Market Efficiency and Microstructure Evolution in U.S. Equity Markets: A High-Frequency Perspective

Jeff Castura, Robert Litzenberger, Richard Gorelick, Yogesh Dwivedi
RGM Advisors, LLC

August 30, 2010

1 Introduction

The impact of high frequency trading (HFT) on the U.S. equity markets has received considerable attention in the wake of the financial crisis of 2008 and the so-called 'flash-crash' of May 6, 2010. It has been suggested that HFT now accounts for over half of U.S. equity share volume [1]. With such a large presence in the market, it is important to understand if there are any adverse effects caused by such activity. While the existence of a causal relationship is not proven, evidence is presented which suggests that the U.S. markets have improved in several respects as HFT activity has grown.

This work presents some evidence showing that the U.S. equity markets appear to have become more efficient with tighter spreads, greater liquidity at the inside, and less mean reversion of mid-market quotes over the past several years; a period that has seen a sizable increase in the prevalence of HFT, and a period during which there has been coincident growth in automation and speed on many exchanges. Furthermore, evidence is presented which shows that exchanges which moved toward greater automation earlier saw earlier improvements in market efficiency metrics.

An important determinant of overall market quality is the total cost of participation which is comprised of a number of components. For smaller market participants who trade small volumes, bid-ask spread is likely the dominant component. For larger market participants two additional components become important. First, the available size to trade becomes a factor as limited size at or near the best price will result in worse execution prices. Second, the mean-reverting component of price impact represents a cost that can be significant for larger investors.

One measure of efficiency investigated in this paper is the bid-ask spread. It is expected that the presence of more participants, algorithmic and otherwise, will drive spreads down due to competition thereby decreasing costs to other investors. The results presented in this paper confirm the results of many other studies, showing that bid-ask spreads have come down over time for a broad range of stocks, coincident with improvements in automation on exchanges.

Another measure of efficiency is liquidity, representing the ability of investors to obtain their desired inventories with minimal price impact. Again, it is expected that more participants implies a greater amount of liquidity in the markets, a benefit to investors. This appears to be the case as this paper confirms the results of other papers demonstrating an increase in available liquidity over time as automation on exchanges has improved.

It was shown by Samuelson that if a stock price is efficient, i.e., the price is fairly valued with all public information, then it must follow a martingale process [2]. As a consequence, an efficient price exhibits no serial autocorrelation, either positive (momentum) or negative (mean-reversion). Fama explored these ideas further and subsequently tested some ideas of market efficiency, providing additional support for this concept of efficiency in markets [3, 4].

A variance ratio test was developed by Lo and Mackinlay which makes use of the fact that in an efficient market, the variance per unit time of the price of a stock should be constant [5]. This allows ratios of variances over different time horizons to be taken and compared with theoretical expectations where, in an efficient market, these tests would show that there is little or no serial autocorrelation in prices. Another advantage of this type of test is that it does not depend on a particular order of serial autocorrelation, only whether any such autocorrelation is present. The application of these tests to high frequency data, a novel contribution of this paper, demonstrates that for all the data-sets investigated there is an overall improvement in efficiency in prices over time, particularly after exchanges have made investments in their capacity to support automation.

The data-sets used in this study are the Russell 1000 components, consisting of 1000 large-cap and mid-cap stocks, and the Russell 2000 components, consisting of 2000 small-cap stocks. The set of components are taken as of Q4 2009, and no attempt is made to correct for survivor bias, though it may be argued that the nature of this study is not sensitive to such effects.

Additionally, each index is partitioned into two sets; NYSE-listed stocks and NASDAQ-listed stocks. For much of the time period studied, NASDAQ-listed stocks traded primarily on automated, electronic exchanges while NYSE-listed stocks have transitioned from being primarily traded manually on the NYSE to being traded on a more competitive, automated group of electronic exchanges. The data essentially represents four distinct subsets of stocks, at least from an historical context: large-cap stocks largely traded automatically (approximately 200 NASDAQ-listed stocks in the Russell 1000), large-cap stocks that have transitioned from being largely traded manually to being largely traded automatically (approximately 800 NYSE-listed stocks in the Russell 1000), small-cap stocks largely traded automatically (approximately 1300 NASDAQ-listed stocks in the Russell 2000), and small-cap stocks that have transitioned from being largely traded manually to being largely traded automatically (approximately 700 NYSE-listed stocks in the Russell 2000). This partition allows comparisons to be made that help more clearly identify the impact of automation and technology advances on the health of the market.

The raw data is sampled at 1 second intervals for each stock during the period January 1, 2006 to June 30, 2010 inclusive, representing 18 quarters of data. The first 10 minutes and last 10 minutes of each day are omitted to prevent opening and closing activities from influencing the results. Inside values are used across the NASDAQ, NYSE, NYSE ARCA and BATS exchanges. This represents a significant fraction of all shares traded in the U.S. and so is taken to be representative of overall market activity.

With this data-set a series of statistical tests and measurements are run, designed to reflect the health of the market. Spreads, available liquidity, and transient volatility in the form of variance ratio tests are presented here as these are commonly cited metrics of market efficiency and market quality.

2 Bid-Ask Spreads

Spreads are a cost to trading and, all else being equal, smaller spreads are evidence of a better cost structure for investors. Conversely, market makers and other liquidity providers earn profits through the spread. To that extent smaller spreads imply not only smaller revenues for market makers but also that these participants, by quoting smaller spreads, are more competitive; a sign of a healthy market.

Bid-ask spreads are presented as the mean absolute spread of each of the components of the index, where the absolute spread is defined as the best ask price less the best bid price. There are other common ways to present bid-ask spread data including the use of relative spreads, defined as the absolute spread divided by the stock price. This formulation is meant to more directly reflect transaction costs for investors caused by the bid-ask spread. Market makers and other liquidity providers commonly adjust their quotes based on market volatility in order to compensate for their increased risk of holding inventory [6]. Therefore a volatility adjustment is commonly done to attempt to mitigate the impact of volatility from spreads, typically making it easier to spot trends in spreads over time. Dollar-value weighting is also sometimes used in an effort to better reflect costs of the spread paid by investors. Equal weighting is chosen here because many of the largest and most liquid stocks are pinned at a spread of one penny.

Figure 1 presents the mean of the absolute spread over time for the Russell 1000 stocks partitioned into its NYSE-listed and NASDAQ-listed components. This is done to try to isolate differences in behavior over the period studied that may be attributable to structural changes on each of these exchanges. Both groups have seen a reduction in spreads over the period investigated, dropping by about 1.5 pennies for the NYSE-listed stocks and about 1 penny for the NASDAQ-listed stocks. By the end of 2009 it appears the the mean spread of the two groups has converged to approximately the same value, something that could not be said previously.

It is known that the rate of adoption of automated trading on NYSE-listed

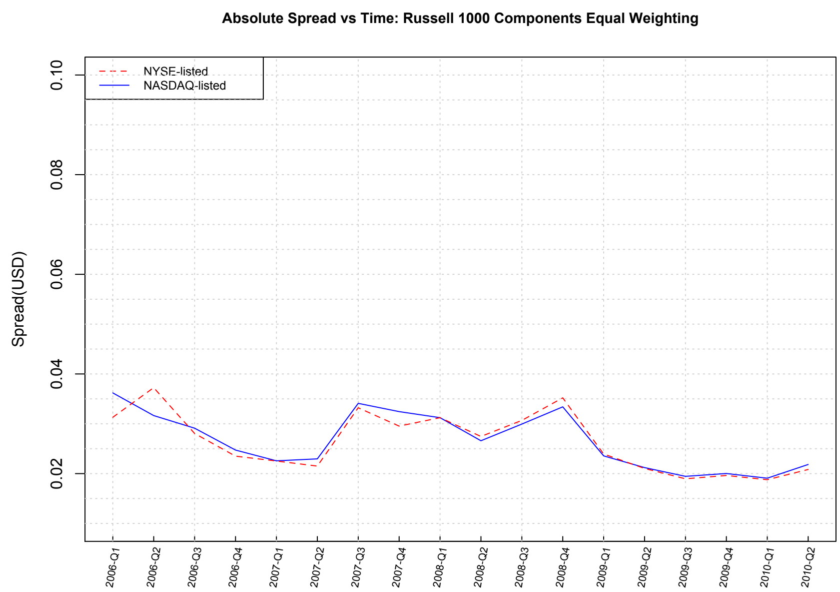

Figure 1: Mean bid-ask spread for Russell 1000

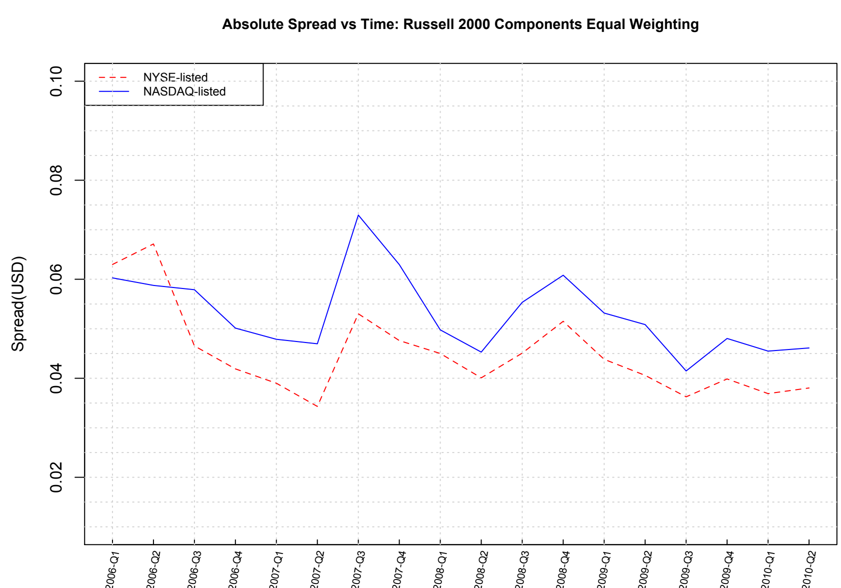

Figure 2: Mean bid-ask spread for Russell 2000

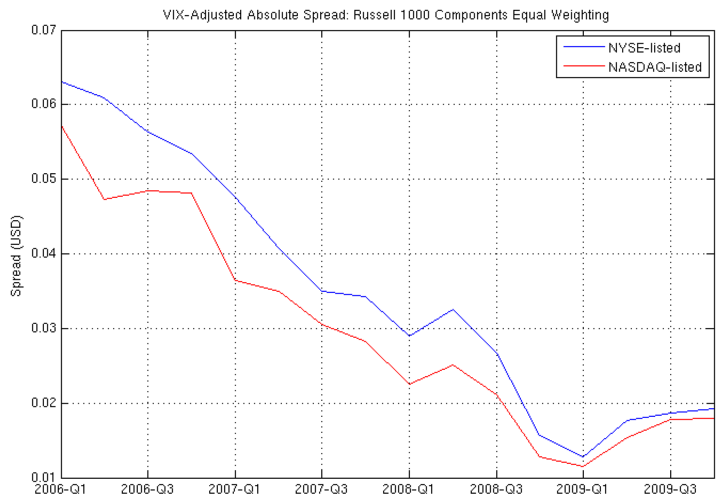

Figure 3: Mean bid-ask spread for Russell 1000, VIX-adjusted

stocks lagged behind that of NASDAQ-listed stocks. As the NYSE moved to an electronic system to catch up technologically with the NASDAQ, and as other electronic venues began taking market share from the NYSE, spreads in the Russell 1000 dropped more dramatically for the NYSE-listed stocks than the NASDAQ-listed stocks. This also suggests a relationship between the entrance of algorithmic trading with a reduction in spreads, something that is noted for the German DAX [7].

The same information for the Russell 2000 index is presented in Figure 2. Like the Russell 1000, these stocks have seen a reduction in mean spreads by about a penny, with the NYSE-listed symbols showing a more dramatic reduction than the NASDAQ-listed symbols.

A clearer perspective on these trends can be seen by adjusting the spreads by the volatility in the market. For this, quarterly VIX-values are used to deflate the bid-ask spreads. VIX-adjusted spread data is presented in Figures 3 and 4 showing the Russell 1000 and Russell 2000 spreads over time.

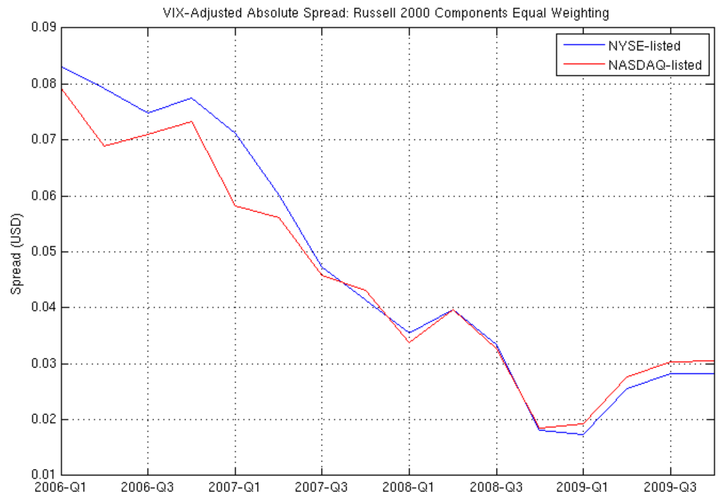

Figure 4: Mean bid-ask spread for Russell 2000, VIX-adjusted

3 Available Liquidity

Liquidity is an important part of a vital market. It is often loosely defined as the ability of participants to trade the amount that they wish at the time they wish. One measure of liquidity is the amount of size offered for sale or for purchase by market makers and other liquidity providers at a given point in time. If more shares are available to be bought and sold at any given time, then market participants have a greater ability to get into or out of positions based on their needs or desires and are less dependent on either waiting for sufficient size to become available or to seek an alternative execution venue.

Available liquidity is measured as the dollar value available to buy or sell at any instant in time at the inside bid and ask, and time averages over an entire quarter are taken. Each stock in an index is weighted by its capitalization reported for the quarter to produce a single capitalization-adjusted available liquidity metric. The motivation for weighting by capitalization is that it more closely reflects the available fraction of a company's total value that can be transacted at any given time which may be more representative of the limitations to investors. Additional available liquidity data is presented in the appendix, including results for the NASDAQ-100.

Figure 5 presents the adjusted available liquidity for the Russell 1000 components partitioned into NYSE-listed and NASDAQ-listed stocks. Between 2006 and the end of 2009, the available liquidity of both groups of stocks increased significantly, by about a factor of two, though all of that gain appears to have taken place in 2009. Similar results are seen for the two groups of stocks in the Russell 2000 which is shown in Figure 6.

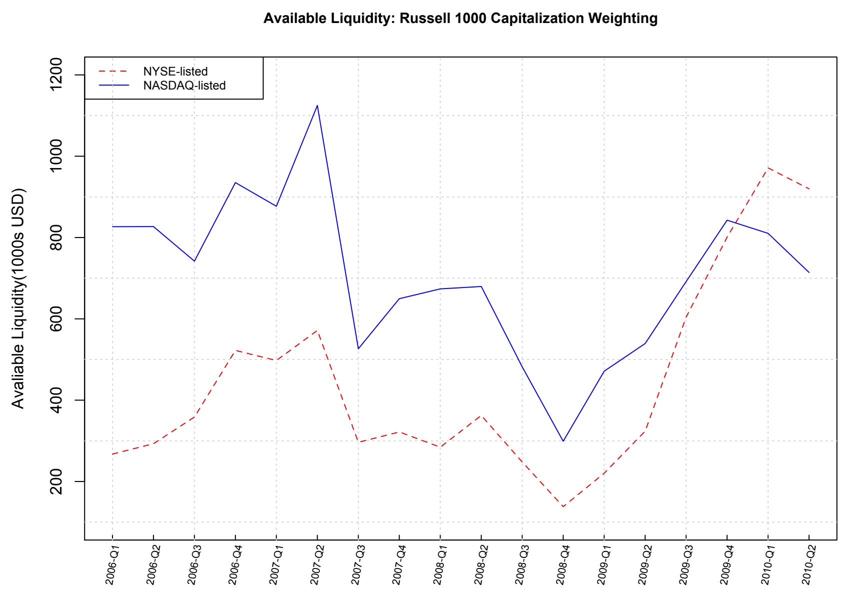

Figure 5: Mean available liquidity for Russell 1000

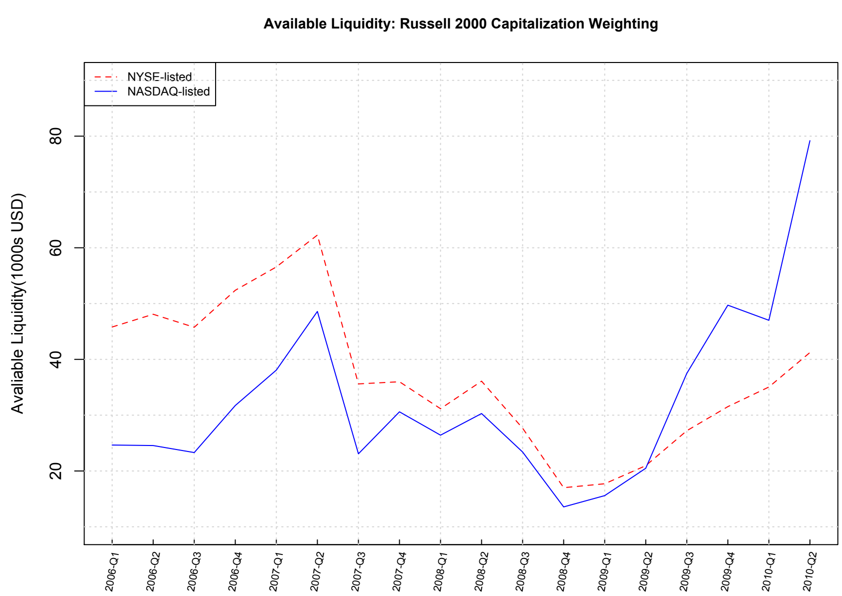

Figure 6: Mean available liquidity for Russell 2000

It is plausible that the increase in liquidity can be explained, at least in part, by the presence of HFT participants. Since the data used in this work is sampled at a high rate, one can also claim that this liquidity measure is representative of the immediacy that is available to market participants. This immediacy is a type of option that is available to market participants providing them with more flexibility than may otherwise be available.

4 Market Efficiency Tests

There exists a large body of research devoted to tests of market efficiency. In this context, efficiency typically refers to the degree to which the price time-series of a stock resembles a random walk. The theoretical foundation for this was developed by Samuelson where he proves that a properly anticipated stock price should fluctuate randomly with no serial autocorrelation [2]. Pioneering empirical work in this area in the form of a variance ratio test was presented by Lo and Mackinlay, who show with some level of statistical confidence that daily NYSE closing stock prices do not appear similar to a random walk, suggesting inefficiency in the markets [5]. The data used in their paper is sampled daily and ends in 1988, prior to a significant number of structural and regulatory changes that have dramatically changed the nature of U.S. equity markets.

Let X_t be the random walk

$$X_t = \mu + X_{t-1} + \epsilon_t, \tag{1}$$

where μ represents the drift and ϵ_t is an independent stochastic disturbance with zero mean. Let σ_a^2 be the true variance of the first difference of the original series. Define a subsequence of this time series with a q-period holding time to be $\{X_1, X_{1+q}, X_{1+2q}, \ldots\}$. Let σ_b^2 be the true variance of the first difference of this subsequence. It can be shown that for the time series above, the ratio $\frac{\sigma_b^2}{q\sigma_a^2}$ equals one. A variance ratio's deviation from unity can then be considered to be proportional to the amount of inefficiency present in that stock or index. Variance ratios greater than one imply a momentum process, equivalently a positive serial autocorrelation, while values less than one imply mean-reversion, or negative serial autocorrelation.

Lo and Mackinlay also defined the following test statistics

$$Z_1(q) = \sqrt{(}Nq)(\frac{\hat{\sigma}_b^2}{q\hat{\sigma}_a^2} - 1)[2(2q-1)(q-1)/3q]^{-\frac{1}{2}} \sim N(0,1), \tag{2}$$

$$Z_2(q) = \sqrt{(}Nq)(\frac{\hat{\sigma}_b^2}{q\hat{\sigma}_a^2} - 1)[V(q)]^{-\frac{1}{2}} \sim N(0,1), \tag{3}$$

where $V(q)$ is defined in [5] and N is the number of observations. $Z_1(q)$ is used when the stochastic disturbances are assumed to homoscedastic and $Z_2(q)$ is used when disturbances are heteroscedastic. These test statistics provide a way to test for a random walk, where the Gaussian $\alpha\%$ significance level could be directly used for hypothesis testing.

Subsequent research extended the variance ratio tests of Lo and Mackinlay to provide alternative methods to test market efficiency. In particular Chow and Denning extend the variance ratio test to provide a more statistically powerful test procedure and it is this "Chow-Denning" test that is used as a metric of market efficiency in this section [8]. To the best of the authors' knowledge, such tests have not previously been applied to intra-day data sampled at a fine resolution as is done here.

The Chow-Denning method tests the null hypothesis that a price time-series is drawn from a random walk, and produces a single test statistic. Compared to the variance ratio test, the Chow-Denning method tests the random walk hypothesis simultaneously over multiple holding periods. Suppose there are m holding periods in a time-series of interest. Define the test statistics $|Z_1^*(q)|$ and $|Z_2^*(q)|$ to be the maximum of the absolute value of all the test statistics from (2) and (3) over all the holding periods, i.e,

$$Z_1^*(q) = max\{|Z_1(q_i)|\}_{i=1}^{m} \tag{4}$$

$$Z_2^*(q) = max\{|Z_2(q_i)|\}_{i=1}^{m} \tag{5}$$

These values can be compared to a threshold for a certain significance level and values exceeding this threshold suggest with some confidence level that the time-series does not resemble a random walk. In this study 5% is used as the significance level. At 5% significance, if this test were run on 100 truly random time-series, one would expect to see about 5 test outcomes reject the null hypothesis. That is to say, due to the statistical nature of this test, it may produce false positives about 5% of the time. The evolution of this test on a data-set over time provides an indication of changes in market efficiency. A value that approaches or drops below the 5% significance level over time implies an improvement in efficiency over that time.

It is important to note that at this sampling rate micro-structural effects are expected to be present. In particular, bid-ask bounce and statistical influences caused by the discrete nature of price values will tend to skew the results toward appearing mean-reverting. These effects are expected at high sampling rates and are expected to decay as the sampling rate is decreased. However, for a given sampling interval, the effect is expected to be roughly constant over time, and thus the interesting aspect of the results is how they have changed over time and whether they have converged toward a value of one. An attempt has been made in the variance calculations to account for the discrete price values and midpoint prices are used rather than last trade prices to minimize the effect of bid-ask bounce. More details are available in the appendix, along with some results based on last trade prices.

Raw variance ratio tests are applied to the Russell 1000 and Russell 2000, partitioned into NYSE-listed and NASDAQ-listed stocks. Three ratios are chosen to be representative of what may be typical HFT holding periods; 10 seconds over 1 second, 60 seconds over 10 seconds, and 600 seconds over 10 seconds.

Figures 7 and 8 show the raw variance ratios of 10 seconds over 1 second for midpoint price data from the Russell 1000 and 2000, respectively. These indexes are partitioned into NYSE-listed and NASDAQ-listed stocks. At this high frequency, it is seen that the Russell 1000, NASDAQ-listed stocks show a high degree of efficiency, and have been relatively efficient throughout the entire period investigated, with some improvement seen over time. As these stocks have largely been traded electronically for the entire period, such results are expected. The NYSE-listed components, by contrast, show a relatively large amount of inefficiency in 2006, but have increased to over 0.95 by 2009 and now appear to be at least as efficient as the NASDAQ-listed stocks.

The Russell 2000 index in Figure 8 shows the same general trends, though the overall efficiency is lower than the Russell 1000. This is to be expected since the smaller-cap stocks of the Russell 2000 do not have the same amount of trading activity as large-cap stocks. The NYSE-listed symbols show a greater degree of improvement in efficiency than the NASDAQ-listed symbols, again coinciding with improvements in automation and increased participation in these stocks by automated trading firms.

A variance ratio of 10 minutes over 10 seconds is presented to provide a picture of market efficiency over larger time-scales. Figures 9 and 10 show the results for the Russell 1000 and Russell 2000, respectively, and the same general trends seen in the previous plots of variance ratios are present in these figures.

Additional results are presented for the variance ratios of 1 minute over 10 seconds in the appendix in Figures 19 and 20 for the Russell 1000 and Russell 2000, respectively, showing the same trends.

The Chow-Denning test was applied over each of the 18 quarters, individually to each stock in each of the data-sets with the input to the test being the logarithm of the midpoint price and sampled at 10 minute and 10 second intervals. Additionally, binomial tests were conducted to construct the confidence interval for the proportion of tests that rejected the null hypothesis. Assuming independence of stocks, the confidence interval would be expected to contain 0.05 at a 5% significance level.

Results for the 10 minute sampled Chow-Denning tests are presented in Figures 11 and 13 for the Russell 1000 and 2000 data-sets, respectively for the NASDAQ-listed stocks. These figures show the fraction of stocks in the index that the Chow-Denning test reported as not being drawn from a random walk at a 5%-significance level. Figure 11 shows that at 10 minute sampling, the number of such occurrences has dropped over time and has largely been below 5% since the beginning of 2009, suggesting that there is no statistically significant inefficiencies at this sampling interval that this test detects. Results for the NYSE-listed stocks are given in Figures 12 and 14 and appear to have a more dramatic improvement, in agreement with the variance ratio results presented above.

Additional results using 10 second sampling are provided in the appendix. Although they show a higher degree of inefficiency than the 10 minute sampling results, the same general trends are present.

An alternative interpretation of these results is that of an increase in the speed of mean-reversion over time. As mentioned, mean-reversion is present in this data due in part to micro-structural effects, and as the rate of trading and

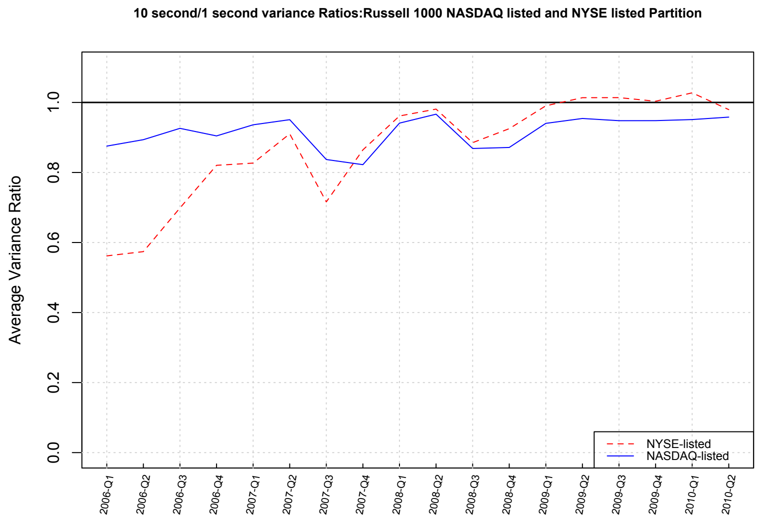

Figure 7: Variance Ratios, Russell 1000, 10 seconds / 1 second

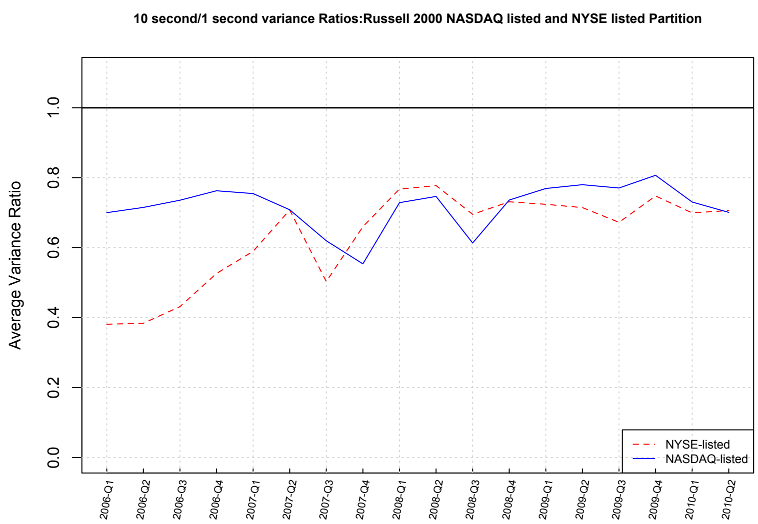

Figure 8: Variance Ratios, Russell 2000, 10 seconds / 1 second

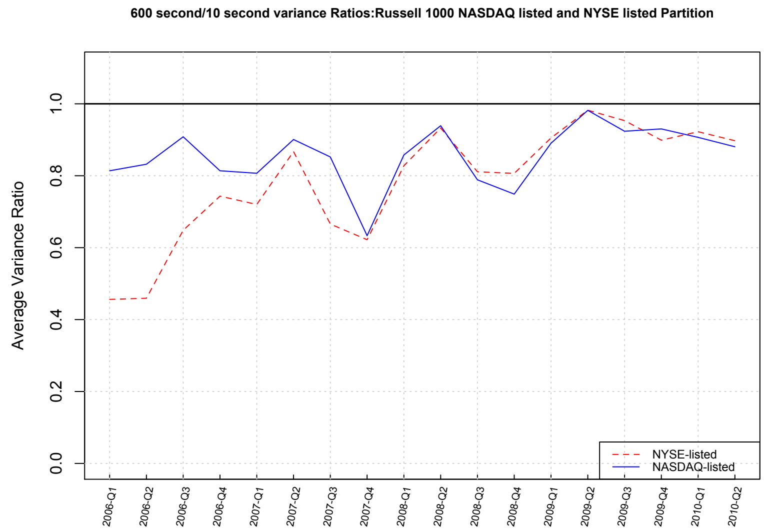

Figure 9: Variance Ratios, Russell 1000, 10 minute / 10 seconds

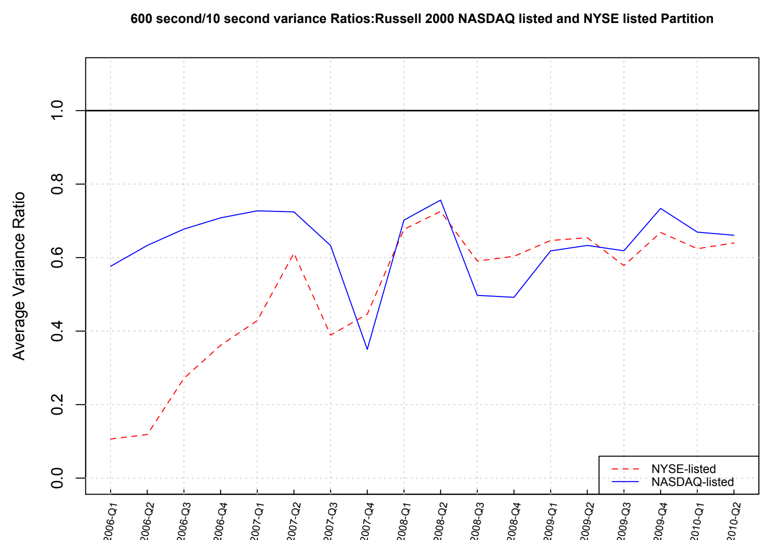

Figure 10: Variance Ratios, Russell 2000, 10 minute / 10 seconds

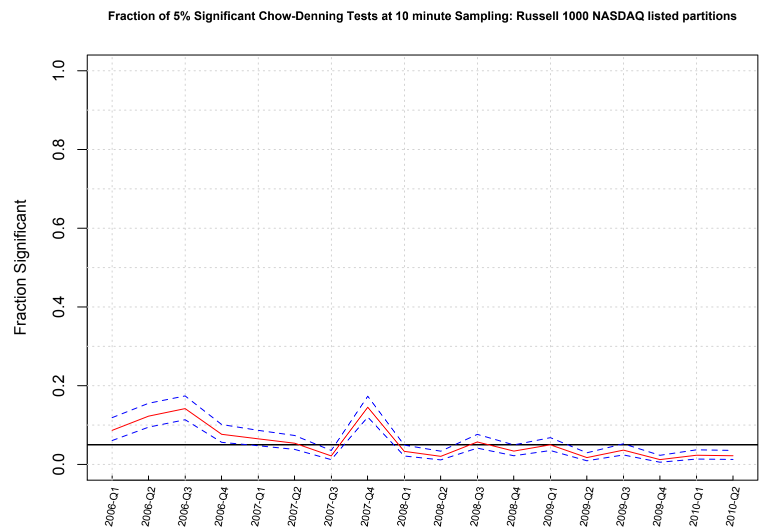

Figure 11: Chow-Denning test results for the Russell 1000 NASDAQ-listed stocks, 10 minute sampling

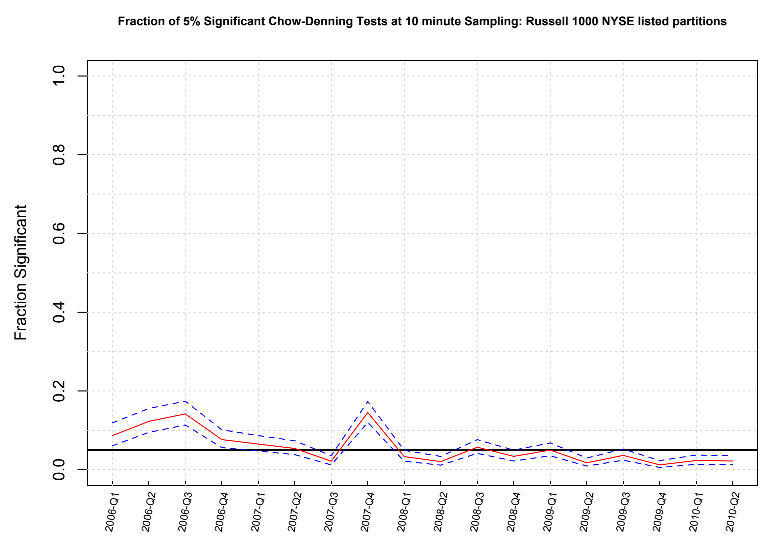

Figure 12: Chow-Denning test results for the Russell 1000 NYSE-listed stocks, 10 minute sampling

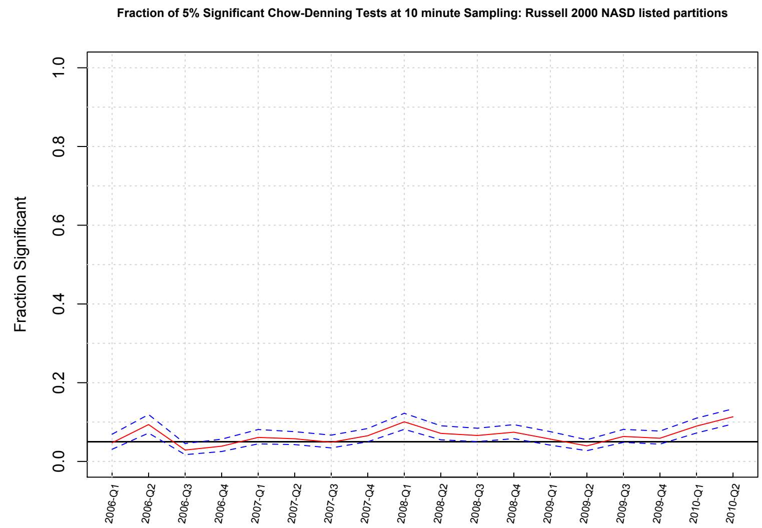

Figure 13: Chow-Denning test results for the Russell 2000 NASDAQ-listed stocks, 10 minute sampling

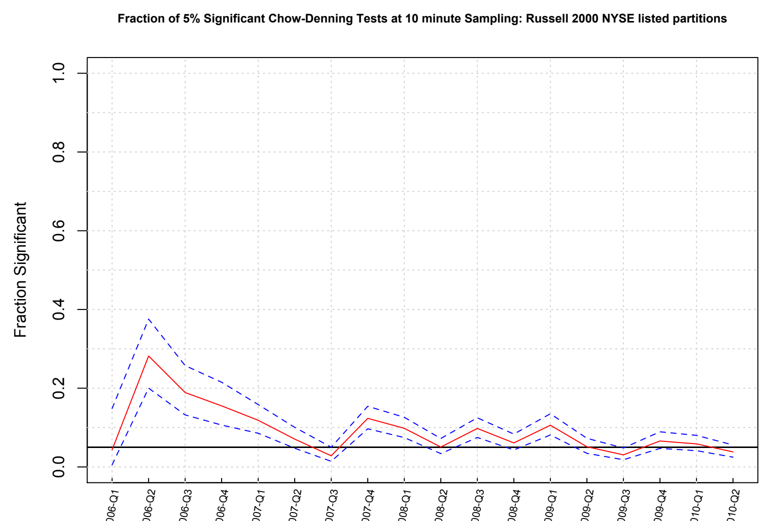

Figure 14: Chow-Denning test results for the Russell 2000 NYSE-listed stocks, 10 minute sampling

market activity increases, the impact of such noise on these variance ratio-based tests become less prevalent. Therefore one can conjecture that the decrease in the Chow-Denning test statistics may be as a result of an increased rate of reversion of prices to their mean. This is also an indication of an increasing competitive landscape and increasing efficiency in the market.

5 Summary

The presented data is suggestive that the U.S. equity markets have become more liquid and efficient over the past four years, despite macro-economic shocks. As the ratio of HFT activity to total market activity has grown, there appears to be no evidence that short-term volatility, liquidity or spreads have risen for the bulk of market participants. To the contrary, the evidence presented here suggests a continued improvement in each of these factors, implying a sympathetic relationship between HFT and the health of the overall markets. From the perspective of total costs to trading, the reduction in bid-ask spreads over the past several years is a benefit to all investors. The reduction in mean reversion coupled without a drop in available liquidity suggests that larger investors have also enjoyed lower overall costs to trading.

The partitioning of data into the Russell 1000 and Russell 2000 shows that there has generally been a larger degree of improvement in efficiency metrics in the Russell 1000. The difference in trends observed between NYSE-listed and NASDAQ-listed stocks also supports the hypothesis that increased automation and the presence of HFT that has come with it has improved the market quality metrics investigated in this paper.

References

[1] The TABB Group, "US equity high frequency trading: Strategies, sizing and market structure", *TABB Group, http://www.tabbgroup.com*

[2] Samuelson, P.A, "Proof that properly anticipated prices fluctuate randomly", *Industrial Management Review*, 1965

[3] Fama, E., "The Behavior of Stock Market Prices", Journal of Business 38, 1965

[4] Fama, E., "Efficient Capital Markets: A Review of Theory and Empirical Work" Journal of Finance, 25, 1970

[5] Lo, A.W. and Mackinlay A.C., "Stock market prices do not follow random walks: Evidence from a simple specification test", *Review of Financial Studies*, 1988

[6] Glosten, L.R. and Harris, L., "Estimating the components of the bid-ask spread", *Journal of Financial Economics*, 1988

[7] Hendershott, T. and Riordan, R., "Algorithmic trading and information", *http://econpapers.repec.org/paper/netwpaper/0908.htm*, 2009

[8] Chow, K.V. and Denning, K.C., "A simple multiple variance ratio test", *Journal of Econometrics*, 1993

[9] Weisstein, E. W., "Sheppard's Correction", *From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/SheppardsCorrection.html*

[10] Gerig, A. and Michayluk, D., "Automated liquidity provision and the demise of traditional market making", `http://www.finance.uts.edu.au/staff/austing/GerigMichayluk2010_AutomatedLiquidity.pdf`

[11] Harris, L., "Trading and exchanges: Market microstructure for practitioners", Oxford University Press, 2002

[12] Brogaard, J., "High frequency trading and its impact on market quality". `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1641387`

[13] Hasbrouck, J., "Measuring the information content of stock trades", Journal of Finance, 46(1):179–207, 1991

[14] Jarnecic, E. and Snape, M., "An analysis of trades by high frequency participants on the London Stock Exchange", `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1628548`

[15] The CME Group, "Algorithmic trading and market dynamics", `http://www.cmegroup.com/education/files/Algo_and_HFT_Trading_0610.pdf`

6 Appendix

6.1 Bid-Ask Spreads

Absolute spreads are computed as follows. An individual stock i has a spread at time t of $S_i(t) = a_i(t) - b_i(t)$. The spread over a quarter q is defined as

$$\langle S_i(q) \rangle = \frac{\sum_{t \in q} S_i(t)}{\sum_{t \in q} 1}.$$

The spread S_q^N over an index N is the weighted average over all components, where w_i represents the weighting of stock i. The spread is then

$$S_q^N = \frac{\sum_{i \in N} w_i \langle S_i(q) \rangle}{\sum_{i \in N} w_i}.$$

The choice of equal weighting sets all $w_i = 1$. Dollar value weighting is determined by setting the weight for each stock to the total dollar value of all transactions for that stock in the quarter.

Relative spread can be computed in a similar manner, with the relative spread $S_R(t)_i = \frac{a(t)_i - b(t)_i}{p(t)_i}$ replacing the absolute spread above, and where $p_i(t)$ represents price. A common adjustment made to bid-ask spreads is a volatility adjustment [6]. The VIX is used for this purpose and its value relative to the mean of its value over the time period studied is chosen as the deflator. The value of the VIX over the period studied is given in Figure 15.

For comparison, spread data is also presented for the NASDAQ-100 index. Absolute spreads, both unadjusted and VIX-adjusted are given in Figure 16. The trend for this index is consistent with that seen in the Russell data-sets. Relative spreads are presented in a number of ways in Figure 17 and these adjustments do not change the overall trends presented in the body of the text.

6.2 Available Liquidity

The available liquidity for a stock i at time t is given as

$$L_i(t) = p_i(t) \left(s_i^a(t) + s_i^b(t)) \right),$$

where $s_i^a(t)$ and $s_i^b(t)$ are the inside size at the ask and bid, respectively. In a quarter q, the average available liquidity of a stock is

$$\langle L_i(q) \rangle = \frac{\sum_{t \in q} L_i(t)}{\sum_{t \in q} 1}.$$

The available liquidity over an index N is the weighted average over all components, such that

$$L_q^N = \frac{\sum_{i \in N} w_i \langle L_i(q) \rangle}{\sum_{i \in N} w_i},$$

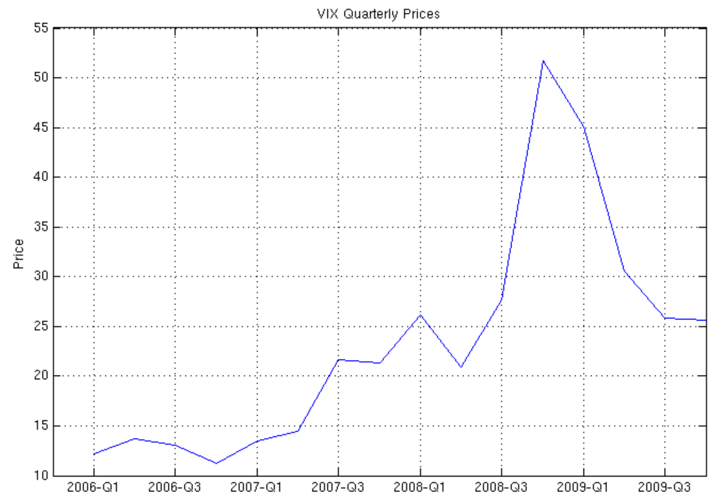

Figure 15: Quarterly VIX prices

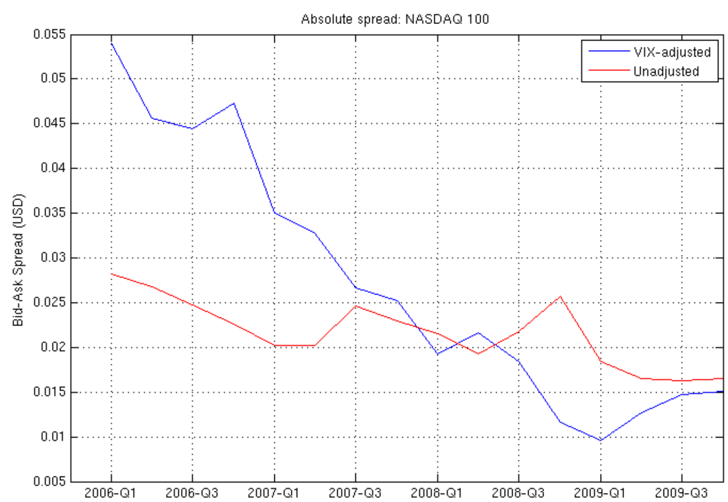

Figure 16: Absolute equal-weighted bid-ask spread for NASDAQ 100

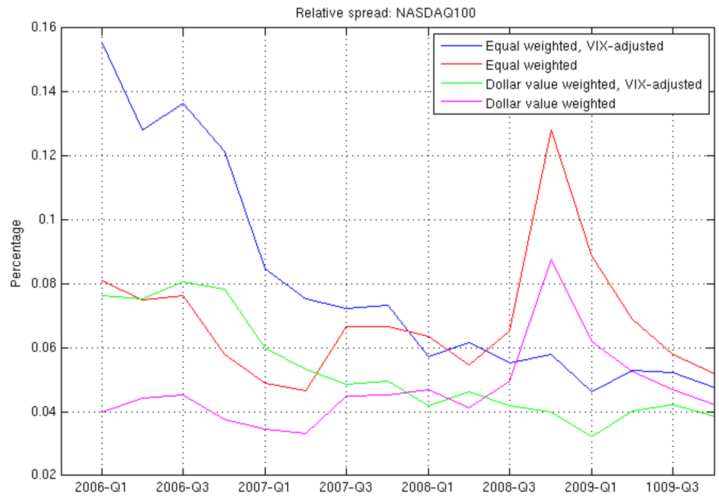

Figure 17: Bid-ask spread for NASDAQ 100

where w_i is the weighting for stock i. A common adjustment made is a capitalization adjustment, which is done by setting w_i to the market capitalization of a stock i in quarter q.

The main body of this paper presents results for the Russell 1000 and Russell 2000. For comparison, the available liquidity for the NASDAQ-100 is presented in Figure 18, showing both a capitalization-weighting and an equal-weighting. In both cases, the general trend of increasing available liquidity over the period studied is seen.

6.3 Market Efficiency

The methodology used to compute the variance ratio values follows that presented in [5]. In particular, equations (12a) and (12b) are used. The raw variance ratio r_i for a stock i with time-ratio D is given by

$$r_i = \frac{v_i^{s_1}}{D v_i^{s_2}},$$

where v^{s_1} is the variance for sampling rate s_1 and v^{s_2} is the variance for sampling rate s_2 and by convention, $\frac{s_1}{s_2} = D > 1$.

In order to gain a sense of the impact of bid-ask bounce and spreads on variance ratios, Figure 21 presents the raw variance ratios for the NASDAQ 100 using the last traded price and the midpoint price in the same figure. From the left panel, showing a fine sampling rate, it is seen that the impact of the bid-ask bounce on last trade prices results in a smaller variance ratio than when midpoint prices are used. As the sampling rate is decreased to longer

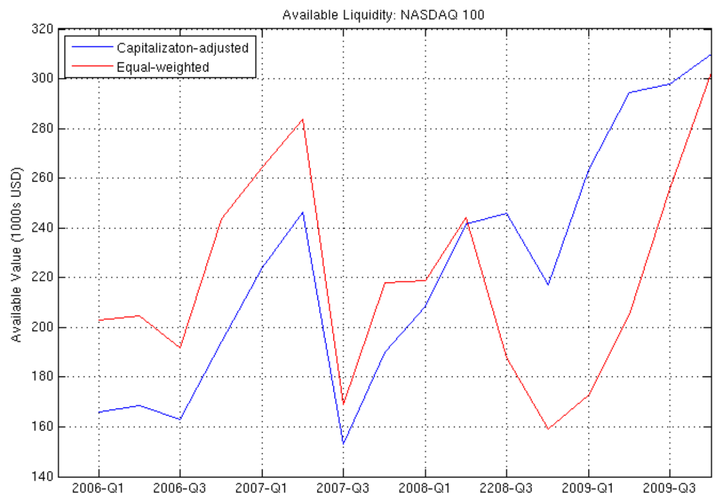

Figure 18: Mean available liquidity for NASDAQ 100

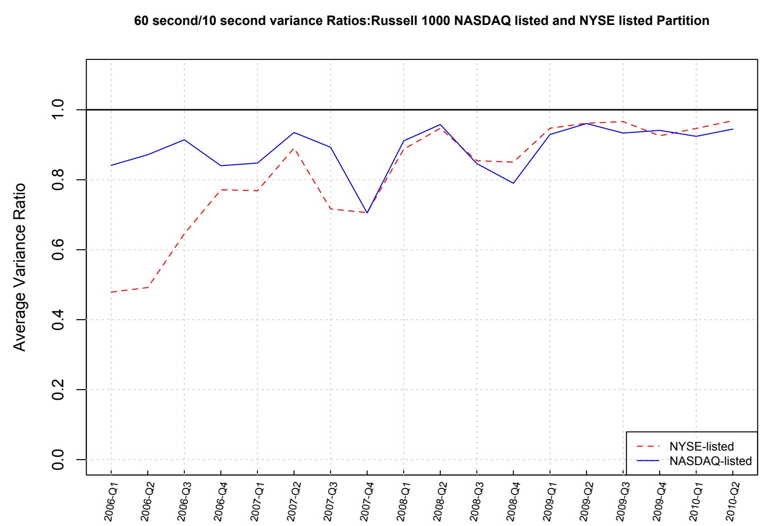

Figure 19: Variance Ratios, Russell 1000, 1 minute / 10 seconds

time periods, the impact of bid-ask bounce becomes less pronounced. This is demonstrated in the right panel of Figure 21, where the difference between the variance ratios using trade prices and midpoint prices is much smaller.

A small sampling interval of 10 seconds is given here for the Chow-Denning tests, and the results of these computations are presented in Figures 22 and 24 for the Russell 1000 and Russell 2000 NASDAQ-listed stocks, respectively. At this sampling rate the impact of microstructural noise is expected to have a more significant impact than at 10 minute sampling. Despite a higher degree of apparent inefficiency, Figures 22 and 23 demonstrate that even at such fine sampling, the Russell 1000 appears to have improved over the four years studied, and that the NYSE-listed symbols have shown a more dramatic improvement in that time, largely converging with the NASDAQ-listed symbols. Similar observations are made for the Russell 2000 index in Figures 24 and 25.

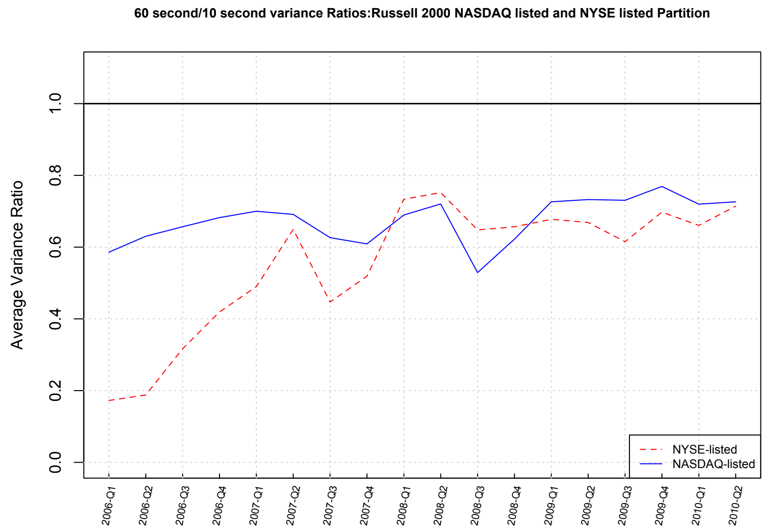

Figure 20: Variance Ratios, Russell 2000, 1 minute / 10 seconds

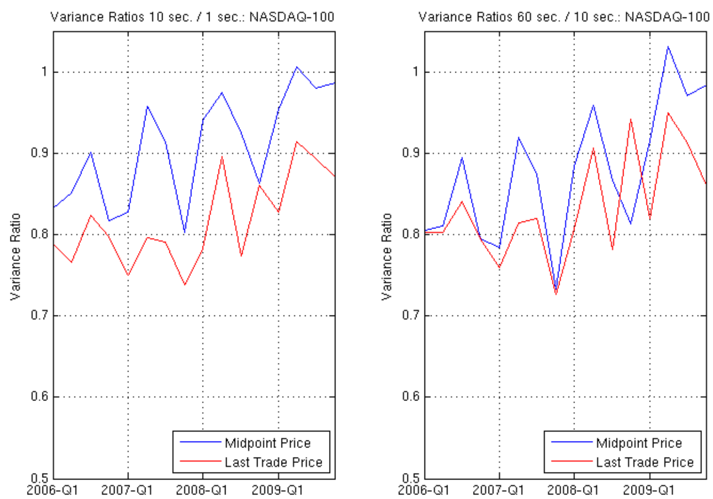

Figure 21: Mean Variance Ratios of Midpoint Prices vs. Trade Prices, NASDAQ 100. Left: 10 seconds / 1 second. Right: 1 minute / 10 seconds

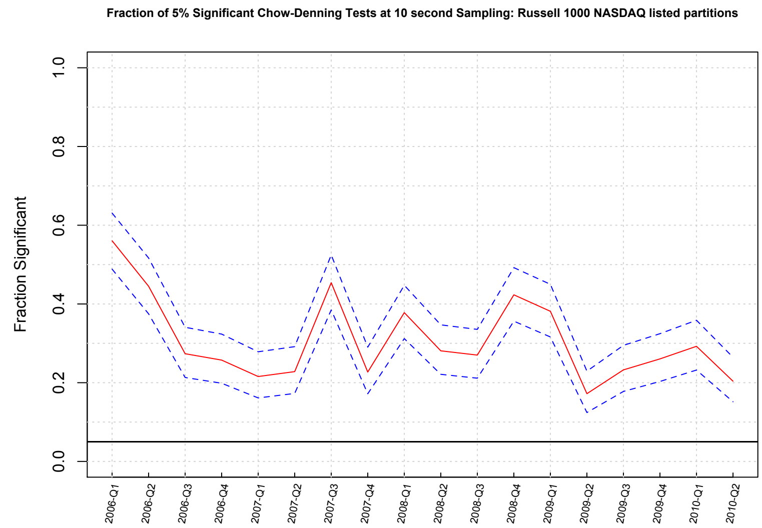

Figure 22: Chow-Denning test results for the Russell 1000 NASDAQ-listed stocks, 10 second sampling

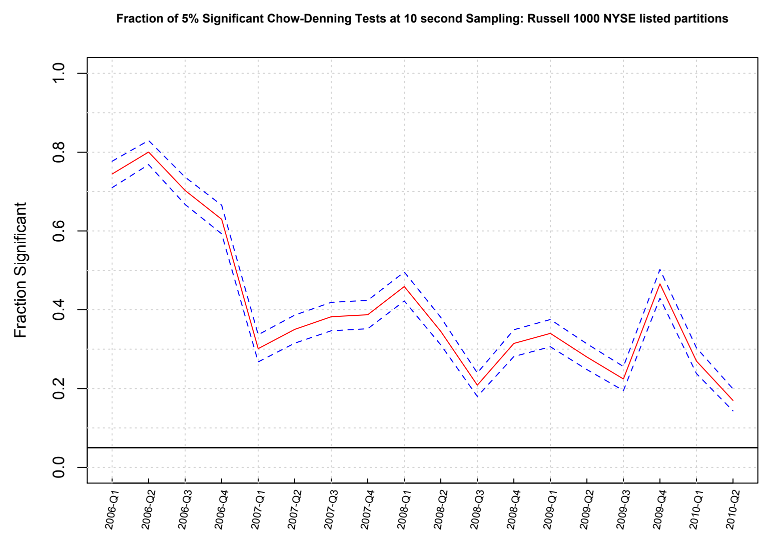

Figure 23: Chow-Denning test results for the Russell 1000 NYSE-listed stocks, 10 second sampling

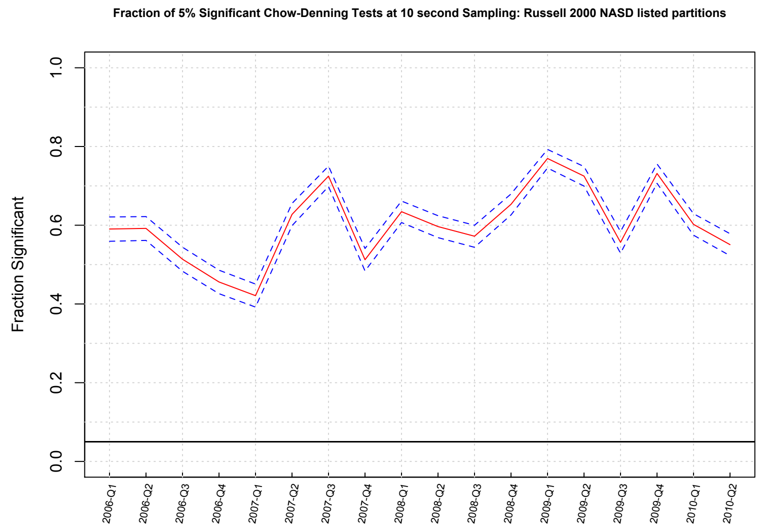

Figure 24: Chow-Denning test results for the Russell 2000 NASDAQ-listed stocks, 10 second sampling

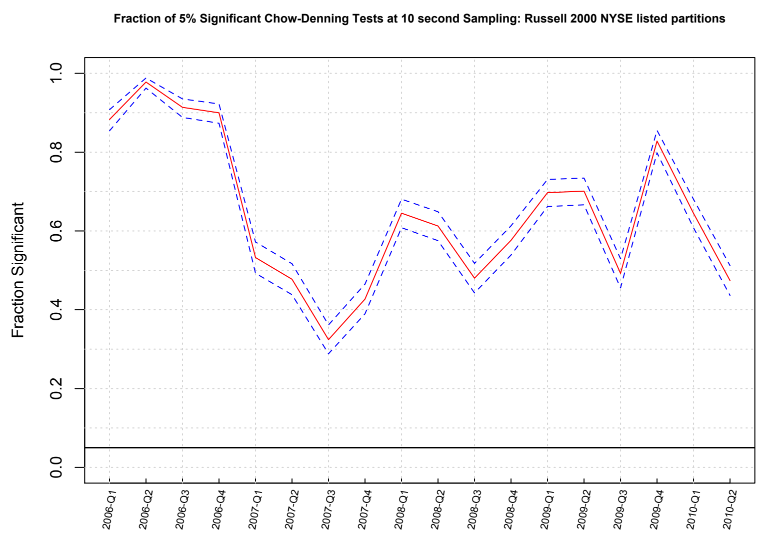

Figure 25: Chow-Denning test results for the Russell 2000 NYSE-listed stocks, 10 second sampling