High-Frequency Quoting: Measurement, Detection and Interpretation

Joel Hasbrouck

Outline

- Background
- Look at a data fragment
- Economic significance
- Statistical modeling
- Application to larger sample
- Open questions

Economics of high-frequency trading

- Absolute speed
  - In principle, faster trading leads to smaller portfolio adjustment cost and better hedging
  - For most traders, latencies are inconsequential relative to the speeds of macroeconomic processes and intensities of fundamental information.
- Relative speed (compared to other traders)
  - A first mover advantage is extremely valuable.
  - Low latency technology has increasing returns to scale and scope.
  - This gives rise to large firms that specialize in high-frequency trading.

Welfare: “HFT imposes costs on other players”

- They increase adverse selection costs.
- The information produced by HFT technology is simply advance knowledge of other players’ order flows.
    - Jarrow, Robert A., and Philip Protter, 2011.
    - Biais, Bruno, Thierry Foucault, and Sophie Moinas, 2012

Welfare: "HFT improves market quality."

- Supported by most empirical studies that correlate HF measures/proxies with standard liquidity measures.
    - Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld, 2010
    - Hasbrouck, Joel, and Gideon Saar, 2011
    - Hendershott, Terrence J., and Ryan Riordan, 2012

"HFTs are efficient market-makers"

- Empirical studies
  - Kirilenko, Andrei A., Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2010
  - Menkveld, Albert J., 2012
  - Brogaard, Jonathan, 2010a, 2010b, 2012
- Strategy: identify a class of HFTs and analyze their trades.
- HFTs closely monitor and manage their positions.
- HFTs often trade passively (supply liquidity via bid and offer quotes)
- But …
  - HFTs don't maintain a continuous market presence.
  - They sometimes trade actively ("aggressively")

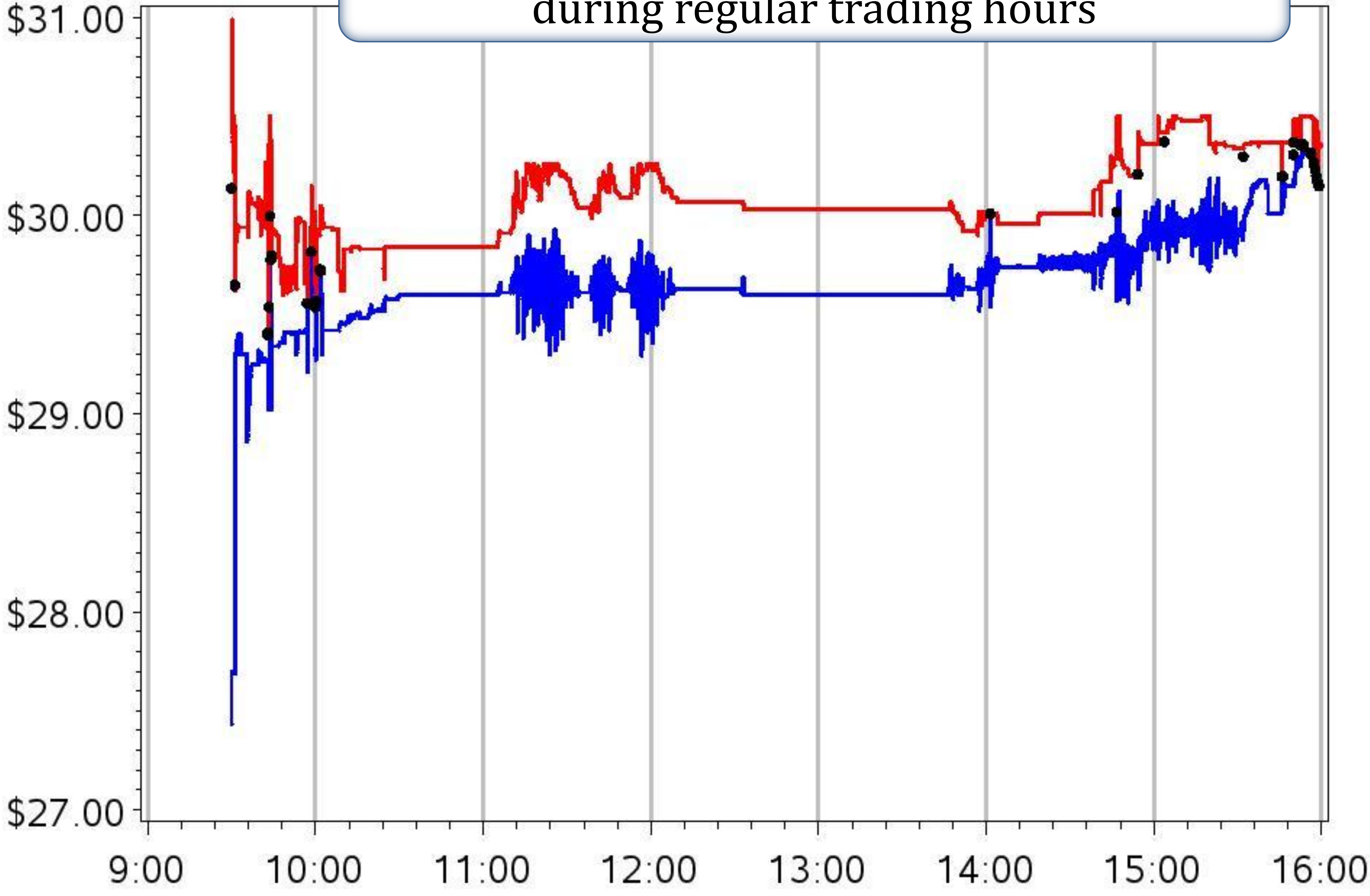
Positioning

- We use the term "high frequency trading" to refer to all sorts of rapid-paced market activity.
- Most empirical analysis focuses on trades.
- This study emphasizes quotes.

High-frequency quoting

- Rapid oscillations of bid and/or ask quotes.
- Example
  - AEPI is a small Nasdaq-listed manufacturing firm.
  - Market activity on April 29, 2011
  - National Best Bid and Offer (NBBO)
    - The highest bid and lowest offer (over all market centers)

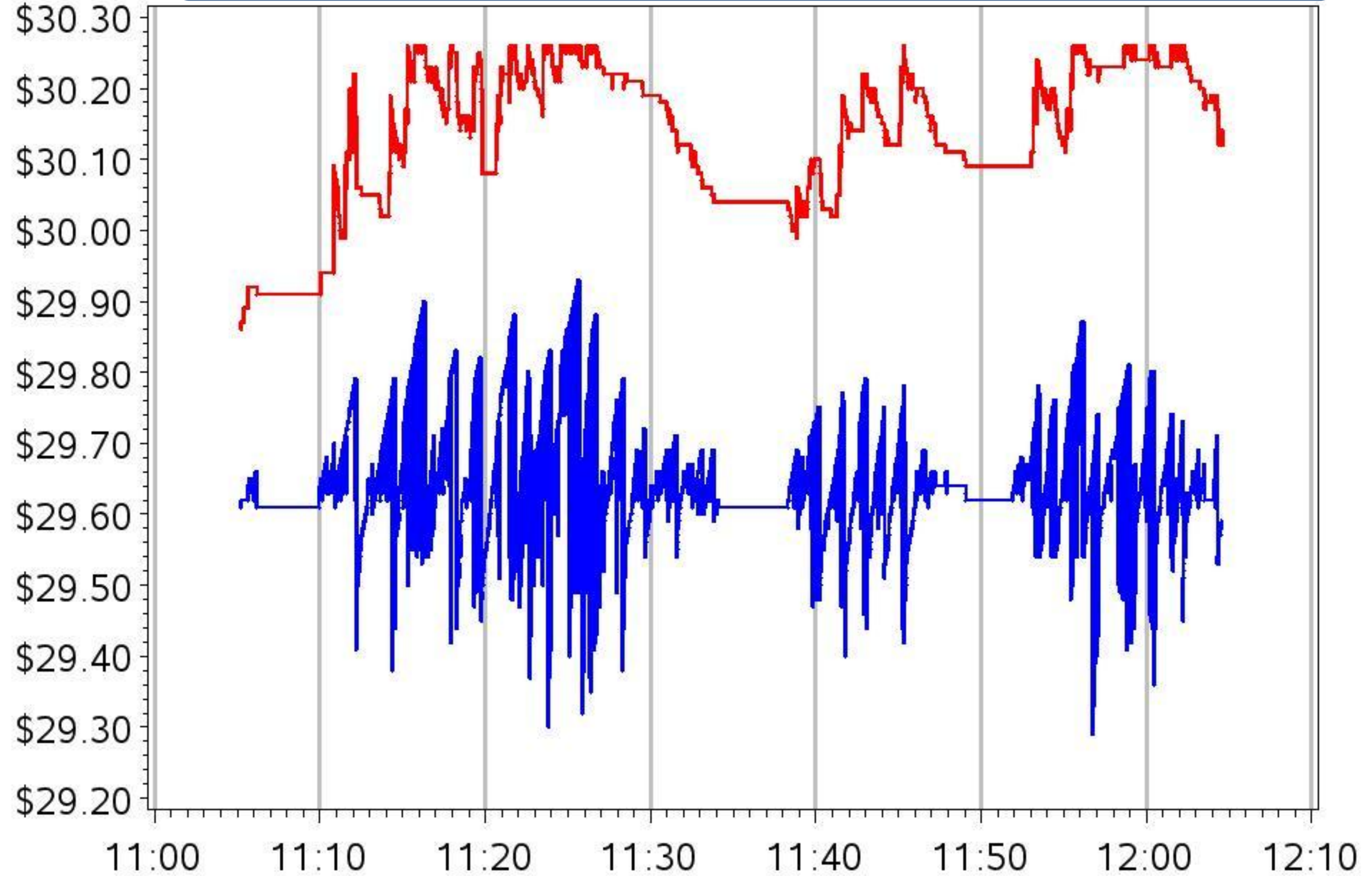
National Best Bid and Offer for AEPI during regular trading hours



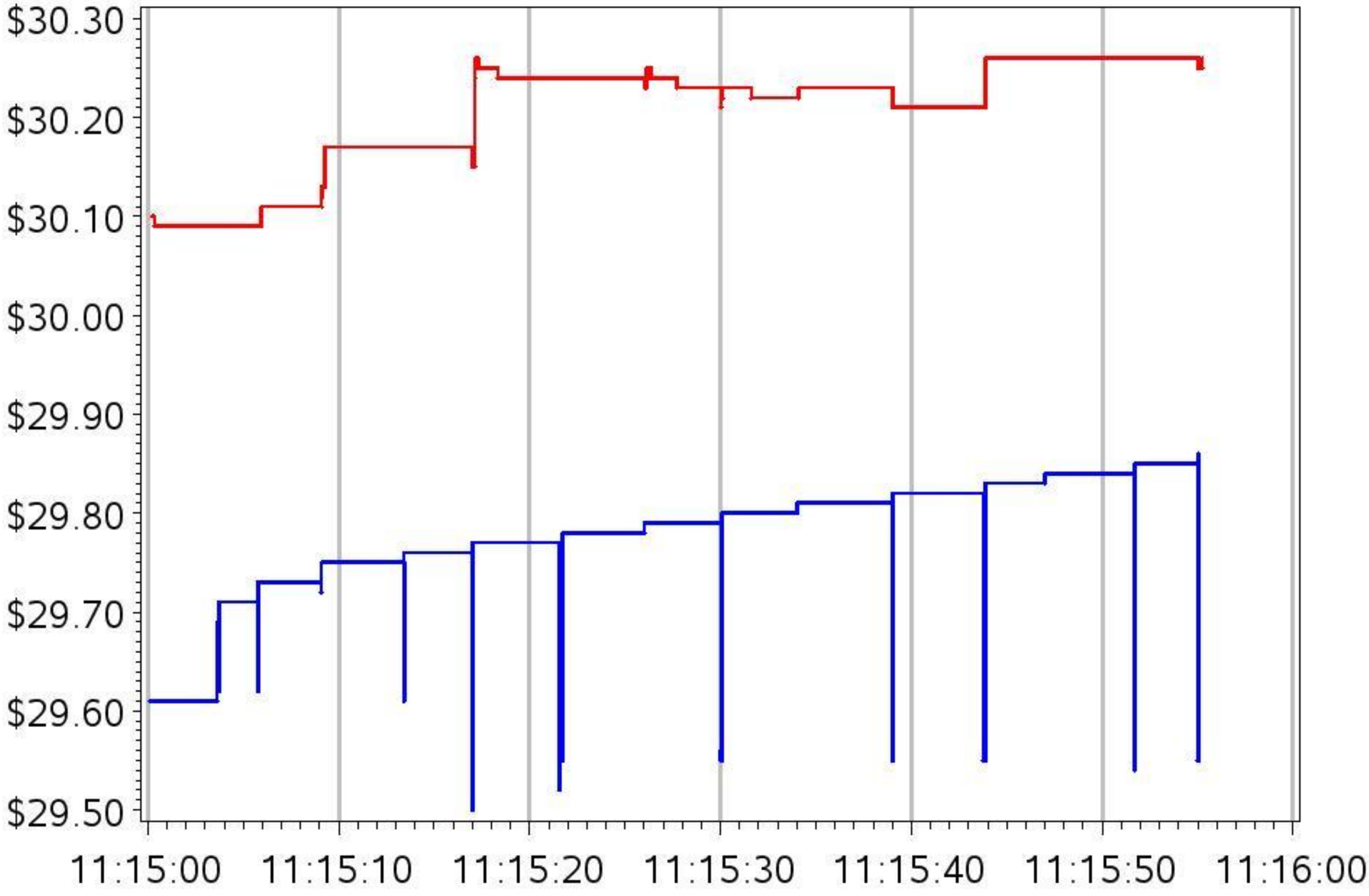
Caveats

- Ye & O'Hara (2011)
  - A bid or offer is not incorporated into the NBBO unless it is 100 sh or larger.
  - Trades are not reported if they are smaller than 100 sh.
- Due to random latencies, agents may perceive NBBO's that differ from the "official" one.
- Now zoom in on one hour …

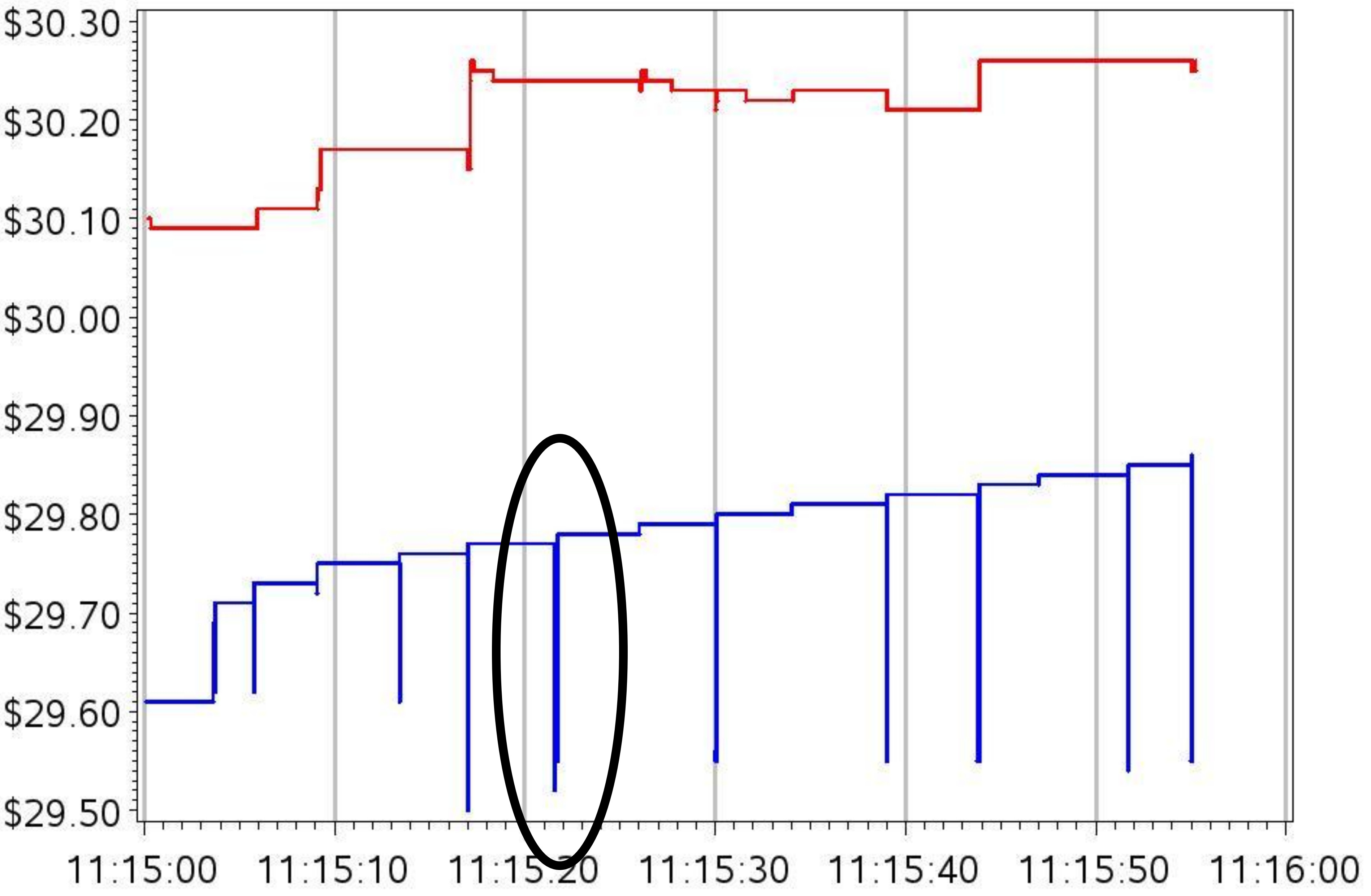
National Best Bid and Offer for AEPI from 11:00 to 12:10

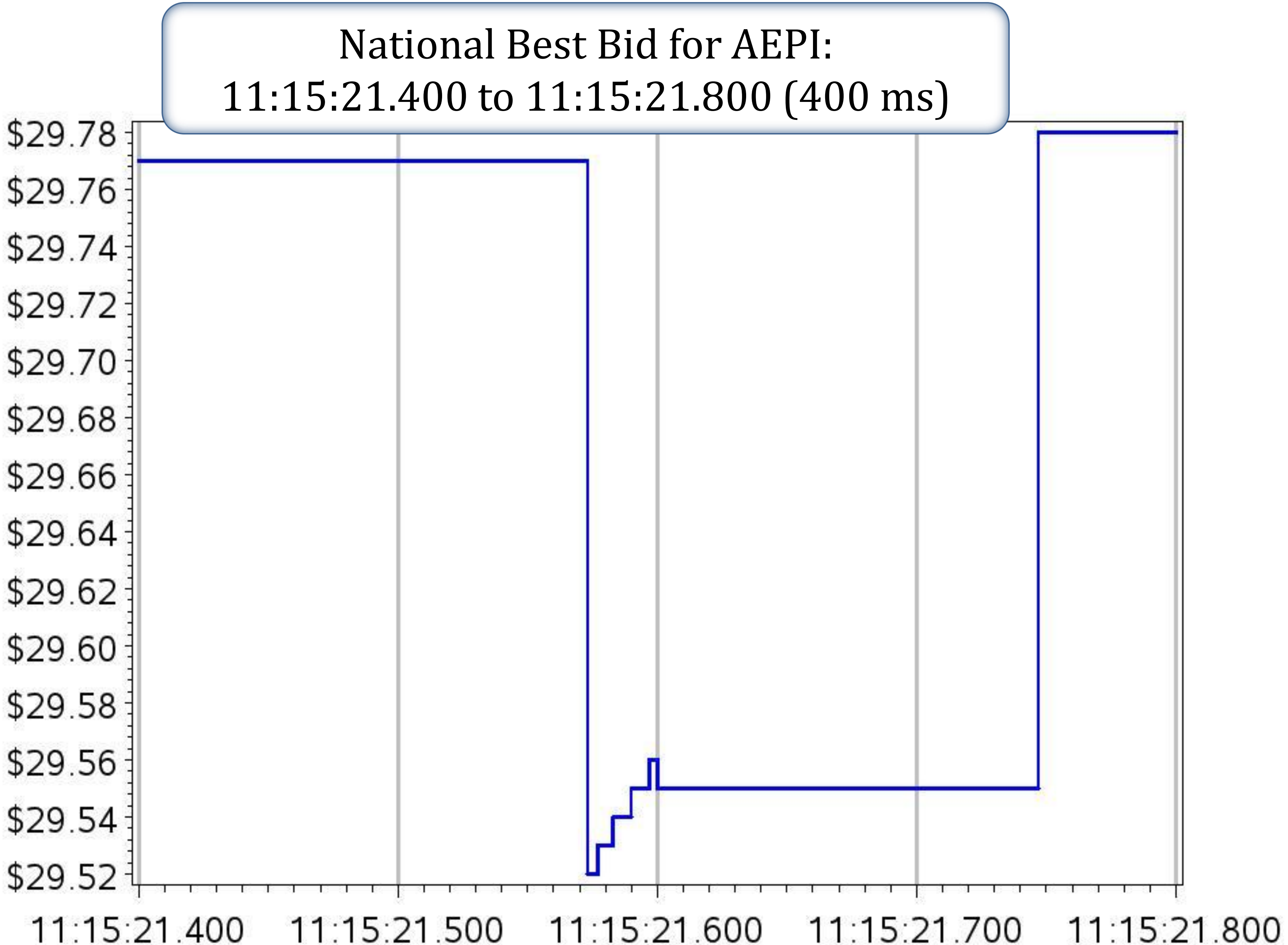


National Best Bid and Offer for AEPI from 11:15:00 to 11:16:00



National Best Bid and Offer for AEPI from 11:15:00 to 11:16:00




National Best Bid for AEPI:
11:15:21.400 to 11:15:21.800 (400 ms)
$29.78
$29.76
$29.74
$29.72
$29.70
$29.68
$29.66
$29.64
$29.62
$29.60
$29.58
$29.56
$29.54
$29.52
11:15:21.400
11:15:21.500
11:15:21.600
11:15:21.700
11:15:21.800

So what?

- HFQ noise degrades the informational value of the bid and ask.
- HFQ aggravates execution price uncertainty for marketable orders.
- And in US equity markets …
  - NBBO used as reference prices for dark trades.
  - Top (and only the top) of a market's book is protected against trade-throughs.

“Dark” Trades

- Trades that don’t execute against a visible quote.
- In many trades, price is assigned by reference to the NBBO.
  - Preferenced orders are sent to wholesalers.
    - Buys filled at NBO; sells at NBB.
  - Crossing networks match buyers and sellers at the midpoint of the NBBO.

Features of the AEPI episodes

- Extremely rapid *oscillations* in the bid.
- Start and stop abruptly
- Possibly unconnected with fundamental news.
- Directional (activity on the ask side is much smaller)

A framework for analysis: the requirements

- Need precise resolution (the data have one ms. time-stamps)
- Low-order vector autoregression?
- Oscillations: spectral (frequency) analysis?
  - Represent a time series as a combination sine/cosine functions.
    - But the functions are recurrent over the full sample.
- AEPI episodes are localized.

Stationarity

- The oscillations are locally stationary.
- Framework must pick up stationary local variation …
  - But not exclude random-walk components.
- Should identify long-run components as well as short-run.

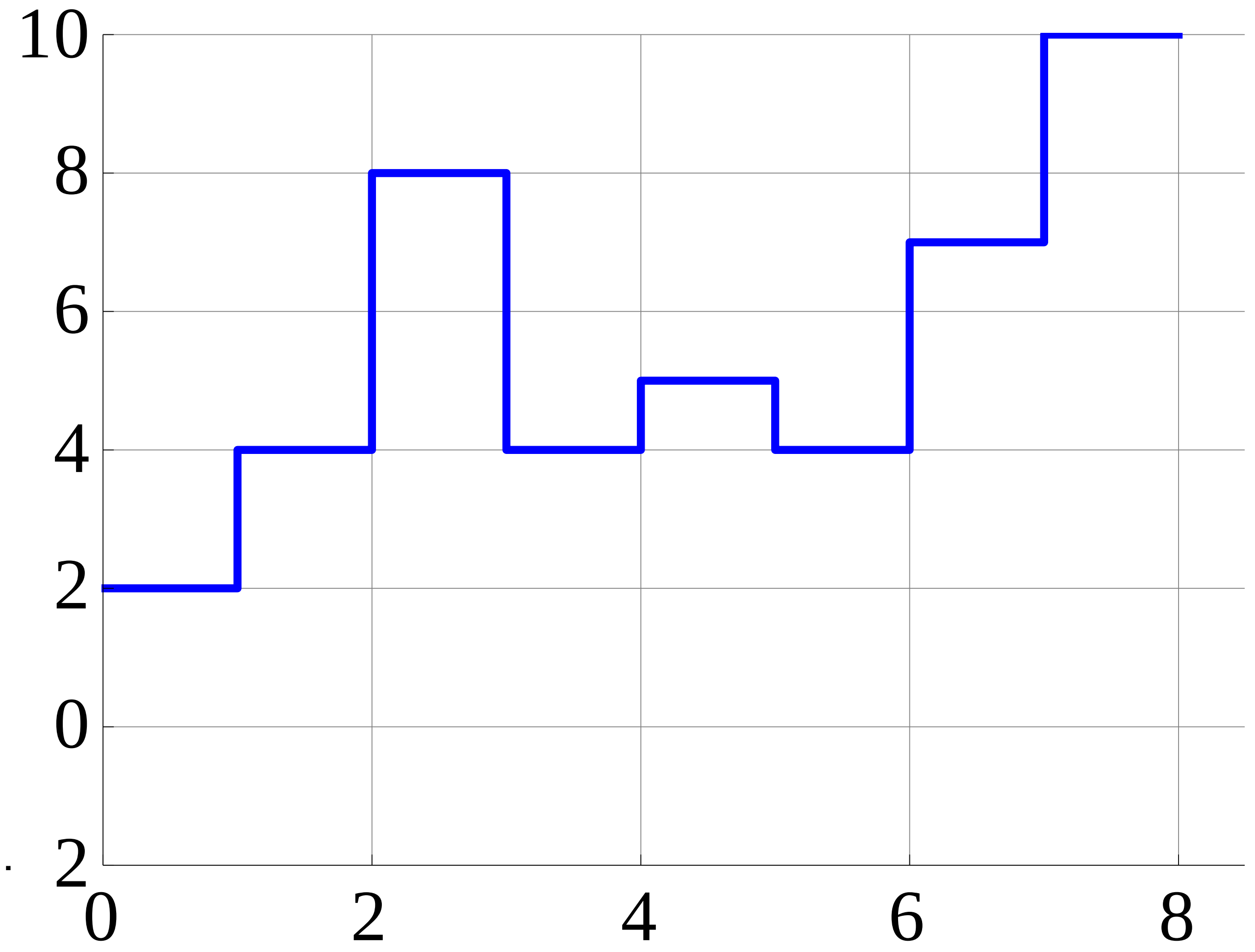
Intuitively, I'd like to …

- Use a moving average to smooth series.
  - Implicitly estimating the long-term component.
- Isolate the HF component as a residual.

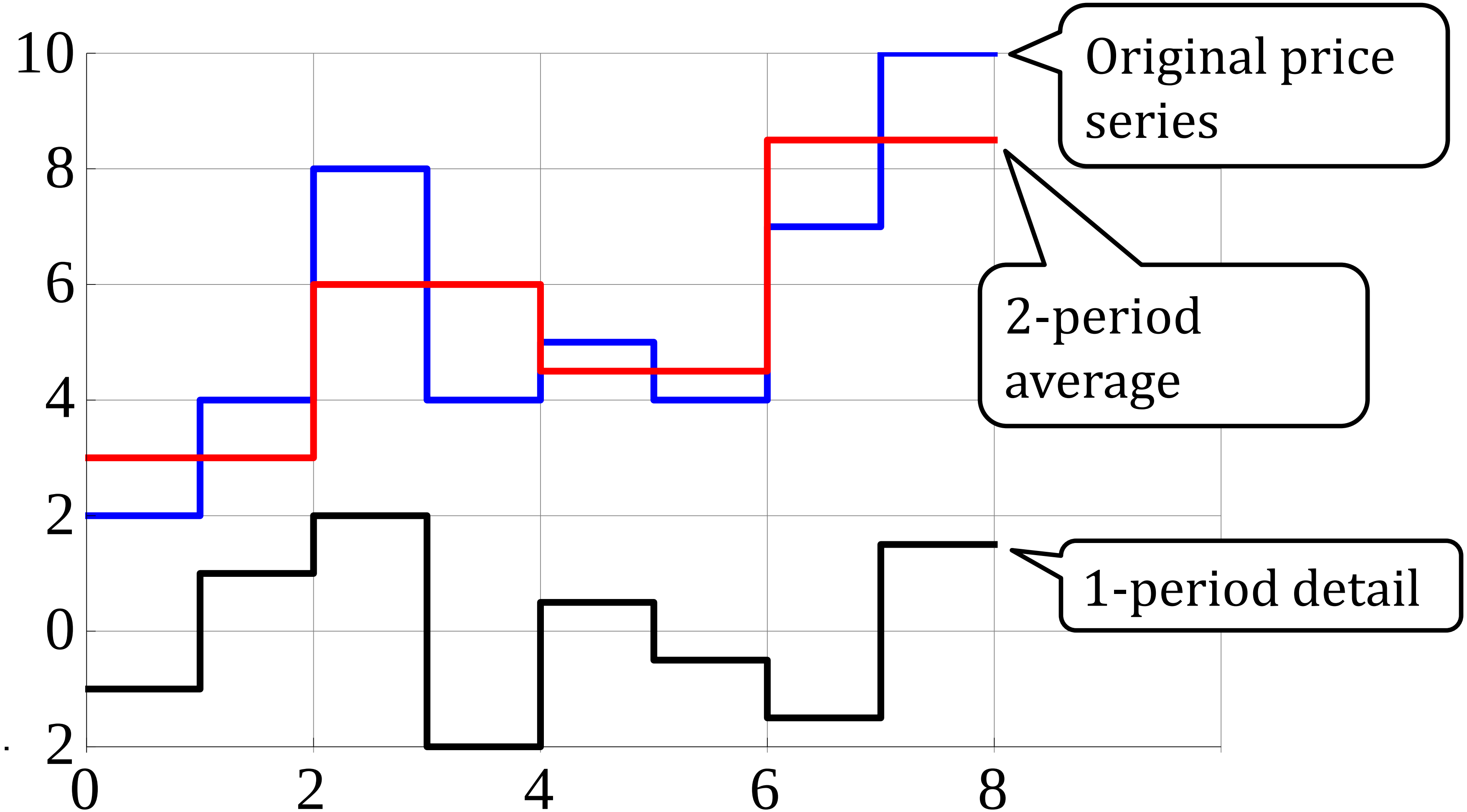
Alternative: Time-scale decomposition

- Represent a time-series in terms of basis functions (wavelets)
- Wavelets:
  - Localized
  - Oscillatory
  - Use flexible (systematically varying) time-scales.
- Accepted analytical tool in diverse disciplines.
  - Percival and Walden; Gencay et. al.

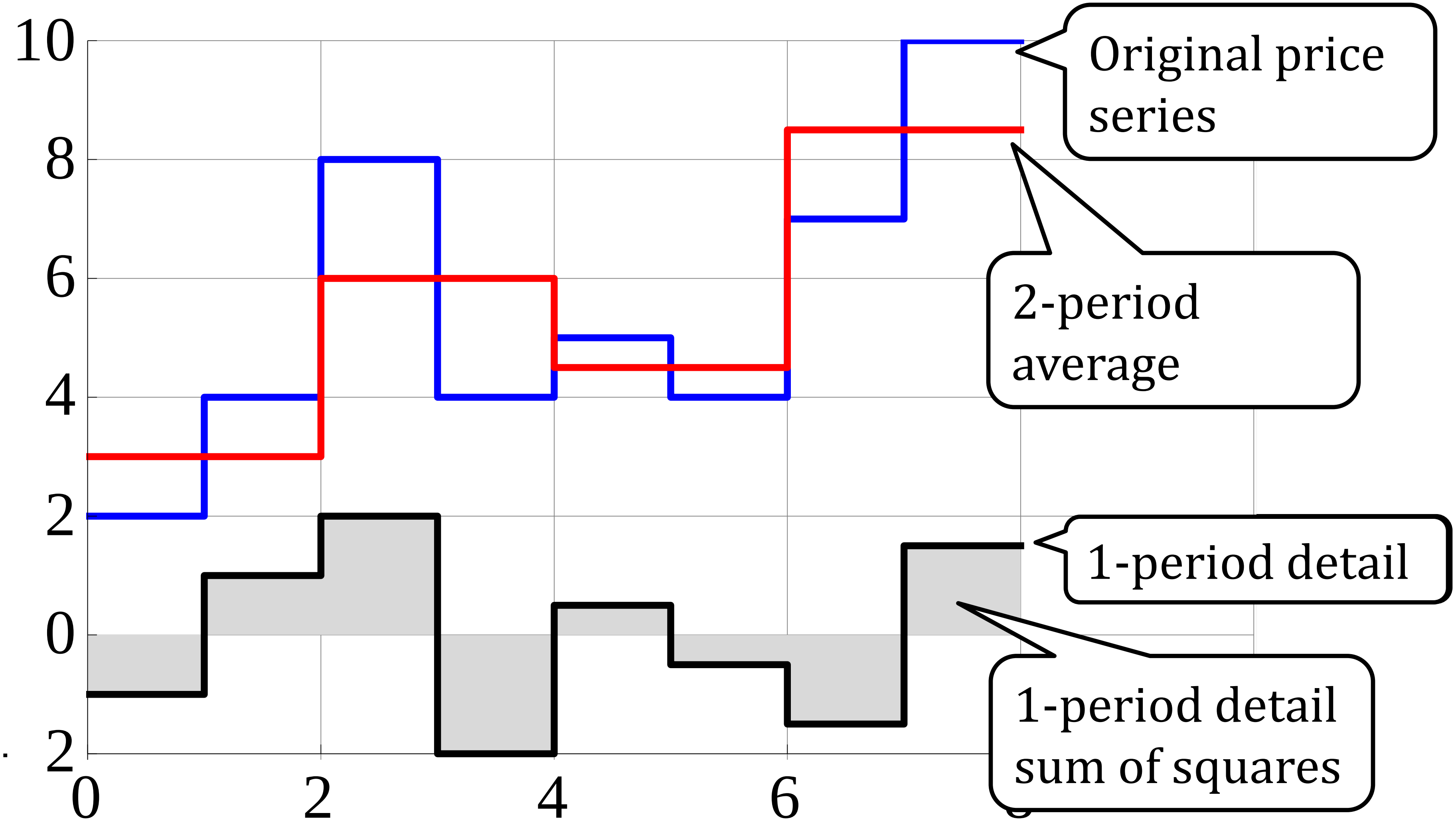
Sample bid path



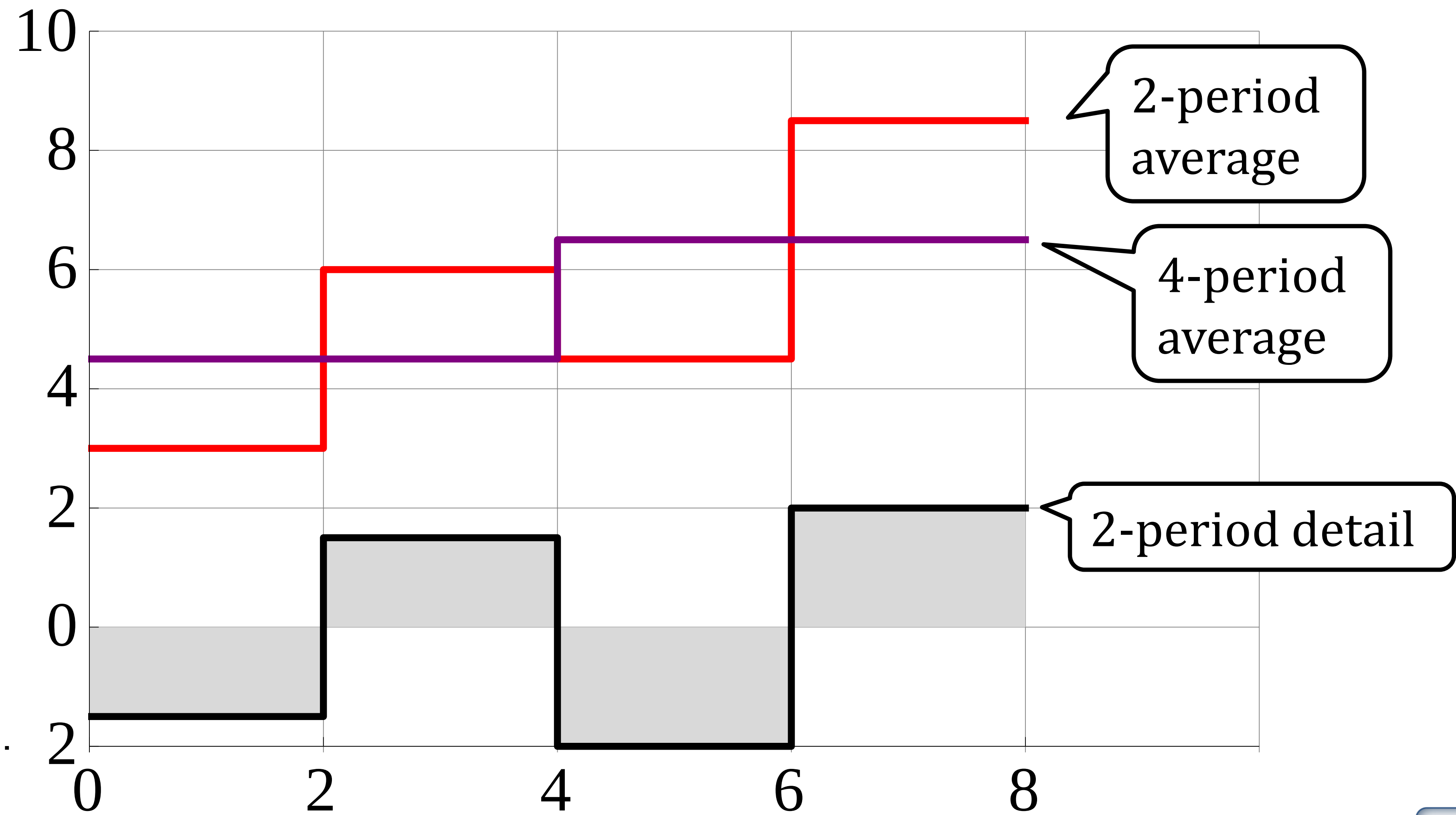
First pass (level) transform



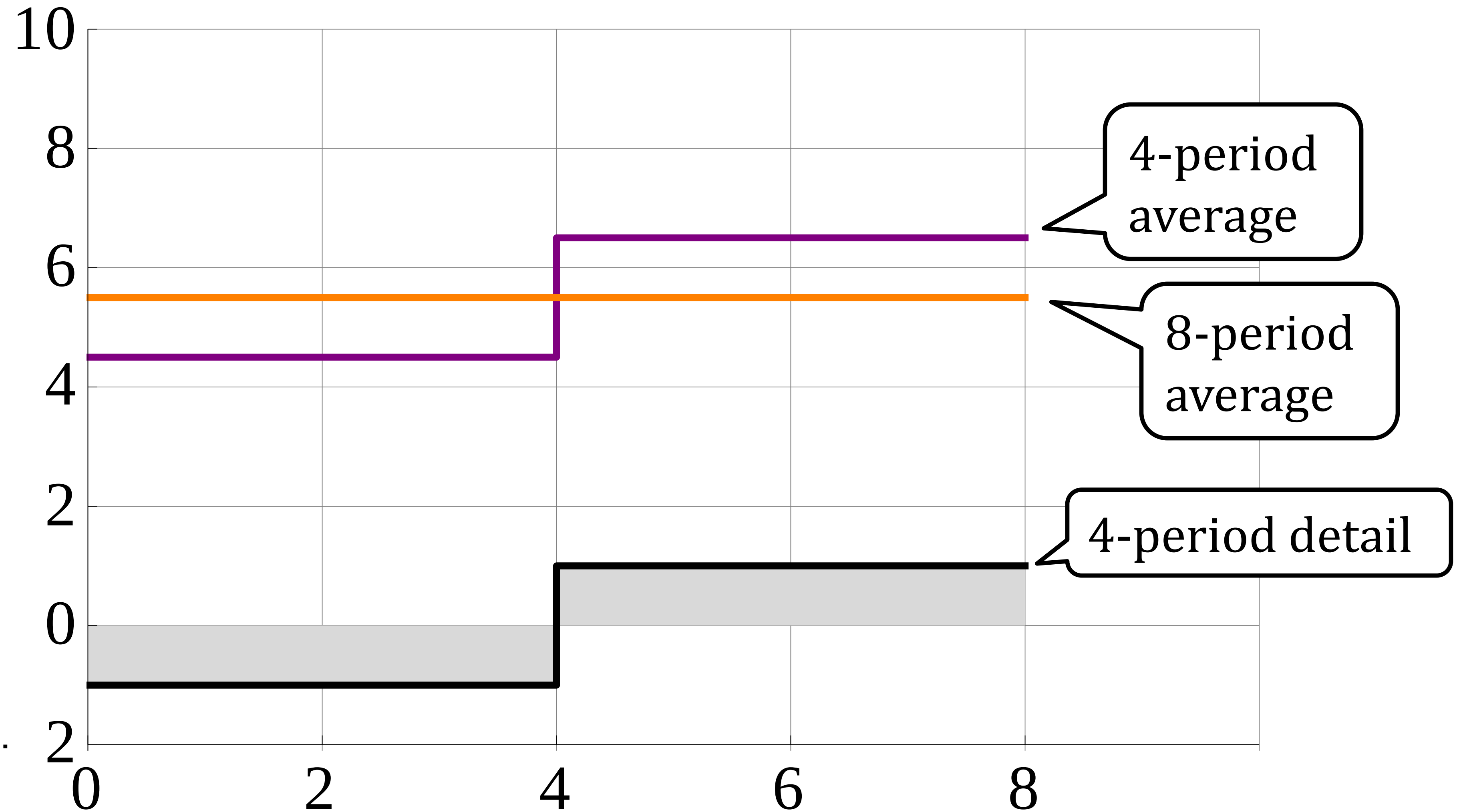
First pass (level) transform



Second pass (level) transform



Third level (pass) transform



For each level $j = 1,2, \ldots$, we have ...

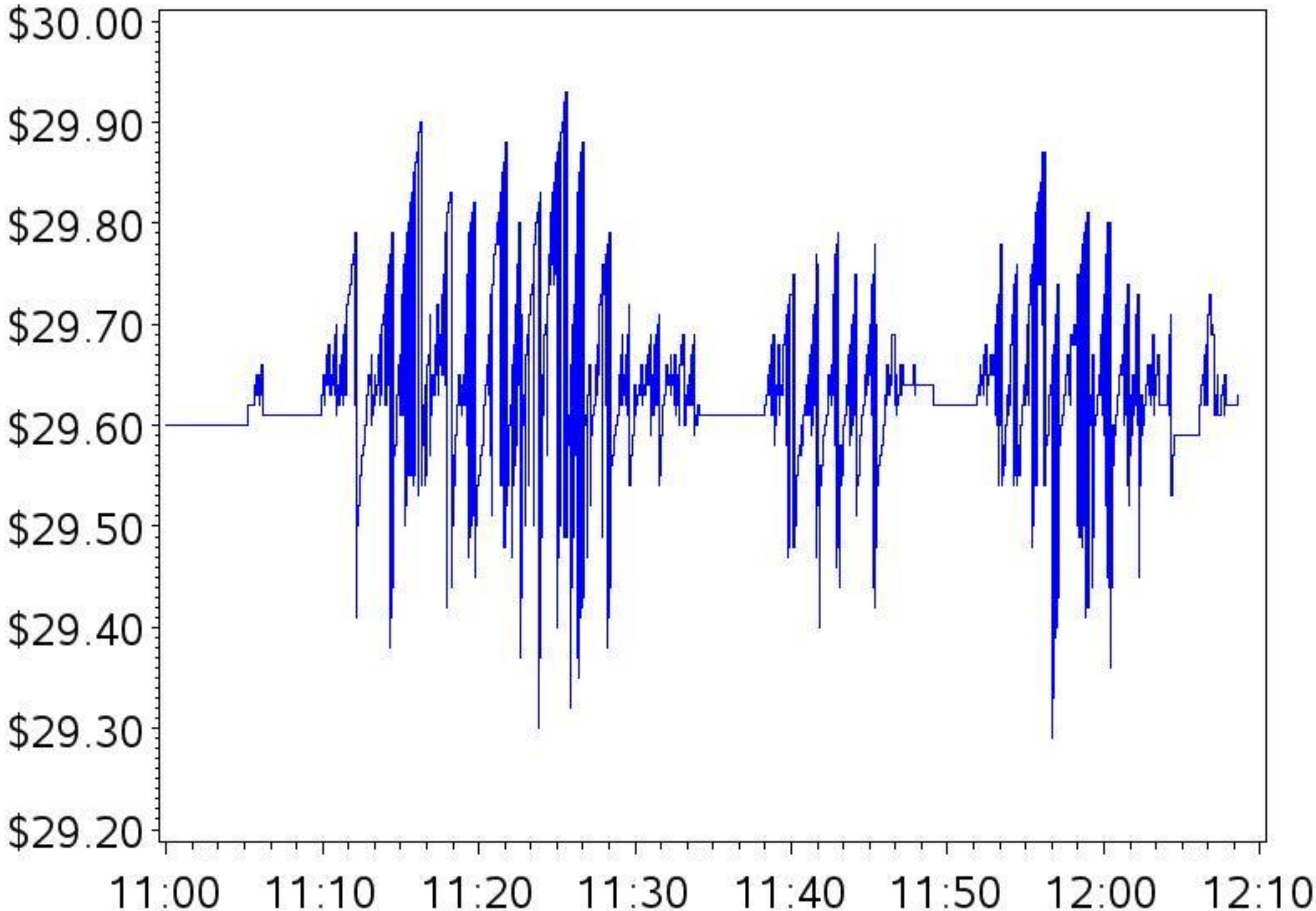
- A time scale, $\tau_j = 2^{j-1}$
  - Higher level → longer time scale.
  - $\tau_j \in \{1, 2, 4, \ldots\}$
  - "the persistence of the level-*j* component"
- A scale-τ_j "detail" component.
  - Centered ("zero mean") series that tracks changes in the series at scale τ_j.
- A scale-τ_j sum of squares.

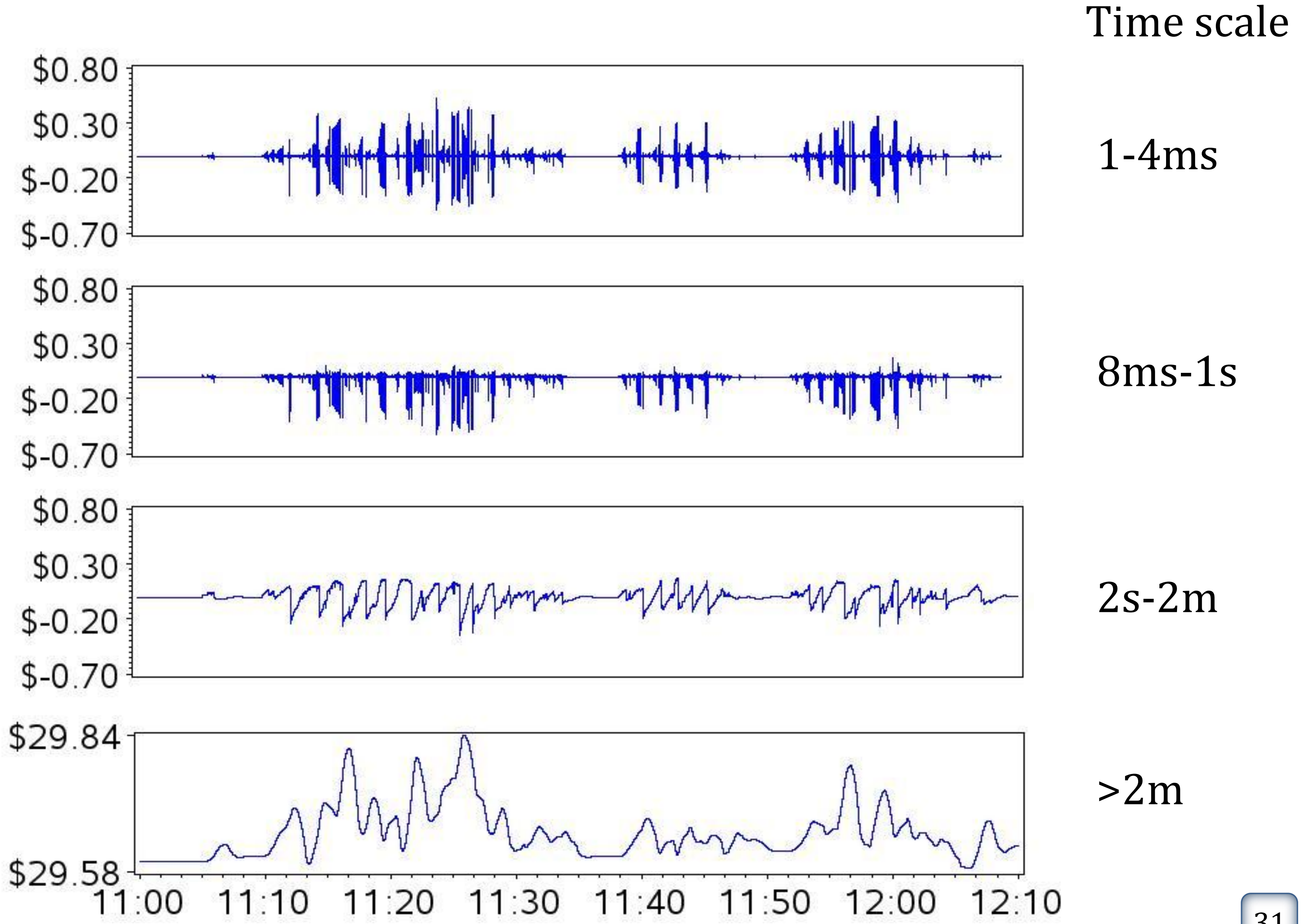
Interpretation

- The full set of scale-τ_j components decomposes the original series into sequences ranging from "very choppy" to "very smooth".
  - Multi-resolution analysis.
- With additional structure, the full set of scale-τ_j sums of squares corresponds to a variance decomposition.

Multi-resolution analysis of AEPI bid

- Data time-stamped to the millisecond.
- Construct decomposition through level $J = 18$.
- For graphic clarity, aggregate the components into four groups.
- Plots focus on 11am-12pm.


$30.00
$29.90
$29.80
$29.70
$29.60
$29.50
$29.40
$29.30
$29.20
11:00
11:10
11:20
11:30
11:40
11:50
12:00
12:10



Time scale
$0.80
$0.30
$-0.20
$-0.70
1-4ms
8ms-1s
2s-2m
$29.84
$29.58
>2m
11:00
11:10
11:20
11:30
11:40
11:50
12:00
12:10

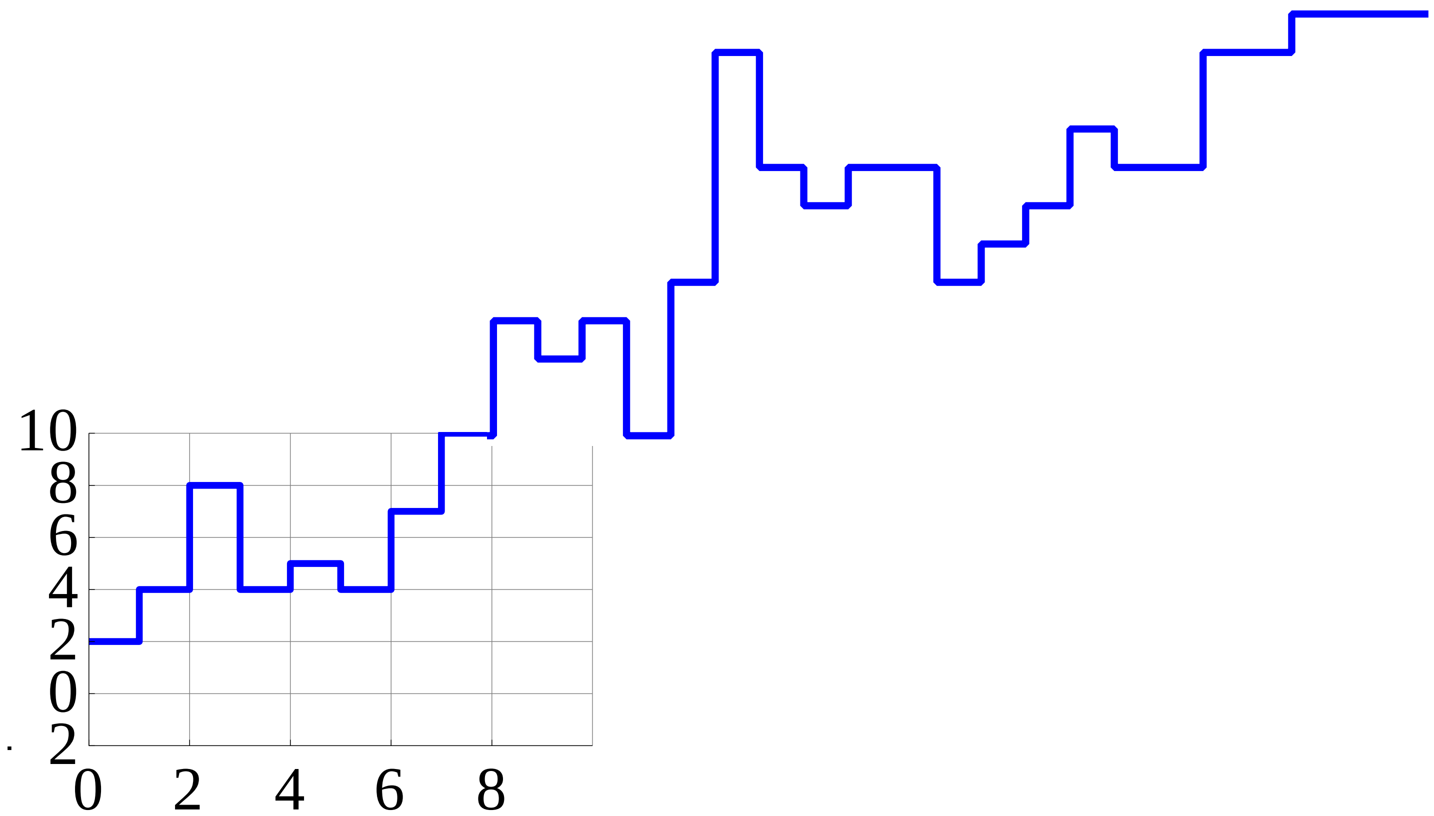
Connection to standard time series analysis

- Suppose p_t is a stochastic process
  - e.g., a random-walk
- The scale-τ_j sum-of-squares over the sample path (divided by n) defines an *estimate* of the *wavelet variance*.
- Wavelet variance (and its estimate) are well-defined and well-behaved assuming that the first differences of p_t are covariance stationary.
- Wavelet decompositions are performed on the *levels* of p_t not the first differences.

The wavelet variance of a random-walk

- $\nu^2(\tau_j) \equiv$ wavelet variance at scale τ_j
- For a random-walk
  - $p_t = p_{t-1} + e_t$
    where $Ee_t = 0$ and $Ee_t^2 = \sigma_e^2$
  - $\nu^2(\tau_j) = \phi(\tau_j)\sigma_e^2$
    where scaling factor $\phi(\tau_j) = \frac{1}{6}\left(\tau_j + \frac{1}{2\tau_j}\right)$
- $\phi(\tau_j) \in \{0.25, 0.38, 0.69, 1.3, 2.7, 5.3, 10.7, \ ...\}$

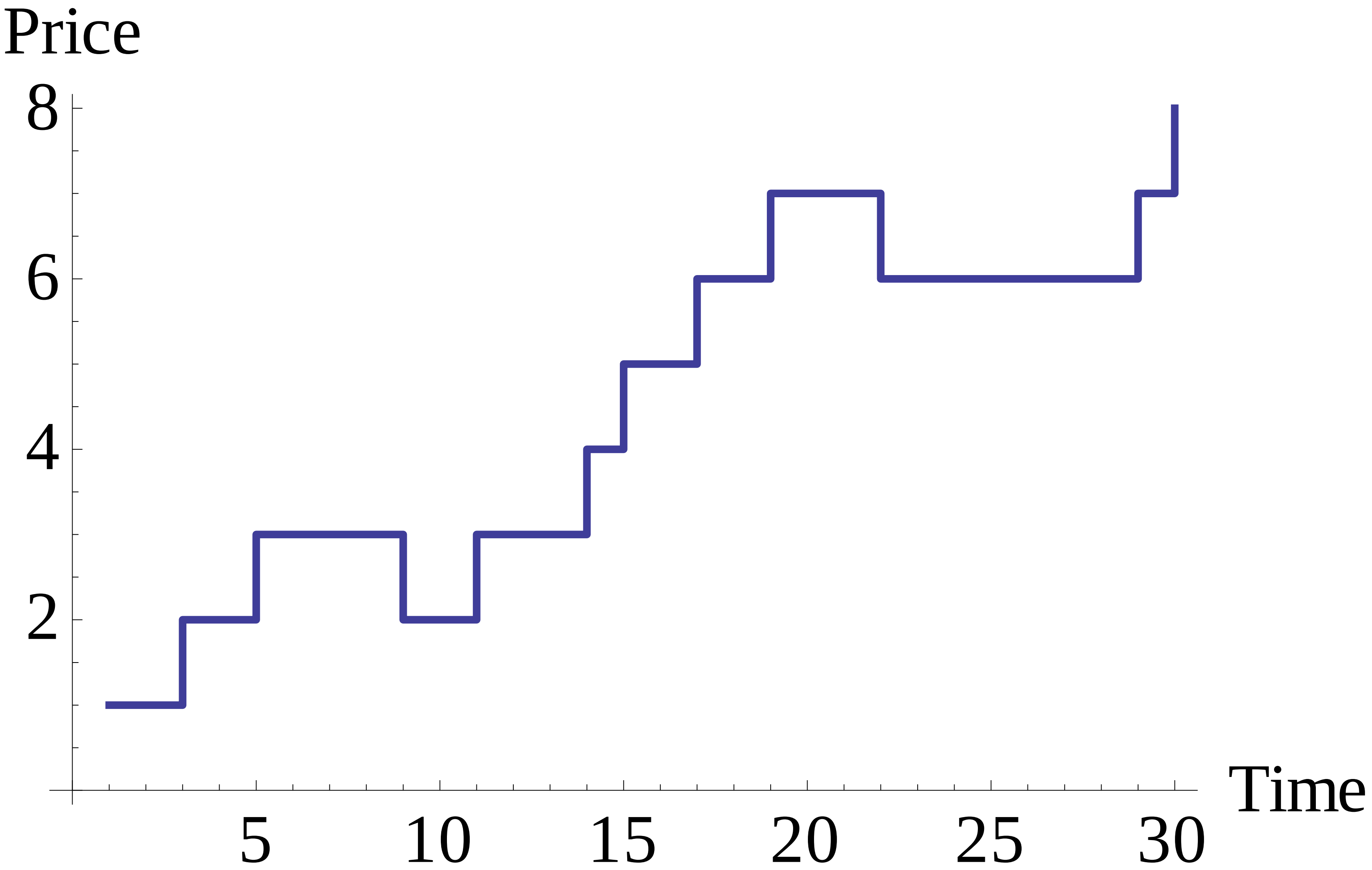
$$\lim_{j \to \infty} \nu^2(\tau_j) = \infty$$



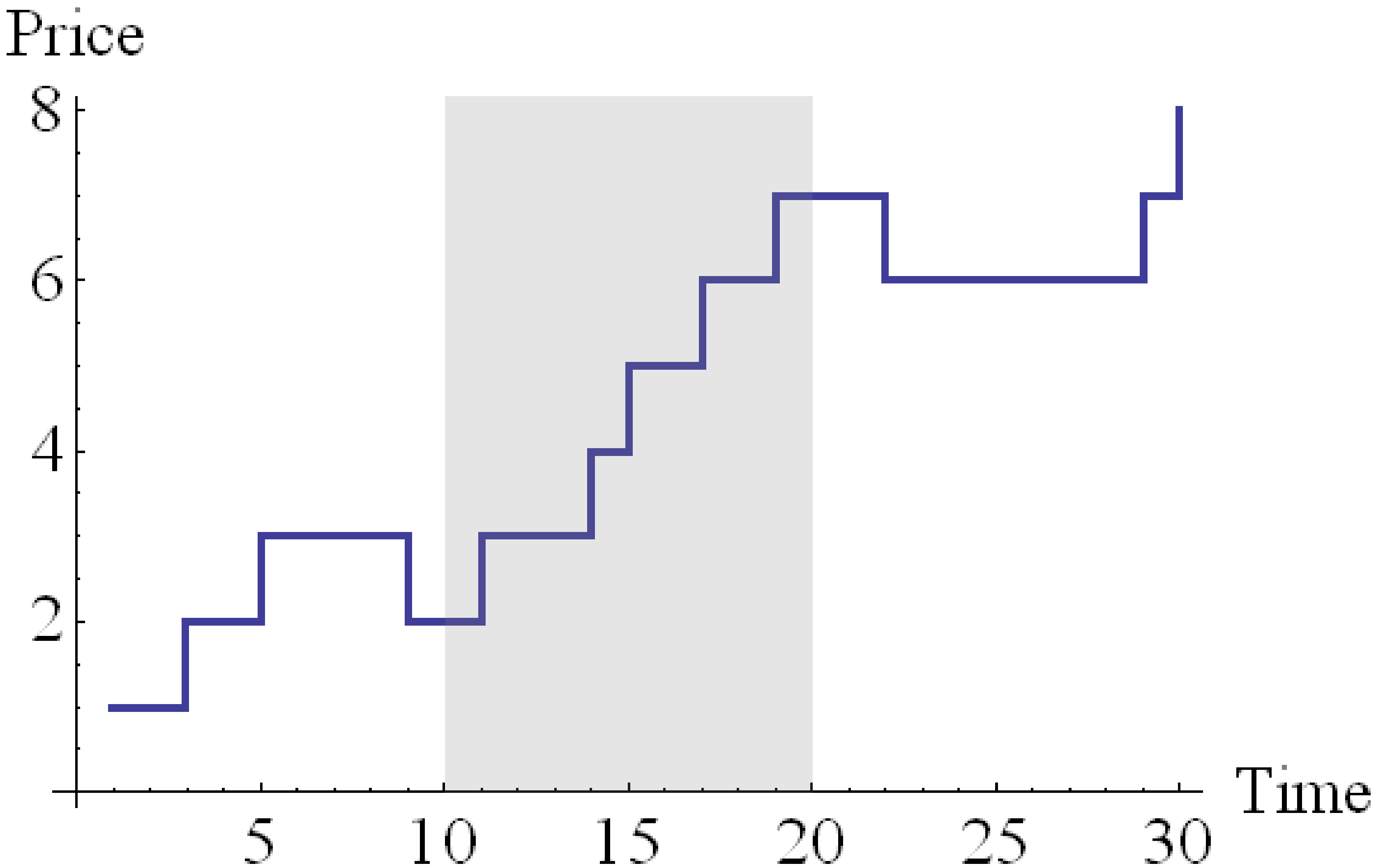
The wavelet variance for the AEPI bid: an economic interpretation

- Orders sent to market are subject to random delays.
  - This leads to arrival uncertainty.
  - For a market order, this corresponds to price risk.
- For a given time window, the wavelet variance measures this risk.

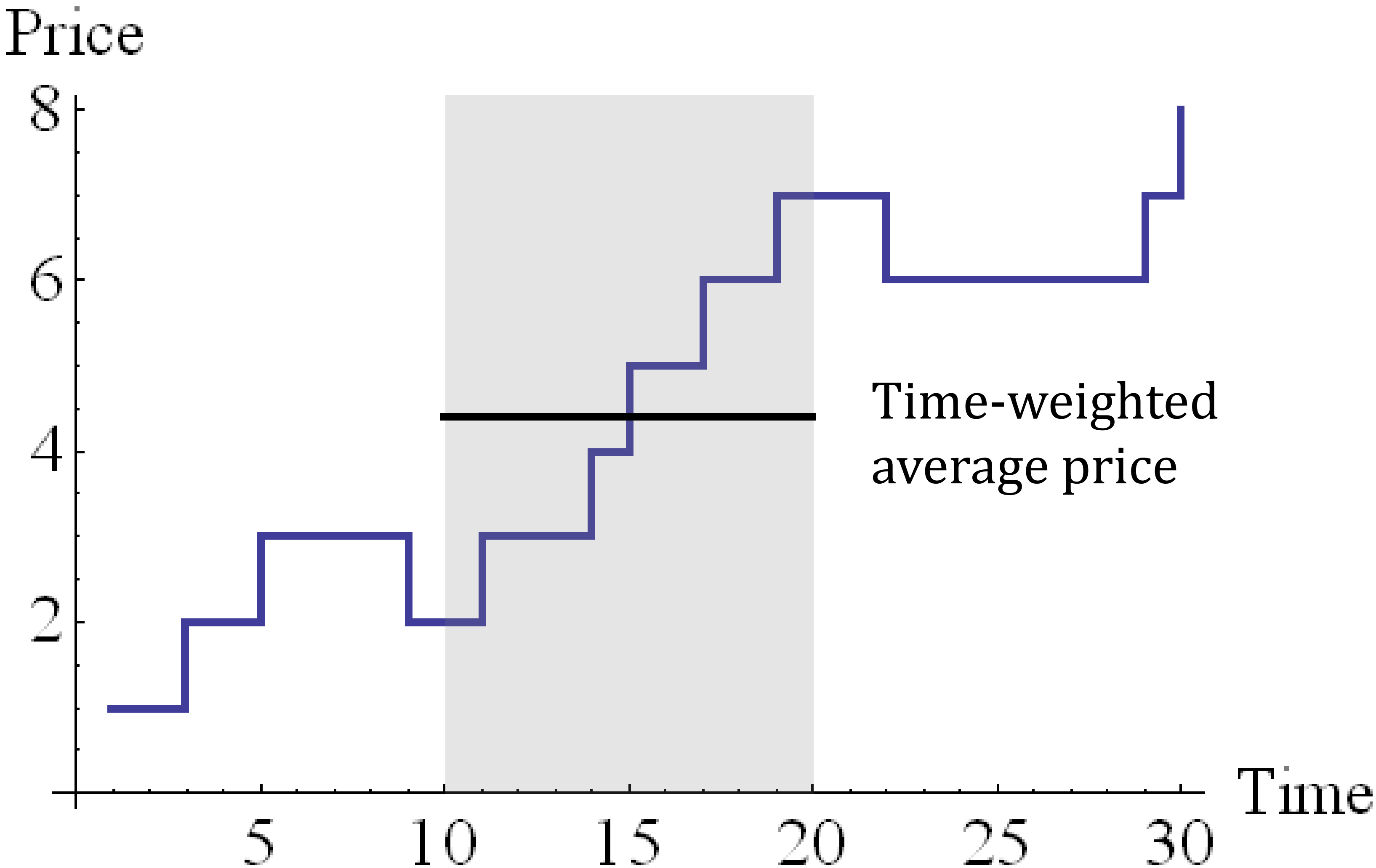
Timing a trade: the price path



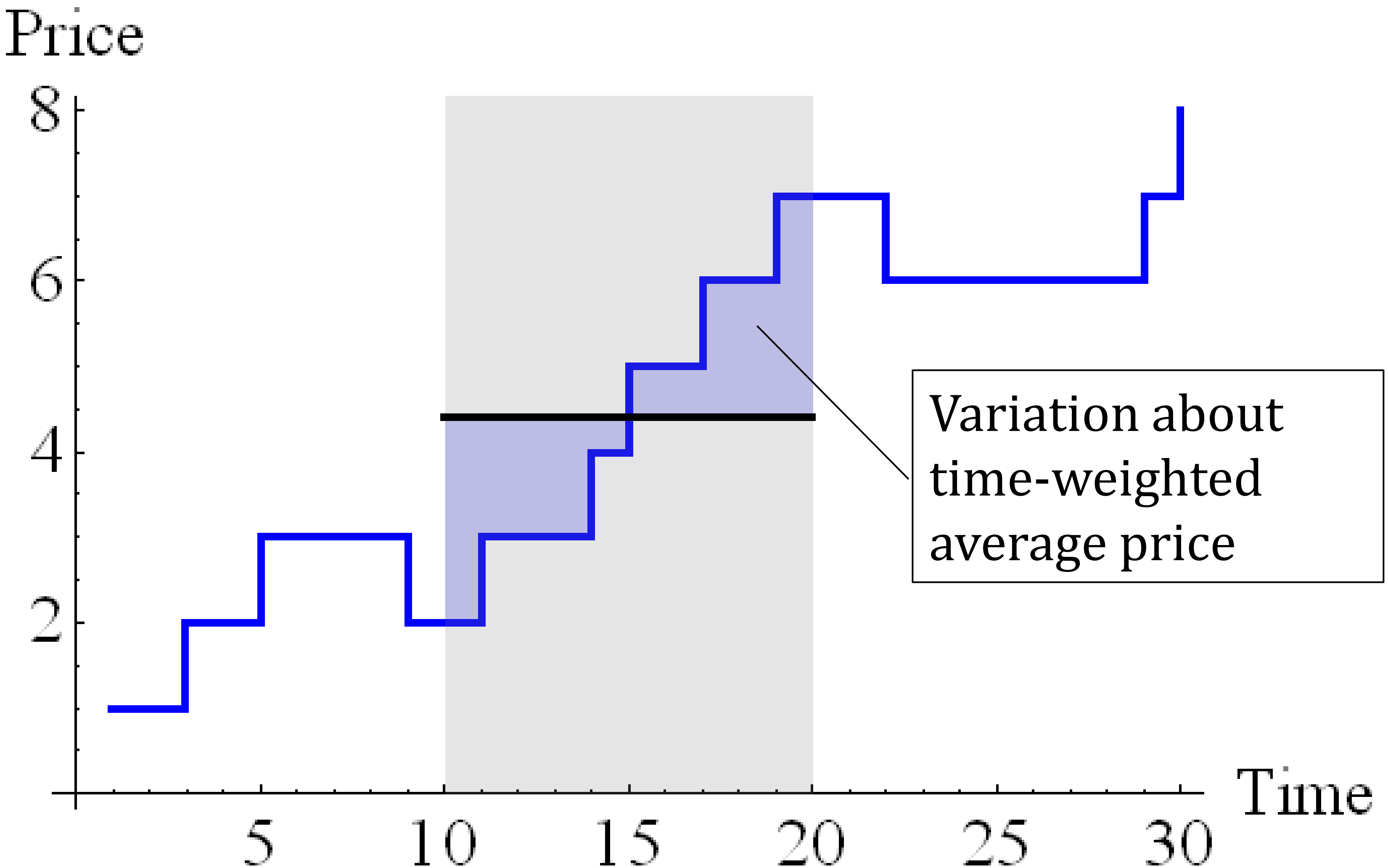
Timing a trade: the arrival window



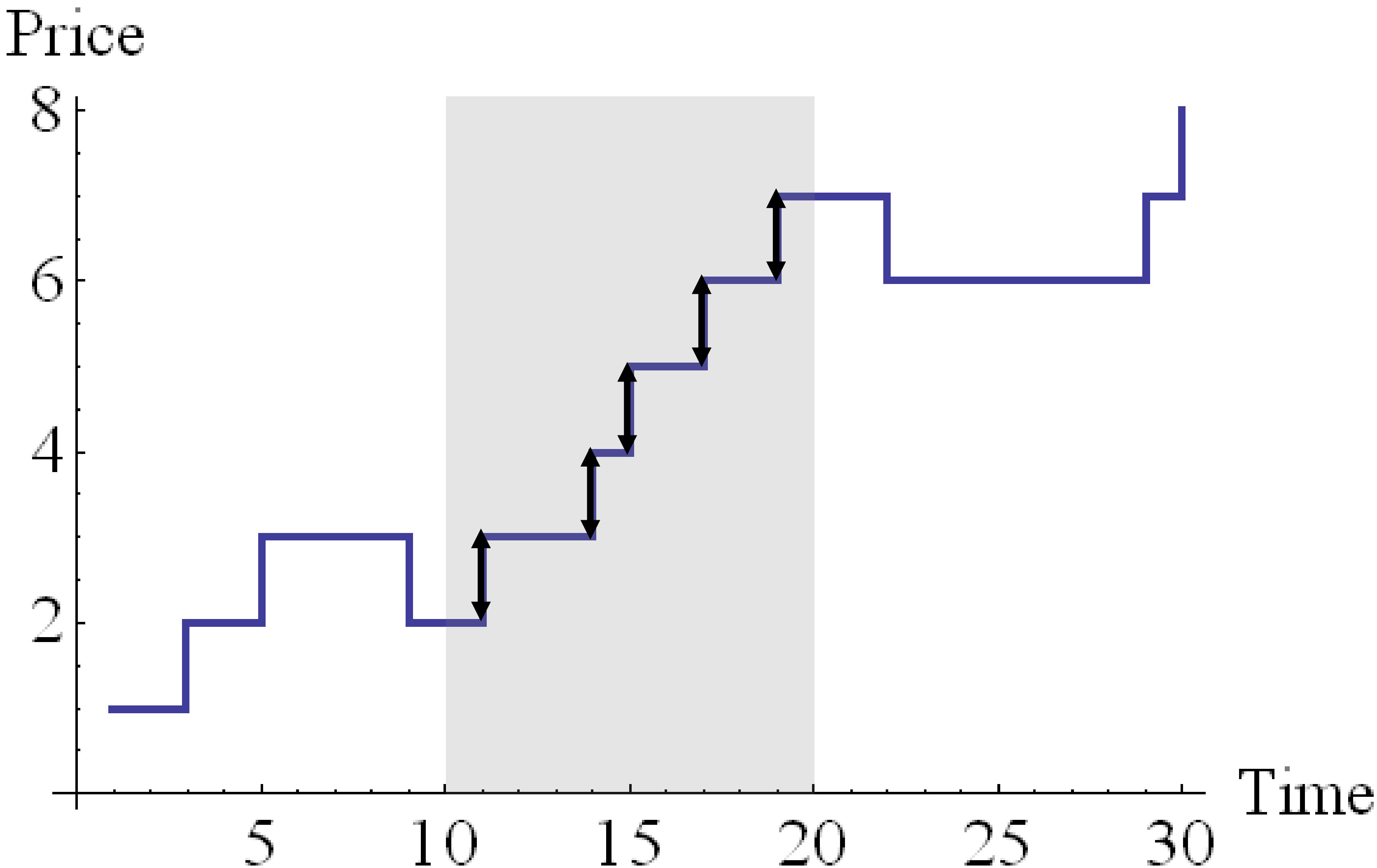
The time-weighted average price (TWAP) benchmark



Timing a trade: TWAP Risk



The wavelet variance: a comparison with realized volatility



Data sample

- 100 US firms from April 2011
- Sample stratified by equity market capitalization
- Alphabetic sorting
- Within each market cap decile, use first ten firms.
- Summary data from CRSP
- HF data from daily (“millisecond”) TAQ

	Median				
	Share price, EOY 2010	Mkt cap, EOY, 2010, $Million	Avg daily dollar vol, 2010, $Thousand	Avg daily no. of trades, April 2011	Avg daily no. of quotes, April 2011
Full sample	$13.75	420	2,140	1,111	23,347
Dollar Volume Deciles					
0 (low)	$4.18	30	20	17	846
1	$3.56	30	83	45	2,275
2	$3.70	72	228	154	5,309
3	$6.43	236	771	1,405	15,093
4	$7.79	299	1,534	468	15,433
5	$17.34	689	3,077	1,233	34,924
6	$26.34	1,339	5,601	2,045	37,549
7	$28.40	1,863	13,236	3,219	52,230
8	$36.73	3,462	34,119	7,243	94,842
9 (high)	$44.58	18,352	234,483	25,847	368,579

Computational procedures

- $10\ hrs \times 60min \times 60sec \times 1{,}000ms = 3.6 \times 10^7$ "observations" (per series, per day)
- Analyze data in rolling windows of ten minutes
- Supplement millisecond-resolution analysis with time-scale decomposition of prices averaged over one-second.
- Use maximal overlap discrete transforms with Daubechies(4) weights.

Example: AAPL (Apple Computer)

- 20 days
- Regular trading hours are 9:30 to 16:00.
  - I restrict to 9:45 to 15:45
  - $20\ days\ \times 6\ hrs \times 60\ min = 7{,}200\ min$
- Compute $\hat{v}^2(\tau_j)$ for $j = 1, \ldots, 18$
- Time scales: 1ms to 131,072ms (about 2.2 minutes)
- Tables report values for odd j (for brevity)

$\sqrt{Wavelet\ variances}$ for AAPL (units: \$0.01/sh)

	$\sqrt{\hat{\nu}^2(\tau_j)}$				$\rho_{Bid,Ask}(\tau_j)$
Time scale	*Median*		*99th percentile*		*Median*
τ_j	*Bid*	*Ask*	*Bid*	*Ask*	
1 ms	0.024	0.024	0.076	0.076	0.088
4 ms	0.038	0.038	0.122	0.120	0.070
16 ms	0.071	0.070	0.200	0.203	0.219
64 ms	0.145	0.143	0.400	0.403	0.375
256 ms	0.303	0.298	0.831	0.842	0.530
1.0 sec	0.573	0.564	1.664	1.685	0.627
4.1 sec	1.119	1.109	3.649	3.600	0.829
16.4 sec	2.043	2.031	7.851	7.842	0.967

$\sqrt{\boldsymbol{Cumulative\ wavelet\ variances}}$

for AAPL *Bid* (units: $0.01/sh)

τ_j	*Median*	99^{th} *percentile*
1 ms	0.024	0.076
4 ms	0.053	0.172
16 ms	0.103	0.302
64 ms	0.205	0.571
256 ms	0.423	1.153
1.0 sec	0.828	2.337
4.1 sec	1.621	4.857
16.4 sec	3.167	10.331

$\sqrt{\textit{Cumulative wavelet variances}}$
for AAPL *Bid* (units: \$0.01/sh)

The price uncertainty for a trader who can only time his marketable trades within a 4-second window has $\sigma = \$0.016$
Compare: current access fees $\approx$ \$0.003

1 ms	[illegible]	[illegible]
4 ms	0.053	[illegible]
16 ms	0.103	[illegible]02
64 ms	0.205	0.571
256 ms	0.423	1.153
1.0 sec	0.828	2.337
4.1 sec	1.621	4.857
16.4 sec	3.167	10.331

The wavelet correlation $\rho_{X,Y}(\tau_j)$

- For two series X and Y, the wavelet variances are $\nu_X^2(\tau_j)$ and $\nu_Y^2(\tau_j)$
- The wavelet covariance is $\nu_{X,Y}(\tau_j)$
- The wavelet correlation is

$$\rho_{X,Y}(\tau_j) = \frac{\nu_{X,Y}(\tau_j)}{\sqrt{\nu_X^2(\tau_j)\nu_Y^2(\tau_j)}}$$

- Fundamental value changes should affect both the bid and the ask.
- The wavelet correlation at scale τ_j indicates the contribution of fundamental volatility.
- Next: wavelet correlation for AAPL bid and ask:

How closely do the wavelet variances for AAPL's bid correspond to a random walk?

Scale	*Wavelet variance estimate*	*Random-walk variance factors*	*Implied random-walk variance*
τ_j	$\hat{v}^2(\tau_j)$	$\phi(\tau_j)$	$\hat{v}^2(\tau_j)\,/\phi(\tau_j)$
1 ms	0.0009	0.1875	0.0050
4 ms	0.0024	0.4849	0.0049
16 ms	0.0075	1.8960	0.0039
64 ms	0.0309	7.5740	0.0041
256 ms	0.1360	30.2935	0.0045
1.0 sec	0.5140	121.1730	0.0042
4.1 sec	2.1434	484.6930	0.0044
16.4 sec	8.6867	1,938.7700	0.0045

Reasonable?

- If 0.005 (cents per share)^2 is the random-walk variance over one ms., the accumulated variance over a 6-hour mid-day period is:
- $0.005 \times 1{,}000 \times 3{,}600 \times 6 = 108{,}000$
- The implied 6-hour standard deviation is about 329 (cents per share).
- AAPL's average price in the sample is about \$340
- $\frac{3.29}{340} \approx 1\%$

Volatility Signature Plots

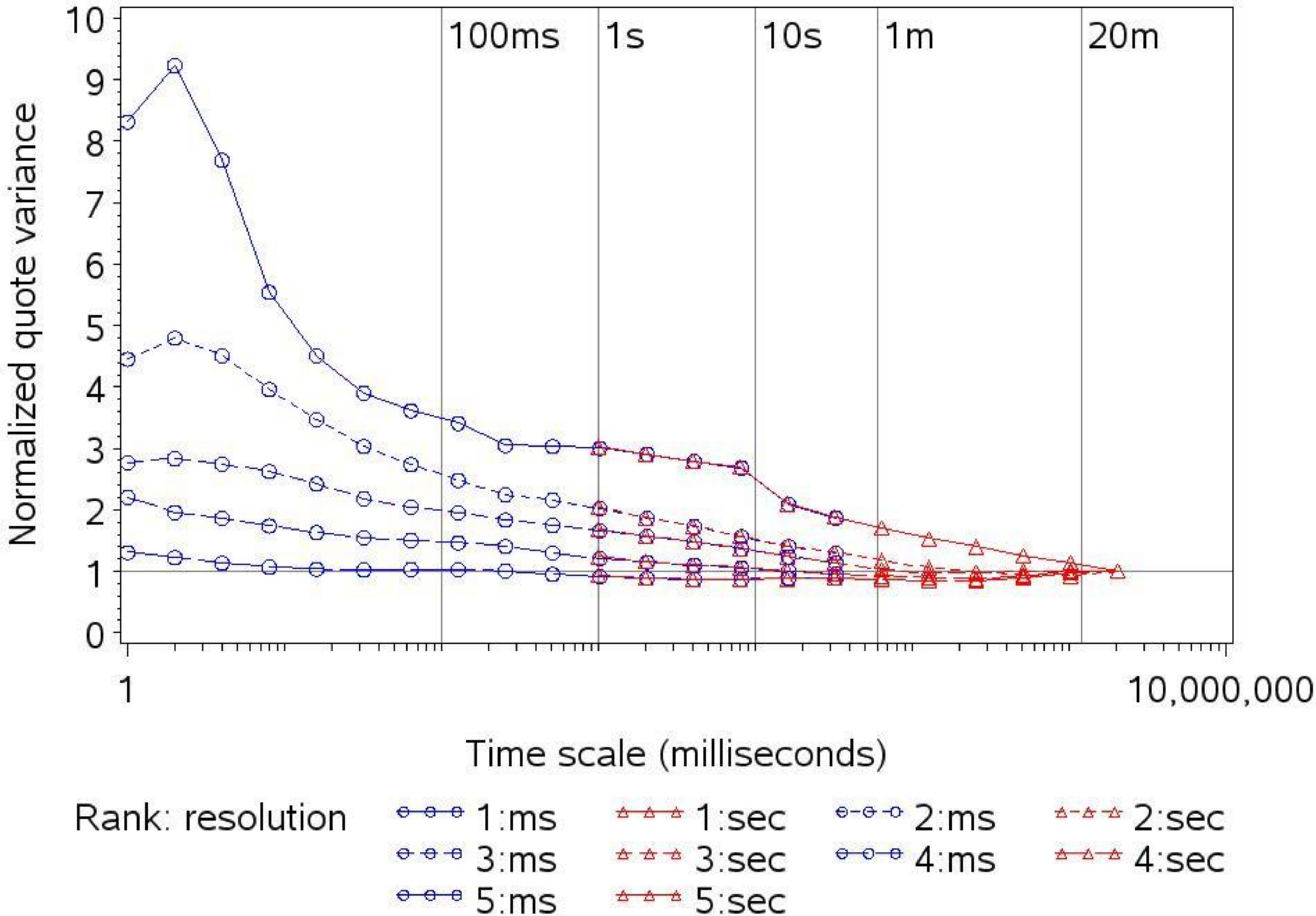
- Suggested by Andersen and Bollerslev.
- Plot realized volatility (per constant time unit)
  vs. length of interval used to compute the realized volatility.
- Basic idea works for wavelet variances.
- "How much is short-run quote volatility inflated, relative to what we'd expect from a random walk?"

Normalization of wavelet variances

- For a given stock, the implied random-walk variance at scale τ_j is $\hat{\nu}^2(\tau_j)\ /\phi(\tau_j)$.
- The longest time scale in the analysis is about 20 minutes.
- The ratio $\dfrac{\hat{\nu}^2(\tau_j)\ /\phi(\tau_j)}{\hat{\nu}^2(20\ min)\ /\phi(20\ min)}$ measures variance at scale τ_j relative to the wavelet variance at 20 minutes, under a random-walk benchmark.
- If the price is truly a random walk, this should be unity for all τ_j.

For presentation …

- Market cap deciles collapsed into quintiles.
- Within each quintile, I average $\dfrac{\hat{\nu}^2(\tau_j)\ /\phi(\tau_j)}{\hat{\nu}^2(20\ min)\ /\phi(20\ min)}$ across firms.
- Results from millisecond- and second-resolution analyses are spliced.
- Next: the (normalized) volatility signature plot.



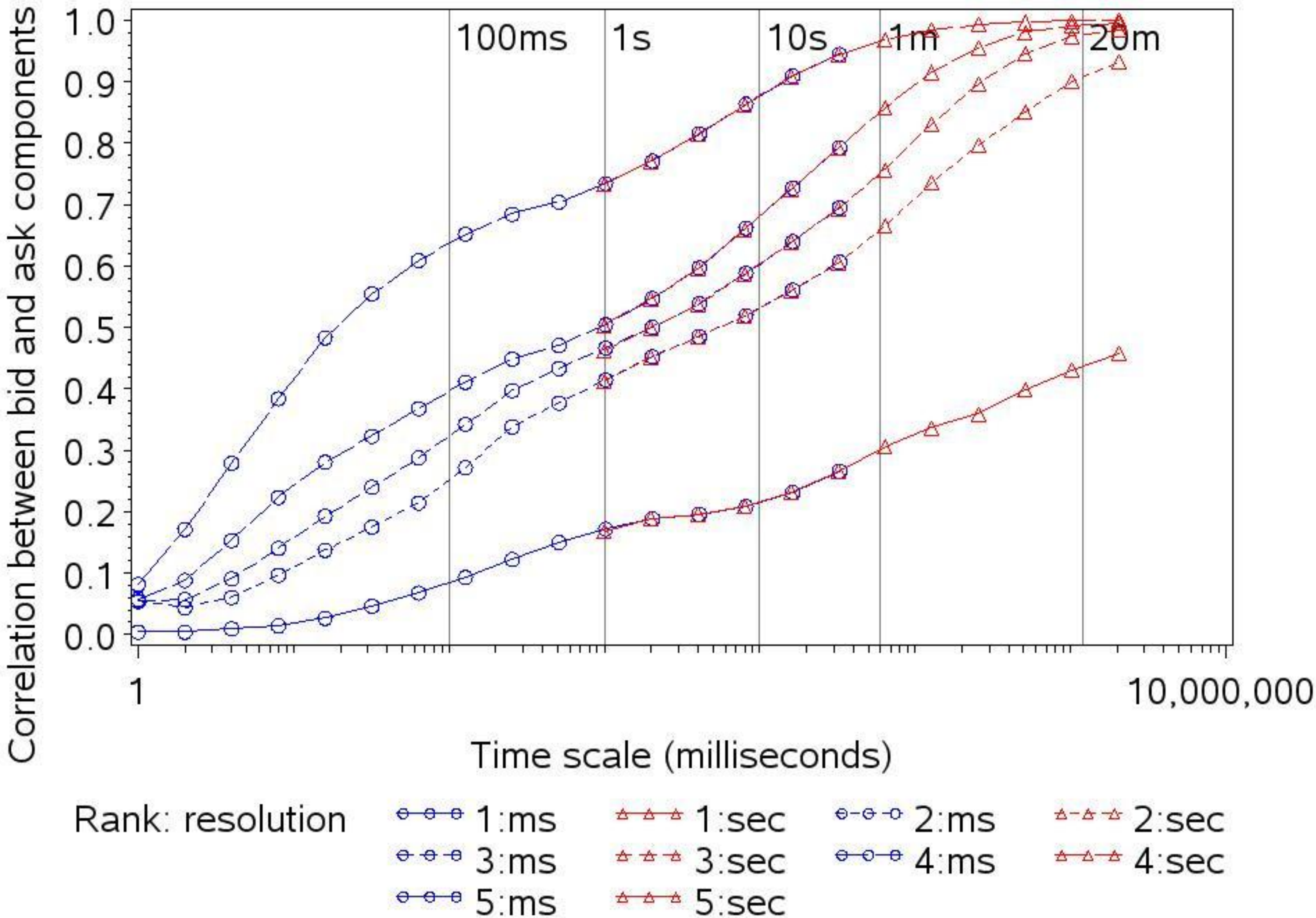
10
9
8
7
6
5
4
3
2
1
0
Normalized quote variance
100ms
1s
10s
1m
20m
1
10,000,000
Time scale (milliseconds)
Rank: resolution
1:ms
1:sec
2:ms
2:sec
3:ms
3:sec
4:ms
4:sec
5:ms
5:sec

The take-away

- For high-cap firms
  - Wavelet variances at short time scales have modest elevation relative to random-walk.
- Low-cap firms
  - Wavelet variances are strongly elevated at short time scales.
  - Significant price risk relative to TWAP.

Sample bid-ask wavelet correlations

- These are already normalized.
- Compute quintile averages across firms.



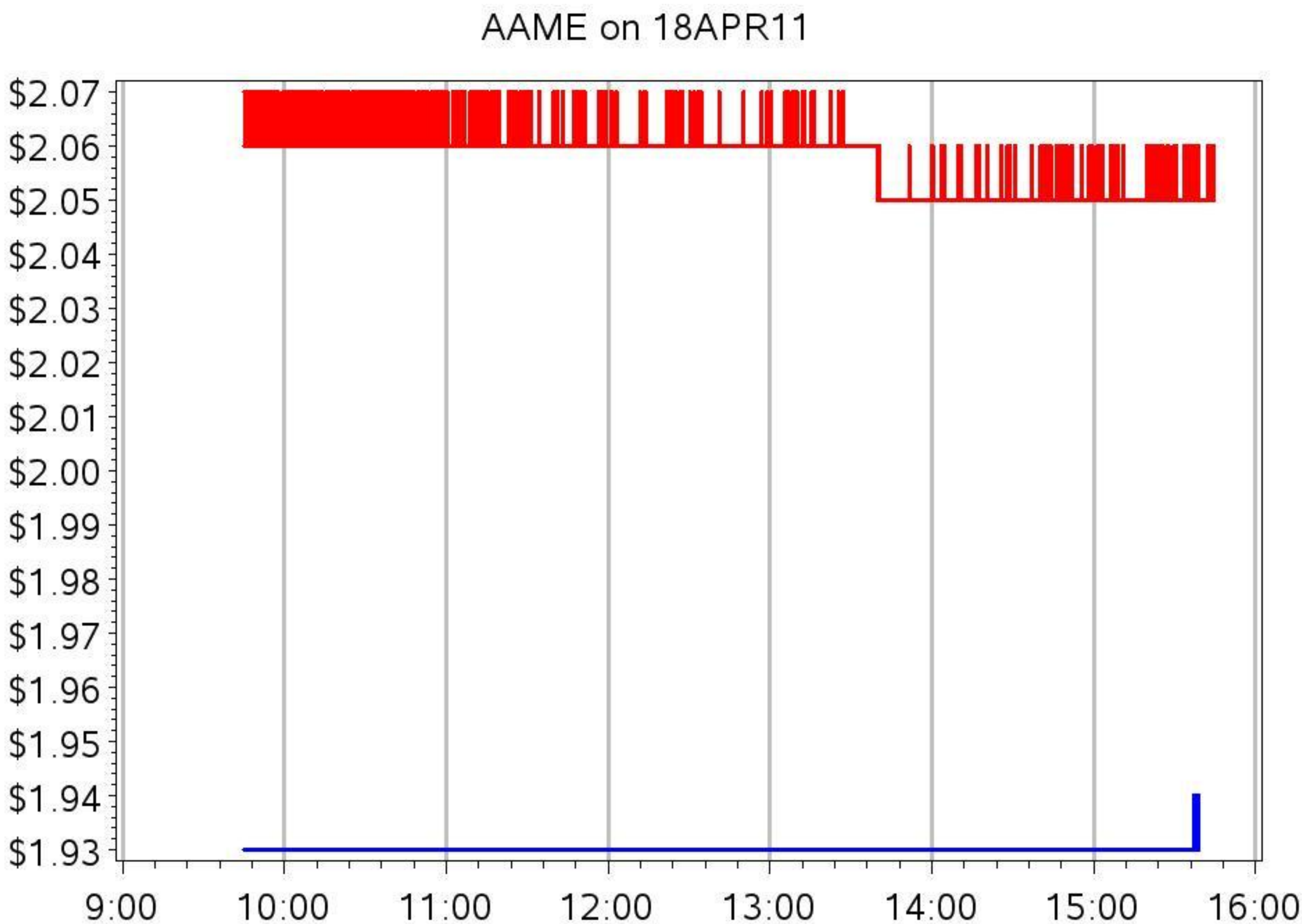
Correlation between bid and ask components
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
100ms
1s
10s
1m
20m
1
10,000,000
Time scale (milliseconds)
Rank: resolution
1:ms
1:sec
2:ms
2:sec
3:ms
3:sec
4:ms
4:sec
5:ms
5:sec

How closely do movements in the bid and ask track?

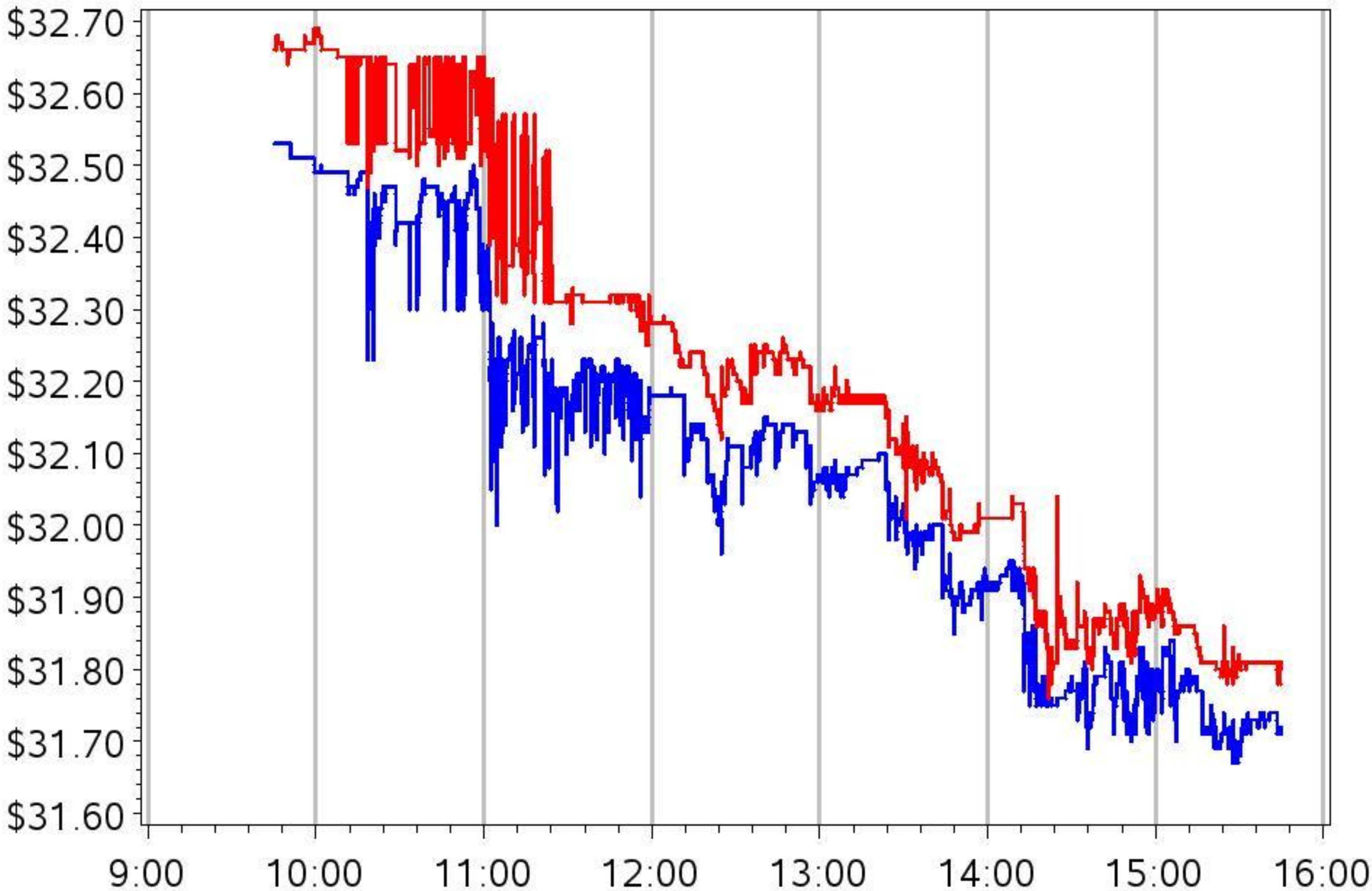
- Positive in all cases (!)
- For high-cap stocks, $\rho \approx 0.7$ (one second) and $\rho > 0.9$ (20 seconds)
- For bottom cap-quintile, $\rho < 0.2$ (one second) and $\rho < 0.5$ (20 minutes)

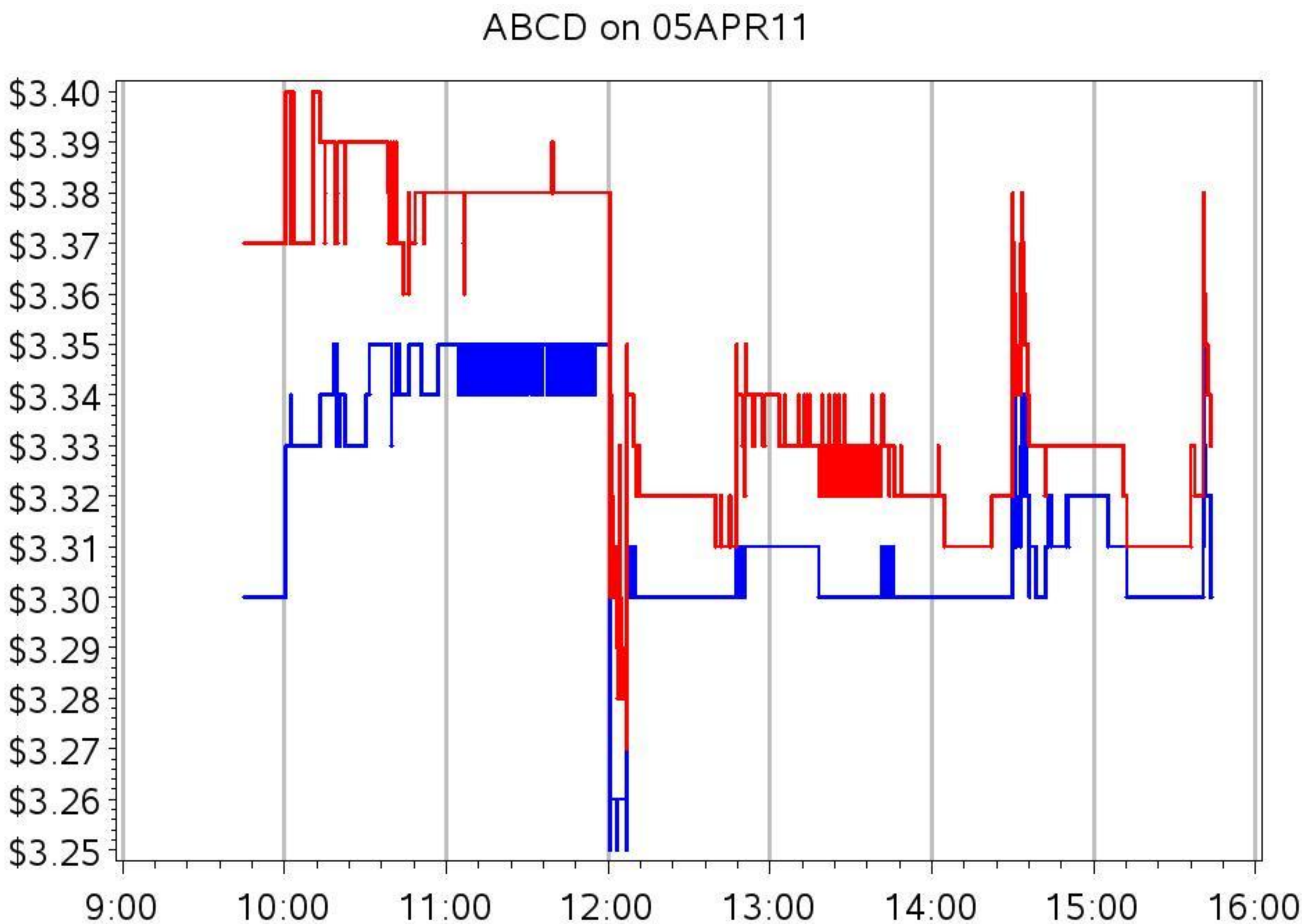
A gallery

- For each firm in mkt cap deciles ≤ 6, I examined the day with the highest wavelet variance at time scales of 1 second and under.
  - HFQ is easiest to see against a backdrop of low activity.
- Next slides … some examples

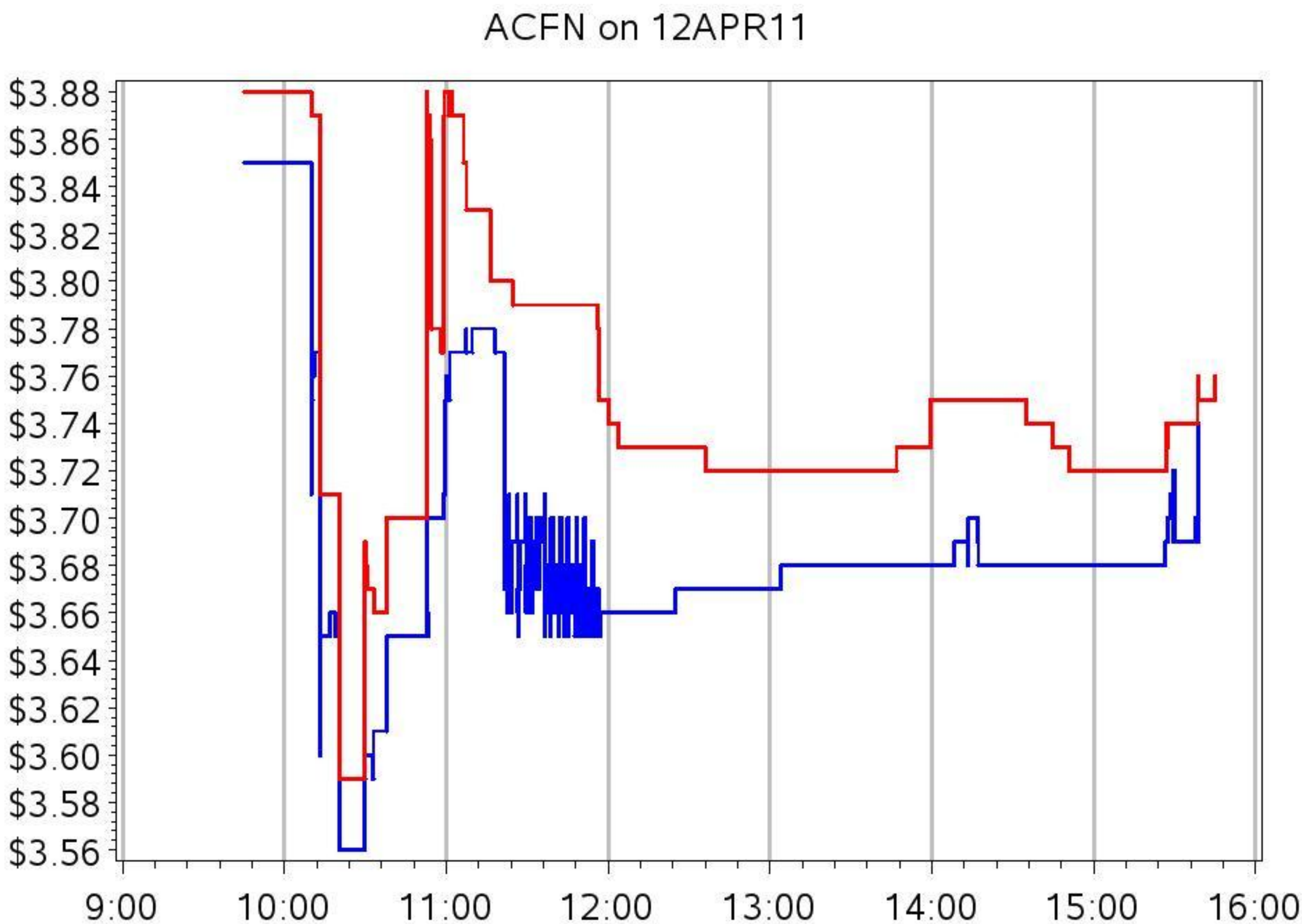

AAME on 18APR11
$2.07
$2.06
$2.05
$2.04
$2.03
$2.02
$2.01
$2.00
$1.99
$1.98
$1.97
$1.96
$1.95
$1.94
$1.93
9:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00

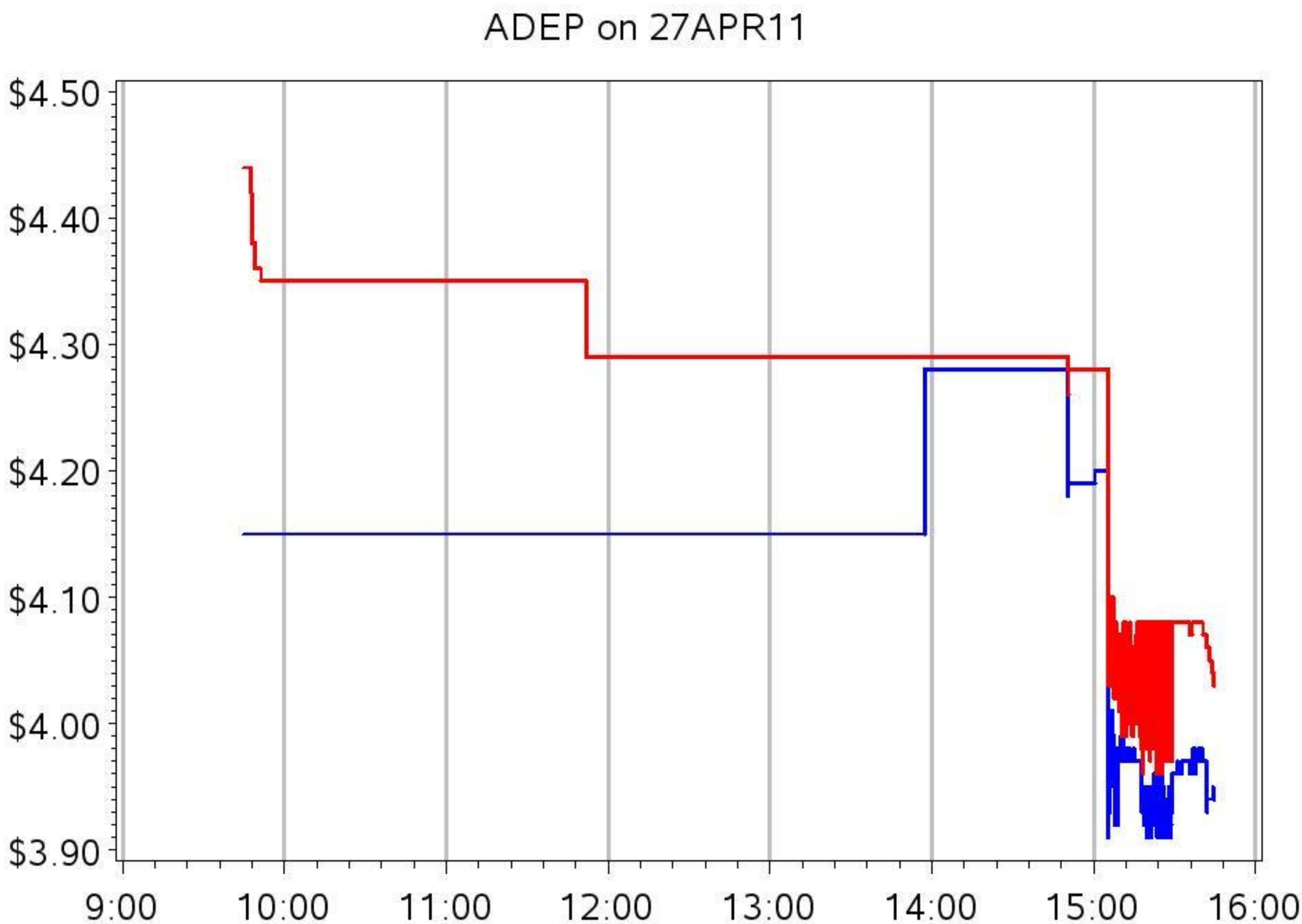
AAON on 08APR11

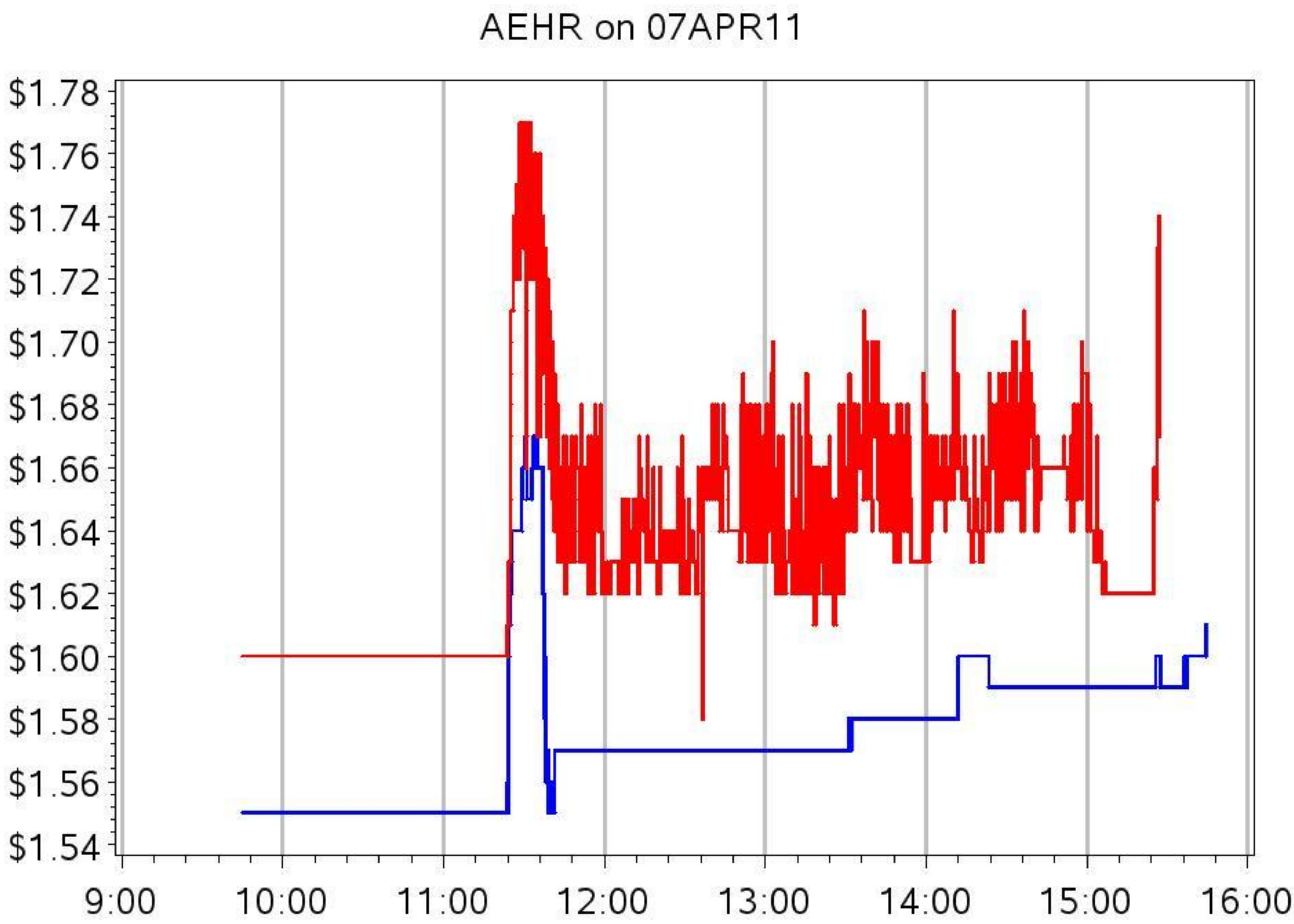


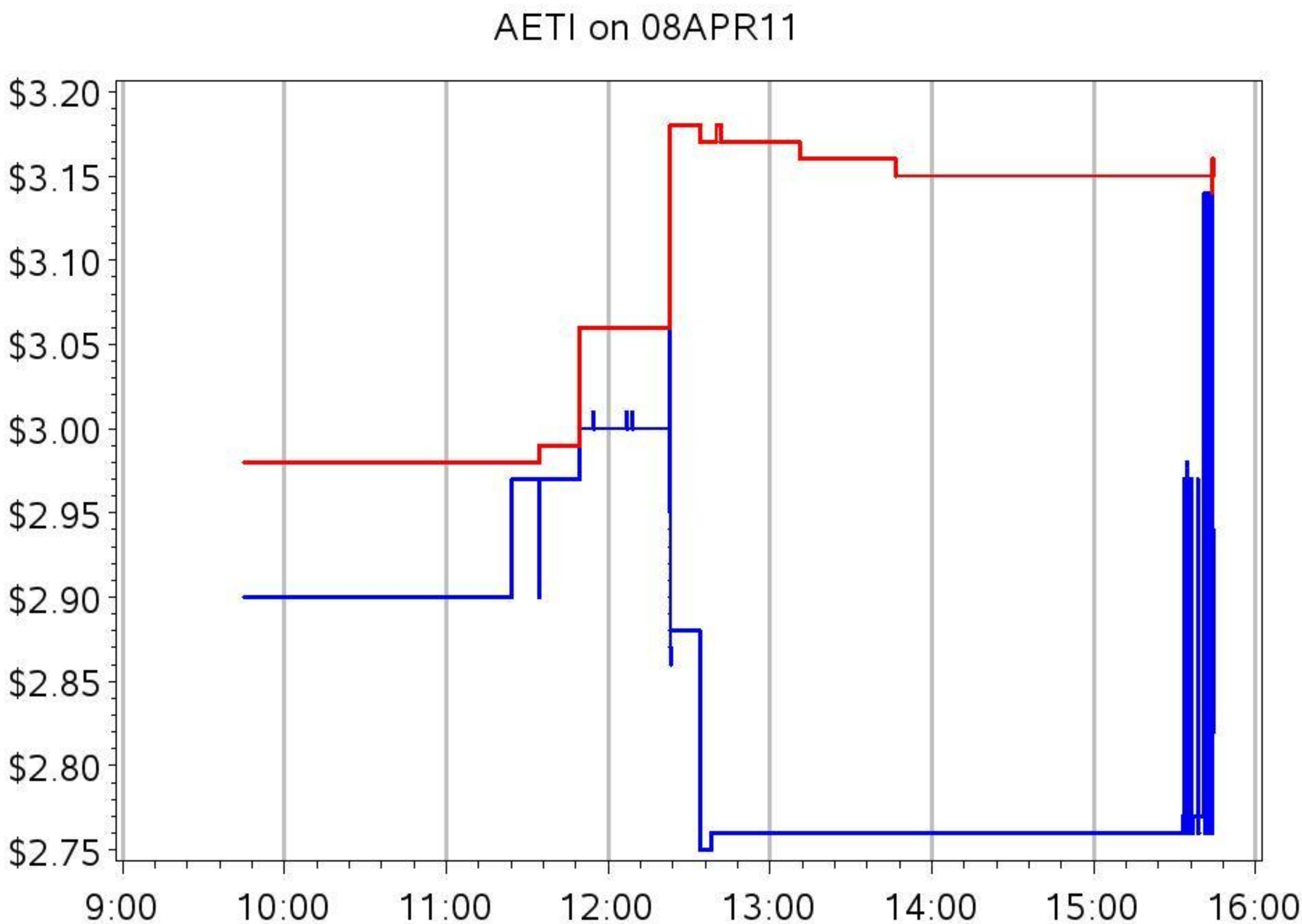


ABCD on 05APR11
$3.40
$3.39
$3.38
$3.37
$3.36
$3.35
$3.34
$3.33
$3.32
$3.31
$3.30
$3.29
$3.28
$3.27
$3.26
$3.25
9:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00


ACFN on 12APR11
$3.88
$3.86
$3.84
$3.82
$3.80
$3.78
$3.76
$3.74
$3.72
$3.70
$3.68
$3.66
$3.64
$3.62
$3.60
$3.58
$3.56
9:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00


ADEP on 27APR11
$4.50
$4.40
$4.30
$4.20
$4.10
$4.00
$3.90
9:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00


AEHR on 07APR11
$1.78
$1.76
$1.74
$1.72
$1.70
$1.68
$1.66
$1.64
$1.62
$1.60
$1.58
$1.56
$1.54
9:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00


AETI on 08APR11
$3.20
$3.15
$3.10
$3.05
$3.00
$2.95
$2.90
$2.85
$2.80
$2.75
9:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00

Conclusions

- High frequency quoting is a real (but episodic) fact of the market.
- Time-scale decompositions are useful in measuring the overall effect.
  - … and detecting the episodes
- Remaining questions …

Why does HFQ occur?

- Why not? The costs are extremely low.
- Testing?
- Malfunction?
- Interaction of simple algos?
- Genuinely seeking liquidity (counterparty)?
- Deliberately introducing noise?
- Deliberately pushing the NBBO to obtain a favorable price in a dark trade?