

Human bias in algorithmic trading

John Broussard, PhD
Rutgers School of Business –Camden
227 Penn St, Camden, NJ, 08102
(856) 225-6647

Andrei Nikiforov, PhD
Rutgers School of Business –Camden
227 Penn St, Camden, NJ, 08102
(856) 225-6594

Abstract

This paper documents a stark periodicity in intraday volume and in the number of trades. We find activity in both variables spikes by about 20% at regular intervals of 5 or 10 minutes throughout the trading day. We argue that this activity is the result of algorithmic trading influenced by human traders/programmers' behavioral bias to transact on round time marks. An alternative explanation, that algorithms choose to concentrate their trades in time to take advantage of lower costs or to protect themselves from better informed traders, is not supported.

v. 4 (December 30, 2013)

Human bias in algorithmic trading

Introduction

In this paper we document pervasive periodicity in intraday trades and volume. We show that an unusually large percentage of trades is completed within just 30 seconds after the end of a “round” time interval such as 9:45 AM or 10:00 AM. We show that throughout the trading day, both the number of trades and the trading volume spike sharply every five or ten minutes only to return to normal levels 30 seconds later (see Figure 1). We investigate this phenomenon and test two competing hypotheses about the nature of such clustering. We conclude that the observed phenomenon is a representation of humans’ tendency towards round numbers, whether these humans are traders or programmers of trading algorithms. We test and do not find support for a “rational” explanation, which argues that traders and algorithms cluster their trades at certain time intervals to receive the benefits of increased liquidity and reduced transaction costs during these short-lived spikes.

Various periodical patterns in stock characteristics have been documented in the literature. Researchers have concentrated on periodicity in, mainly, returns on different scales: from the very long term (DeBondt and Thaler, 1985, 1987), to the medium term (Jagadeesh, 1990, Jagadeesh and Titman, 1993, 2001), to the relatively short-term such as weekly and daily (Gibbons and Hess, 1981), to intraday (Heston, Korajczyk, and Sadka, 2010). We report a completely new phenomenon not previously described in the literature, the sharp spikes in trading activity on 5 minute time marks within a trading day. The purpose of this paper is to evaluate various explanations for the observed intraday recurring periodicity in trading activity. We investigate two primary hypotheses to explain the observed phenomenon. The first

hypothesis appeals to human nature's tendency to use round numbers when faced with decisions in an unstructured and uncertain environment (e.g Loomes, 1988). We hypothesize that buy-side institutional traders faced with a decision to space the execution of orders throughout the trading day, will tend to pick round numbers and intervals for when the trade should be executed. Even if the execution is implemented through a trading algorithm that breaks down a large order into smaller ones, the default settings for the timing of the smaller trades are set by programmers and are thus subject to the same "round –number-preference" bias exhibited by all humans.

The second hypothesis states that the short bursts of clustered trading activity (within 30-second time buckets) every 5 minutes are the outcomes of a semi-spontaneous market coordination game where the players, or traders represented by optimizing algorithms, come to trade every five or ten minutes because these are the time periods with the highest liquidity and lowest trading costs. This hypothesis indicates traders, or their algorithms, rationally choose to trade in short bursts on a periodic schedule to improve the execution of their trades.

We evaluate these two hypotheses using a unique dataset courtesy of Nanex. The time period of analysis covers over two and one half years (2004-2006), and incorporates all trades executed from all exchanges in 33 securities. Although we focus solely on underlying equity transactions, the 33 securities were chosen from those having high option trading activity over the period investigated. We do not find support for the rational hypothesis, that trading is concentrated due to liquidity benefits. Our measures of liquidity, such as depth and spreads, remain unchanged during the short spikes of trading following every five minute mark. Our measure of price impact, based on Amihud's (2002) illiquidity ratio, actually worsens following full hour, half hour, and ten minute marks. Thus, there appears to be no liquidity advantage during the clustered 30 second spikes over any other 30 second period during the trading day. In

contrast we find that the recurring periodicity of crowded trading activity increases as the “roundness” of time marks increases. One hour marks are the most attractive for clustered activity, followed by half-hour marks and then by 10-minute and finally, 5 minute marks. This preference of increasing roundness points to the human nature of the observed phenomenon.

Although algorithmic execution of trades has become the dominant way of conducting financial transactions over the past several years, very little is known about the influence of this shift on market behavior. We contribute to the literature on algorithmic trading by connecting the recurring periodicity in trading activity to human behavioral preferences, as well as to rationally motivated explanations. We show that even as financial trading moves away from direct human execution, it is still influenced by human behavioral preferences.

One alternative that could somewhat explain the observed phenomenon is the interaction of human biases and algorithmic optimization. In this view, the algorithms utilize the human bias toward round numbers as a coordinating device to conduct transactions. It is not clear, however, what the algorithms are optimizing since price impact worsens and the bid ask spread is not reduced. We do not pursue this hypothesis in the current article.

The paper proceeds as following. The next section describes the two hypotheses with corresponding literature review. Section three describes the data and methodology. Section four presents univariate and multivariate analyses to test our main hypotheses. Finally, section five offers discussions and conclusions.

Hypotheses Development

Round numbers hypothesis

People prefer to use round numbers when faced with a task of estimating a quantity in an uncertain environment (Loomes, 1988). This preference is manifest in many financial markets as the uneven distribution of price endings (e.g. Ikenberry and Weston (2007) show that in the stock market, most price quotes end on .05 and .10). Rounding preference is also documented in the clustering of trade size (Alexander and Peterson, 2007), when trades cluster on hundreds such as 100 shares, 200 shares, 500 shares, 1,000 shares etc., and bank rates ((Kahn, Pennachhi, and Sopranzetti (1999)). The reasons behind such clustering are not entirely clear. Ball et al (1985), point out that prices can only be quoted as precisely as value can be estimated. When value is estimated with high precision, a finer pricing grid is favored. When value is highly uncertain, a coarser pricing grid is used. Ball et al (1985) posit an “informational equilibrium,” in the sense of Grossman and Stiglitz (1980), where information is incorporated in price up to the point where the marginal cost of acquiring and using information equals the marginal benefit of doing so. The degree of price resolution and fineness of the pricing grid depends on the extent to which information is incorporated to precisely estimate price. Thus, Ball et al (1985) hypothesize that the tendency to round estimated values and, therefore, employ only a subset of available price quotes, varies with the amount of information in a market, the flow of information into the market, and the cost of introducing information into the market.

Motivated by Ball et al. (1985), other authors have investigated the relationship between clustering and price volatility in a variety of markets. The focus on this relation is based on two assumptions. First, the underlying value of the asset investigated is always uncertain. Second, the psychology and marketing literature show that humans tend to round estimates when dealing

with uncertain values [see e.g. Dehaene and Mehler (1992) and Schindler and Kirby (1997)]. The relation between price uncertainty and clustering is sometimes estimated, as in Ball et al. (1985), in a time series for a single asset. More often, the relation is estimated using a cross-section of assets within a market, as in Harris (1991), or with cross-market comparisons, as in Grossman et al. (1997).

The difficulty of explaining clustering increases when one considers information processing and difficulty in communicating indicators of value. Goodhart and Curcio (1991) propose that prices obtained from continuously distributed underlying values are rounded to the closest “round” number to enhance storage and communication (decision making) capabilities. The rounding depends on the attractiveness of the closest number and distance to the rounded value.

Aitken et al (1996) argue that clustering is the result of uncertain values (haziness) coupled with the natural attraction of round numbers. They test “the attraction hypothesis” that people are most attracted to integers (or zero endings), then 5s, then 2s, and so on. They find that clustering increases with the price of the stock (scaling effect) and with proxies for haziness (stock market volatility, own stock volatility, trade size, and bid-ask spread). Clustering decreases with trading frequency, for stocks with options, and for stocks that can be sold short (because trading increases the efficiency of price discovery and therefore decreases haziness). Mitchell (2001) reviews the literature on clustering and psychological barriers and argues that clustering is rooted in anchoring and simplifying algorithms for traders to produce quicker, more cost-effective (decision making) results.

Brown et al (2002) look at the cultural effects on clustering. They conjecture (following Van Newman and Mortensen (1953)) that humans operate in a sphere of haziness and they make

decisions about equivalence intervals rather than point estimates (MacCrimmon and Smith (1986)). However, they store these intervals as a single rounded value, where the rounded number from the interval serves as an anchor of value. The degree of roundness depends on the riskiness of the decision (Loomes (1988)), and is based on the “attraction hypothesis.” That is, the rounded anchor is based on two conditions: 1) being close to the center the interval, and 2) being the most preferred number. The wider the interval, the more likely a rounded number is selected as an anchor for the interval.

There are no studies investigating human preference for round numbers on time scales. However, day-to-day experience suggests a natural order of preference for time marks on the time line. When forced to come up with a quick estimates of certain activities people naturally chose the following order of preference: full hour, half-hour, quarter-hour, ten minutes, and five minutes. We take this preference order as an assumption when we evaluate our hypothesis.

To summarize, a diverse literature from financial economics, psychology, and marketing suggests that because humans have limited capacity to process quantitative information, they may “chunk” or “lump” activity by using round numbers. The tendency to round is so pervasive that it has become a default mechanism when dealing with quantitative information. From a financial market trading perspective, the direct result of such tendency is that even when trading is done through a computer algorithm, the preference for rounding enters via default settings from the algorithm’s programmers. The round numbers hypothesis thus predicts a preference for rounding in trading activity. We provide support for this hypothesis by showing rounded time periodicity in trading activity.

Rational concentration hypothesis

The second hypothesis proposes that the observed clustering of trades and volume is a rational reaction of market participants to an increasingly “toxic” trading environment when information advantage can be utilized in a matter of milliseconds (e.g. Easley, de Prado, and O’Hara (2011)). By clustering their trades together, liquidity (or uninformed) traders protect themselves from better informed rivals. In this setting, clustering increases liquidity by decreasing the price impact of trading.

Admati and Pfleiderer (1988) use the Kyle (1985) and Glosten and Milgrom (1985) frameworks to develop a model that explains why clustering of volume might occur during some intraday periods. They use a notion of discretionary liquidity traders, a type of liquidity traders who can time their transactions throughout the trading day. They show that such traders prefer to execute their transactions when the market is “thick,” when their trading has little or no impact on prices. Given this incentive they will naturally “band” together to increase overall liquidity and to decrease the cost of trading with informed traders. While informed traders prefer “thick” markets as well, the competition among them drives the value of their common private information down further reducing the disadvantage of liquidity trading. Admati and Pfleiderer’s model predicts that in equilibrium, liquidity costs, as represented by Kyle’s λ , would be lower when volume is higher. Their model was used to explain the much higher volume at the opening and the closing of a trading day.

Foster and Viswanathan (1990) use a similar model to explain a lower trading volume on Mondays. They argue that the longer the market stays closed the bigger the advantage of informed traders over liquidity traders when the markets open. When discretionary liquidity traders can delay their trading, they will rationally choose to avoid Mondays. This incentive

leads to volume concentration on Tuesdays which results in lower price impact of their trading. While the time horizon in Forster and Viswanathan's model is longer than in Admati and Pfleiderer, they both point to the same mechanism of volume concentration: the presence of discretionary liquidity traders who seek to trade during periods of low price impact.

Back and Pedersen (1998) introduced long-lived information into Admati and Pfleiderer's model and show that clustering in liquidity trades induces clustering of information use and volatility. However, they find no systematic pattern in the price impacts of orders and argue that the timing of long-term information arrival is unimportant. Brooks and Su (1997) develop a model showing that the best time to trade for the small liquidity trader is at the opening call auction when both types, informed and uninformed investors, mix their trades. Admati and Pfleiderer is the most appropriate rational framework to explain the observed pattern of spikes, because it can be easily adjusted to intervals shorter than one trading day. In effect, Admati and Pfleiderer string together a sequence of one period Kyle-style models, each period representing one day. We can shorten these periods to represent 5 minute intervals. One concern for shortening the intervals to just 5 minutes is that to generate the volatility (and volume) patterns in the model, private information needs to vary inside the five minute interval. We do not see this as a problem. In fact, given that during the study period (2004-2006) the amount of high-frequency trading reached over 50% of overall trading volume (Brogaard, Hendershott, and Riordan (2012)), it is quite reasonable to assume that the amount of private information (e.g. about the order flow in each stock) used by high-frequency traders for trading decisions would change in a matter of seconds or even milliseconds.

To summarize, the rational concentration hypothesis based on an adjusted Admati and Pfleiderer's model predicts that the chief reason for trade concentration is that liquidity will improve significantly during the observed spikes in volume and trades.

Data and Methodology

Data description

The data used in this study are trade and quote data for 30 liquid U.S. stocks and 3 ETFs over the period April 17, 2003 to October 18, 2006, or a total of 882 trading days. The data were obtained from Nanex, which provides real-time option and stock price data to its customers via its NxCore product. The data were archived by Nanex as they arrived from the exchanges at Nanex's server, and time-stamped by Nanex to 25 millisecond precision as they arrived. The data come from all U.S. exchanges where a given instrument is traded. For trades, transaction price, size, exchange code, and various other information are available. For quotes, exchange-level best quotes and volumes are available. The data include each instance when any exchange adjusts its best quote or quoted volume, even if this change does not change the national best bid and offer (NBBO). The main reason for limiting the sample size to 33 stocks and ETFs was a data storage limitation. The size of the compressed database exceeds 1,400 GB.

Methodology

To investigate the round numbers hypothesis, e.g. clustering due to the default preference for round numbers, we examine whether clustering is progressively more intense as the markers on the time line become more and more round. The "most round" number is a whole hour. There are 6 such markers during the trading day: 10:00 AM, 11:00 AM, 12:00 PM, 1:00 PM, 2:00 PM, and 3:00 PM. We remove the first and last 5 minutes of trading from our analysis because the price formation and trading mechanism during the first and last minutes differ from those during

the rest of the day. The next level of rounding would correspond to the half-hour marks with the first being 10:30 AM followed by 11:30, 12:30 PM, 1:30 PM, 2:30 PM and, finally, 3:30 PM. The next level of rounding are non-overlapping 10 minutes endings and the last level of clustering that we test is the non-overlapping 5 minutes endings. We conduct both univariate and multivariate analyses on the rounded time marks.

To investigate the rational concentration hypothesis, e.g. that clustering of volume and trades is caused by rational optimizing behavior of discretionary liquidity traders, we calculate several measures of liquidity and check if those measures indicate a significant improvement in liquidity on the round time marks. First, we calculate the measure of price impact using Amihud's (2002) illiquidity measure. We also report the measures of transaction costs measured by spreads, absolute and proportional, and depth, the quoted number of shares. The measures are standard in the literature (e.g. Chordia, Roll and Subrahmanyam (2000)) and are defined as follows:

QSPR – quoted spread defined as $P_{ASK} - P_{BID}$ and measured in dollars,

PQSPR – proportional quoted spread defined as $QSPR / P_{Midpoint}$, where $P_{Midpoint}$ is the midpoint price between P_{ASK} and P_{BID} ,

ESPR – effective spread defined as $2 \times |P_T - P_{Midpoint}|$, where P_T is the trade price,

PESPR – proportional effective spread defined as $ESPR / P_T$

DEP – Depth defined as $\frac{1}{2} \times (Q_{ASK} + Q_{BID})$ and measured in number of shares, with $Q_{ASK, BID}$ being the shares offered for trading at the ask price and at the bid price respectively.

ALIQ is the measure based on Amihud's illiquidity measure (Amihud (2002)) calculated as the ratio of returns for each individual stock over 30-second intervals to the dollar

trading volume during the same interval. The ratio is then averaged across all stocks for each time interval.

All variables are winsorized at 0.5% and 99.5% to exclude extreme observations. We conduct both univariate and multivariate analyses on the liquidity measures to evaluate the rational concentration hypothesis.

Results

Spike Periodicity

Figure 1 shows the typical (average) volume over a trading day for our entire sample. The volume is aggregated across 33 securities for each 30-second interval for each day and then averaged across all days in our sample. Volume exhibits the familiar U-shaped pattern with very high opening activity that declines to a minimum point around lunch time. Activity increases again until closing at 4 PM. While the U-shaped pattern is not surprising, the sharp spikes in volume that appear with stark periodicity every 5 to 10 minutes have not been reported in prior research. The robust periodicity of spikes implies a systematic origin of this phenomenon, which is the main focus of this paper.

To show that the spikes are indeed highly regular and not just a case of random patterns, we transform the average 30-second aggregate volume into percent changes over the previous 30 second time interval. The results are presented in Figure 2. As one can see, volume spikes occur at highly regular intervals. All large spikes (except for the last half hour) are exactly 30 minutes apart. During the last half hour of trading day, the spikes are exactly 10 minutes apart. The smaller spikes are exactly either 5 or 10 minutes apart. The sharp downward spikes correspond to the 30 second time buckets immediately following the sharp upward spikes and represent the return of volume back to normal levels.

The next two figures, Figure 3 and Figure 4, present the same information for the number of trades. Figure 3 shows that the pattern in the number of trades mirrors that of volume, with sharply higher number of transactions at the beginning and the end of the trading day. Figure 4 shows that the trading spikes are not caused by previously generated trades executing on the round time marks.

Table 1 shows the basic characteristics of trading activity in our sample. On average, there are 1,195 trades conducted during any given 30-second time bucket for all 33 securities. Each trade is preceded by an average of 13 quote messages from all major exchanges. The average volume of transactions inside a 30-second time bucket is over 632 thousand shares. These securities are very liquid, with average quoted spread and average effective spread of 4.5 cents, and 2.1 cents, respectively. In terms of percentages, the quoted spread as a percentage of security's price is about one sixth of a penny, with the proportional effective spread half that amount (0.07 of one penny). An average of over 6,800 shares quoted at the bid and ask also indicates highly liquid securities, these are very liquid securities. The Amihud price impact ratio ranges from 3.97 (very low price impact) to 98.99 with an average of 21.24 and the standard deviation of 11.62.

Testing hypothesis I (preference for round numbers)

In this section we test our first hypothesis, that the spike periodicity in volume and trades, is driven by humans' preference for round numbers, even if trading has become increasingly automated. Brogaard, Hendershott, and Riordan (2012) state that close to 70% of NASDAQ dollar trading volume was attributed to high-frequency trading firms alone. Hasbrouck and Saar (2012) further highlight that NASDAQ reporting almost entirely omits the activity generated by proprietary trading desks of large banks as well as by direct access brokers. Algorithmic trading

is indicated in our sample of 33 actively traded securities during the 2004 -2006 period by the fact that activity is spaced by such precise time intervals (multiple of 5 minutes). It is difficult to explain the persistent recurring periodicity of direct human trading activity without relying on an assumption that the activity is computer-generated. Since computer algorithms are developed by human beings (as far as we know), the algorithms generating trading activity must have some sort of time-spacing default setting. We argue that since humans have a strong bias toward utilizing round numbers, the programmers will have a tendency to incorporate rounding into their algorithms. The most obvious manifestation of the human influence on the algorithmic activity is transacting on round time marks.

To test this hypothesis, we regress the volume and trade variables on various non-overlapping time dummies. Time dummies correspond to round hour (10 AM through 3 PM) to half-hours (starting from 10:30, 11:30, up to 15:30), fifteen minutes that are not also hours and half hours (i.e. 9:45, 10:15, 10:45 up to 15:15), ten minute dummies that are not also hours and half-hours (9:40, 9:50, 10:10, 10:20, 10:40 up to 15:20), and five minute dummies (9:55, 10:05, 10:25, 10:35, 10:55 up to 15:25). The results of the regression analysis are presented in Table 2.

The percent increase columns show how the average volume (or the number of trades) increases on the corresponding time marks. The most obvious pattern in the table is the monotonic increase in average volume and in the number of trades as we progress from 5 minute time marks (e.g. 9:55) through 10- 15- 30- and full hour marks. The increase for the five minute marks barely registers at half a percentage point and it is not statistically significant at conventional levels of significance. However, that increase jumps tenfold for the 10-minute marks (e.g. 9:50) and then jumps again for the fifteen and thirty minute marks reaching as high as 15.7% higher on the full hour marks for volume and 18.3% higher for the number of trades.

All ten-, fifteen, thirty-, and whole hour jumps are highly statistically and economically significant. The robust monotonicity in the increase of volume and the number of trades that mirrors the human tendency to prefer round numbers points to the biased-based mechanism for the origin of this periodicity. Regressions with the changes in volume and the number of trades over the previous 30-second intervals produce similar results and are not reported here.

Testing hypothesis II (higher liquidity)

The second hypothesis is based on Admati and Pfleiderer (1988) and states that the concentration of trading activity on round time marks is a rational response to a “toxic” environment. Discretionary liquidity traders “band together” to protect themselves from the predatory trading of insiders. They show that such concentration of discretionary liquidity traders has the benefits of higher liquidity.

We test this hypothesis by first, calculating 5 measures of transaction costs and one measure of price impact and, second, running regressions where we regress each liquidity variable on various time dummies. We use both the levels of liquidity measures and their first differences. Table 3 shows the results.

Looking at the five measures of transaction costs (the first five columns), one can see that none of the coefficients is statistically different from zero. Moreover, there is no clear pattern in the coefficients’ signs. For example, depth is decreasing and all spreads are increasing (albeit insignificantly) on the hour marks precisely when trading activity is at it’s highest. The picture reverses, however, for the half hour and for the fifteen minutes marks where the signs for all measures are pointing toward improving liquidity. Then again, the picture reverses to show declining activity for the ten- and five minute marks. Overall, the evidence in the table tell us

that liquidity, the foundation of the rational hypothesis, is in general not affected on the round time marks, despite sharply increasing trading activity.

The last column of Table 3 presents results for our measure of price impact. Contrary to the hypothesis' predictions, the price impact significantly increases on the full hour, half hour and ten minute marks. The impact increases by over 10% for the full hour and ten minute marks and by over 8% for the half hour mark. However, the price impact improves (decreases) for fifteen and five minute marks by between 2.3 and 4.5%. Therefore, on balance, the price impact as measured by our version of Amihud ratio increases, leading to worse terms of trade during the spikes.

The negative results for this hypothesis are in agreement with other empirical work that investigates the relation between trading activity and the cost of transacting. For example, given the U-shape of trading activity inside a trading day, Admati and Pfleiderer's model predicts an inverted U-shape for transaction cost and price impact. However, Madhavan, Richardson and Roomans (1998) show that the bid-ask spread is also U-shaped and Ferguson et al (1996) find no relationship between volume and execution costs.

To summarize our testing, we are inclined to attribute the sharp spikes in trading activity on round time marks to the influence of human preference for rounding rather than to a rational optimizing activity of traders and their algorithms. Trading activity is a result of human preference for regularity, not necessarily a result of trading in an increased liquidity environment.

Discussion and Conclusion

In this paper we document a startling periodicity in trading activity that occur in the financial markets. We show that trading activity explodes by double digit percent on round time marks. An explosion of activity can be detected in much higher number of trades and trading volume that manifest every five or ten minutes. We investigate and test two hypotheses about the origin of this phenomenon. We conclude that even though trading is executed automatically, it is the influence of the human preference toward using round numbers that drives this periodicity. We consider, but do not find support for a rational explanation of the observed pattern. The rational hypothesis states that improved liquidity during those short lived moments attracts traders to transact on specific time marks. In this paper, we show that five commonly used measures of liquidity do not change in any significant way during the spikes.

The findings in Alexander and Peterson (2007) point toward the behavioral nature of clustering. Even though they look at a different dimension of trading, trade size clustering, they find that when institutions are engaged in stealth trading and break large orders into smaller orders, they disproportionately prefer the round numbers. For example, they find that there significantly more trades for 3,500 shares than for either 3,400 or 3,600 shares.

An intriguing alternative that could be investigated in the future is the combination of the two explanations. If only some algorithms have rigid round time marks for executions as their default settings, they can nonetheless influence the behavior of “smarter” algorithms that do not have such constraints. The reason for such influence could be the commonly used VWAP quality of execution benchmark. If some (maybe just one) algorithm trades on round time marks, it will attract the counterparties to trade with it. This increasing activity attracts ever more participants driven by the demands of VWAP performance resulting in spikes. In effect, a single algorithm

designed to execute on round time marks or a single large human trader will lead to a sunspot – type equilibrium (Cass and Shell (1983)) when the whole market starts trading on round clock marks. We leave this investigation to future efforts

References

- Admati, Anat, and Paul Pfleiderer, A theory of intraday patterns: Volume and price variability, *Review of Financial Studies* 1, 1988, 3–40.
- Aitken, P. Brown, C. Buckland, H.Y. Izan, T. Walter Price clustering on the Australian Stock Exchange, *Pacific-Basin Finance Journal*, 4, 1996, pp. 297–314
- Alexander, G. J. and M. A. Peterson, An analysis of trade-size clustering and its relation to stealth trading, *Journal of Financial Economics* 84, 2007, 435–71.
- Amihud, Y., Illiquidity and stock returns cross-section and time-series effects. *Journal of Financial Markets*, 5, 2002, pp. 31–56.
- Ball, C. A., Torus, W. N. and A. Tschoegl, ‘The Degree of Price Resolution: the Case of the Gold Market,’ *Journal of Futures Markets*, Vol. 5, 1985, pp. 29–43.
- Beaver, W. H., ‘The Information Content of Annual Earnings Announcements,’ *Empirical Research in Accounting, Journal of Accounting Research* (Supplement), 1968, 67-92.
- Brogaard, J., Hendershott, T., Riordan, R., 2012. *High Frequency Trading and Price Discovery*. Working paper, University of Washington.
- Brooks, R.M. and T. Su A simple cost reduction strategy for liquidity traders: trade at the opening. *Journal of Financial and Quantitative Analysis*, 32, 1997, pp. 525–540.
- Brown, P., Chua, A., Mitchell, J., The influence of cultural factors on price clustering: Evidence from Asia-Pacific stock market. *Pacific-Basin Finance Journal*, 10, 2002, 307-332.
- Cass, D. and K. Shell Do sunspots matter? *Journal of Political Economy*, 91, 1983, pp. 193–227
- Chordia, T., R. Roll, and A. Subrahmanyam Commonality in liquidity, *Journal of Financial Economics*, 56, 2000, pp. 3–28
- Chordia, T, Sarkar, A., and Subrahmanyam, A., ‘An Empirical Analysis of Stock and Bond Market Liquidity,’ *Review of Financial Studies*, vol. 18, 2005, pp. 85-130.
- De Bondt, W., Thaler, R., Does the stock market overreact? *Journal of Finance*, 40, 1985, 793-808.
- De Bondt and Thaler, W. De Bondt, R. Thaler Further evidence of investor overreaction and stock market seasonality. *Journal of Finance*, 42, 1987, pp. 557–581
- Dehaene, S., and J. Mehler, ‘Cross-Linguistic Regularities in the Frequency of Number Words,’ *Cognition*,’ vol. 43, 1992, pp. 1–29.

Easley, D., M. Lopez de Prado, and M. O'Hara, The Exchange of Flow Toxicity, *The Journal of Trading*, Vol. 6, No. 2, 2011, pp. 8-13.

Ferguson, M., Mann, S., Schneck, L., Depth and tightness. 1996, Working paper.

Gibbons, M. R. and Hess, P. J., Day of the week effects and asset returns. *Journal of Business*, 54, 1981, 579–596.

Glosten, Larry R. and Paul Milgrom, Bid, ask and transactions prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14, 1985, pp. 71–100.

Goodhart, C.A.E. and R. Curcio, The clustering of bid/ask prices and the spread in the Foreign exchange market, 1991, Manuscript, (London School of Economics).

Grossman, Sanford J., and Joseph E. Stiglitz, On the impossibility of informationally efficient markets, *American Economic Review* 70, 1980, 393–408.

Grossman, Sanford J., Merton Miller, Daniel Fischel, Kenneth Cone, David Ross Clustering and competition in asset markets, ReportLexecon Inc, Chicago, IL (1995).

Gwilym, O. A., Clare, A., Thomas, S., Extreme price clustering in the London equity index futures and options markets. *Journal of Banking and Finance* 22, 1998, 1193-1206.

Harris, L., 'Price Clustering and Discreteness,' *Review of Financial Studies*, vol. 4, 1991, pp. 389–415.

Hasbrouck, J., and G. Saar, 2012. Low latency trading. Working paper, New York University.

Heston, S.L., R.A. Korajczyk, R. Sadka Intraday patterns in the cross-section of stock returns *The Journal of Finance*, 65, 2010, pp. 1369–1408.

Ikenberry D.L. and Weston, J. P., 'Clustering in US Stock Prices after Decimalization', *European Financial Management*, vol. 14, 2007, pp. 30-54.

Jegadeesh, N. , Evidence of predictable behavior of security returns, *Journal of Finance*, 45, 1990, pp. 881–898.

Jegadeesh, N. S. Titman Returns to buying winners and selling losers, *Journal of Finance*, 48, 1993, pp. 65–91.

Kahn, C., G. Pennachi, B. Sopranzetti, Bank deposit clustering: theory and empirical evidence, *Journal of Finance*, 54, 1999, pp. 2185–2214.

Kyle, A., Continuous auctions and insider trading, *Econometrica*, 53, 1985, pp. 1315-1335.

Loomes, G. 1988. Different Experimental Procedures For Obtaining Valuations Of Risky Actions: Implications For Utility Theory. *Theory and Decision* 25: 1-23.

MacCrimmon, Kenneth and Maxwell Smith., Imprecise Equivalences: Preference Reversals in Money and Probability, 1986, University of British Columbia Working Paper 1211.

Madhavan, A., Richardson, M., Roomans, M., Why do security prices change? A transaction-level analysis of NYSE stocks. *Review of Financial Studies*, 1998.

Mitchell, Jason, Clustering and the psychological barriers: The importance of numbers, *Journal of Futures Markets*, 21, 2001, 395–428.

Niederhoffer, Victor, Clustering of stock prices, *Operations Research*, 13, 1965, 258–265.

Niederhoffer, Victor, A new look at clustering of stock prices, *Journal of Business*, 39, 1966, 309–313.

Osborne, M. F., ‘Periodic Structure in the Brownian Motion of Stock Prices,’ *Operations Research*, 1960, pp. 345–79.

Schindler, R. and Kirby, P., ‘Patterns of Rightmost Digits Used in Advertised Prices: Implications for Nine-Ending Effects,’ *Journal of Consumer Research*, vol. 24, 1997, pp. 192-201.

Schwartz, A., Van Ness, B., and R. Van Ness, ‘Clustering in the Futures Market: Evidence From S&P 500 Futures Contracts,’ *Journal of Futures Markets*, vol. 24, 2004, pp. 413-428.

Sopranzetti, B. and Datar, V., ‘Price Clustering in Foreign Exchange Spot Markets,’ *Journal of Financial Markets*, vol. 5, 2002, pp. 411-417.

Von Neumann, John and Morgenstern, Oskar., Theory of games and economic behavior. Princeton, NJ, Princeton University Press. Third Edition 1953.

Figure 1. Intraday periodicity in volume traded over 30 second intervals.

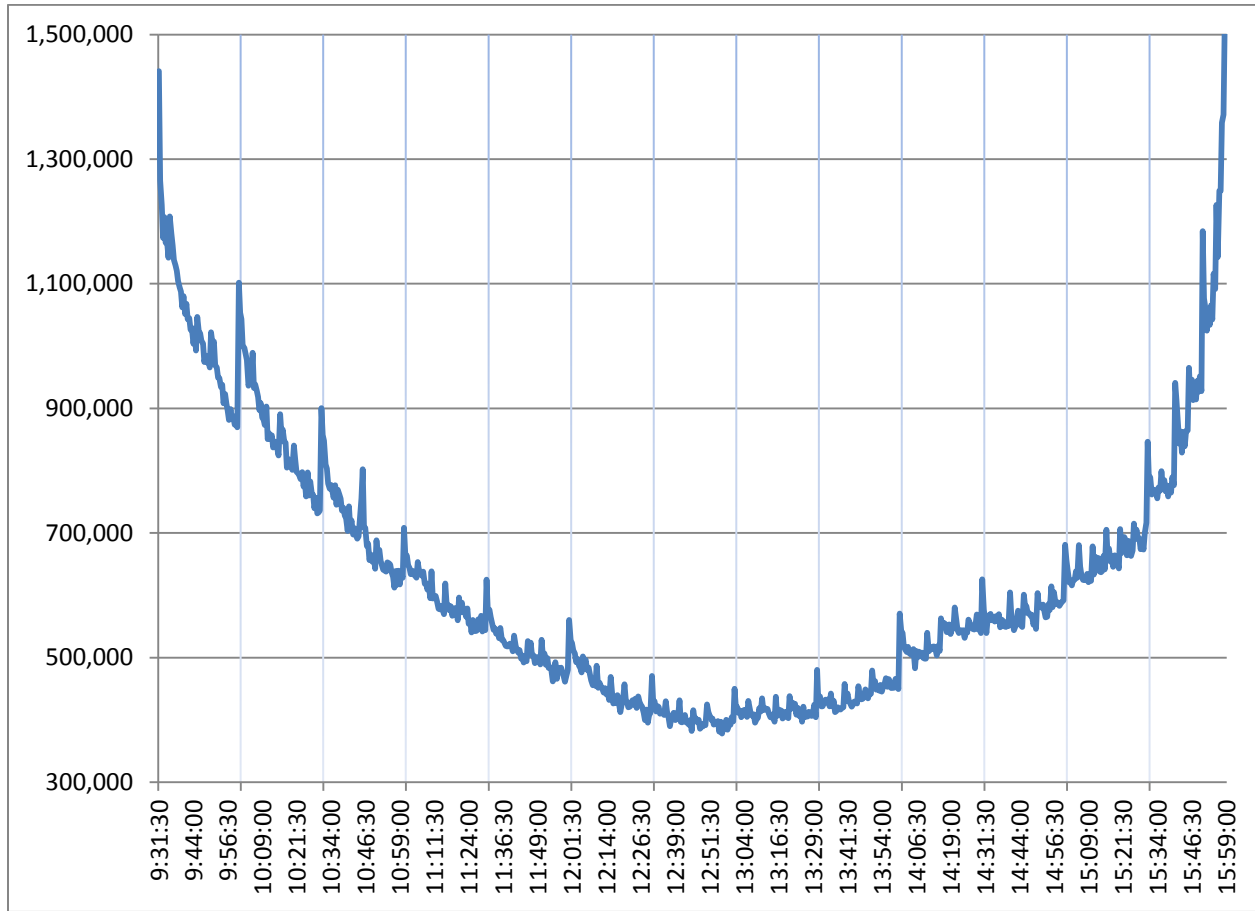
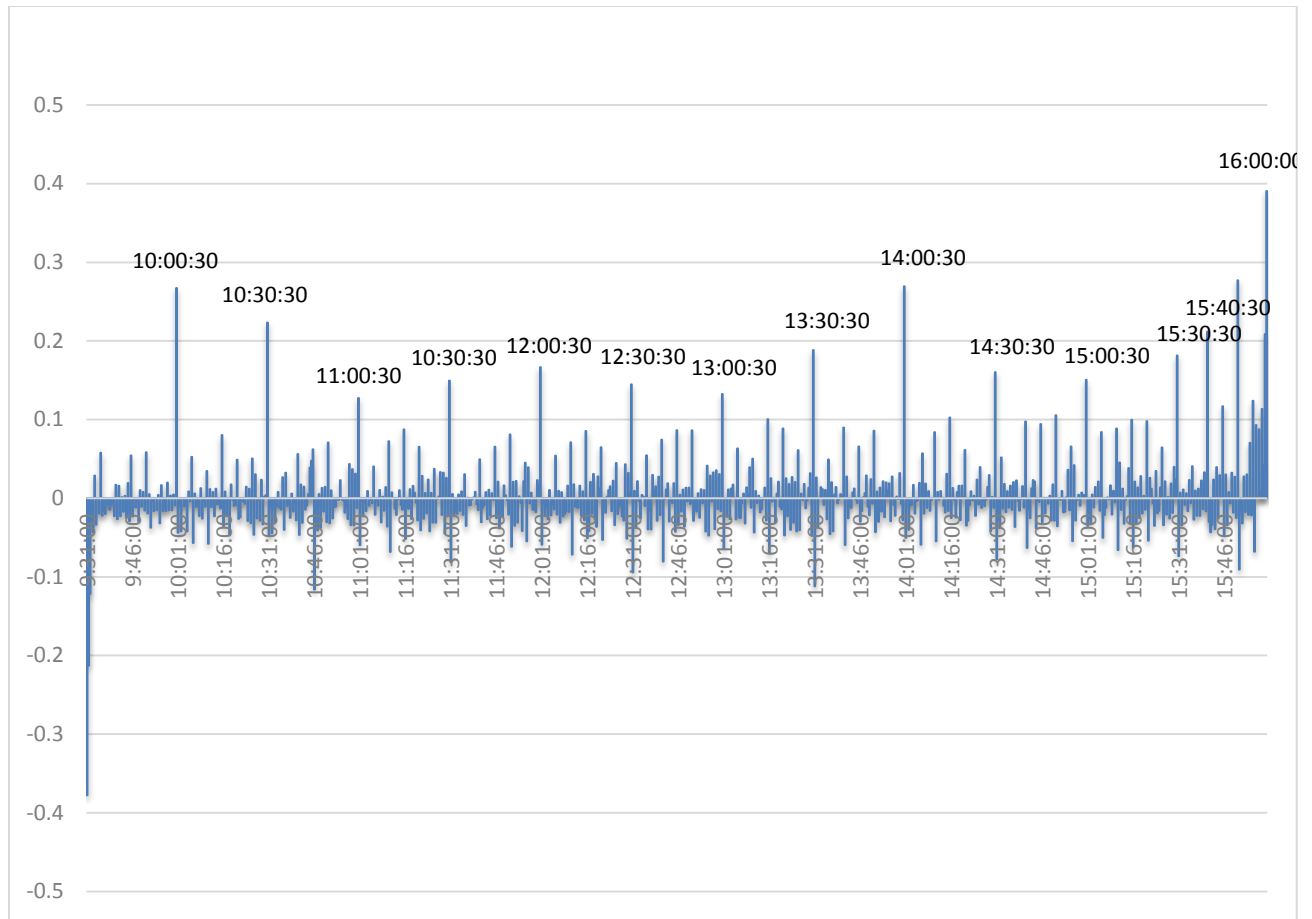


Figure 1 shows intraday periodicity in trading volume throughout the trading day. The volume in shares traded is calculated for 30 second intervals and averaged across all 33 securities and all trading days. The vertical gridlines separate the half hours starting from 9:30 AM.

Figure 2. Changes in Volume for each 30-second bucket, in percent



The figure shows the percentage change in aggregate 30-second volume from the prior 30-second interval. The largest spikes are labeled with the corresponding ending time for that bucket.

Figure 3. Intraday periodicity in the number of trades transacted over 30 second intervals.

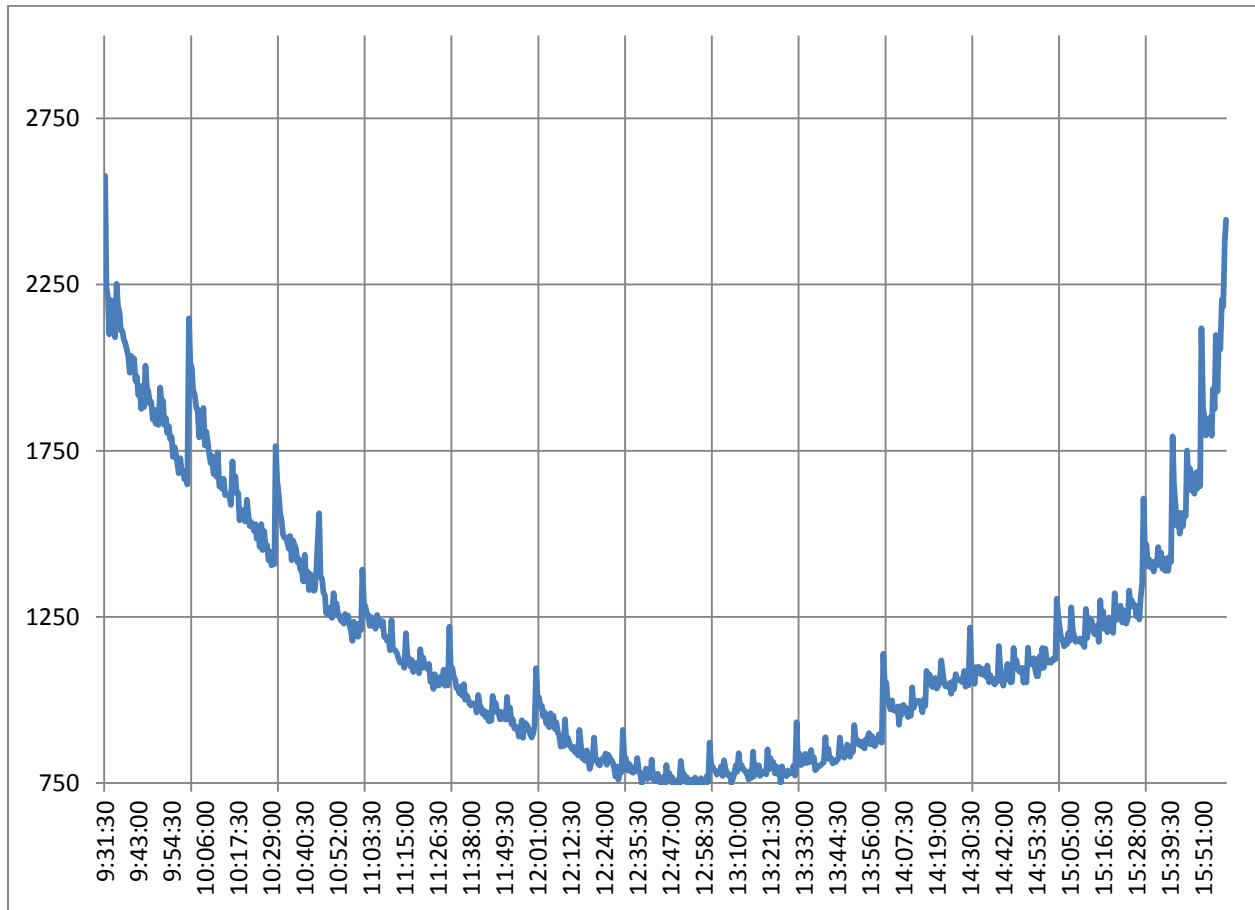
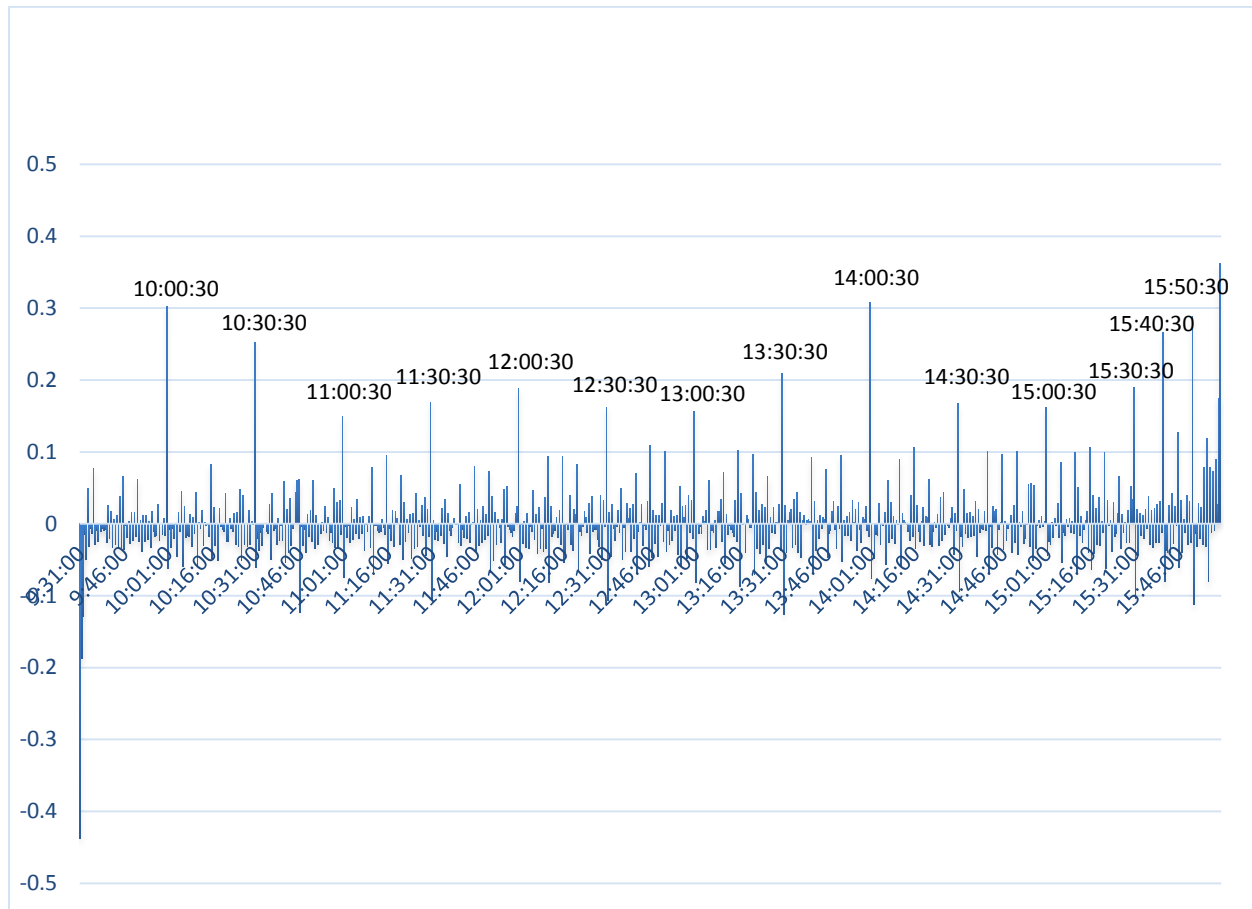


Figure 2 shows intraday periodicity in the number of trades throughout the trading day. The number of transactions is aggregated for 30 second intervals and averaged across all 33 securities and all trading days. The vertical gridlines separate the half hours starting from 9:30 AM.

Figure 4. Changes in Number of Trades for each 30-second bucket, %



The figure shows the percent changes in the number of transactions for each 30-second bucket over the previous 30-second bucket. All large spikes are labeled with the corresponding end time for the specific 30-second bucket.

Table 1. Descriptive statistics for the 30 second intervals

Variable	N	Mean	Std. Dev.	Min.	Max.
Trading Volume	672,042	632,468	377,880	100,245	5,494,7636
Number of Trades	672,042	1,195	713	245	10,560
Depth	617,322	6,887	10,623	100	50,000
Quoted Spread	671,978	0.04	0.11	0.01	1.75
Proportional Quoted Spread	672,042	0.00	0.00	0.00	0.05
Effective Spread	672,037	0.02	0.03	0.00	0.32
Proportional Effective Spread	672,042	0.00	0.00	0.00	0.01
Amihud Illiquidity Measure	672,042	21.24	11.62	3.97	98.99

The table shows the statistics for an average 30 second time intervals. The averaging is over all 33 securities on all trading days in the sample (from April 17, 2003 to October 18, 2006, a total of 882 trading days). The liquidity variables are defined as in Chordia et al (2000). The Amihud measure is defined as in Amihud (2002). All variables are winsorized at 0.5% and 99.5% and the total sample is the intersection of all remaining values (97.7% of the original sample).

Table 2. Regression analysis with time period dummies for volume and trade variables

	Volume	% Increase	Trades	% Increase
Intercept	586,888 (0.00)		1,120.79 (0.00)	
HOUR	91,991 (0.00)	15.67%	204.86 (0.00)	18.28%
HALF HOUR	71,145 (0.00)	12.12%	154.57 (0.00)	13.79%
FIFTEEN	40,821 (0.00)	6.96%	86.84 (0.00)	7.75%
TEN	34,196 (0.00)	5.83%	68.77 (0.00)	6.14%
FIVE	3,190 (0.19)	0.54%	5.57 (0.24)	0.50%
N	577,189		577,189	

The table shows the results of regression analysis of volume and trade variables on non-overlapping time dummies. Time dummies correspond to round hour (10 AM through 3 PM) to half-hours (starting from 10:30, 11:30, up to 15:30), fifteen minutes that are not also hours and half hours (i.e. 9:45, 10:15, 10:45 up to 15:15), ten minute dummies that are not also hours and half-hours (9:40, 9:50, 10:10, 10:20, 10:40 up to 15:20), and five minutes (9:55, 10:05, 10:25, 10:35, 10:55 up to 15:25). Volume and trade variables are 30-second aggregate volumes averaged over all trading days. Percent increase columns show the increase in volume and the number of trades relative to the overall daily average (the corresponding intercepts). The liquidity variables are defined as in Chordia et al (2000) and are calculated for each 30-seconds interval aggregated for each security and averaged across all days. p-values are in parentheses.

Table 3. Regression analysis with time dummies for liquidity variables

	Depth	ESPR	PESPR	QSPR	PQSPR	AMIHU
Intercept	6,767.28 (0.00)	0.0217 (0.00)	0.0007 (0.00)	0.0474 (0.00)	0.0015 (0.00)	21.52 (0.00)
HOUR	-232.37 (0.13)	0.0007 (0.31)	0.0000 (0.56)	0.0011 (0.61)	0.0001 (0.38)	2.23 (0.00)
HALF HOUR	6.9603 (0.96)	-0.0010 (0.16)	0.0000 (0.54)	-0.0032 (0.15)	-0.0001 (0.21)	1.87 (0.00)
FIFTEEN	30.5004 (0.78)	-0.0003 (0.57)	0.0000 (0.05)	-0.0020 (0.21)	-0.0001 (0.28)	-0.98 (0.00)
TEN	72.2285 (0.35)	0.0002 (0.63)	0.0000 (0.42)	0.0003 (0.77)	0.0000 (0.99)	2.36 (0.00)
FIVE	-73.4442 (0.35)	0.0001 (0.85)	0.0000 (0.62)	0.0000 (0.97)	0.0000 (0.56)	-0.51 (0.00)
N	577,189	577,189	577,189	577,189	577,189	577,189

The table shows the results of regression analysis of five liquidity variables on non-overlapping time dummies. Time dummies correspond to round hour (10 AM through 3 PM) to half-hours (starting from 10:30, 11:30, up to 15:30), fifteen minutes that are not also hours and half hours (i.e. 9:45, 10:15, 10:45 up to 15:15), ten minute dummies that are not also hours and half-hours (9:40, 9:50, 10:10, 10:20, 10:40 up to 15:20), and five minutes (9:55, 10:05, 10:25, 10:35, 10:55 up to 15:25). p-values are in parentheses.