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

Joel Hasbrouck*

November 13, 2012

This version: February 22, 2013

I have benefited from the comments of Ramo Gençay, Dale Rosenthal, Gideon Saar, Mao Ye, seminar participants at the Emerging Markets Group (Cass Business School, City University London), l'Institute Louis Bachelier, Jump Trading, SAC Capital, the University of Illinois at Chicago, the University of Illinois at Champaign and Utpal Bhattacharya's doctoral students at the University of Indiana. All errors are my own responsibility. DISCLAIMER: This research was not specifically supported or funded by any organization. During the period over which this research was developed, I taught (for compensation) in the training program of a firm that engages in high frequency trading, and served as a member (uncompensated) of a CFTC advisory committee on high frequency trading.

I am grateful to Jim Ramsey for originally introducing me to time scale decompositions.

*Department of Finance, Stern School of Business, New York University, 44 West 4th Street, New York, NY 10012 (Tel: 212-998-0310, jhasbrou@stern.nyu.edu).

High-Frequency Quoting:

Short-Term Volatility in Bids and Offers

Abstract

High frequency changes, reversals, and oscillations induce volatility in a market's bid and offer quotes. This volatility degrades the informational content of the quotes, exacerbates execution price risk for marketable orders, and impairs the reliability of the quotes as reference marks for the pricing of dark trades. This paper examines variance on time scales as short as fifty milliseconds for the National Best Bid and Offer (NBBO) in the US equity market. On average, in a 2011 sample, NBBO variance at the fifty millisecond time scale is approximately four times larger than can be attributed to long-term fundamental price variance. The historical picture is complex. There is no marked upward trend in short-term quote volatility over 2001-2011. Its character, though, has changed. In the early years (and especially prior to Reg NMS) quote volatility is driven by large spikes in bids and offers. In later years it is more a consequence of high frequency oscillations comparable to the bid-offer spread in magnitude.

I. Introduction

Recent developments in market technology have called attention to the practice of high frequency trading. The term is used commonly and broadly in reference to all sorts of fast-paced market activity, not just "trades", but trades have certainly received the most attention. There are good reasons for this, as trades signify the actual transfers of income streams and risk. Quotes also play a significant role in trading process, however. This paper examines short-term volatility in bids and offers of US equities, a consequence of what might be called high frequency quoting.

By way of illustration, Figure 1 depicts the bid and offer for AEP Industries (a NASDAQ-listed manufacturer of packaging products) on April 29, 2011.[1] In terms of broad price moves, the day is not a particularly volatile one, and the bid and offer quotes are stable for long intervals. The placidity is broken, though, by several intervals where the bid undergoes extremely rapid changes. The average price levels, before, during and after the episodes are not dramatically different. Moreover, the episodes are largely one-sided: the bid volatility is associated with an only moderately elevated volatility in the offer quote. Nor is the volatility associated with increased executions. These considerations suggest that the volatility is unrelated to fundamental public or private information. It appears to be an artifact of the trading process.

It is not, however, an innocuous artifact. Bids and offers in all markets represent price signals, and, to the extent that they are firm and accessible, immediate trading opportunities. From this perspective, the noise added by quote volatility impairs the informational value of the public price. Most agents furthermore experience latency in ascertaining the market center with the best price and in timing of their order delivery. Elevated short-term volatility increases the execution price risk associated with these delays. In US equity markets the bid and offer are particularly important, because they are used as

[1] The bid is the National Best Bid (NBB), the maximum bid across all exchanges. The offer is the National Best Offer (NBO), the minimum offer. They are often jointly referred to as the NBBO. Unless otherwise noted, or where clarity requires a distinction, "bid" and "offer" indicate the NBBO.

benchmarks to assign prices in so-called dark trades, a category that includes roughly thirty percent of all volume.[2]

In the context of the paper's data sample, the AEPI episode does not represent typical behavior. Nor, however, is it a singular event. It therefore serves to motivate the paper's key questions. What is the extent of short-term volatility? How can we distinguish fundamental (informational) and transient (microstructure) volatility? Finally, given the current public policy debate surrounding low-latency activity, how has it changed over time?

These questions are addressed empirically in a broad sample of US equity market data using summary statistics that are essentially short-term variances of bids and offers. Such constructions, though, inevitably raise the question of what horizon constitutes the "short term" (a millisecond? a minute?). The answer obviously depends on the nature of the trader's market participation, as a collocated algorithm at one extreme, for example, or as a remotely situated human trader at the other. The indeterminacy motivates empirical approaches that accommodate flexible time horizons. This analysis uses time scale variance decompositions to measure bid and offer volatility over horizons ranging from under 50 ms to about 27 minutes.

The next section establishes the economic and institutional motivation for the consideration of local bid and offer variances with sliding time scales. Section III discusses the statistical framework. The paper then turns to applications. Section IV presents an analysis of a recent sample of US equity data featuring millisecond time stamps. To extend the analysis to historical samples in which time stamps are to the second, Section V describes estimation in a Bayesian framework where millisecond time stamps are simulated. Section VI applies this approach to a historical sample of US data from 2001 to 2011. Connections to high frequency trading and volatility modeling are discussed in Section VII. A summary concludes the paper in Section VIII.

[2] Dark mechanisms do not publish visible bids and offers. They establish buyer-seller matches, either customer-to-customer (as in a crossing network) or dealer-to-customer (as in the case of an internalizing broker-dealer). The matches are priced by reference to the NBBO: generally at the NBBO midpoint in a crossing network, or at the NBB or the NBO in a dealer-to-customer trade.

II. The economic effects of quote volatility.

High frequency quote volatility may be provisionally defined as the short-term variance of the bid or offer, the usual variance calculation applied to the bid or offer *level* over a relatively brief window of time. This section is devoted to establishing the economic relevance of such a variance in a trading context. The case is a simple one, based on the function and uses of the bid and offer, the barriers to their instantaneous availability, the role of the time-weighted price mean as a benchmark, and the interpretation of the variance about this mean as a measure of risk.

In current thinking about markets, most timing imperfections are either first-mover advantages arising from market structure or delays attributed to costly monitoring. The former are exemplified by the dealer's option on incoming orders described in Parlour and Seppi (2003), and currently figure in some characterizations of high frequency traders (Biais, Foucault and Moinas (2012); Jarrow and Protter (2011)). The latter are noted by Parlour and Seppi (2008) and discussed by Duffie (2010) as an important special case of inattention which, albeit rational and optimal, leads to infrequent trading, limited participation, and transient price effects (also, Pagnotta (2009)).

As a group these models feature a wide range of effects bearing on agents' arrivals and their information asymmetries. An agent's market presence may be driven by monitoring decisions, periodic participation, or random arrival intensity. Asymmetries mostly relate to fundamental (cash-flow) information or lagged information from other markets. Agents in these models generally possess, however, timely and extensive market information. Once she "arrives" in a given market, an agent accurately observes the state of that market, generally including the best bid and offer, depth of the book and so on. Moreover, when she contemplates an action that changes the state of the book (such as submitting, revising or canceling an order), she knows that her action will occur before any others'.

In reality, of course, random latencies in receiving information and transmitting intentions combine to frustrate these certainties about the market and the effects of her orders. The perspective of this paper is that for some agents these random latencies generate randomness in the execution prices, and that short-term quote variances can meaningfully measure this risk. Furthermore, although all agents incur random latency, the distributions of these delays vary among participants. An agent's latency distribution can be summarized by time scale, and this in turn motivates time scale decompositions of bid and offer variances.

While random latencies might well affect strategies of all traders, the situation is clearest for someone who intends to submit a marketable order (one that seeks immediate execution) or an order to a dark pool. In either case, ignoring hidden orders, an execution will occur at the bid, the offer or at an average of the two. A trader whose order arrival time is uniformly distributed on a given interval faces price risk over that interval. For a marketable sell order, the variance of the bid over the interval quantifies the price risk, relative to a benchmark equal to the average bid.[3]

The situations discussed to this point involve a single trader and single market. In a fragmented market, the number of relevant latencies may be substantially larger. In the US there are presently about 17 "lit" market centers, which publish quotes. A given lit market's quotes are referenced by the other lit markets, dark pools (currently around 30 in number), by executing broker-dealers (approximately 200), and by data consolidators (U.S. Securities and Exchange Commission (2010)). The National Best Bid and Offer (NBBO) is in principle well-defined. The NBBO perceived by any given market center, consolidator or other agent, however, comprises information subject to random transmission delays that differ across markets and receiving agents. These delays introduce noise into the determination. Local time-averaging (smoothing) can help to mitigate the effects of this noise, while the local variance indicates the magnitude of the noise.

If the execution price risk associated with quote volatility is zero-mean and diversifiable across trades, it might appear to be economically trivial. In general, however, agents do not have symmetric

---

[3] The use of an average price in the presence of execution timing uncertainty is a common principle in transaction cost analysis. Perold's implementation shortfall measure is usually operationally defined for a buy order as the execution price (or prices) less some hypothetical benchmark price (and for a sell order as the benchmark less the execution price, Perold (1988)). As a benchmark price, Perold suggests the bid-offer midpoint prevailing at the time of the decision to trade. Many theoretical analyses of optimal trading strategies use this or a similar pre-trade benchmark. Practitioners, however, and many empirical analyses rely on prices averaged over some comparison period. The most common choice is the value-weighted average price (VWAP), although the time-weighted average price (TWAP) is also used. One industry compiler of comparative transaction cost data notes, "In many cases the trade data which is available for analysis does not contain time stamps. …. When time stamps are not available, pension funds and investment managers compare their execution to the volume weighted average price of the stock on the day of the trade" (Elkins-McSherry (2012)). This quote attests to the importance of execution time uncertainty, although a day is certainly too long to capture volatility on the scale of transmission and processing delays. Average prices are also used as objectives by certain execution strategies. A substantial portion of the orders analyzed by Engle, Ferstenberg and Russell (2012) target VWAP, for example.

exposure to this risk. Market-order traders with faster technology possess a systematic advantage relative to those with slower technology. This can be viewed as an information asymmetry that leads (in the usual fashion) to a transfer of wealth from the slower to the faster participants.

Asymmetric exposure to quote volatility is also likely to place customers at a disadvantage relative to their dealers. The recent SEC concept release notes that virtually all retail orders are routed to OTC market-makers, who execute the orders by matching the prevailing NBBO (U.S. Securities and Exchange Commission (2010)). Stoll and Schenzler (2006) note that these market-makers have flexibility in delaying executions to obtain favorable reference prices. They describe this as a look-back option, and find support for this behavior in a 1999 sample. Dark trading venues also face this sort of problem. A customer sending a sell order to a dark pool or crossing network can submit a *buy* order to a lit market center that will briefly boost quote midpoint, thereby achieving a better price if he receives a dark execution of his sell order. This practice (a form of "spoofing") is forbidden in the Dodd-Frank framework, but remains difficult to detect and prove in the presence of timing uncertainties.

The SEC's Reg NMS ruling on trade-through protection recognized the problem of "flickering quotes", and mandated a one-second grace period: "… pursuant to Rule 611(b)(8) trading centers would be entitled to trade at any price equal to or better than the least aggressive best bid or best offer, as applicable, displayed by the other trading center during that one-second window." Sub-second intervals were considered, but the benefits were not believed sufficient to justify the costs (U.S. Securities and Exchange Commission (2005)). Clearly, quote volatility within the one-second window weakens the trade-through protection.[4]

[4] The SEC has recently mandated a consolidated audit trail intended to track all events in an order's life cycle (such as receipt, routing, execution and cancellation) (U.S. Securities and Exchange Commission (2012)). In the final rule, the Commission recognized the importance of accurately sequencing the events, and mandated time-stamps at least to the granularity of the millisecond, and, "to the extent that the order handling and execution systems of any SRO or broker-dealer utilize time stamps in increments finer than the minimum required by the NMS plan time stamps, such SRO or member must use time stamps in such finer increments when reporting data to the central repository."

III. Time scale decompositions

This section develops the empirical framework for the analysis. Consider a price process, p_t, in discrete time. The *n*-period mean ending at time *t* is $S(n,t) = n^{-1}\sum_{i=0}^{n-1} p_{t-i}$. The deviations about this mean are $R(n,t,s) = p_s - S(n,t)$ for $t-n+1 < s \le t.$ The mean square of this deviation is $MSD(n,t) = n^{-1}\sum_{s=t-n+1}^{t}\left[R(n,t,s)\right]^2$. These are simply computational definitions, summarizing calculations that might be performed on any segment of the price path. To interpret these quantities in a probabilistic framework, we now assume that the first differences of the price process constitute a stationary (but not necessarily uncorrelated) stochastic process. With this assumption, the expectation of $MSD(n,t)$ is finite, time-invariant and equal to the variance of *R*: $E\left[MSD(n,t)\right] = Var\left[R(n,t,s)\right] \equiv \sigma_n^2$.

Given her technology and a contemplated order, an individual trader might well focus on the particular horizon associated with her order arrival uncertainty period. In the present analysis, though, it more useful to consider a broad range of averaging periods that spans horizons of general interest. Any increasing sequence of *n* might be chosen (such as, 1 sec, 5 sec, 20 sec). For various mathematical reasons, though, it is convenient to let *n* be dyadic (increasing by powers of two): $n_j = n_0 2^j$ for $j = 0,1,\ldots$, where $n_0 \ge 1$ defines the starting resolution for the analysis. The corresponding sequence of *R* variances $\sigma_{n_j}^2$ is denoted σ_j^2 for $j = 0,1,\ldots$ It will also be useful to define $\nu_j^2 = \sigma_j^2 - \sigma_{j-1}^2$ as the incremental variance associated with moving from averaging interval n_{j-1} to n_j.

The random-walk special case is useful both as an illustration of the calculations and also as benchmark in interpreting the empirical results. Suppose that the price evolves in continuous time with variance per unit time σ_u^2 (not necessarily a Gaussian diffusion) and that the prices are initially averaged over successive intervals of length M_0 units of time. It is shown in the appendix that $\sigma_j^2 = 2^{j-1} M_0 \sigma_u^2 / 3$ and $\nu_j^2 = 2^{j-2} M_0 \sigma_u^2 / 3$.

The characterization of a time series according to time scale was historically based on Fourier (frequency domain) analysis, in which a series is decomposed as a sum of sine and cosine basis functions. Modern approaches use broader classes of basis functions, called wavelets. The essential distinction is that while trigonometric functions cycle repeatedly over the full sample, wavelets are localized, and are therefore better suited to picking up phenomena like the AEPI movements that are very concentrated in time.

Percival and Walden (2000) is a comprehensive textbook discussion of wavelets that emphasizes connections to conventional time series analysis. The present notation mostly follows their conventions. *S* and *R* denote smoothed and rough components (or simply "smooths" and "roughs"). The dyadic convention facilitates the use of standard computationally efficient methods (wavelet transforms) to compute the σ_j^2 and ν_j^2 estimates. The quantity ν_j^2 is formally considered a wavelet variance. It is important to emphasize, though, that it can be defined (as above) and even computed (albeit suboptimally) without wavelet transforms. Denoting it as the wavelet variance simply places it in an extensive and well-developed literature.[5]

Figure 2 illustrates the decomposition for a simple eight-observation price series. As a first step (level one), the eight observations are partitioned into four pairs. S_1 and R_1 are the pairwise means and deviations from these means. At the second level, S_2 and R_2 are the means and deviations when the observations are partitioned into two sets, each of length four. At the third level, the full sample is used. Explicitly graphing the rough (residual) components R_j is useful because it illustrates the changes in rough components as the averaging period grows. The successive differences in roughs are called the

[5] Wavelet transformations, also known as time scale or multi-resolution decompositions are widely used across many fields. Gençay, Selçuk and Whitcher (2002) discuss economic and financial applications in the broader context of filtering. Nason (2008) discusses time series and other applications of wavelets in statistics. Ramsey (1999) and Ramsey (2002) provides other useful economic and financial perspectives. Walker (2008) is clear and concise, but oriented more toward engineering applications.

Studies that apply wavelet transforms to the economic analysis of stock prices loosely fall into two groups. The first set explores time scale aspects of stock comovements. A stock's beta is a summary statistic that reflects short-term linkages (like index membership or trading-clientele effects) and long-term linkages (like earnings or national prosperity). Wavelet analyses can characterize the strength and direction of these horizon-related effects (for example, Gençay, Selçuk and Whitcher (2005); In and Kim (2006)). Most of these studies use wavelet transforms of stock prices at daily or longer horizons. A second group of studies uses wavelet methods to characterize volatility persistence (Dacorogna, Gencay, Muller, Olsen and Pictet (2001); Elder and Jin (2007); Gençay, Selçuk, Gradojevic and Whitcher (2010); Gençay, Selçuk and Whitcher (2002); Høg and Lunde (2003); Teyssière and Abry (2007)). These studies generally involve absolute or squared returns at minute or longer horizons. Wavelet methods have also proven useful for jump detection and jump volatility modeling Fan and Wang (2007). Beyond studies where the focus is primarily economic or econometric lie many more analyses where wavelet transforms are employed for ad hoc stock price forecasting (Atsalakis and Valavanis (2009); Hsieh, Hsiao and Yeh (2011), for example). An early draft of Hasbrouck and Saar (2011) used wavelet analyses of message count data to locate periods of intense message traffic on NASDAQ's Inet system.

detail components, defined by $D_1 = R_1$ and $D_j = R_j - R_{j-1}$ for $j \geq 2$. The details are graphed in the third column. The wavelet variances are identical to the detail variances.

The detail component and wavelet variance possess an economic interpretation. Consider a stylized situation in which the full price path is known, but agents face timing uncertainties. A slow trader can only decide whether to trade in the first or second half of the sample. Within each half, his actual arrival time is uniformly random. His expected prices are given by S_2. A slow buyer can expect to pay a lower price in the first half; a slow seller can expect to receive a higher price in the second half. A faster trader can pick a quarter (and within each quarter her arrival is random). Her expected prices are those shown in S_1. The detail component $D_2 \equiv R_2 - R_1 = S_1 - S_2$ represents her incremental advantage in the following sense. A slow buyer (trading in the first half) expects to pay 4.5. If upon arrival he is given the opportunity to pick the first or second quarter, he will buy in the first quarter, expecting to pay 3. The difference, 1.5, is the magnitude of D_2 in the first half. A slow seller (trading in the second half) can expect to receive 6.5. If he is now allowed to pick the third or fourth quarter, he'll sell in the fourth quarter and expect to receive 8.5. The difference, 2, is the magnitude of D_2 in the second half. The detail (wavelet) variance $\nu_2^2 = 3.125$ can be viewed as the expected squared differential improvement over the full sample. This example is unrealistic, particularly in its assumption of a known price path, but by focusing purely on timing uncertainty, it illustrates the incremental profitability associated with lower latency. The same numbers and setup could be used to illustrate the customer savings associated with restricting the dealer look-back option from four periods to two.[6]

The rough variance σ_j^2 reflects variation over all time scales shorter than the averaging period $n_j = n_0 2^j$. The wavelet variance ν_j^2, though, is defined as an incremental quantity, and reflects variation only at time scale $\tau_j = n_0 2^{j-1}$. The market-order trader exposed to timing risk at a particular time scale would generally also be exposed to risk at all shorter time scales, and so would prefer to assess uncertainty by the rough variance. The detail variance, though, may be useful (as above) in identifying incremental effects. Furthermore, there may be economic or technological reasons why an effect can't

[6] Relative advantage does not necessarily imply the existence of an arbitrage opportunity. The price path in the example is intended to be either a bid or an offer, not a price that is available to one agent for trade in both directions.

occur on a shorter time scale. Intermarket feedback effects, for example, can't operate over scales shorter than the intermarket transmission time. Time scale decompositions that exclude these components may therefore potentially offer better characterizations of intermarket contagion effects.

Excluding short-term components in security price analyses may also be preferable in situations where the aim is an estimate of fundamental volatility. A long horizon wavelet variance, for example, is essentially free of shorter-horizon noise. Fan and Gencay (2010) apply this principle to unit root tests based on time scale decompositions. Gencay and Signori (2012) explore the use of variance ratios at different time scales to test for serial correlation. In the present application, multi-scale variance ratios can be used to assess the excess high frequency volatility relative to what would be implied by a random-walk calibrated to a low frequency. Consider the variance ratio

$$V_{j,J} = 2^{J-j}\nu_j^2 \big/ \nu_J^2$$

Here, J denotes the largest index (longest time scale) considered in the study, j $\left(0 \le j \le J\right)$ generally denotes a shorter time scale, and 2^{J-j} is a scaling factor. If the process follows a random walk with uncorrelated increments, $V_{j,J} = 1$. To the extent that $V_{j,J}$ exceeds unity, there is excess short-term volatility. This variance ratio is defined in terms of the wavelet variances. A similar normalization can be defined for the rough variances as

$$VR_{j,J} = 2^{J-j-1}\sigma_j^2 \big/ \nu_J^2$$

Note that while the variance in the numerator is a rough variance, the denominator is a wavelet variance. This term reflects variation only at the longest time scale. In principle it is stripped of all shorter-term variation. It is in this respect preferable to simply using a variance based on long-term price differences.

There is a long tradition of variance ratios in empirical market microstructure (Amihud and Mendelson (1987); Barnea (1974); Hasbrouck and Schwartz (1988)).[7] Microstructure effects are generally thought to induce transitory mispricing, which generally inflates short-term variances relative to

---

[7] Return variance ratios are also used more broadly in economics and finance to characterize deviations from random-walk behavior over longer horizons (Charles and Darné (2009); Faust (1992); Lo and MacKinlay (1989)).

long-term variances. Ratios constructed from wavelet variances give a more precise and nuanced characterization.

The wavelet covariance between two processes is defined analogously to the wavelet variance. Of particular importance is the covariance between the bid and offer, denoted $\nu^2_{bid,offer,j}$. The wavelet correlation, denoted $\rho_{bid,offer,j} = \nu^2_{bid,offer,j} \Big/ \sqrt{\nu^2_{bid,j}\nu^2_{offer,j}}$, is used to assess the extent to which the bid and offer co-move at different time scales.

Percival and Walden characterize the asymptotic distributions of wavelet variance estimates. By most standards, the number of observations in the present application is more than sufficient to rely on asymptotic results. (With a 50 ms observation interval, a six-hour trading day contains 432,000 observations.) The data exhibit, however, bursts of activity, long periods with no changes ("too many zeroes," as some have noted), and other features that suggest convergence to the asymptotic results might be very slow. Accordingly, the present analysis relies on cross-firm means and standard errors of these means.

IV. A cross-sectional analysis

From a trading perspective, stocks differ most significantly in their general level of activity (volume measured by number of trades, shares or values). The first analysis aims to measure the general level of high frequency quote volatility and to relate the measures to trading activity in the cross-section for a recent sample of firms.

IV.A. Data and sample construction.

The analyses are performed for a subsample of US firms using quote data from April, 2011 (the first month of my institution's subscription.) The subsample is constructed from all firms present on the CRSP and TAQ databases from January through April of 2011 with share codes of 10 or 11, with closing prices between two and one thousand dollars, and with a primary listing on the New York, Amex or NASDAQ exchanges. [8] I compute the daily average dollar volume based on trading in January through

[8] The American Stock Exchange merged with NYSE Euronext in 2008, and was renamed NYSE Amex LLC. In May, 2012, the name was changed to NYSE MKT LLC. For the sake of clarity, it is identified here simply as "Amex".

March, and randomly select 15 firms from each decile. For brevity, reported results are grouped into quintiles.

The U.S. equity market is highly fragmented, but all exchanges post their quotes to the Consolidated Quote System (CQS).[9] The CQ and NBBO files from the NYSE's daily TAQ dataset used here are definitive transcripts of the consolidated activity, time-stamped to the millisecond.[10] A record in the consolidated quote (CQ) file contains the latest bid and offer originating at a particular exchange. If the bid and offer establish the NBBO this fact is noted on the record. If the CQ record causes the NBBO to change for some other reason, a message is posted to another file (the NBBO file). Thus, the NBBO can be obtained by merging the CQ and NBBO files. It can also be constructed (with a somewhat more involved computation) directly from the CQ file. Spot checks confirm that these two approaches are consistent.

Studies involving TAQ data have traditionally used error filters to suppress quotes that appear spurious. Recent daily TAQ data, though, appear to be much cleaner than older samples. In particular, the NBBO construction provided by the NYSE clearly defines what market participants would have perceived. Some quotes present in the CQ file are not incorporated into the NBBO because they are not firm, indicative or otherwise deemed "not NBBO-eligible". Beyond these exclusions, however, I impose no additional filters for the estimates discussed in this section. Error filters are used, however, in the subsequent historical analysis, and will be discussed in greater detail at that point.

Table 1 reports summary statistics. Post-Reg NMS US exchanges have become more similar in structures and trading mechanisms. With respect to listing characteristics, though, differences persist. The NYSE "classic" has the largest proportion of high-volume stocks, NYSE Amex has the smallest, and NASDAQ falls in the middle.

---

[9] At the same time that an exchange sends a quote update to the consolidated system, it can also transmit the update on its own subscriber line. For subscribers this can reduce the delay associated with consolidation and retransmission (which is on the order of about five milliseconds). Thus, while the CQS is a widely-used single-source of market data, it is not the fastest. Moreover, bids and offers with sizes under 100 shares are not reported.

[10] The "daily" reference in the Daily TAQ dataset refers to the release frequency. Each morning the NYSE posts files that cover the previous day's trading. The Monthly TAQ dataset, more commonly used by academics is released with a monthly frequency and contains time stamps in seconds.

Market event counts (trades, quotes, and so forth) display some interesting patterns. There are large numbers of quote records, since one is generated when any market center changes its best bid, best offer, or size at the bid or offer. If the action establishes the bid and offer as the NBBO this fact is noted on the quote record. But if the action causes some other change in the aggregate prices or sizes at the NBBO, an NBBO record is generated. Since many quote records don't induce such a change, there are substantially fewer NBBO records. Finally, many actions might change one of sizes or one side of the quote. Thus, the numbers of NBB and NBO changes are smaller yet.

Volatility and spreads tend to be elevated at the start and end of trading sessions (9:30 to 16:00). To remove the effect of these deterministic effects, I confine the variance estimates to the 9:45 to 15:45 subperiod. The estimates are computed using the maximal overlap Haar transform.[11] Estimates are computed separately for the bid and offer, and then averaged for convenience in presentation. Reported means are generally computed across firms, and the standard errors of these means are constructed in the usual fashion, assuming independence across observations. Due to volatility commonalities, this is likely to bias the standard errors downwards. Market commonalities of all sorts weaken at shorter horizons, however, and this is likely to be especially true of the extremely brief intervals considered here.

To facilitate economic interpretation, I report the time scale variances in three ways: mils ($0.001) per share, basis points relative to average price, and as a short/long-term variance ratio. The mils per share scaling is useful because many trading fees (such as commissions and clearing fees) are assessed on a per share basis. Access fees, the charges levied by exchanges on taker (aggressor) sides of executions are also assessed per share. US SEC Regulation NMS caps access fees at 3 mils ($0.003) per share, and in practice most exchanges are close to this level. Practitioners regard access fees as significant

[11] A maximal overlap transform mitigates alignment problems. In the example discussed in Section III and depicted in Figure 2, the components are always aligned on dyadic boundaries. A maximal overlap transform essentially averages over all possible alignments. The computation of a one-second variance, for example, would involve not only periods aligned exactly on the one-second boundaries (1, 2, 3, …), but also one-second periods aligned on half-second boundaries (1.5, 2.5, 3.5, …). I assume no overlap across days, and discard boundary values affected by wrap-around.

The computations were performed in *Matlab* using the WMTSA package (Cornish (2006)). These routines conform closely to Percival and Walden. Although *Matlab* has its own wavelet toolbox, the data structures and other conventions differ significantly from those of Percival and Walden. I also find the *Mathematica* wavelet functions to be consistent with Percival and Walden.

to the determination of order routing decisions, and this magnitude therefore serves an approximate threshold of economic importance. Basis point scaling is meaningful because most analyses involving investment returns or comparison across firms assume that share normalizations are arbitrary. Variance ratios provide a summary measure of short-term variance inflation relative to what would be expected from a random-walk calibrated to long-term variance.

IV.B. Results

Table 2 summarizes the averages for all time scales of wavelet and rough variances under all three normalizations. For example, a trader facing arrival time uncertainty of 50 milliseconds is exposed to a price risk standard deviation of 0.39 mils per share (from column (1)), or 0.22 bp (from column (2)). The entry in column (3), 3.99, implies that the price risk is roughly four times what would be implied by a random-walk calibrated to longest time scale in the analysis (27.3 minutes). At 400 ms, the rough volatility risk crosses the one mil threshold (1.05, column (2)). At 1,600 ms, it is on the order of one basis point. The variance ratios (columns (3) and (6)) increase monotonically in moving to shorter time scales. Column (7) of Table 2 reports the wavelet correlations between bids and offers. If the bid and offer always moved in lock step, this correlation would be unity at every time scale. At longer time scales this correlation is indeed quite high, but at shorter time scales it is only moderately positive.

Table 3 reports results for a subset of the measures and time scales, but provides standard errors and detail across dollar volume subsamples. Panels A and B report estimates of rough variances in mils per share and basis points, respectively. Stocks in the two lowest dollar volume quintiles have sharply higher short-term volatility. In comparing the two normalizations, it is apparent that volatility in mils per share (Panel A) at the shorter scales is somewhat more stable across dollar volume quintiles than volatility in basis points (Panel B). In moving from lowest to highest quintiles, short time scale volatilities in mils per share approximately double; while volatilities in basis points decline by a factor of four. This decline appears to be mostly caused by the increase in share prices across the quintiles (cf. Table 1). Put another way, it appears that quote volatility is best characterized as a "mils per share" phenomenom, perhaps due to the tick size effects or the use of per-share cost schedules in assessing trading fees.

Table 3 Panel C reports selected rough variance ratios across dollar volume quintiles. Figure 3 graphs the averages wavelet variances. From this graph, for the highest volume quintile, the excess variance seems to be about 30% at the shortest time scales. For the lowest volume quintile, however, the excess is, at ten or above, substantially higher. The wavelet bid-offer correlations are reported in Table 3 Panel D, and graphed in Figure 4. These also exhibit marked variation across dollar volume. For the highest quintile, they are close to unity at a time scale of 25.6 seconds; for the lowest, the correlation at 27.4 minutes is a modest 0.51. This indicates a pronounced de-coupling of the bid and offer for smaller firms.

Hansen and Lunde note that to the extent that volatility is fundamental, we would expect bid and offer variation to be perfectly correlated, that is, that a public information revelation would shift both prices by the same amount (Hansen and Lunde (2006)). Against this presumption, the short-term correlation estimates are striking. At time scales of 200 ms or lower, the correlation is below 0.7 for all activity quintiles. For the shortest time scales and lower activity quintiles, the correlation is only slightly positive. This suggests that substantial high frequency quote volatility is of a distinctly transient nature.

V. Analysis with truncated time stamps.

The analysis in the preceding section relies on a recent one-month sample of daily TAQ data. For addressing policy issues related to low-latency activity, it would be useful to conduct a historical analysis, spanning the period over which low-latency technology was deployed. Extending the analysis backwards, however, is not straightforward. Millisecond time-stamps are only available in the daily TAQ data from 2006 onwards. Monthly TAQ data (the standard source used in academic research) is available back to 1993 (and the precursor ISSM data go back to the mid-1980s). These data are substantially less expensive than the daily TAQ, and they have a simpler logical structure.

The time stamps on the Monthly TAQ and ISSM datasets are reported only to the second. At first glance this might seem to render these data useless for characterizing sub-second variation. This is unduly pessimistic. It is the purpose of this section to propose, implement and validate an approach for estimating sub-second characteristics of the bid and offer series using the second-stamped data. This is possible because the data generation and reporting process is richer than it initially seems.

Specifically, the usual sampling situation in discrete time series analysis involves either aggregation over periodic intervals (such as quarterly GDP) or point-in-time periodic sampling (such as the end-of-day S&P index). In both cases there is one observation per interval, and in neither case do the data support resolution of components shorter than one interval. In the present situation, however, quote updates occur in continuous time and are disseminated continuously. The one second time-stamps arise as a truncation (or equivalently, a rounding) of the continuous event times. The Monthly TAQ data include all quote records, and it is not uncommon for a second to contain ten or even a hundred quote records.

Assume that quote updates arrive as a Poisson process of constant intensity. If the interval $(0, t)$ contains n updates, then the update times have the same distribution as the order statistics in a sample of n independent random variables uniformly distributed on the interval $(0,t)$, (Ross (1996), Theorem 2.3.1). Within a one-second interval containing n updates, therefore, we can simulate continuous arrival times by drawing n realizations from the standard uniform distribution, sorting, and assigning them to quotes (in order) as the fractional portions of the arrival times. These simulated time-stamps are essentially random draws from true distribution. This result does not require knowledge of the underlying Poisson arrival intensity.

We make the additional assumption that the quote update times are independent of the updated bid and offer prices. (That is, the "marks" associated with the arrival times are independent of the times.) Then all estimates based on the simulated time stamp series constitute draws from their corresponding posterior distributions. This procedure can be formalized in a Bayesian Markov-Chain Monte Carlo (MCMC) framework. To refine the estimates, we would normally make repeated simulations ("sweeps") over the sample, but due to computational considerations and programming complexity, I make only one draw for each CQ record.

It is readily granted that few of the assumptions underlying this model are completely satisfied in practice. For a time-homogeneous Poisson process, interevent durations are independent. In fact, inter-event times in market data frequently exhibit pronounced serial dependence, and this feature is a staple of the autoregressive conditional duration and stochastic duration literature (Engle and Russell (1998); Hautsch (2004)). In NASDAQ Inet data, Hasbrouck and Saar (2011) show that event times exhibit intra-second deterministic patterns. Suboordinated stochastic process models of security prices suggest that transactions (not wall-clock time) are effectively the "clock" of the process (Shephard (2005)).

We can assess the reliability of the randomization approach, however, by a simple test. The time-stamps of the data analyzed in the last section are stripped of their millisecond remainders. New millisecond remainders are simulated, the random-time-stamped data are analyzed, and we examine the correlations between the two sets (original and randomized) of estimates. Let $\nu^2_{bid,i,j,d}$ denote the bid wavelet variance estimate for firm *i* on day *d* at level *j* based on the original time stamps, and let $\tilde{\nu}^2_{bid,i,j,d}$ denote the estimate based on the simulated time stamps. (Results are similar for estimates on the offer side.) Table 4, Panel A reports estimates across firms and days of $Corr\left(\nu^2_{bid,i,j,d},\tilde{\nu}^2_{bid,i,j,d}\right)$.The agreement between original and randomized estimates is very high for all time scales and in all subsamples. Even at the time scale of less than fifty ms, the mean correlation is 0.952. At time scales above one second, the agreement is nearly perfect.

Given the questionable validity of some of the assumptions, and the fact that only one draw is made for each second's activity, this agreement might seem surprising. It becomes more reasonable, however, when one considers the extent of averaging underlying the construction of both original and randomized estimates. There is explicit averaging in that each wavelet variance estimate is formed over a sample of roughly six hours. As long as the order is maintained, a small shift in a data point has little impact over the overall estimate.[12]

In parallel fashion let $\nu^2_{bid,offer,i,j,d}$ and $\tilde{\nu}^2_{bid,offer,i,j,d}$ denote the bid-offer covariance estimates based (respectively) on original and simulated millisecond time stamps. Table 4, Panel B reports estimates across firms and days of $Corr\left(\nu^2_{bid,offer,i,j,d},\tilde{\nu}^2_{bid,offer,i,j,d}\right)$. The agreement is somewhat weaker than in the case of the variances. The correlation of under-50 ms components is 0.775 (in the full sample), although this climbs to 0.979 at a time scale of 200 ms. The reason for the relatively poorer performance of the randomized covariance estimates is simply that the wavelet covariance between two series is sensitive to alignment. For a given CQ record, the bid and offer quotes are paired, but in a typical record sequence the NBB and NBO are not changed in the same record. When a bid change is shifted even by a small amount relative to the offer, the inferred pattern of co-movement is distorted.

Across dollar volume quintiles, the correlations generally improve for all time scales. This is true for both wavelet variances and covariances, but is more evident in the latter. This is a likely consequence of the greater incidence, in the higher quintiles, of multiple quote records within the same second.

---

[12] Also, inherent in the wavelet transformation is an (undesirable) averaging across time scales known as leakage, wherein the variance at one time scale affects to a small degree the estimate at neighboring time scale (Percival and Walden, p. 303).

Specifically, for a set of *n* draws from the uniform distribution, the distribution of any order statistic tightens as *n* increases. (For example, the distribution of the first order statistic in a sample of five hundred is tighter than the distribution of the first order statistic in a sample of one.) Essentially, an event time can be located more precisely within the second if the second contains more events. This observation will have bearing on the analysis of historical samples with varying numbers of events.

In working with Monthly TAQ data, Holden and Jacobsen (2012, HJ) suggest assigning sub-second time stamps by evenly-spaced interpolation. If there is one quote record in the second, it is assigned a millisecond remainder of 0.500 seconds; if two records, 0.333 and 0.667 seconds, and so on. HJ show that interpolation yields good estimates of effective spreads. It is not, however, equivalent to the present approach. Consider a sample in which each one-second interval contains one quote record. Even spacing places each quote at its half-second point. As a result, the separation between each quote is one second. For example, a sequence of second time stamps such as 10:00:01, 10:00:02, 10:00:03 … maps to 10:00:01.500, 10:00:02.500, 10:00:03.500, and so on. The interpolated time stamps are still separated by one second, and therefore the sample has no information regarding sub-second components. In contrast, a randomized procedure would sweep the space of all possibilities, including 10:00:01.999, 10:00:02.000, …, which provides for attribution of one-millisecond components. Of course, as the number of events in a given one-second interval increases, the two approaches converge: the distribution of the *kth* order statistic in a sample of *n* uniform observations collapses around its expectation, $k/(n+1)$ as *n* increases.[13]

[13] For one class of time-weighted statistics in this setting, interpolated time stamps lead to unbiased estimates. Consider a unit interval where the initial price, p_0, is known, and there are *n* subsequent price updates $p_i, i=1,\ldots,n$ at occurring at times $0<t_1<\cdots<t_n<1$. The time-weighted average of any price function $f(p)$ is $Avg^{TW}=\sum_{i=0}^{n} f(p_i)(t_{i+1}-t_i)$, where $t_0\equiv 0$ and $t_{n+1}\equiv 1$. Assuming a time-homogeneous Poisson arrival process, the t_i are distributed (as above) as uniform order statistics. This implies $Et_i=i/(n+1)$, the linear interpolated values. If the marks (the p_i) are distributed independently of the t_i, $E\left[Avg^{TW}\right]=(n+1)^{-1}\sum_{i=0}^{n} f(p_i)$. This result applies to time-weighted means of prices and spreads (assuming simultaneous updates of bids and offers). It also applies to wavelet transforms and other linear convolutions. It does not apply to variances (or wavelet variances), however, which are nonlinear functions of arrival times.

VI. Historical evidence

This section describes the construction and analysis of variance estimates for a sample of US stocks from 2001 to 2011. In each year, I construct variance estimates for a single representative month (April) for a subsample of firms.

The period covers significant changes in market structure and technology. Decimalization had been mandated, but was not completely implemented by April, 2001. Reg NMS was proposed, adopted, and implemented.[14] Dark trading grew over the period. Market information and access systems were improved, and latency emerged as a key concern of participants. The period also includes many events related to the financial crisis, which are relatively exogenous to equity market structure.

The regulatory and technological shifts over the period caused changes in the fundamental nature of bid and offer quotations. Markets in 2001 were still dominated by what would later be called "slow" procedures. Quotes were often set manually. Opportunities for automated execution against these quotes were few (cf. the NYSE's odd-lot system, and NASDAQ's Small Order Execution System). Trade-through protection was limited and weakly enforced. Quotes for 100 shares or less were not protected. With the advent of Reg NMS, the bids and offers became much more accessible (for automated execution). These considerations are important in interpreting the results that follow.

VI.A. Data

The data for this phase of the analysis are drawn from CRSP and *Monthly* TAQ datasets. The sample selection procedure in each year is essentially identical to that described for the 2011 cross-sectional sample. In each year, from all firms present on CRSP and TAQ in April, with share codes in (10 and 11), and with primary listings on the NYSE, Amex and NASDAQ exchanges, I draw fifteen firms from each dollar trading volume decile.[15] Quote data are drawn from TAQ.

Table 5 reports summary statistics. The oft-remarked increase in the intensity of trading activity is clearly visible in the trends for median number of trade and quote records. From 2001 to 2011, the

---

[14] Reg NMS was proposed in February, 2004) and adopted in June 2005 with an effective date of August 2005. It was implemented in stages, mostly over 2006.

[15] As of April, 2001, NASDAQ had not fully implemented decimalization. For this year, I do not sample from stocks that traded in sixteenths.

average annual compound growth rate is about 25% percent for trades, and about 36% for quotes. As described in the last section, all of a firm's quote records in a given second are assigned random, but order preserving, millisecond remainders. The NBBO is constructed from these quote records. This yields a NBBO series with (simulated) millisecond time stamps. The 2011 numbers differ slightly from those reported in Table 1. These differences are a consequence of the randomization, and (as will be indicated) the use of error filters.

Prior to the construction of the NBBO the bid and offer are filtered for extreme values. The following quotes (bids or offers) are eliminated: those with zero size and/or zero price; those quotes priced at 20% or lower of the smallest closing price reported on CRSP in the month; those priced at 500% or higher of highest closing price. Quotes that crossed the market are only eliminated if the crossing is a dollar or more, or more than 10 percent of the midpoint price. Other filters use the previously prevailing bid and offer midpoint as a benchmark. For stocks priced at ten dollars or less, the bid and offer has to be within forty percent of the benchmark; for stocks between ten and one hundred dollars, the cutoff is twenty percent; for stocks between one hundred and 250, ten percent; above 250, five percent.[16] These filters do not eliminate all suspicious bids and offers, a point to which the discussion will subsequently return.

VI.B. Results

In analyzing 2001-2011, it is best to begin with the wavelet variance ratios. By construction they are normalized with respect to long-term variance, and over this period there are large swings in market-wide long-term volatility (evident from a cursory examination of the VIX). These would be expected to affect the short term variances as well. Table 6 Panel A reports the mean normalized wavelet variances for shorter time scales in the analysis. As in the 2011 sample, there is substantial variance inflation relative to the random-walk in all years. Perhaps surprisingly, though, the excess variance is high in all years, including the early years of the decade. The pattern does not suggest an increasing trend.

---

[16] The error filters are applied uniformly for the Monthly TAQ data in all years 2001-2011. For 2011 this causes a small apparent discrepancy in the counts for NBB and NBO changes, between Tables 1 and 5. The inputs to Table 5 are filtered, and hence have slightly fewer NBB and NBO changes relative to the unfiltered inputs to Table 1.

Given the recent media attention devoted to low-latency activity and the undeniable growth in quote volume, the absence of a strong trend in quote volatility seems surprising. There are several possible explanations. In the first place, "flickering quotes" drew comment well before the start of the sample, in an era when quotes were dominated by human market makers (Harris (1999); U.S. Commodities Futures Trading Commission Technology Advisor Committee (2001)). Also an artifact of this era is the specialist practice of "gapping" the quotes to indicate larger quantities at worse prices (Jennings and Thirumalai (2007)). In short, the quotes may have in reality been less unwavering than popular memory holds. The apparent discrepancy between quote volatility and quote volume can be explained by appealing to the increase in market fragmentation and consequent growth in matching quotes.

Bid-offer plots for firm-days in each year that correspond to extreme realizations of the variances exhibit an interesting pattern. In later years, these outlier plots tend to resemble the initial AEPI example, with rapid oscillations of relatively low amplitude. In the earlier years, they are more likely to feature small number of prominent spikes associated with a sharply lower bid or elevated offer that persists for a minute or less.

As an example, Figure 5 (Panel A) depicts the NBBO for PRK (Park National Corporation, Amex-listed) on April 6, 2001. At around 10:00 there is a downward spike in the NBB. Shortly after noon there is a sharp drop in the NBB of roughly three dollars and a sharp rise in the NBO of about one dollar. To better document this behavior, Table 7 details the CQ records in the vicinity of the noon episode. There are multiple exchanges active in the market, but Amex (*A*) is the apparent price leader. At 12:02:22, *A* establishes the NBB at 86.74. At 12:03:11, *A* bids 83.63, exposing the previous *T* (NASDAQ) bid of 86.68 as the new NBB. At 12:03:16, *T* backs off, leaving *A*'s 83.63 as best. Within half a minute, however, the NBB is back at 86.50. The lower bid is not marketed by any special mode flag. It is not a penny ("stub") bid. The size of the bid at two (hundred shares) is typical for the market on that day. A similar sequence of events sends the NBO up a dollar for about one second.

These quotes are not so far off the mark as to be clearly erroneous. We must nevertheless question whether they were "real"? Did they reliably indicate the consensus market values at those instances? Were they accessible for execution? Were they truly the best in the market? There were no trades between 11:38 and 12:13, but if a market order had been entered, would it in fact have been executed at the NBBO?[17] These are meaningful questions because they bear directly on market quality. Ultimately, though, the record is unlikely to provide clear answers. The US equity market in 2001 reflected a blend of human and automated mechanisms, practices and conventions that defies detailed description even at a distance of only twelve years.

Discerning whether or not quote volatility increased over the period, therefore, requires that we sharpen the question. The quote volatility in the initial AEPI example is of high frequency, but low amplitude. This is visually distinct from the spikes of high frequency and high amplitude found in PRK. The latter is sometimes called "pop" noise, in reference to its sound in audio signals (Walker (2008)). As in the de-noising of audio signals, the goal is to remove the pops from the signals of lower amplitude. The wavelet literature has developed many denoising approaches (see Percival and Walden, Gençay et al, and Walker). When the stochastic properties of the noise and signal processes are known, optimal methods can often be established. In the present case, though, I adopt a simpler method.

As indicated in Section III and Figure 2, wavelet transforms facilitate the direct computation of smooth and rough components. This process, known as multiresolution analysis, isolates components at different time scales. As an example, Panel B of Figure 5 plots the rough component of the PRK bid at a time scale of 51.2 seconds. It is zero mean by construction, and the spikes are cleanly resolved. On the principle that high frequency quoting (as in the AEPI example) should not be substantially larger than the bid-offer spread in magnitude, I set acceptance bands at $\pm Min\left(1.5\times\left(average\ spread\right),\$0.25\right)$. The minimum of \$0.25 is set to accommodate stocks with very tight spreads. For PRK, the bands are approximately ±\$0.33, and they are indicated in the figure by horizontal black lines. Values lying outside of the band are set to the band limits. This clips the high-amplitude peaks, while leaving the low-amplitude components, some of which are highly oscillatory, untouched. The signal (bid or offer) is reconstituted using the clipped rough, and analysis proceeds on this denoised signal. I recompute all estimates for all firms using the denoised bids and offers.

[17] The Amex (like the NYSE) had specialists in 2001. Specialists generally had affirmative price continuity obligations that would have discouraged (though not expressly forbidden) trades occurring at prices substantially different from those prevailing immediately before and immediately after. A broker-dealer, however, would not have been subject to this restriction.

Table 6 Panel B reports the wavelet variance ratios for the denoised quotes. The results are striking. In the early years, the variance ratios computed from the denoised quotes are much lower than those computed from the raw data. In later years, however, the reduction associated with the denoising is small. For the 200 ms variance ratio, for example, the 2001 drop is from 5.28 (for the raw quotes) to 1.56 (for the denoised quotes), but the 2011 value only drops from 3.74 to 3.57. These results are consistent with the view that over the decade, the nature of quote volatility, if not the overall level, has changed. In the early years, the volatility was of relatively high amplitude but non-oscillatory. It is removed by the pop-denoising procedure. The procedure does not attenuate the low-amplitude highly oscillatory components, however, which drive quote volatility in the later years. The difference between the raw and denoised ratios generally declines throughout the decade, but the largest drops occur around the Reg NMS period.

To further examine historical patterns, Table 8 presents rough wavelet volatilities in mils per share (Panel A) and rough variance ratios (Panel B). These estimates are based on bids and offers that are denoised as described above. For the rough volatilities there is substantial variation across years, but at subsecond time scales (800 ms and under) the highest values are generally found in 2009 and 2010. The 2011 values, however, are low. The rough variance ratios (Panel B) appear to exhibit a somewhat clearer pattern. The subsecond ratios are generally low in the first years (2001 and 2002), and high in the last years (2010 and 2011). The highest values, however, are found in the middle years of 2004-2006. The pattern can be partially explained by the long-term (fundamental) wavelet volatility used in the denominator of the ratios. Over the period, the behavior of this series closely follows the long-term rough volatility reported in the last row of Panel A, where the middle years are relatively low.

In summary, neither the rough volatilities nor the variance ratios display the clear trend that might be expected from ongoing enhancements to trading technology. Quote volatility almost certainly has not climbed in equal measure with quote traffic.

VII. Discussion

From an economic perspective, high frequency quote volatility is connected most closely to other high frequency and low latency phenomena in modern markets. From a statistical perspective, it is connected to volatility modeling. I discuss both in turn.

VII.A. High frequency quoting and high frequency trading

Most definitions of algorithmic and high frequency trading encompass many aspects of market behavior (not just executions), and would be presumed to cover quoting as well.[18] Executions and quotations are nevertheless very different events. It is therefore useful to consider their relation in the high frequency context.

Quote volatility is not necessarily associated with high frequency executions. One can envision regimes where relatively stable quotes are hit intensively when fundamental valuations change, and periods (such as Figure 1) where frenetic quoting occurs in the absence of executions. Nevertheless, the same technology that makes high frequency executions possible also facilitates the rapid submission, cancellation and repricing of the nonmarketable orders that define the bid and offer. One might expect this commonality of technology to link the two activities in practice.

Executions are generally emphasized over quotes when identifying agents as high frequency traders. For example, Kirilenko, Kyle, Samadi and Tuzun (2010) select on high volume and low inventory. The low inventory criterion excludes institutional investors who might use algorithmic techniques to accumulate or liquidate a large position. The NASDAQ HFT dataset uses similar criteria (Brogaard (2010); Brogaard, Hendershott and Riordan (2012)). Once high frequency traders are identified, their executions and the attributes of these executions lead to direct measures of HF activity in panel samples.

[18] A CFTC draft definition reads: “High frequency trading is a form of automated trading that employs: (a) algorithms for decision making, order initiation, generation, routing, or execution, for each individual transaction without human direction; (b) low-latency technology that is designed to minimize response times, including proximity and co-location services; (c) high speed connections to markets for order entry; and (d) high message rates (orders, quotes or cancellations)” (U.S. Commodities Futures Trading Commission (2011)).

In some situations, however, identifications based on additional, non-trade information are possible. Menkveld (2012) identifies one Chi-X participant on the basis of size and prominence. The Automated Trading Program on the German XETRA system allows and provides incentives for designating an order as algorithmic (Hendershott and Riordan (2012)). Other studies analyze indirect measures of low-latency activity. Hendershott, Jones and Menkveld (2011) use NYSE message traffic. Hasbrouck and Saar (2011) suggest strategic runs (order chains) of cancel and replace messages linked at intervals of 100 ms or lower.

Most of these studies find a positive association between low-latency activity and market quality. Low-latency activity, for example, tends to be negatively correlated with as posted and effective spreads, which are inverse measures of market quality. Most also find a zero or negative association between low-latency activity and volatility, although the constructed volatility measures usually span intervals that are long relative to those of the present paper. With respect to algorithmic or high frequency activity: Hendershott and Riordan (2012) find an insignificantly negative association with the absolute value of the prior 15-minute return; Hasbrouck and Saar (2011) find a negative association with the high-low difference of the quote midpoint over 10-minute intervals; Brogaard (2012) finds a negative relation with absolute price changes over intervals as short as ten seconds.

The time-scaled variance estimates used here clearly aim at a richer characterization of volatility than the high/low or absolute return proxies used in the studies above. The present study does not, on the other hand, attempt to correlate the variance measures with intraday proxies for high frequency trading. One would further suspect, of course, that the ultimate strategic purpose of high frequency quoting is to facilitate a trade or to affect the price of a trade. The mechanics of this are certainly deserving of further research.

The discussion in Section II associates short-term quote volatility with price uncertainty for those who submit marketable orders, use dark mechanisms that price by reference, or face monitoring difficulties. From this perspective, quote volatility is an inverse measure of market quality. Although the present study finds evidence of economically significant and elevated quote volatility, it does not establish a simple connection to technological trends associated with low latency activity.

VII.B. High frequency quoting and volatility modeling

Security prices at all horizons are a mix of integrated and stationary components. The former are usually identified with persistent fundamental information innovations; the latter, with transient microstructure effects. The former are important to long-term hedging and investment; the latter, to trading and market-making. The dichotomy is sometimes reflected in different statistical tools and models.

Between the two approaches, the greatest common concerns arise in the analysis of realized volatility (Andersen, Bollerslev, Diebold and Ebens (2001); Andersen, Bollerslev, Diebold and Labys (2003a); Andersen, Bollerslev, Diebold and Labys (2003b)). RVs are calculated from short-term price changes. They are useful as estimates of fundamental integrated volatility (IV), and typically serve as inputs to longer-term forecasting models. RVs constructed directly from trade, bid and offer prices are typically noisy, however, due to the presence of microstructure components. Local averaging moderates these effects. (See Hansen and Lunde (2006) and the accompanying comments. Other approaches are discussed in Ait-Sahalia, Mykland and Zhang (2011); Zhang (2006); Zhang, Mykland and Aït-Sahalia (2005).) There is a methodological connection here, in that long-term wavelet variances are computed from short-term averages, much like the pre-averaged inputs to realized volatility.

The present study draws on several themes in the RV literature. The volatility ratio plots in Figure 3 serve a purpose similar to the volatility signature plots introduced by Fang (1996) and used in Andersen, Bollerslev, Diebold and Ebens (2002) and Hansen and Lunde (2006). Hansen and Lunde also articulate the connection between bid-offer comovement and fundamental volatility: since the bid and offer have economic fundamentals in common, divergent movements must be short-term, transient, and unconnected to fundamentals.

One strand in the RV literature emphasizes analysis of multiple time-scales. Zhang, Mykland and Aït-Sahalia (2005) posit a framework consisting of a Brownian motion with time-varying parameters, $dX_t = \mu_t dt + \sigma_t dz,$ and a discretely-sampled noisy observation process, $Y_{t_i} = X_{t_i} + \varepsilon_{t_i}$. The Y_{t_i} are viewed as transaction prices, and ε_{t_i} constitute i.i.d. microstructure noise. The objective is estimation of the integrated volatility $\int \sigma_t^2 dt$ over a sample. Zhang et al propose a two-scale variance estimator in which a long-scale estimate is corrected for bias with an adjustment based on properties of the noise estimated at a short scale. While the present analysis also features multiple time scales, there are major differences in the perspective. In the present situation, execution price risk is caused by volatility in the observed process (the quote, not the underlying latent value, X_t); the quote process is right-continuous (and continuously observable); the noise is not necessarily i.i.d. (cf. the AEPI episodes in Figure 1); and, the noise is possibly correlated with the X_t increments.

The paper also departs from the RV literature in other respects. The millisecond time scales employed in this paper are several orders of magnitude shorter than those typically encountered. Most RV studies also focus on relatively liquid assets (index securities, Dow-Jones stocks, etc.). The low-activity securities included in the present paper's samples are important because, due to their larger spreads and fewer participants, they are likely to exhibit relatively strong, persistent and distinctive microstructure-related components.

VIII. Conclusion and outstanding questions

High frequency volatility in the bid and offer quotes induces price risk for agents who experience delay in communicating with the market. The risk may be quantified as the price variance over the interval of delay, relative to the average price over the interval. This volatility degrades the informational value of the quotes. Furthermore, because the bid and offer are often used as reference prices for dealer trades against customers, the volatility increases the value of a dealer's look-back option and exacerbates monitoring problems for customers, exchanges, and regulators.

This study is a preliminary analysis of short-term quote volatility in the US equity market. Estimates of sub-second high frequency variance for the National Best Bid and Offer (NBBO) are well in excess of what would be expected relative to random-walk volatility estimated over longer intervals. The excess volatility is more pronounced for stocks that have lower average activity. The sub-second volatility is comparable in magnitude to access fees and other transaction expenses. Furthermore, the correlations between bids and offers at these time scales are positive, but low. That the bid and offer are not moving together also suggests that the volatility is not fundamental.

The paper proposes a simulation approach to measuring millisecond-level volatility in US equity data (like the Monthly TAQ) that possess all quote records, but are time-stamped only to the second. In data time-stamped to the millisecond I compare two sets of estimates: one set based on the original time-stamps; the other based on simulated time stamps. I find high correlations between the two estimates, establishing the reliability of the simulation procedure.

With these results, the paper turns to a longer US historical sample, 2001-2011, with one-second time-stamps. Despite the current public scrutiny of high frequency trading, the rapid growth in the number of quote records, and the presumption that low-latency technology is a new and recent phenomenon, the excess short-term quote volatility found in the 2011 data also appears in earlier years.

The nature of the volatility has changed over the decade, however. Prior to Reg NMS, the volatility appears attributable to spikes associated with bids and offers that are neither clearly erroneous nor reliably valid. Post Reg NMS, the volatility is more attributable to oscillatory low-amplitude changes: rapid movements not substantially larger than the spread. The generating mechanisms and strategic purposes of these movements are deserving of further study.

Appendix: Deviations about averages of random walks

Consider a price series that evolves as $p_t = p_{t-1} + u_t$ where u_t is a white-noise process with unit variance. Without loss of generality, we initialize $p_0 = 0$ and consider the mean-squared deviations about the mean over the first n observations.

$$MSD(n) = \frac{1}{n}\sum_{i=1}^{n} p_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} p_i\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{i} u_i\right)^2 - \left(\frac{1}{n}\sum_{i=1}^{n}\left(\sum_{j=1}^{i} u_i\right)\right)^2$$

Taking expectations (noting that $Eu_i u_j = 1$ for $i = j$, and zero otherwise) and simplifying the sums gives

$$\sigma_n^2 \equiv E\left(MSD(n)\right) = \frac{n+1}{2} - \frac{(n+1)(2n+1)}{6n} = \frac{n^2-1}{6n}$$

For the sequence of averaging periods $n_j = n_0 2^j$ for $j = 0,1,2,\ldots$, the corresponding sequence of variances is

$$\sigma_j^2 = \frac{4^j n_0^2 - 1}{3n_0 2^{j+1}}$$

In moving from $j-1$ to j the incremental change in variance (also known as the wavelet variance) is

$$\nu_j^2 = \sigma_j^2 - \sigma_{j-1}^2 = \frac{4^j n_0^2 + 2}{3n_0 2^{j+2}}$$

We now reinterpret these results in a slightly expanded framework. Suppose that the original time subscript t indexes periods of Δ time units (“milliseconds”) and that the variance per unit time of the u_t process is σ_u^2. Let M denote the averaging period measured in units of time, and correspondingly, $M_j = M_0 2^j$ for $j = 0,1,\ldots$ Then the rough and wavelet variances become

$$\sigma_j^2 = \frac{(4^j M_0^2 - \Delta^2)\sigma_u^2}{3M_0 2^{j+1}} \text{ and } \nu_j^2 = \frac{(4^j M_0^2 + 2\Delta^2)\sigma_u^2}{3M_0 2^{j+2}}.$$

In the continuous time limit, as $\Delta \to 0$, that $\sigma_j^2 = 2^{j-1} M_0 \sigma_u^2 / 3$ and $\nu_j^2 = 2^{j-2} M_0 \sigma_u^2 / 3$. These results suffice to define and characterize the variances considered in the paper.

References

Ait-Sahalia, Yacine, Per A. Mykland, and Lan Zhang, 2011, Ultra high frequency volatility estimation with dependent microstructure noise, *Journal of Econometrics* 160, 160-175.

Amihud, Yakov, and Haim Mendelson, 1987, Trading mechanisms and stock returns: An empirical investigation, *Journal of Finance* 42, 533-553.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, The distribution of realized stock return volatility, *Journal of Financial Economics* 61, 43-76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2003a, The distribution of realized exchange rate volatility (vol 96, pg 43, 2001), *Journal of the American Statistical Association* 98, 501-501.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2003b, Modeling and forecasting realized volatility, *Econometrica* 71, 579-625.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and H. Ebens, 2002, Great realizations, *Risk* 13, 105-108.

Atsalakis, George S., and Kimon P. Valavanis, 2009, Surveying stock market forecasting techniques - Part II: Soft computing methods, *Expert Systems with Applications* 36, 5932-5941.

Barnea, Amir, 1974, Performance evaluation of New York Stock Exchange specialists, *Journal of Financial and Quantitative Analysis* 9, 511-535.

Biais, Bruno, Thierry Foucault, and Sophie Moinas, 2012, Equilibrium High frequency Trading, *SSRN eLibrary*.

Brogaard, Jonathan, 2010, High frequency trading and its impact on market quality, *SSRN eLibrary*.

Brogaard, Jonathan, 2012, High frequency trading and volatility, *SSRN eLibrary*.

Brogaard, Jonathan, Terrence J. Hendershott, and Ryan Riordan, 2012, High frequency trading and price discovery, *SSRN eLibrary*.

Charles, Amélie, and Olivier Darné, 2009, Variance-ratio tests of random walk: an overview, *Journal of Economic Surveys* 23, 503-527.

Cornish, Charlie, 2006, The WMTSA wavelet toolkit for MATLAB (Department of Atmospheric Sciences, University of Washington, Seattle).

Dacorogna, Michel M., Ramazan Gencay, Ulrich A. Muller, Richard B. Olsen, and Olivier B. Pictet, 2001. *High-Frequency Finance* (Academic Press, New York).

Duffie, Darrell, 2010, Presidential Address: Asset Price Dynamics with Slow-Moving Capital, *The Journal of finance* 65, 1237-1267.

Elder, John, and Hyun J. Jin, 2007, Long memory in commodity futures volatility: A wavelet perspective, *Journal of Futures Markets* 27, 411-437.

Elkins-McSherry, 2012, Methodology, https://www.elkinsmcsherry.com/EM/methodology.html, accessed on November 2, 2012.

Engle, Robert F., and Jeffrey R. Russell, 1998, Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica* 66, 1127-1162.

Engle, Robert, Robert Ferstenberg, and Jeffrey Russell, 2012, Measuring and modeling execution cost and risk, *Journal of Portfolio Management* 38, 14-28.

Fan, Jianqing, and Yazhen Wang, 2007, Multi-scale jump and volatility analysis for high-frequency financial data, *Journal of the American Statistical Association* 102, 1349-1362.

Fan, Yanqin, and Ramazan Gencay, 2010, Unit root tests with wavelets, *Econometric Theory* 26, 1305-1331.

Fang, Y., 1996, Volatility Modeling and Estimation of High-Frequency Data with Gaussian Noise, (MIT Sloan School).

Faust, Jon, 1992, When are variance ratio tests for serial dependence optimal?, *Econometrica* 60, 1215-1226.

Gençay, Ramazan, Faruk Selçuk, Nikola Gradojevic, and Brandon Whitcher, 2010, Asymmetry of information flow between volatilities across time scales, *Quantitative Finance* 10, 895-915.

Gençay, Ramazan, Faruk Selçuk, and Brandon Whitcher, 2005, Multiscale systematic risk, *Journal of International Money and Finance* 24, 55-70.

Gençay, Ramazan, Frank Selçuk, and Brandon Whitcher, 2002. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics* (Academic Press (Elsevier), San Diego).

Gencay, Ramazan, and Danielle Signori, 2012, Multi-scale tests for serial correlation, (Simon Fraser University).

Hansen, Peter R., and Asger Lunde, 2006, Realized variance and market microstructure noise, *Journal of Business & Economic Statistics* 24, 127-161.

Harris, Lawrence E., 1999, Trading in pennies: a survey of the issues, (Marshall School, University of Southern California).

Hasbrouck, Joel, and Gideon Saar, 2011, Low-latency trading, (SSRN eLibrary).

Hasbrouck, Joel, and Robert A. Schwartz, 1988, Liquidity and execution costs in equity markets, *Journal of Portfolio Management* 14, 10-16.

Hautsch, Nikolaus, 2004. *Modelling Irregularly Spaced Financial Data: Theory and Practice of Dynamic Duration Models* (Springer).

Hendershott, Terrence J., and Ryan Riordan, 2012, Algorithmic trading and the market for liquidity, *SSRN eLibrary*.

Hendershott, Terrence, Charles M. Jones, and Albert J. Menkveld, 2011, Does algorithmic trading improve liquidity?, *Journal of Finance* 66, 1-33.

Høg, Esben, and Asger Lunde, 2003, Wavelet estimation of integrated volatility, (Aarhus School of Business, Department of Information Science).

Holden, Craig W., and Stacey E. Jacobsen, 2012, Liquidity measurement problems in fast, competitive markets: expensive and cheap solutions, (Kelley School, University of Indiana).

Hsieh, Tsung-Jung, Hsiao-Fen Hsiao, and Wei-Chang Yeh, 2011, Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm, *Applied Soft Computing* 11, 2510-2525.

In, Francis, and Sangbae Kim, 2006, The hedge ratio and the empirical relationship between the stock and futures markets: a new approach using wavelet analysis, *The Journal of Business* 79, 799-820.

Jarrow, Robert A., and Philip Protter, 2011, A Dysfunctional Role of High Frequency Trading in Electronic Markets, *SSRN eLibrary*.

Jennings, Robert H., and Ramabhadran S. Thirumalai, 2007, Advertising for liquidity on the New York Stock Exchange, (Kelley School, University of Indiana).

Kirilenko, Andrei A., Albert S. Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2010, The Flash Crash: the impact of high frequency trading on an electronic market, *SSRN eLibrary*.

Lo, Andrew W., and A. Craig MacKinlay, 1989, The size and power of the variance ratio test in finite samples, *Journal of Econometrics* 40, 203-238.

Menkveld, Albert J., 2012, High frequency trading and the new market makers, *SSRN eLibrary*.

Nason, Guy P., 2008. *Wavelet Methods in Statistics with R* (Springer Science+Business Media, LLC, New York).

Pagnotta, Emiliano S., 2009, Trading strategies at optimal frequencies, (Stern School).

Parlour, Christine A., and Duane Seppi, 2008, Limit order markets: a survey, in Anjan V. Thakor, and Arnaud W. A. Boot, eds.: *Handbook of Financial Intermediation and Banking* (Elsevier, Amsterdam).

Parlour, Christine A., and Duane J. Seppi, 2003, Liquidity-based competition for order flow, *Review of financial studies* 16, 301-303.

Percival, Donald B., and Andrew T. Walden, 2000. *Wavelet Methods for Time Series Analysis* (Cambridge University Press, Cambridge).

Perold, Andre, 1988, The implementation shortfall: paper vs. reality, *Journal of Portfolio Management* 14, 4-9.

Ramsey, James B., 1999, The contribution of wavelets to the analysis of economic and financial data, *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 357, 2593-2606.

Ramsey, James B., 2002, Wavelets in economics and finance: past and future, *Studies in Nonlinear Dynamics and Econometrics* 6.

Ross, Sheldon M., 1996. *Stochastic Processes* (John Wiley and Sons, Inc., New York).

Shephard, Neil, 2005, General introduction, in Neil Shephard, ed.: *Stochastic Volatility* (Oxford University Press, Oxford).

Stoll, H. R., and C. Schenzler, 2006, Trades outside the quotes: Reporting delay, trading option, or trade size?, *Journal of Financial Economics* 79, 615-653.

Teyssière, Gilles, and Patrice Abry, 2007, Wavelet analysis of nonlinear long-range dependent processes: applications to financial time series, in Gilles Teyssière, and Alan P. Kirman, eds.: *Long Memory in Economics* (Springer, Berlin).

U.S. Commodities Futures Trading Commission, 2011, Presentation (Working Group 1, Additional Document), Technology Advisor Committee.

U.S. Commodities Futures Trading Commission Technology Advisor Committee, 2001, Market Access Committee Interim Report.

U.S. Securities and Exchange Commission, 2005, Regulation NMS (Final Rule Release No. 34-51808; June 9, 2005), http://www.sec.gov/rules/final/34-51808.pdf, accessed on.

U.S. Securities and Exchange Commission, 2010, Concept release on equity market structure.

U.S. Securities and Exchange Commission, 2012, Consolidated Audit Trail (Rule 613, Final Rule).

Walker, James S., 2008. *A Primer on Wavelets and Their Scientific Applications* (Chapman and Hall/CRC (Taylor and Francis Group), Boca Raton).

Zhang, Lan, 2006, Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach, *Bernoulli* 12, 1019-1043.

Zhang, Lan, Per A. Mykland, and Yacine Aït-Sahalia, 2005, A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data, *Journal of the American Statistical Association* 100, 1394-1411.

Table 1. Sample Summary Statistics

Source: CRSP and Daily TAQ data, April 2011. The sample is 150 firms randomly selected from CRSP with stratification based on average dollar trading volume in the first quarter of 2011, grouped in quintiles by dollar trading volume. NBB is the National Best Bid; NBO the National Best Offer. Except for the counts (first four rows), all table entries are cross-firm medians.

		Dollar trading volume quintile				
	Full sample	1 (low)	2	3	4	5 (high)
No. of firms	150	30	30	30	30	30
NYSE	47	0	5	7	16	19
Amex	6	2	2	0	1	1
NASDAQ	97	28	23	23	13	10
Avg. daily CT records (trades)	1,331	31	431	1,126	3,478	16,987
Avg. daily CQ records (quotes)	23,928	967	7,706	24,026	53,080	181,457
Avg. daily NBBO records	7,138	328	3,029	7,543	16,026	46,050
Avg. daily NBB changes	1,245	120	511	1,351	2,415	4,124
Avg. daily NBO changes	1,164	103	460	1,361	2,421	4,214
Avg. price (bid-offer midpoint)	$15.62	$4.87	$5.46	$17.86	$27.76	$51.60
Market capitalization of equity, $ Million	$683	$41	$202	$747	$1,502	$8,739

Table 2. Time scale volatility estimates for US equities in 2011

Estimates of time scale variances and related measures for 150 US firms during April, 2011. The wavelet variances, ν_j^2, are estimates of the price variance at the time scale $\tau_j = 50 \times 2^{j-1}$. The rough variances, σ_j^2, measure cumulative variation at all time scales $\le \tau_j$. For presentation, I report the square-roots of the variances, in mils ($0.001) per share and basis points (0.01%). The wavelet variance ratio is $V_{j,J} = 2^{J-j} \nu_j^2 / \nu_J^2$ where $J = 16$ is the longest time-scale in the analysis, and the rough variance ratio is similarly defined as $VR_{j,J} = 2^{J-j} \sigma_j^2 / \nu_J^2$. For random-walk, both ratios would be unity at all horizons. The bid-offer correlation is the wavelet correlation (correlation between detail components) at the indicated time scale. All entries are cross-firm means. The National Best Bid and Offer are computed from TAQ data; the bid and offer are separately transformed using the Haar basis; the reported variance estimates are averages of the bid and offer variances. The data are time stamped to the millisecond. Prior to transformation, I take the average of the bid or offer over non-overlapping 50 millisecond intervals. Entries for $j = 0$ are variances within the 50 ms intervals.

		Rough variances			Wavelet variances			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Level, j	Time scale	$\sqrt{\sigma_j^2}$ (mils per share)	$\sqrt{\sigma_j^2}$ (basis pts)	Ratio, $VR_{j,J}$	$\sqrt{\nu_j^2}$ (mils per share)	$\sqrt{\nu_j^2}$ (basis pts)	Ratio, $V_{j,J}$	Correlation $\rho_{bid,offer,j}$
	< 50 ms	0.28	0.16	4.22				
1	50 ms	0.39	0.22	3.99	0.27	0.15	3.76	0.32
2	100 ms	0.55	0.31	3.79	0.38	0.21	3.58	0.36
3	200 ms	0.76	0.43	3.53	0.53	0.30	3.27	0.41
4	400 ms	1.05	0.59	3.21	0.73	0.41	2.88	0.44
5	800 ms	1.46	0.83	2.90	1.01	0.57	2.59	0.47
6	1,600 ms	2.02	1.14	2.64	1.40	0.79	2.38	0.51
7	3.2 sec	2.80	1.58	2.40	1.94	1.09	2.16	0.55
8	6.4 sec	3.90	2.18	2.12	2.71	1.49	1.84	0.60
9	12.8 sec	5.43	2.99	1.88	3.77	2.04	1.65	0.64
10	25.6 sec	7.54	4.10	1.70	5.23	2.79	1.51	0.69
11	51.2 sec	10.48	5.61	1.54	7.25	3.82	1.39	0.74
12	102.4 sec	14.53	7.68	1.42	10.04	5.22	1.29	0.79
13	3.4 min	20.12	10.51	1.32	13.87	7.14	1.21	0.83
14	6.8 min	27.88	14.40	1.23	19.22	9.78	1.15	0.86
15	13.7 min	38.55	19.70	1.16	26.45	13.33	1.08	0.88
16 (=J)	27.3 min	52.84	26.79	1.08	35.73	17.91	1.00	0.90

Table 3. Time scale volatility estimates for US equities in 2011, across dollar trading volume quintiles.

Estimates of time scale variances and related measures for 150 US firms during April, 2011, for quintiles constructed on dollar trading volume. The wavelet variances, ν_j^2, are estimates of the price variance at the time scale $\tau_j = 50 \times 2^{j-1}$. The rough volatilities, σ_j^2, measure cumulative variation at all time scales $\le \tau_j$. For presentation, I report the rough volatilities (square-roots of the rough variances) in mils ($0.001) per share (Panel A) and basis points (0.01%, Panel B).The rough wavelet variance ratio is $VR_{j,J} = 2^{J-j} \sigma_j^2 / \nu_J^2$ where $J = 16$ is the longest time-scale in the analysis (Panel C). For a random-walk $VR_{j,J}$ would be unity at all horizons. The bid-offer correlation (Panel D) is the wavelet correlation (correlation between bid and offer components) at the indicated time scale. Table entries are cross-firm means with standard errors in parentheses. The National Best Bid and Offer are computed from TAQ data; the bid and offer are separately transformed using the Haar basis; the reported variance estimates are averages of the bid and offer variances. The data are time stamped to the millisecond. Prior to transformation, I take the average of the bid or offer over non-overlapping 50 millisecond intervals. Entries for $j = 0$ are variances within the 50 ms intervals. Transforms are performed through level $J = 16$; for brevity only a subset of time-scales are reported.

Panel A. Rough volatility, $\sqrt{\sigma_j^2}$, mils per share

			Dollar trading volume quintiles				
Level, j	Time scale	Full sample	1 (low)	2	3	4	5 (high)
0	< 50 ms	0.28	0.15	0.19	0.29	0.37	0.40
		(<0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
1	50 ms	0.39	0.21	0.27	0.40	0.51	0.56
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
3	200 ms	0.76	0.40	0.50	0.76	0.99	1.11
		(0.01)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
5	800 ms	1.46	0.76	0.95	1.45	1.91	2.14
		(0.02)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
7	3.2 sec	2.80	1.46	1.79	2.75	3.70	4.19
		(0.04)	(0.07)	(0.06)	(0.06)	(0.08)	(0.09)
10	25.6 sec	7.54	3.63	4.51	7.05	10.12	12.02
		(0.10)	(0.15)	(0.15)	(0.15)	(0.23)	(0.27)
14	6.8 min	27.88	11.83	15.29	24.76	38.35	47.61
		(0.39)	(0.47)	(0.50)	(0.52)	(0.94)	(1.14)
16	27.3 min	52.84	20.94	28.09	46.87	74.69	90.49
		(0.78)	(0.86)	(0.91)	(1.03)	(2.04)	(2.26)

Table 3. Time scale volatility estimates for US equities in 2011 across dollar trading volume quintiles (continued).

Panel B. Rough volatility, $\sqrt{\sigma_j^2}$, basis points

			Dollar trading volume quintiles				
Level, *j*	Time scale	Full sample	1 (low)	2	3	4	5 (high)
0	< 50 ms	0.16	0.26	0.21	0.14	0.11	0.08
		(<0.01)	(0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)
1	50 ms	0.22	0.37	0.29	0.20	0.16	0.11
		(<0.01)	(0.02)	(0.01)	(<0.01)	(<0.01)	(<0.01)
3	200 ms	0.43	0.70	0.56	0.38	0.31	0.22
		(0.01)	(0.03)	(0.01)	(0.01)	(<0.01)	(<0.01)
5	800 ms	0.83	1.34	1.07	0.73	0.59	0.43
		(0.01)	(0.05)	(0.02)	(0.01)	(0.01)	(0.01)
7	3.2 sec	1.58	2.56	2.03	1.41	1.15	0.85
		(0.02)	(0.08)	(0.04)	(0.02)	(0.01)	(0.01)
10	25.6 sec	4.10	6.32	5.17	3.65	3.13	2.42
		(0.04)	(0.17)	(0.09)	(0.06)	(0.04)	(0.03)
14	6.8 min	14.40	20.19	17.85	12.96	11.92	9.64
		(0.14)	(0.52)	(0.35)	(0.22)	(0.16)	(0.15)
16	27.3 min	26.79	35.62	33.19	24.41	23.20	18.38
		(0.27)	(0.95)	(0.72)	(0.43)	(0.35)	(0.30)

Panel C. Rough variance ratio, $VR_{j,J}$

			Dollar trading volume quintiles				
Level, *j*	Time scale	Full sample	1 (low)	2	3	4	5 (high)
0	< 50 ms	4.22	12.72	3.45	2.62	1.76	1.37
		(1.28)	(6.96)	(0.18)	(0.07)	(0.04)	(0.02)
1	50 ms	3.99	12.01	3.23	2.44	1.69	1.35
		(1.25)	(6.81)	(0.16)	(0.06)	(0.04)	(0.02)
3	200 ms	3.53	10.40	2.83	2.20	1.57	1.30
		(1.06)	(5.77)	(0.11)	(0.05)	(0.03)	(0.02)
5	800 ms	2.90	7.82	2.50	2.02	1.43	1.21
		(0.66)	(3.56)	(0.08)	(0.04)	(0.03)	(0.02)
7	3.2 sec	2.40	5.87	2.17	1.82	1.32	1.15
		(0.38)	(2.08)	(0.06)	(0.04)	(0.02)	(0.02)
10	25.6 sec	1.70	3.06	1.70	1.49	1.19	1.17
		(0.12)	(0.64)	(0.04)	(0.03)	(0.02)	(0.02)
14	6.8 min	1.23	1.58	1.24	1.17	1.04	1.16
		(0.02)	(0.12)	(0.03)	(0.03)	(0.02)	(0.02)
16	27.3 min	1.08	1.19	1.08	1.06	1.01	1.06
		(0.01)	(0.06)	(0.03)	(0.02)	(0.02)	(0.02)

Table 3. Time scale volatility estimates for US equities in 2011 across dollar trading volume quintiles (continued).

Panel D. Wavelet bid-offer correlations.

			Dollar trading volume quintiles				
Level, *j*	Time scale	Full sample	1 (low)	2	3	4	5 (high)
1	50 ms	0.32	0.05	0.23	0.31	0.41	0.56
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
3	200 ms	0.41	0.11	0.33	0.42	0.49	0.65
		(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
5	800 ms	0.48	0.15	0.40	0.51	0.56	0.72
		(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
7	3.2 sec	0.55	0.19	0.47	0.59	0.66	0.82
		(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
10	25.6 sec	0.70	0.27	0.61	0.75	0.85	0.95
		(0.01)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
14	6.8 min	0.86	0.44	0.88	0.97	0.99	1.00
		(0.02)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
16	27.3 min	0.90	0.51	0.96	0.99	1.00	1.00
		(0.02)	(0.04)	(0.05)	(0.03)	(0.04)	(0.03)

Table 4. Correlations between statistics based on actual vs. simulated millisecond time stamps.

This table assesses the reliability of wavelet estimates based on simulated millisecond time stamps. For the sample of 150 US firms during April, 2011, I compute two sorts of estimates. Let $\nu^2_{bid,j,i,d}$ denote the bid wavelet variance estimate for firm *i* on day *d* at level *j* based on the original millisecond time stamps. For the alternative estimates, I round the timestamps down to the next second (thereby stripping the millisecond portions) and assign randomly-generated millisecond remainders. Let $\tilde{\nu}^2_{bid,j,i,d}$ denote the estimate from the data with simulated time stamps. Panel A reports the estimated $Corr\left(\nu^2_{bid,i,j,d}, \tilde{\nu}^2_{bid,i,j,d}\right)$ computed over all days and firms in the full sample (or the indicated dollar volume quintile subsample). Results for the offer side are similar and omitted for brevity. Now let $\nu^2_{bid,offer,i,j,d}$ denote the estimate of the wavelet covariance between the bid and offer for firm *i* on day *d* at level *j* based on the original millisecond time stamps; $\tilde{\nu}^2_{bid,offer,i,j,d}$ denotes the corresponding estimate based on the simulated millisecond time stamps. Panel B reports the estimated $Corr\left(\nu^2_{bid,offer,i,j,d}, \tilde{\nu}^2_{bid,offer,i,j,d}\right)$ computed over all days and firms in the full sample (or the indicated dollar volume quintile subsample).

Panel A. Correlations for bid wavelet variance estimates, $Corr\left(\nu^2_{bid,i,j,d}, \tilde{\nu}^2_{bid,i,j,d}\right)$

			Dollar trading volume quintiles				
Level, *j*	Time scale	Full sample	1 (low)	2	3	4	5 (high)
0	< 50 ms	0.952	0.948	0.960	0.958	0.916	0.979
1	50 ms	0.953	0.944	0.952	0.952	0.937	0.982
3	200 ms	0.975	0.965	0.969	0.975	0.977	0.988
5	800 ms	0.994	0.991	0.989	0.995	0.996	0.998
7	3.2 sec	0.999	0.999	0.999	1.000	1.000	1.000
10	25.6 sec	1.000	1.000	1.000	1.000	1.000	1.000
14	6.8 min	1.000	1.000	1.000	1.000	1.000	1.000
16	27.3 min	1.000	1.000	1.000	1.000	1.000	1.000

Panel B. Correlations for bid-offer wavelet covariance estimates, $Corr\left(\nu^2_{bid,offer,i,j,d}, \tilde{\nu}^2_{bid,offer,i,j,d}\right)$

			Dollar trading volume quintiles				
Level, *j*	Time scale	Full sample	1 (low)	2	3	4	5 (high)
0	< 50 ms	0.775	0.333	0.768	0.896	0.919	0.943
1	50 ms	0.900	0.662	0.926	0.965	0.972	0.978
3	200 ms	0.979	0.921	0.986	0.995	0.995	0.998
5	800 ms	0.999	0.998	0.999	1.000	1.000	1.000
7	3.2 sec	1.000	1.000	1.000	1.000	1.000	1.000
10	25.6 sec	1.000	1.000	1.000	1.000	1.000	1.000
14	6.8 min	1.000	1.000	1.000	1.000	1.000	1.000
16	27.3 min	1.000	1.000	1.000	1.000	1.000	1.000

Table 5. Summary statistics for US equities, 2001-2011

From the CRSP file, for each year, 2001-2011 and all stocks present in January through April of that year with share codes equal to 10 or 11, I draw 150 firms in a random sample stratified by dollar trading volume in January through March. NBB is the National Best Bid; NBO, the National Best Offer; CT, Consolidated Trade; CQ, Consolidated Quote. Trade and quote counts are from the Monthly TAQ database (one-second time stamps). Except for the number of firms, table entries are cross-firm medians.

	2001	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
No. firms	146	146	150	150	150	150	150	150	150	150	150
NYSE	108	46	51	44	48	44	55	53	56	54	47
Amex	22	4	11	12	8	15	14	7	5	14	6
NASDAQ	16	96	88	94	94	91	81	90	89	82	97
Avg. daily CT records (trades)	142	122	187	393	425	605	970	1,209	1,790	1,141	1,331
Avg. daily CQ records (quotes)	1,078	534	1,299	3,850	5,828	7,307	12,521	16,328	39,378	23,249	23,928
Avg. daily NBB changes	103	127	203	509	596	761	772	1,144	1,618	1,466	1,210
Avg. daily NBO changes	103	129	213	537	729	751	789	1,119	1,731	1,457	1,126
Avg. price (bid-offer midpoint)	$18.85	$17.94	$14.83	$16.53	$16.10	$21.14	$15.81	$14.01	$10.72	$16.32	$15.62
Market capitalization of equity, $ Million	$745	$302	$189	$345	$325	$411	$480	$405	$316	$478	$683

Table 6. Wavelet variance ratios for US firms, 2001-2011

In each year 2001-2011, 150 US firms are randomly selected from CRSP (stratified by average daily dollar trading volume during the first quarter of the year). Quote records for April are taken from the NYSE Monthly TAQ database. Within each second, quotes are randomly assigned order-preserving millisecond fractional portions. Wavelet variances, ν_j^2, are estimates of the price variance at the time scale. The wavelet variance ratio is $V_{j,J} = 2^{J-j}\nu_j^2 / \nu_J^2$ where $J = 16$ is the longest time-scale in the analysis. For random-walk, the ratio would be unity at all horizons. All entries are cross-firm means. The National Best Bid and Offer are computed from TAQ data; the bid and offer are separately transformed using the Haar basis; variance estimates are formed as the average of the bid and offer variances. Estimates in Panel A are constructed from bids and offers that were filtered for errors, but not otherwise adjusted. Estimates in Panel B are constructed from denoised bids and offers (with short-term peaks clipped).

Panel A. Wavelet variance ratios, $V_{j,J=16}$, computed from raw bids and offers

Level, j	Time scale	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
1	50 ms	5.22	7.16	6.03	10.28	6.69	8.57	6.96	6.06	4.52	7.08	4.70
2	100 ms	5.44	6.58	5.28	9.69	6.51	8.07	6.27	5.38	4.12	6.26	4.32
3	200 ms	5.28	6.28	5.13	9.03	6.22	7.34	5.33	4.64	3.68	5.40	3.74
4	400 ms	4.59	5.23	5.00	8.16	5.75	6.30	4.25	3.84	3.21	4.53	3.07
5	800 ms	3.12	4.04	3.93	5.57	5.03	5.10	3.41	3.11	2.76	3.71	2.56
6	1,600 ms	2.11	2.55	3.25	4.11	4.14	4.05	2.89	2.59	2.43	3.04	2.23
7	3.2 sec	1.98	2.24	2.93	3.38	3.48	3.37	2.56	2.29	2.17	2.53	2.01
8	6.4 sec	1.94	2.11	2.62	2.91	2.93	2.92	2.35	2.08	1.95	2.16	1.82

Panel B. Wavelet variance ratios, $V_{j,J=16}$, computed from denoised bids and offers

Level, j	Time scale	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
1	50 ms	1.60	2.37	3.15	7.02	6.09	8.24	6.56	5.83	4.20	6.79	4.46
2	100 ms	1.57	2.32	3.09	6.82	5.89	7.76	5.89	5.17	3.83	6.00	4.07
3	200 ms	1.56	2.27	3.03	6.48	5.61	7.04	4.99	4.45	3.41	5.18	3.57
4	400 ms	1.55	2.23	2.94	5.90	5.16	6.02	3.96	3.68	2.97	4.36	3.00
5	800 ms	1.57	2.19	2.83	5.00	4.47	4.82	3.13	2.98	2.56	3.58	2.52
6	1,600 ms	1.64	2.20	2.71	3.99	3.60	3.79	2.63	2.51	2.27	2.94	2.20
7	3.2 sec	1.81	2.30	2.62	3.44	3.02	3.16	2.33	2.23	2.04	2.46	2.00
8	6.4 sec	2.11	2.51	2.59	3.20	2.65	2.75	2.15	2.04	1.86	2.11	1.82

Table 7. Consolidated quote record for PRK, April 6, 2001.

The table contains the consecutive records from the monthly TAQ consolidated quote file for the Park National Corporation. The first five columns are directly from the CQ file. The national best bid and offer (NBBO), and the exchange(s) at the NBBO are inferred. The NBBO columns contain entries only when there is a change. The size is units of 100 shares. ("4x2" means that 400 shares are bid for and 200 shares are offered.) The exchange codes are "A" (the American Stock Exchange, the primary listing exchange [presently named "NYSE MKT LLC"]); "M," Midwest; "C," Cincinnati; "T," NASDAQ.

Time	Bid	Offer	Size	Ex	Mode	NBB		NBO	
12:01:33	86.73	86.90	4x2	A	12	86.73	A	86.90	A
12:01:34	86.63	87.00	1x1	M	12				
12:01:35	86.35	87.28	1x1	C	12				
12:01:35	86.67	86.96	1x1	T	12				
12:01:35	86.67	86.96	1x1	T	12				
12:02:22	86.74	86.90	3x2	A	12	86.74	A		
12:02:23	86.64	87.00	1x1	M	12				
12:02:25	86.68	86.96	1x1	T	12				
12:02:25	86.68	86.96	1x1	T	12				
12:03:11	83.63	86.90	1x2	A	12	86.68	T		
12:03:13	83.53	87.00	1x1	M	12				
12:03:15	83.25	87.28	1x1	C	12				
12:03:15	83.60	86.90	2x2	A	12				
12:03:16	83.57	86.96	1x1	T	12	83.60	A		
12:03:16	83.57	86.96	1x1	T	12				
12:03:16	83.50	87.00	1x1	M	12				
12:03:21	83.54	86.96	1x1	T	12				
12:03:21	83.54	86.96	1x1	T	12				
12:03:27	83.81	86.90	1x2	A	12	83.81	A		
12:03:29	83.71	87.00	1x1	M	12				
12:03:30	83.43	87.28	1x1	C	12				
12:03:30	83.81	86.90	1x2	A	12				
12:03:32	83.75	86.96	1x1	T	12				

Time	Bid	Offer	Size	Ex	Mode	NBB		NBO	
12:03:37	83.75	86.96	1x1	T	12				
12:03:38	86.50	86.90	2x2	A	12	86.50	A		
12:03:39	86.40	87.00	1x1	M	12				
12:03:40	86.50	86.90	2x2	A	12				
12:03:40	86.12	87.28	1x1	C	12				
12:03:45	86.44	86.96	1x1	T	12				
12:03:45	86.44	86.96	1x1	T	12				
12:03:46	86.50	88.00	2x9	A	12			86.96	T
12:03:48	86.40	88.10	1x1	M	12				
12:03:49	86.12	88.38	1x1	C	12				
12:03:51	86.44	88.06	1x1	T	12			88.00	A
12:03:51	86.44	88.06	1x1	T	12				
12:03:52	86.50	86.90	2x2	A	12			86.90	A
12:03:54	86.40	87.00	1x1	M	12				
12:03:55	86.12	87.28	1x1	C	12				
12:03:58	83.00	86.90	2x2	A	12	86.44	T		
12:03:58	86.44	86.96	1x1	T	12				
12:03:58	86.44	86.96	1x1	T	12				
12:04:00	82.90	87.00	1x1	M	12				
12:04:01	82.62	87.28	1x1	C	12				
12:04:01	82.94	86.96	1x1	T	12	83.00	A		
12:04:01	82.94	86.96	1x1	T	12				
12:04:06	86.50	86.90	2x2	A	12	86.50	A		

Table 8. Time scale volatility estimates for US equities, 2001-2011

In each year 2001-2011, 150 US firms are randomly selected from CRSP (stratified by average daily dollar trading volume during the first quarter of the year). Quote records for April are taken from the NYSE Monthly TAQ database. Within each second, quotes are randomly assigned order-preserving millisecond fractional portions. The wavelet variances, ν_j^2, are estimates of the price variance at the time scale $\tau_j = 50 \times 2^{j-1}$. The rough variances, σ_j^2, measure cumulative variation at all time scales $\le \tau_j$. For presentation, I report the square-roots of the rough variances, in mils ($0.001) per share (Panel A). The rough wavelet variance ratio (Panel B) is $VR_{j,J} = 2^{J-j} \sigma_j^2 / \nu_J^2$ where $J = 16$ is the longest time-scale in the analysis (for a random-walk $VR_{j,J}$ would be unity at all horizons). Table entries are cross-firm means with standard errors in parentheses. The National Best Bid and Offer are computed from TAQ data; the bid and offer are separately transformed using the Haar basis; the reported estimates are averages across bid and offer sides. Transforms are performed through level $J = 16$; for brevity only a subset of time-scales are reported. All estimates are constructed from denoised bids and offers (with short-term peaks clipped).

Panel A. Rough variances, mils per share

Level, *j*	Time scale	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
1	50 ms	0.38	0.33	0.27	0.37	0.35	0.40	0.29	0.37	0.43	0.45	0.30
		(<0.01)	(<0.01)	(<0.01)	(<0.01)	(<0.01)	(0.01)	(<0.01)	(0.01)	(<0.01)	(0.01)	(<0.01)
3	200 ms	0.99	0.87	0.70	0.97	0.90	1.03	0.72	0.91	1.06	1.12	0.75
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.03)	(0.01)
5	800 ms	2.07	1.80	1.44	1.95	1.80	2.02	1.38	1.75	2.09	2.16	1.47
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.05)	(0.02)
7	3.2 sec	4.21	3.57	2.80	3.67	3.35	3.75	2.59	3.29	3.97	3.96	2.80
		(0.06)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)	(0.05)	(0.09)	(0.03)
10	25.6 sec	12.35	9.84	7.37	9.40	8.59	9.57	6.84	8.58	10.48	9.85	7.53
		(0.26)	(0.20)	(0.10)	(0.14)	(0.08)	(0.10)	(0.10)	(0.10)	(0.12)	(0.19)	(0.10)
14	6.8 min	45.20	34.42	25.02	32.00	29.82	33.50	24.64	30.91	38.31	34.17	27.87
		(0.58)	(0.39)	(0.28)	(0.39)	(0.31)	(0.38)	(0.38)	(0.33)	(0.47)	(0.59)	(0.39)
16	27.3 min	83.77	63.12	44.73	57.53	53.24	60.50	45.22	57.81	71.28	62.69	52.83
		(1.08)	(0.67)	(0.51)	(0.71)	(0.61)	(0.72)	(0.70)	(0.63)	(0.89)	(1.07)	(0.78)

Table 8. Time scale volatility estimates for US equities, 2001-2011 (continued)

Panel B. Rough variance ratios

Level, *j*	Time scale	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
1	50 ms	1.60	2.37	3.15	7.02	6.09	8.24	6.56	5.83	4.20	6.79	4.46
		(0.02)	(0.06)	(0.11)	(0.56)	(0.39)	(0.75)	(0.40)	(0.46)	(0.31)	(0.38)	(1.42)
3	200 ms	1.57	2.30	3.06	6.65	5.76	7.42	5.47	4.86	3.65	5.65	3.84
		(0.02)	(0.06)	(0.10)	(0.51)	(0.36)	(0.64)	(0.30)	(0.37)	(0.26)	(0.32)	(1.16)
5	800 ms	1.56	2.23	2.91	5.61	4.94	5.72	3.87	3.58	2.91	4.25	2.94
		(0.03)	(0.07)	(0.09)	(0.37)	(0.27)	(0.43)	(0.17)	(0.23)	(0.16)	(0.23)	(0.69)
7	3.2 sec	1.71	2.25	2.71	4.11	3.64	3.94	2.78	2.63	2.31	3.02	2.28
		(0.09)	(0.15)	(0.10)	(0.23)	(0.15)	(0.22)	(0.09)	(0.10)	(0.09)	(0.12)	(0.35)
10	25.6 sec	2.36	2.70	2.60	3.16	2.42	2.53	2.03	1.89	1.74	1.93	1.70
		(0.37)	(0.60)	(0.29)	(0.54)	(0.07)	(0.09)	(0.05)	(0.05)	(0.04)	(0.05)	(0.12)
14	6.8 min	1.37	1.41	1.50	1.58	1.49	1.52	1.37	1.29	1.32	1.30	1.23
		(0.04)	(0.06)	(0.03)	(0.06)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
16	27.3 min	1.12	1.13	1.16	1.18	1.16	1.17	1.12	1.09	1.11	1.10	1.08
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)

Figure 1. The bid and offer for AEPI, April 29, 2011

National best bid and offer (NBBO) from the NYSE Daily TAQ dataset. The National best bid (bottom line, in blue) is the maximum bid, taken over all market centers reporting to the Consolidated Tape Association; the National best offer (top line, red) is the minimum offer.

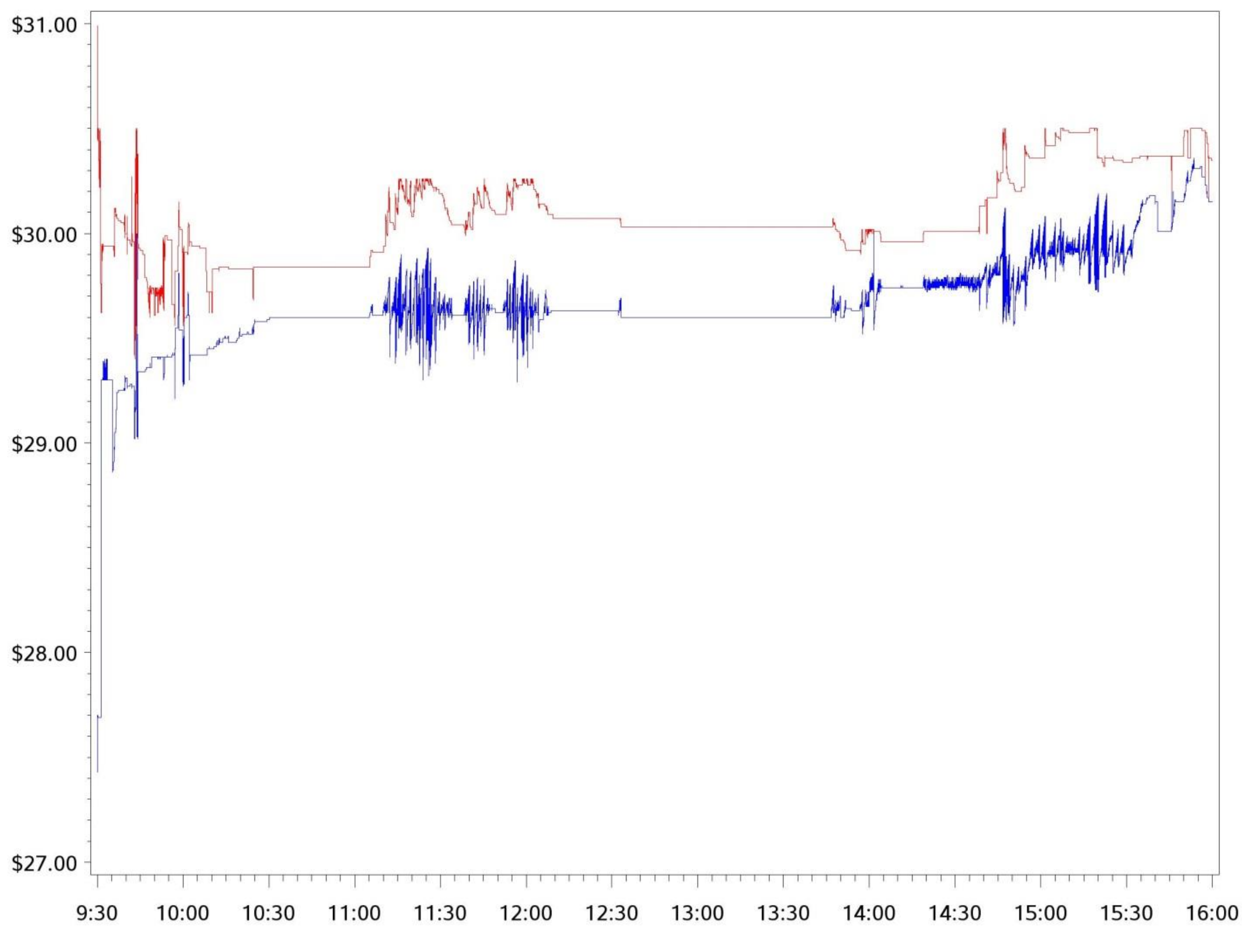


Figure 2. An illustrative time scale decomposition.

The figure depicts a hypothetical price series P , defined over eight intervals as {2,4,8,4,5,4,7,10}, and its time scale decomposition using the Haar wavelet transform. $S_j, R_j,$ and D_j are (respectively) the smooth, rough, and detail components at level *j*. The decomposition may be expressed additively in terms of smooths and roughs as $P = S_j + R_j$ for $j = 1,2,3.$ Alternatively, in terms of smooths and details, $P = S_1 + D_1 = S_2 + D_1 + D_2 = S_3 + D_1 + D_2 + D_3$. The correspondence between details and roughs is given by $R_1 = D_1$; $R_2 = D_1 + D_2$; $R_3 = D_1 + D_2 + D_3$ and $D_j = R_j - R_{j-1}$ for $j = 2,3.$ Graph labels indicate the mean and variance of the depicted component. Means of rough and detail components are zero by construction.

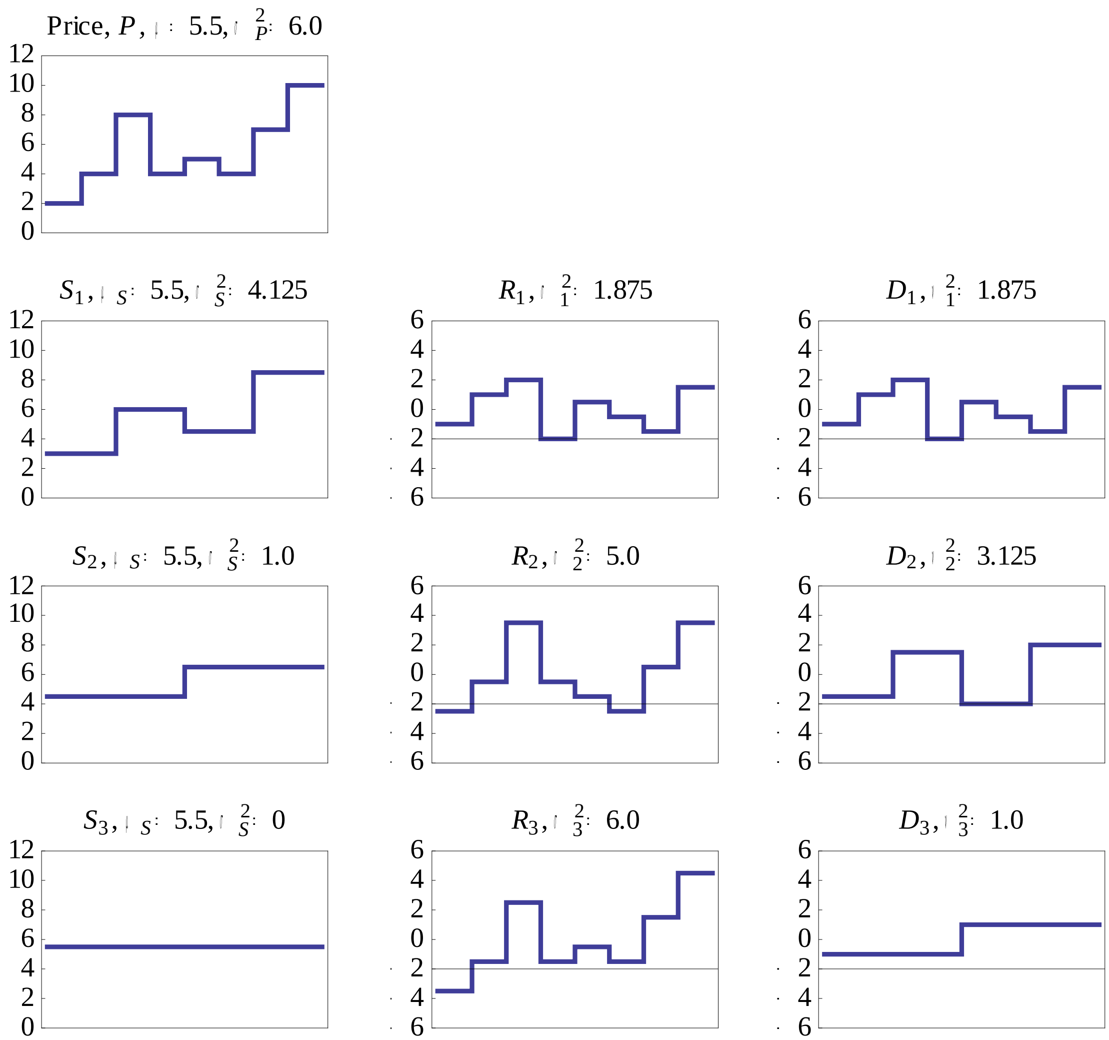


Figure 3. Variance ratios for US equities in 2011

Estimates of time scale variance ratios for 150 US firms during April, 2011. The wavelet variances, ν_j^2, are estimates of the price variance at the time scale $\tau_j = 50 \times 2^{j-1}$. The wavelet variance ratio is $V_{j,J} = 2^{J-j}\,\nu_j^2 / \nu_J^2$ where $J = 16$ is the longest time-scale in the analysis. For random-walk, both ratios would be unity at all horizons. Plotted points are means (across firms) of estimated variance ratios in quintiles constructed on dollar trading volume. The National Best Bid and Offer are computed from TAQ data; the bid and offer are separately transformed using the Haar basis; the reported variance estimates are averages of the bid and offer variances. The data are time stamped to the millisecond. Prior to transformation, I take the average of the bid or offer over non-overlapping 50 millisecond intervals.

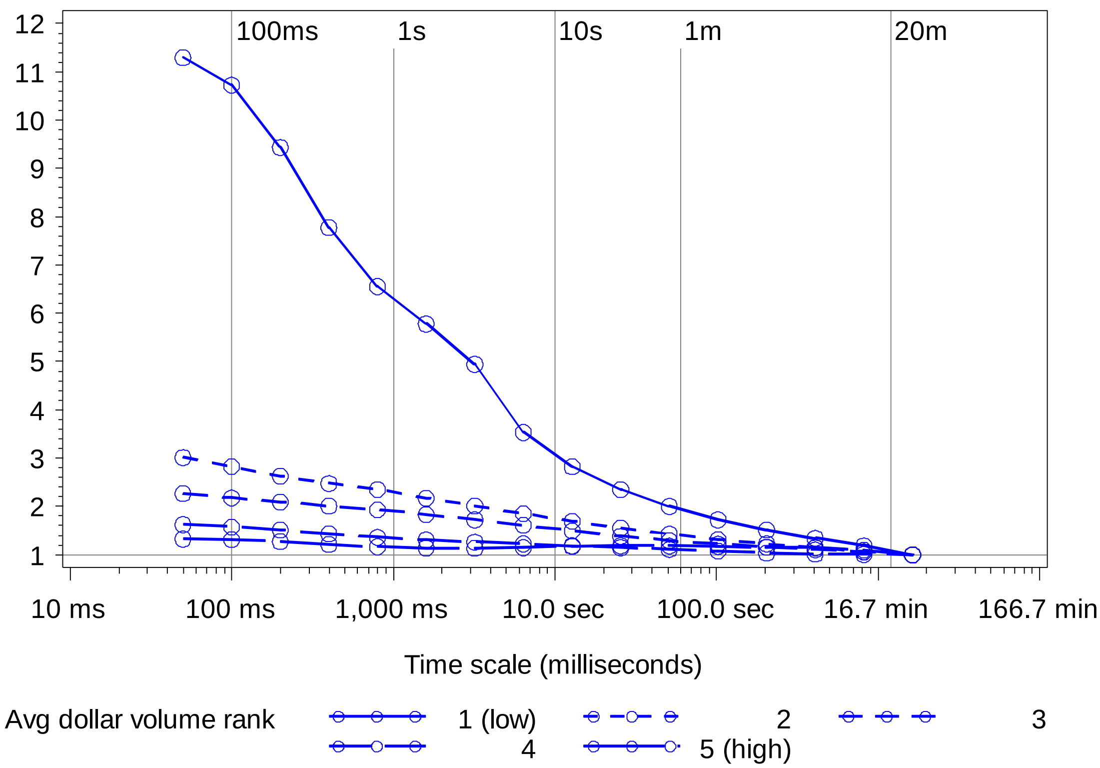


Figure 4. Wavelet correlations between the National Best Bid and National Best Offer

Estimates of bid-offer wavelet correlations for 150 US firms during April, 2011. The wavelet correlation between the bid and offer at level j (and time scale $\tau_j = 50 \times 2^{j-1}$) is defined as $\rho_{bid,offer,j} = \nu^2_{bid,offer,j} \Big/ \sqrt{\nu^2_{bid,j}\nu^2_{offer,j}}$ where $\nu^2_{bid,j}$, $\nu^2_{offer,j}$ and $\nu^2_{bid,offer,j}$ denote the bid wavelet variance, the offer wavelet variance, and the bid-offer wavelet covariance. Plotted points are means (across firms) of estimated correlations in quintiles constructed on dollar trading volume. The National Best Bid and Offer are computed from TAQ data; the bid and offer are separately transformed using the Haar basis; the reported variance estimates are averages of the bid and offer variances. The data are time stamped to the millisecond. Prior to transformation, I take the average of the bid or offer over non-overlapping 50 millisecond intervals. The sample is 150 randomly chosen U.S. stocks, over April, 2011.

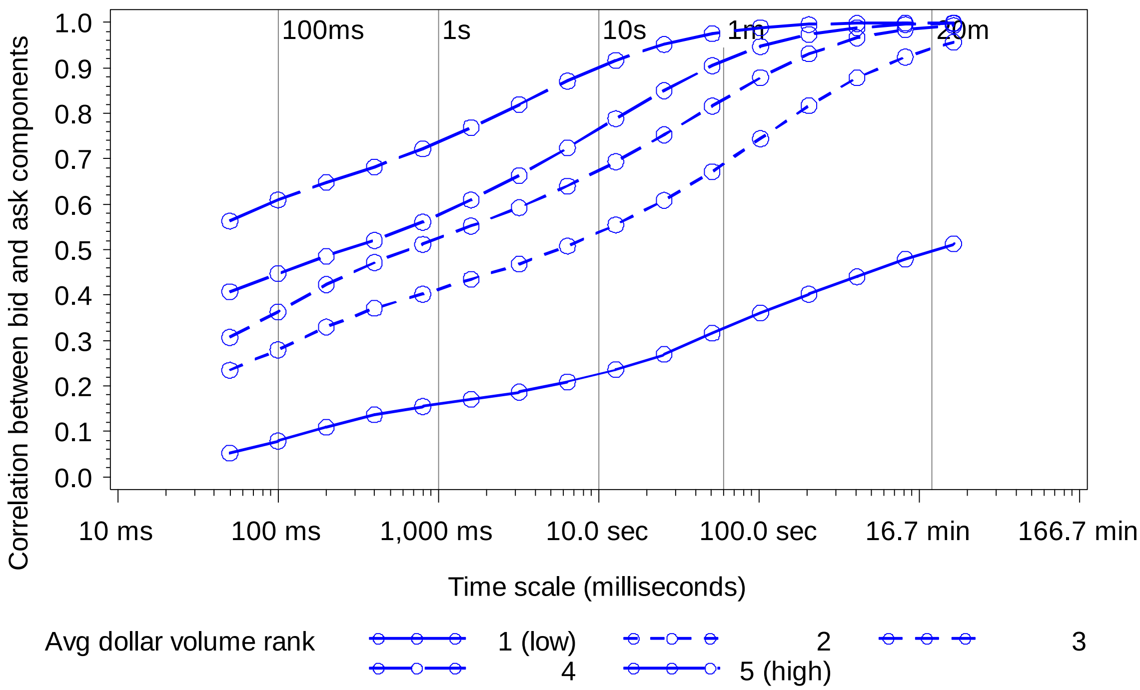
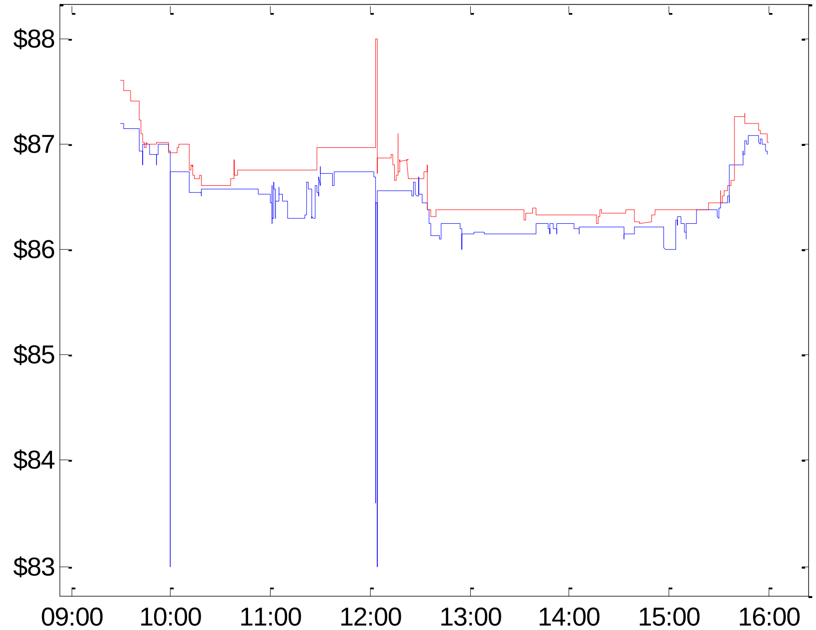


Figure 5. Bid and offer for PRK (Park National Corporation) on April 6, 2001.

Panel A. National best bid and offer (NBBO) from the NYSE Daily TAQ dataset. The National best bid (bottom line, in blue) is the maximum bid, taken over all market centers reporting to the Consolidated Tape Association; the National best offer (top line, red) is the minimum offer.



Panel B. Rough component of the National Best bid, constructed from a Haar wavelet transform and comprising components at time scales of 51.2 seconds and lower. The bands demarcate ±$0.33, approximately 150% of the average bid-ask spread for the day.