility) are the most reliable indicators of trends.
Table 6. Wavelet variance ratios for bids and offers, 2001-2011

Panel A: Computed from *unadjusted* bids and offers

<table>
<thead>
<tr>
<th>Time scale</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 ms</td>
<td>5.29</td>
<td>7.36</td>
<td>5.96</td>
<td>10.31</td>
<td>6.56</td>
<td>8.57</td>
<td>6.96</td>
<td>6.07</td>
<td>4.53</td>
<td>7.09</td>
<td>4.71</td>
</tr>
<tr>
<td>100 ms</td>
<td>5.52</td>
<td>6.75</td>
<td>5.20</td>
<td>9.71</td>
<td>6.38</td>
<td>8.07</td>
<td>6.27</td>
<td>5.39</td>
<td>4.12</td>
<td>6.27</td>
<td>4.33</td>
</tr>
<tr>
<td>200 ms</td>
<td>5.35</td>
<td>6.44</td>
<td>5.05</td>
<td>9.06</td>
<td>6.10</td>
<td>7.34</td>
<td>5.33</td>
<td>4.65</td>
<td>3.68</td>
<td>5.41</td>
<td>3.75</td>
</tr>
<tr>
<td>400 ms</td>
<td>4.65</td>
<td>5.35</td>
<td>4.92</td>
<td>8.18</td>
<td>5.64</td>
<td>6.30</td>
<td>4.25</td>
<td>3.84</td>
<td>3.21</td>
<td>4.54</td>
<td>3.07</td>
</tr>
<tr>
<td>800 ms</td>
<td>3.16</td>
<td>4.12</td>
<td>3.86</td>
<td>5.59</td>
<td>4.93</td>
<td>5.10</td>
<td>3.41</td>
<td>3.11</td>
<td>2.76</td>
<td>3.71</td>
<td>2.56</td>
</tr>
<tr>
<td>1,600 ms</td>
<td>2.13</td>
<td>2.56</td>
<td>3.19</td>
<td>4.11</td>
<td>4.06</td>
<td>4.05</td>
<td>2.89</td>
<td>2.59</td>
<td>2.42</td>
<td>3.04</td>
<td>2.23</td>
</tr>
<tr>
<td>3.2 sec</td>
<td>2.00</td>
<td>2.25</td>
<td>2.91</td>
<td>3.39</td>
<td>3.42</td>
<td>3.37</td>
<td>2.56</td>
<td>2.28</td>
<td>2.16</td>
<td>2.53</td>
<td>2.01</td>
</tr>
<tr>
<td>6.4 sec</td>
<td>1.95</td>
<td>2.12</td>
<td>2.61</td>
<td>2.91</td>
<td>2.88</td>
<td>2.92</td>
<td>2.35</td>
<td>2.08</td>
<td>1.94</td>
<td>2.16</td>
<td>1.82</td>
</tr>
</tbody>
</table>
Summary 2001-2011

- Quote volatility is surprisingly high in the early years.
- This reflects large temporary shifts in bids and offers (a consequence of manual markets).
- When the bid and offer series are filtered, volatility is lower in the early years.
- But over 2001-2011 no evidence of a broader trend.
Follow-up questions

- What strategies give rise to the episodic oscillations?
- Are the HFQ episodes unstable algos?
- Are they sensible strategies to detect and access liquidity?