Liquidity and Market Efficiency
by
Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam
January 1, 2006

Abstract

The capacity of an asset market to accommodate order imbalances, a measure of market efficiency, is inversely related to the predictability of returns from previous order flows. For a comprehensive sample of NYSE stocks that traded every day from 1993 through 2002, we find that very short-term (five-minute) return predictability is strong on average but has undergone a secular decline interspersed with sizable fluctuations. The evidence suggests that liquidity facilitates efficiency, in the sense that the market’s capacity to accommodate order flow is larger during periods when the market is more liquid. We also find that informational efficiency increases with exogenous decreases in the tick size suggesting that more private information is incorporated into prices as spreads decline.

Contacts

<table>
<thead>
<tr>
<th>Chordia</th>
<th>Roll</th>
<th>Subrahmanyam</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice: 1-404-727-1620</td>
<td>1-310-825-6118</td>
<td>1-310-825-5355</td>
</tr>
<tr>
<td>Fax: 1-404-727-5238</td>
<td>1-310-206-8404</td>
<td>1-310-206-5455</td>
</tr>
<tr>
<td>E-mail: <a href="mailto:Tarun_Chordia@bus.emory.edu">Tarun_Chordia@bus.emory.edu</a></td>
<td><a href="mailto:Rroll@anderson.ucla.edu">Rroll@anderson.ucla.edu</a></td>
<td><a href="mailto:asubrahm@anderson.ucla.edu">asubrahm@anderson.ucla.edu</a></td>
</tr>
<tr>
<td>Address: Goizueta Business School Emory University Atlanta, GA 30322</td>
<td>Anderson School UCLA Los Angeles, CA 90095-1481</td>
<td>Anderson School UCLA Los Angeles, CA 90095-1481</td>
</tr>
</tbody>
</table>

We thank Doron Avramov, Ekkehart Boehmer, Soeren Hvidkjaer, Vojislav Maksimovic, Lemma Senbet, Haluk Unal, and seminar participants at ISCTE, Faculdade Economica, Universid Porto, University of California at Irvine and University of Maryland, for valuable comments. We also thank Morgan Stanley for the market microstructure grant that funded this project.
Liquidity and Market Efficiency

Abstract

The capacity of an asset market to accommodate order imbalances, a measure of market efficiency, is inversely related to the predictability of returns from previous order flows. For a comprehensive sample of NYSE stocks that traded every day from 1993 through 2002, we find that very short-term (five-minute) return predictability is strong on average but has undergone a secular decline interspersed with sizable fluctuations. The evidence suggests that liquidity facilitates efficiency, in the sense that the market’s capacity to accommodate order flow is larger during periods when the market is more liquid. We also find that informational efficiency increases with exogenous decreases in the tick size suggesting that more private information is incorporated into prices as spreads decline.
Introduction

As recognized at least since Hillmer and Yu (1979), Epps (1979), and Patell and Wolfson (1984), a market that is quite efficient (Fama (1970)) over a daily horizon might not have completely unpredictable returns from trade to trade or from minute to minute. Investors must have time to absorb new information about fundamental values and to expunge any serial dependence remaining after prices adjust to new equilibrium levels.

More recently, Chordia, Roll, and Subrahmanyam (2005) consider the market’s capacity to accommodate order imbalances as an indicator of market efficiency. They assume that this capacity is inversely related to the intraday predictability of future mid-quote returns from past order flows. They find that order imbalances do indeed predict very short-term returns, suggesting investor reaction time is finite, though brief.

Market efficiency is engendered, at least in part, by offsetting or arbitrage trades in response to order imbalances, and the effectiveness of such trades depends on frictions such as the transaction costs faced by traders. This observation notwithstanding, the relation between intraday market efficiency and trading costs has not yet been considered in the literature.¹ Recent financial crises, however, suggest that, at times, market conditions can be so turbulent that effective trading costs jump dramatically and liquidity declines or even disappears.²


¹ An exception is Boehmer, Kelly and Pirinsky (2005) who consider the impact of institutional investors on intra-day efficiency.
² “One after another, LTCM’s partners, calling in from Tokyo and London, reported that their markets had dried up. There were no buyers, no sellers. It was all but impossible to maneuver out of large trading bets.” - Wall Street Journal, November 16, 1998.
We explore a different role for liquidity in this paper. Specifically, we consider the hypothesis that high liquidity is associated with an increased ability of the market to absorb order imbalances and hence an enhanced degree of intraday market efficiency. We ask also whether market efficiency has varied over time. Since successive reductions in the minimum tick size have decreased bid-ask spreads (see, for example, Chordia, Roll and Subrahmanyam (2001), Bessembinder, 2003, or Chordia, Sarkar, and Subrahmanyam, 2005), they may have improved the market’s ability to accommodate imbalances. Finally, previous literature on intraday regularities in spreads (McInish and Wood, 1992) provides evidence that markets are most liquid during the middle of the day, so the market’s ability to accommodate imbalances may be greatest at these times; we test this hypothesis as well.

Other studies have also explored the return/order flow relation, but around specific events or over short periods of time. Sias (1997) analyzes order imbalances in the context of institutional buying and selling of closed-end funds; Lauterbach and Ben-Zion (1993) and Blume, MacKinlay, and Terker, (1989) study order imbalances around the October 1987 crash; and Lee (1992) does the same around earnings announcements. Chan and Fong (2000) investigate the impact of order imbalances on the contemporaneous relation between stock volatility and volume, using data for about six months. Hasbrouck and Seppi (2001) and Brown, Walsh, and Yuen (1997) study order imbalances for thirty and twenty stocks, over one and two years, respectively.

In contrast to the above studies, we construct a comprehensive sample of intraday returns for all stocks that traded every day on the NYSE from 1993 to 2002. We then examine the relation between short-horizon (five-minute) mid-quote returns and lagged order imbalance measures. The results on intraday market efficiency are: First, the degree of market efficiency varies with firm size; the market for larger firms is more efficient. Second, the ability of the market to accommodate imbalances has improved over time. Third, there is a time-of-day regularity in the degree of market efficiency, which matches previously-documented regularities in liquidity. Specifically, for smaller firms, the market’s ability to absorb imbalances is markedly diminished during the morning and late afternoon, which is consistent with liquidity being lower at such
times. Finally, intraday market efficiency is intimately linked to daily liquidity. The market’s capacity to accommodate order flow is larger on days when the market is more liquid.

The notion of efficient markets in the sense of Fama (1970), and, as used in Chordia, Roll, and Subrahmanyam (2005), emphasizes a lack of return predictability, but the microstructure literature also considers informational efficiency, i.e., the amount of private information reflected in prices (viz. Kyle, 1985). We would expect more liquid markets to reflect more private information (Admati and Pfleiderer, 1988). To test this idea, we study how the ratio of open-to-close return variance relative to close-to-open return variance and how the first-order daily return autocorrelations vary across three tick size regimes. The motivation, based on French and Roll (1986), is to ascertain whether the decrease in spreads across the tick size regimes leads to greater informational efficiency in terms of enhanced trading on private information during trading hours.

It turns out that variance ratios increased and first order autocorrelations decreased (strongly for the smaller firms) as the tick size fell. As French and Roll (1986) suggest, an increase in the open-to-close/close-to-open variance ratio could be caused either by more mispricing or by increased informational efficiency. Since increased mispricing after a tick size reduction should be associated with higher serial correlation of returns, while the empirical facts are the opposite, there seems to have been an increase in the impact of private information trading following the exogenous reduction in spreads that accompanied a reduced tick size. This suggests that the liquidity increase following reduction in the tick size has not only been accompanied by an increase in the efficiency with which prices incorporate order flows, but also by enhanced informational efficiency (i.e., by an increase in the degree of private information reflected in prices).

I. A Theoretical Framework
To support the subsequent empirical analysis of intraday market efficiency, this section presents a theory of how markets accommodate order flows. Risk averse market makers are assumed to dynamically accommodate liquidity demands (that are allowed to be autocorrelated, as per the empirical evidence, viz., Chordia, Roll, and Subrahmanyam, 2002). The framework is similar to
Grossman and Miller (1988), except it is dynamic and orders exhibit positive serial correlation.

Consider a security traded at each of two dates, 1 and 2, that pays off a random amount $\theta + \varepsilon$ at date 3; $\theta > 0$ and $E(\varepsilon) = 0$. Risk-averse market-making agents absorb orders emanating from liquidity traders. The market makers have exponential utility with risk aversion coefficient $\gamma$. There are two types of liquidity orders. The first type arrives solely at date 2 and is denoted $z_2$. The total quantity of the second type of liquidity orders is $z_1$, but a fraction $k$ of this order arrives at date 1 and the balance arrives at date 2. The total mass of the market makers is normalized to a number $N$ and a mass $M$ arrives at date 1 with the balance $(N-M)$ arriving at date 2. The additional mass at date 2 may be construed as additional liquidity suppliers such as floor brokers and limit order traders who arrive to absorb the demands at that date. The expected payoff $\theta$ is learned at date 2 while no information about the asset’s payoff is available at date 1. In the analysis below, $\nu_X$ denotes $\text{var}(X)$, except that $z_1$ and $z_2$ have the same variance $\nu_z$. All random variables are normally distributed with mean zero and are mutually independent.

Let $P_i$ denote the price at date $i$. Standard mean-variance analysis indicates that the date 2 demands of the market makers can be written as

$$x_2 = \frac{\theta - P_2}{\gamma \nu_\varepsilon}.$$

From the date 2 market clearing condition, $Nx_2 + z_1 + z_2 = 0$, we find that

$$P_2 = \theta + \frac{\gamma \nu_\varepsilon}{N} (z_1 + z_2).$$

Consider now the date 1 equilibrium. The appendix shows that the date 1 demands of the market makers are given by

$$x_1 = \frac{E(P_2 | z_1) - P_1}{\gamma S_1} + E(x_2 | z_1) \frac{S_1 - S_2}{S_1}.$$  \hspace{1cm} (1)

where $S_1$ and $S_2$ are the first and second elements, respectively, in the first row of the matrix given by
with the first matrix in the expression above simply representing $\text{cov}(P_2, \theta|z_1)^{-1}$. The expression in Eq. (1) represents a component to exploit the expected price appreciation across dates 1 and 2 as well as a second term to account for intertemporal hedging.

The market clearing condition at date 1 is given by $Mx_1 + kz_1 = 0$. Substituting for $x_1$ and solving for $P_1$ gives

$$P_1 = \frac{\gamma z_1[kN\{\gamma^2 v_\epsilon v_z(v_\epsilon + v_\theta) + N^2 v_\theta\} + MNv_\epsilon]}{M(\gamma^2 v_\epsilon v_z + N^2)}.$$

Let $Q_1 = z_1$ represent the date 1 imbalance. Then

$$\text{cov}(P_2 - P_1, Q_1) = \frac{\gamma v_z[M\gamma^2 v_\epsilon^2 v_z - kN\{\gamma^2 v_\epsilon^2 v_z(v_\epsilon + v_\theta) + N^2 v_\theta\}]}{MN(\gamma^2 v_\epsilon v_z + N^2)}.$$  \hspace{1cm} (2)

Thus, there is a positive predictive relation between price changes and imbalances so long as the numerator of equation (2) is positive, which will be the case if, for example, the mass of market makers at date 1 is sufficiently large. The notion is that autocorrelation in imbalances, coupled with inventory effects, cause the predictive relation. Intuitively, an imbalance in one direction predicts price changes because it is accompanied by further imbalances, and consequently further price pressure in the same direction. The predictive relation disappears as the risk aversion of the market makers ($\gamma$) approaches zero, so that inventory effects become negligible. The predictive relation also disappears as the market making capacity of the market becomes large, i.e., as $M$ and $N$ go to infinity.\(^3\) Hence tests of a predictive relation between imbalances and future returns are direct tests of inventory effects. Consequently, this notion of market efficiency essentially captures the efficacy of market making and liquidity providing for imbalances from outside investors.

\(^3\) Suppose a fraction $b$ of the market making agents are present at date 1, so that $M = bN$. Substituting for $M$ in expression (2) for $\text{cov}(P_2 - P_1, z_1)$, it is easy to see that this covariance goes to zero as $N \to \infty$. 

6
What is the relation between efficiency and liquidity? In our framework, illiquidity is represented by the conditional risk premia, which are the coefficients of the demands $z_1$ and $z_2$ in $P_1$ and $P_2$.\footnote{The risk premium is also related to traditional liquidity indicators. Indeed, as Whitcomb (1988) points out, the “bid-ask spread” in the inventory model of Grossman and Miller (1988) can be interpreted as the premium market makers demand to bear inventory risk. This justifies the use of the spreads as liquidity indicators in empirical analyses.} It can easily be shown that the premium in each period goes to zero as $M$ and $N$ become unboundedly large. One would expect market efficiency to be higher when the risk-bearing capacity of market makers is large and, in our model, predictability of returns from order flow arises solely from the limited risk-bearing capacity of market makers. In reality, other institutional frictions such as inventory financing costs could also influence market efficiency. In general, inventory issues impinge on the market’s ability to accommodate order flows, and illiquidity (the risk premium required at a particular time) measures the importance of these inventory issues.

In sum, an important prediction of the model is that illiquidity decreases the market’s ability to absorb order flow, thus leading to a positive predictive relationship between order imbalances and price changes. In the rest of the paper we will explicitly test whether markets are indeed more efficient during periods of enhanced liquidity.

II. The Data

A. Return and Imbalance Data

Since the serial dependence in returns is close to zero for active stocks over a daily horizon,\footnote{See Chordia, Roll, and Subrahmanyam (2005), p. 273.} an investigation of the efficiency-creating process must focus on intraday trading. To measure the timing of efficiency creation as precisely as possible, it seems sensible to examine frequently-traded stocks for which very short term serial dependence can actually be observed. This suggests that small stocks should be excluded because of the inherent difficulty in measuring serial dependence when trading is infrequent.

Another research choice involves the length of the intraday return interval. We decided to focus on five-minute intervals for the following two reasons. First, though shorter intervals are
technically possible, assigning trades as either buyer- or seller-initiated is prone to error that can be mitigated to some extent by aggregating over time. Second, though longer intervals are feasible, ephemeral market inefficiencies would probably be less conspicuous; a predictive relation between order imbalances and future returns is unlikely to last very long. Five minutes seemed like a good compromise.

The sample is selected as follows: At the beginning of each year from 1993 through 2002 all NYSE firms listed on CRSP are sorted on market capitalization and the largest 500 are noted. Of these, 193 traded every day on the New York Exchange over the entire period and are retained in the sample. These years are selected because (a) transactions data are available from the TAQ (Trade and Automated Quotations) database recorded by the Exchange, and (b) they span significant changes in the minimum tick size, which was reduced from $1/8 to $1/16 during 1997 and was reduced further to one cent by January 2001. We hope to discern changes in the price formation process during years preceding and following these events. Future investigations should extend the investigation to smaller firms, and other years, exchanges, and countries.

Every transaction for all of these stocks during the ten years is recovered from the TAQ database, which provides trade prices as well as bid and ask quotes. This allows us to use the Lee/Ready (1991) trade assignment algorithm to estimate whether a particular trade was buyer- or seller-initiated. Order imbalance for each stock over any time interval can then be calculated variously as the number of buyer- less the number of seller-initiated trades divided by the total number of trades (OIB#), or the dollars paid by buyer-initiators less the dollars received by seller-initiators divided by the total dollars traded (OIB$).

The first of these order imbalance measures (OIB#) disregards the size of the trade, counting small orders equally with large orders. The second measure (OIB$) weights large orders more

---

6 Ball and Chordia (2001) show that for the largest stocks, more than half the bid-ask spread in 1996 was due to the impact of rounding onto the tick grid. Also, Chordia and Subrahmanyam (1995) argue that a reduction in the tick size would result in more competition and less payment-for-order-flow, thus, causing orders to flow to the least cost providers of market making services.

7 The Lee/Ready algorithm classifies a trade as buyer- (seller-) initiated if it is closer to the ask (bid) of the prevailing quote. If the trade is exactly at the mid-point of the quote, a “tick test” is used whereby the trade is classified as buyer- (seller-) initiated if the last price change prior to the trade is positive (negative). Note that a limit order is most often the passive side of the trade; i.e., the non-initiator.
heavily. The distinction is important; we hope to shed light on how arbitrageurs and other liquidity providers make the market more efficient over very short horizons and presume that these agents undertake relatively large trades to quickly exploit deviations of prices from fundamentals.

Short-horizon returns are computed from prices closest to the end of five-minute time intervals within the trading day. When calculations involve lagged values, the first five-minute interval of each trading day is discarded because it would have been related to a lagged interval from the previous trading day. There is admittedly some imperfection involved in computing very short-term returns because trades do not necessarily occur at the exact end of each interval. If the closest price to the end of an interval was more than 150 seconds away, either before or after, the return for that interval was not used in the calculations. Within the large stock sample, the average time between transactions was 19 seconds (averaged across the years).

Order imbalances are computed over all trades within each time interval. For example, contemporaneous OIB# during the five minutes ending at 9:50 a.m. contains the number of buyer-initiated trades less the number of seller-initiated trades between 9:45:01 a.m. and 9:50:00 a.m. The corresponding lagged five-minute OIB# involves the accumulation between 9:40:01 a.m. and 9:45:00 a.m.

**B. Liquidity Data**

Measures of aggregate (market-wide) illiquidity are constructed from individual firm bid/ask spreads. First, individual stock intraday data are cleaned. Spreads are discarded when corresponding trades are out of sequence, trades are recorded before the open or after the close, or trades have special settlement conditions (because they might be subject to distinct liquidity considerations). Negative bid-ask spreads, transaction prices, or quoted depth disqualifies the observation. Second, a preliminary investigation revealed that auto-quotes (passive quotes by secondary market dealers) have not been eliminated from the TAQ database. This causes the quoted spread to be artificially inflated. Since there is no reliable way to filter out auto-quotes,
only BBO (best bid or offer)-eligible primary market (NYSE) quotes are used. Third, the raw intraday data reveal a number of anomalous records that appear to be keypunching errors, so transaction records that satisfied the following conditions are also deleted:

1. Quoted spread > $5
2. Effective spread / Quoted spread > 4.0
3. Proportional effective spread\(^8\) / Proportional quoted spread > 4.0
4. Quoted spread/Mid-point of bid-ask quote > 0.4

The above filters removed less than 0.02% of all stock transaction records. Finally, no data are available for the period from September 11 through September 14, 2001 because of market closure owing to the events of 9/11.

Each bid-ask quote included in the sample is matched to a transaction; following Lee and Ready (1991), from 1993-1998 inclusive, the matching quote is the first quote at least five seconds prior to the trade. Due to a generally accepted decline in reporting errors in recent times (see, for example, Madhavan et al., 2002), after 1998, the matching quote is simply the first quote prior to the trade. Then for each stock, two average daily quotes are computed from trade-to-trade data. QSPR is the daily average quoted spread, i.e., the difference between the asked and the bid quotes, averaged over the trading day. The effective spread, ESPR, is twice the absolute distance between the transaction price and the mid-point of the prevailing quote, again averaged over the day.

The individual stock average daily spreads are then value-weighted and averaged across stocks (with market capitalization at the end of the previous year used to calculate weights) to obtain market-wide illiquidity measures. For convenience, the same variable names are used for the aggregated illiquidity measures.

The sample spans three distinct liquidity sub-periods: (i) January 4, 1993 through June 23, 1997, (ii) June 24, 1997 through January 28, 2001, and (iii) January 29, 2001, through December 31, 2002. These periods correspond to the 1/8, 1/16, and 0.01 tick size regimes, respectively. Table 1 presents summary statistics associated with the illiquidity and imbalance measures. While

\(^8\) The proportional spread is the unscaled spread divided by the mid-point of the prevailing bid-ask quote.
illiquidity declined sharply in the third sub-period (the decimal regime) the relative imbalance measures remained quite stable throughout the period. For instance, the effective spread was 12 cents in the eighths regime, 8 cents in the sixteenths regime and 3 cents in the decimal regime. The dollar order imbalance, which seems to have increased over time, was 0.045, 0.069 and 0.086 in the three regimes.9

C. Construction of Five-Minute Return/Imbalance Portfolios

In constructing portfolios, non-synchronous trading is a potential problem. Stocks trading with a lag could induce a spurious forecasting relation between portfolio imbalances and future portfolio returns. Individual security data would not be plagued by non-synchrony but could be excessively noisy. Thus, to alleviate non-synchrony while retaining the benefits of portfolio diversification, we exclude stocks that did not trade at time t-1 from the portfolio constructed at time t.

Table 2 presents correlations among five-minute returns and order imbalances. Returns exhibit high contemporaneous correlations with both imbalance measures, and the correlations are reasonably stable across the three liquidity sub-periods.

III. The Intraday Evidence

A. The Basic Evidence

Table 3 reports a simple regression of five-minute returns on lagged five-minute order imbalances measured by the number of transactions (OIB#) and by dollars (OIB$). Consistent with Chordia, Roll, and Subrahmanyam (2005), lagged OIB is a significant predictor of five-minute returns. This result holds for both measures of imbalance, though the coefficient is larger for OIB#. A one-standard deviation move in OIB$ increases five-minute returns by about 0.5%, which is substantial, given the time interval (five minutes.) The explanatory power of the regression is small, however, with the R² being only about 0.9%.

---

9 The sample sizes are equal to the number of five-minute trading intervals during each period, excluding missing observations for all stocks (e.g., the market was closed after the terrorist attack on September 11, 2001), or, occasionally, when one of the three size groups (to be reported upon later) is missing an interval; 119 intervals were missing over the entire sample, approximately 0.06%. The retained intervals are kept identical in every calculation.
To gain insight on how the ability of the market to absorb imbalances has changed over time, predictive regressions identical to those in Table 3 are computed for each calendar month in the ten-year sample. Figure 1 plots the resulting $R^2$’s and t-statistics for lagged OIB$. Clearly, the degree of NYSE market efficiency has improved over this decade. The $R^2$ has moved from a peak of about 11% in the eighths regime to virtually zero by the end of the sample period. The t-statistic on imbalance has also gone from around 12 at the start of the sample to considerably less than two towards the end.

We now turn to the interaction of day-to-day liquidity with market efficiency (measured as the efficacy with which the market accommodates imbalances). Previous literature (Chordia and Subrahmanyam, 2004, Chordia, Roll, and Subrahmanyam, 2005) and our theoretical model in Section I suggest that order imbalances predict returns because of two reasons: (i) if market makers face inventory costs in accommodating autocorrelated order flows, prices may not be set to remove all predictability in order flows, and (ii) orders to offset order imbalances may not arrive in a timely manner. Periods of high liquidity are likely to correspond to low inventory holding costs (e.g., Amihud and Mendelson, 1980) and are also likely to facilitate arbitrage trades that counteract the effect of order imbalances. Overall, liquid periods should be associated with a higher degree of market efficiency.

For parsimony, results are reported henceforth only for dollar imbalances (OIB$) and the effective spread (ESPR) measure of illiquidity. These are probably the most informative choices because OIB$ provides the economic magnitude of the order imbalance and the effective spread is closer to the actual transaction costs incurred by traders.

Figure 2 plots the time-series of the effective spread (ESPR). Clearly, ESPR has experienced three distinct regimes corresponding to sub-periods for the eighth, sixteenth and decimal minimum tick sizes. Moreover, ESPR appears to be trending to some extent within each sub-period. To examine the interaction of liquidity with return predictability, we need to identify periods of illiquidity within each regime. A very simple approach involves separating observations into two groups based on trading days with low and high liquidity. So ESPR is first

---

10 Results are similar using OIB# and QSPR.
de-trended within each of the three tick size sub-periods separately by regressing ESPR on time. Then, “low liquidity” days are defined as those when the de-trended effective spread is at least one standard deviation above the mean\(^\text{11}\) and “high liquidity” days are all others.

Table 4 reports summary statistics for effective spreads on low and high liquidity days. The numbers of illiquid days as a percentage of the total number of trading days are 16%, 13%, and 19%, in the eighth, sixteenth, and decimal regimes, respectively, and these proportions are reasonably close to each other. The mean effective spread on illiquid days is 120% and 141% of that of liquid days in the sixteenth and decimal regimes, respectively, but this ratio is only about 105% for the eighth regime. This is likely due to the fact that the larger tick sizes may have caused the measured effective spread to be higher than its true value due to the impact of rounding,\(^\text{12}\) leaving less scope for a significant upward movement of the spread on illiquid days relative to normal days.

To estimate the influence of liquidity on market efficiency (i.e., the ability of the market to accommodate order imbalances) a low-liquidity dummy (ILD) is interacted with the basic explanatory variable (lagged OIB$) already used in the regressions of Table 3. The interaction explanatory variable equals OIB$_{t-1}$ on days with low liquidity and zero otherwise.\(^\text{13}\)

Table 5 presents the results. For all sample firms aggregated, the coefficients of both OIB$_{t-1}$ and OIB$_{t-1}$*ILD are positive and significant in all the three sub-periods. The coefficient estimates of OIB$_{t-1}$ have decreased steadily from 0.04 in the eighths regime to 0.01 in the decimal regime. The significantly positive coefficients of OIB$_{t-1}$*ILD suggests that the ability of lagged OIB$ to predict returns increases during periods of illiquidity. Thus, illiquidity decreases efficiency. Indeed, during the sixteenth and decimal sub-periods, the coefficients of OIB$_{t-1}$*ILD are more than twice as large as those of OIB$_{t-1}$ alone; hence, in these later sub-periods, which are generally more efficient on average, illiquidity has a relatively stronger inhibiting influence. For instance, during the eighths regime a one standard deviation increase in

---

\(^\text{11}\) As the de-trended ESPR is actually a residual from a regressing ESPR on time, it has mean of zero.

\(^\text{12}\) See Ball and Chordia (2001).

\(^\text{13}\) The use of an indicator variable is suggested by the likely diminished ability of the market to accommodate imbalances on days of pronounced illiquidity (see for example, Blume, MacKinlay, and Terker, 1989, or Barclay, Hendershott, and Jones, 2003).
lagged OIB$ causes a 1.04% increase in returns when markets are liquid and a 1.6% increase in returns when markets are illiquid. During the decimal regime the impact on returns of a one standard deviation increase in lagged OIB$ is 0.22% during liquid periods and 0.76% during illiquid periods. Overall, this evidence is consistent with a decreased ability of the market to accommodate order imbalances during periods of illiquidity; i.e., liquidity promotes market efficiency.

Table 5 also presents results for three size-based sub-samples formed by ranking firms by market capitalization at the beginning of the year, dividing the ranked firms into thirds, and then calculating value-weighted returns, imbalances, and de-trended effective spreads within each tierce. Again, the coefficients of OIB$t-1 and OIB$t-1*ILD are positive and significant in each of the nine cases corresponding to the three firm size groups and the three tick size regimes. Note that the impact of OIB$t-1*ILD relative to that of OIB$t-1 increases much more for the large size group as the tick size decreases.

The coefficient estimates of lagged OIB$ across the size groups for the eighth regime decline from 0.035 for the large firms to 0.018 for the small firms. Across the tick size regimes, the coefficients for the large firms show a much greater decline in the lower tick size regimes relative to the smaller firms, and are about three to six times larger for the smaller size groups relative to those for the large caps in the sixteenth and decimal regimes. The coefficient patterns suggest that efficiency of the large cap sector has increased after the reduction in minimum tick size and also imply that in the smaller tick size regimes, the market’s capacity to absorb imbalances is greater for larger than for smaller firms. Perhaps this finding is explained by the eighth tick size being more binding on the larger firms (Ball and Chordia, 2001); so that decreased spreads accompanying the tick size reduction had a greater impact on market efficiency for the larger firms.

In interpreting the interaction coefficient of OIB and illiquidity, one possible conundrum is that an exogenous shock might cause extreme order imbalances and simultaneously reduce liquidity. If the market has difficulty accommodating these imbalances, we could be picking up the effect of the shock rather than capturing the role of illiquidity in hampering the establishment of market
efficiency. To investigate this, we compute absolute order imbalances within liquid and illiquid periods for each of the three size groups, across the tick size regimes. In each of the nine cases, the absolute imbalances differ only minimally across liquid and illiquid periods (the difference is less than 5% in each case). This suggests that the identified illiquid periods are not capturing periods of abnormally high (absolute) order imbalances. Moreover, there is no reason to believe that illiquidity is jointly determined with signed imbalances. Thus, when markets are illiquid, one may expect to see fewer buy as well as sell orders, but this has no implication for net order flow, which is our right-hand variable.

A decline in the significance and explanatory power of the Table 5 regressions accompanies the general improvement in liquidity over the sample period. Adjusted R-squares drop across the three tick size regimes from around four to five percent in the eighths regime to less than one percent in the decimal regime. Overall, the evidence suggests that the secular changes in spreads accompanying tick size reductions (Bessembinder, 2003) have been accompanied by a marked increase in the degree of market efficiency.

Since autocorrelated imbalances are the source of the predictability of future returns from imbalances, it is worth considering how imbalance autocorrelations have behaved across the three tick size regimes. For example, the decrease in the predictive power of imbalances could potentially be due to a decrease in imbalance autocorrelation. To shed light on this issue, we find average imbalance autocorrelations of 0.280, 0.210, and 0.211 during the eighth, sixteenth, and decimal regimes, respectively. While there is some decrease in autocorrelations during the smaller tick regimes, it is not enough to reduce the $R^2$ from more than 10% at the beginning of the sample (Figure 1) to virtually zero in the decimal period. Also, from the sixteenth to the decimal period, the $R^2$ and predictive t-statistic dropped further with no reduction in the autocorrelation of imbalance.
B. Variations in the Degree of Market Efficiency Within the Trading Day

To provide additional confirmation for the hypothesis that liquidity impacts market efficiency, we now relate efficiency to liquidity at different times of the day. Previous literature suggests that the market’s ability to accommodate order imbalances could vary by time of day. Brock and Kleidon (1992) and McInish and Wood (1992) document that spreads are higher at the beginning and at the end of the day. Since there is a distinct pattern to liquidity over the trading day, this pattern should have an impact on the predictability of returns due to lagged OIB$ over the trading day. To investigate, we employ dummies for the morning and evening periods. The first dummy, “morn” is unity if the five minute interval begins and ends between 9:30am and 12 noon, whereas the second, “eve,” is unity if the interval is between 2pm and 4pm. Each is interacted with the variables in Table 5 and Table 6 gives the results.

Consider first the interaction of OIB itself with the time of day dummies. Accumulated inventories overnight and towards the end of the trading day could influence the market’s ability to accommodate order flows. As shown in Table 6, during the eighth and sixteenth regimes, both interaction variables are significant for the smallest firms, suggesting that the market is less efficient at accommodating order flows during the morning and evening periods for such firms. While the morning effect is present for the two larger firm groups, the evening effect is not significant. In the decimal regime, neither morning nor evening effects are present in the two larger firm groups, though there is evidence of a morning effect for the smallest firm group. Evidently, increased liquidity accompanying the tick size reduction to decimals has left only the morning effect for the smaller firm group. This is consistent with an enhanced capacity of the market to accommodate order imbalances.

Turning now to the interaction dummies involving both liquidity and time of day, the morning effect is present in all size groups in the eighth regime, but only in the two smaller size groups in the sixteenth regime, and in none of the size groups during the decimal regime. Combining the two sets of interaction coefficients, we may conclude that while there is a diminished ability of

---

14 McInish and Wood (1992) suggest that the observed pattern in spreads could be linked to a U-shaped pattern in return volatility linked to news arrivals during a trading day. Brock and Kleidon (1992) suggest that dealer inventory problems at the close and open could be the cause.
the market to accommodate order flows in the morning relative to the middle of the day, the diminution is greater on illiquid days during the eighth regime. The phenomenon has disappeared in the decimal regime.

With regard to the evening dummy involving the interaction of OIB with both ILD and time of day, there is an evening effect for smaller firms in all three regimes but there is no evidence of an evening effect for the two larger size groups in the eighth and sixteenth regimes. In the decimal regime, however, an evening effect is present in all size groups. Again, combining results for both sets of interactive variables, one may surmise that in the decimal regime, the impact of the evening period on efficiency creation is present only on illiquid days (and not on days with normal levels of liquidity). This evidence suggests that on illiquid days with a decimal regime, market makers have relatively more difficulty managing their inventory than when tick sizes are larger. Perhaps this is because the smaller pricing grid restricts the degree to which market makers can move prices in the presence of competition by floor traders/brokers. However, the low explanatory power of the regressions during the decimal period indicates that such an impact of the evening period is minimal.

Overall, based on all of the evidence, there is a compelling picture that efficiency-creating activity is enhanced when the market is liquid and vice versa. Thus, efficiency is intimately linked with market liquidity.

IV. Informational Efficiency

The previous section focused on the market’s capacity to absorb order imbalances within a day as a measure of efficiency, and implicitly, as a measure of the quality of a financial market. This approach is based on Chordia, Roll, and Subrahmanyam (2005) who relied on Fama’s (1970) measure of market efficiency, a lack of return predictability. The microstructure literature, however, also considers the concept of informational efficiency, i.e., the degree to which prices reflect private information, as an indicator of a financial market’s quality. Based on this literature, we would expect more liquid markets to be accompanied by greater informational
efficiency. This is because greater liquidity would allow more informed agents to acquire and trade on information, which would increase informational efficiency, and further increase liquidity due to increased competition between the informed (Admati and Pfleiderer, 1988). To test this notion, we consider how the ratio of open-to-close and close-to-open return variances change across the three tick size regimes, which are characterized by successive decreases in bid-ask spreads. These ratios have been related to the amount of information incorporated into prices by French and Roll (1986).

French and Roll (1986) find that the variance of open-to-close returns is much greater than that of close-to-open returns. They offer three potential explanations: (i) incorporation of private information during trading hours, (ii) mispricing caused by investor misreaction or market frictions and microstructure noise induced by bid-ask bounce, and (iii) greater incorporation of public information into prices during trading hours. They reject (iii) because the variance ratios are not significantly different on business days when the stock market is closed. They conclude that the other two components help explain the higher ratio during market trading hours, with (i) being the dominant factor. Relying on their conclusions, the variance ratio measures either mispricing or the incorporation of private information into prices during trading hours. Thus, the variance ratio is possibly linked to the informativeness or efficiency of the pricing system in the sense of Kyle (1985).

We seek to discern whether variance ratios have changed across the three tick size regimes, and if so, whether the pattern of changes reveals which of the two components identified by French and Roll (1986) is relatively more important.\(^\text{15}\) Note that returns based on midquotes eliminate one element of mispricing, bid-ask bounce, so it should not be significant here.

Panel A of Table 7 presents the (open-close)÷(close-open) per hour variance ratios for the three tick size regimes and the three size groups. To obtain per hour variances, open-close and close-

\(^{15}\) We also estimate variance ratios across liquid and illiquid days (defined as in the previous section) and find that return variances are consistently higher on illiquid days. On the one hand, this is consistent with high illiquidity (perhaps during days with high levels of private information) increasing asset price fluctuations (Subramanyam, 1994). However, there is an endogeneity issue because exogenous volatility shocks (unrelated to private information) could cause reduced liquidity due to greater inventory risk. Because of the resulting interpretational problem, the empirical results are not presented in the text.
open returns were first used separately to compute raw variances. Then, each such computed variance was divided by the total number of calendar hours over the sub-sample in question. For example, in the eighths sub-period, there are 1130 trading days, each 6.5 hours in length and there are 1634 calendar days. Hence, there are 1130(6.5)=7345 trading hours and 1634(24)-7345=31871 non-trading hours.

As shown in the table, all per hour variance ratios are much higher than unity. To test statistical significance, a non-parametric sign test was computed from paired open-close and close-open squared returns, each normalized by the number of hours in their respective periods. The computed t-statistics ranged from 16.3 to 24.9 over the nine comparisons corresponding to the variance ratios in Table 7, thereby indicating a very high level of statistical probability that open-close volatility far exceeds close-open volatility. This finding is consistent with that of French and Roll (1986).

Table 7 also reveals that all of the variance ratios in the sixteenth and decimal regimes are greater than the corresponding ratios in the eighths regime. A bootstrapping analysis was employed to check the statistical significance of these differences. In the bootstrap, pseudo regimes, with the same numbers of observations as in the original data, were formed by randomly sorting actual dates, with replacement, then computing a ratio of ratios; i.e., \((\text{open-close}) \div (\text{close-open})\) variance ratio in a lower pseudo tick size regime divided by the \((\text{open-close}) \div (\text{close-open})\) variance ratio in the pseudo eighths regime. This procedure is repeated 5000 times, and the actual observed ratio of the variance ratios is compared to the relevant fractile of the bootstrapped distribution. The tests indicate that the per hour \((\text{open-close}) \div (\text{close-open})\) variance ratios for midcap firms during the sixteenth and decimal regime are statistically greater than the corresponding ratios in eighths regime (p-values less than 0.000); the same conclusion applies for smallcap firms in the sixteenth regime vis-à-vis the eighths regime.

Based on French and Roll (1986), the higher variance ratios can be attributed either to an increase in mispricing or to an increase in privately informed trading that results in more
information being incorporated into prices when the market is open.\textsuperscript{16} To provide some suggestive evidence that distinguishes between these possibilities, we consider the first order daily return autocorrelations (based on the midpoint of the last quote matched with a transaction during a trading day) across the three tick regimes. As French and Roll (1986) argue, the absolute levels of these autocorrelations should be positively related to microstructural frictions and misreactions of investors to new information, and thus capture mispricing.\textsuperscript{17} Note that non-synchronous trading is not an issue here because firms were included only if they traded every day and, as mentioned earlier, bid-ask bounce concerns are allayed by using midquotes.

The autocorrelations, presented in Panel B of Table 7, generally decrease across the regimes. For example, none of the autocorrelations are significantly different from zero in the decimal regime, whereas five of the six autocorrelations in the other two regimes are positive and significant. In addition, the autocorrelations in the sixteenth and decimal regimes for the smallest firms in our sample are statistically different from the corresponding one in the eighth regime. The positive autocorrelations are consistent with investors underreacting to information (Barberis, Shleifer, and Vishny, 1998), and/or with incomplete adjustment to autocorrelated order flow owing to market maker risk aversion (Chordia and Subrahmanyam, 2004). Either of these two phenomena can be termed “mispricing,” in the sense of French and Roll (1986). Such mispricing may arise due to behavioral biases (the former case) or insufficient market making capacity (the latter one).

Regardless of the cause, if such mispricing were driving the increase in variance ratios across time, autocorrelations should have increased along with variance ratios as the tick size decreased; but there is no evidence of this. In fact, there is reliable evidence that the opposite transpired for smaller firms. Consequently, the evidence is consistent with private information being more effectively incorporated into prices in the lower tick regimes, especially for smaller firms. With lower bid-ask spreads, informed traders might find it worthwhile to trade on private information

\textsuperscript{16} The magnitude of the open-to-close variances also increases monotonically across the three tick size regimes.  
\textsuperscript{17} In Kyle (1985), where market makers have unlimited market making capacity, autocorrelations are zero. However, under the friction of limited risk bearing capacity (Grossman and Miller, 1988), returns exhibit nonzero autocorrelation. Such autocorrelations also obtain in irrational investor models such as Daniel, Hirshleifer, and Subrahmanyam (1998).
even with modest profit potential.\textsuperscript{18} Since more private information is incorporated into prices as spreads decline, the informational efficiency of the market improves with liquidity.

V. Conclusions

The debate about financial market efficiency has been relatively quiet about the behavior of actual traders. Efficiency cannot happen instantaneously. For stock prices to fluctuate randomly, the market must absorb order imbalances in a timely manner. Further, it stands to reason that market efficiency is not immune to frictions, such as illiquidity, that are barriers to efficiency-creating arbitrage activity.

We examine how the capacity of the equity market to absorb imbalances varies through time and across different liquidity regimes. The analysis relies on a sample of all NYSE stocks that traded every day during the 1993-2002 decade. There is ample evidence that order imbalances do indeed predict future returns over very short intervals, more so for the smaller firms. But the extent of this predictability declines markedly over the sample period. Liquidity plays an important role in efficiency creation. The market’s ability to absorb order imbalances is greater in liquid periods than in illiquid ones.

As distinct from the notion of Fama (1970) that efficiency implies a lack of return predictability, the microstructure literature also considers the notion of informational efficiency, which is defined as the amount of private information revealed in prices in the sense of Kyle (1985). We also shed light on this concept of efficiency by considering per hour open-to-close/close-to-open variance ratios. Variance ratios generally have increased while first order return autocorrelations have declined as the minimum tick size was reduced; this pattern is particularly strong for the smaller firms. Taken together, these findings suggest that the observed increase in open-to-close/close-to-open variance ratios was not due to increased mispricing in the lower tick size regimes, but is consistent with an increase in the amount of private information that is incorporated in prices following the secular decrease in bid-ask spreads accompanying the

\textsuperscript{18} The increased informed trading could potentially reduce liquidity in the post-decimal period. However, as pointed out earlier, Admati and Pfleiderer (1988) suggest that with more informed agents competing for profits, liquidity may actually increase.
lowering of the tick size. In sum, it appears that improved liquidity engenders a higher degree of informational efficiency as there is more trading on private information following reduction in the tick size.

An extension of this analysis would be to study the liquidity-efficiency relation for fixed-income and currency markets. Considering the relation prior to important announcements (when the market might be particularly illiquid because of asymmetric information due to news leakage) would also be an interesting exercise.
References


*American Economic Review* 70, 393-408.


Appendix

Proof of Eq. (1): The expression we derive holds for multiple classes of agents with differential information sets and exponential utility within a dynamic model, not just the market makers modeled here. The following lemma, a standard result for multivariate normal random variables (see, for example, Brown and Jennings [1989]), will prove useful below:

**Lemma 1** Let \( Q(\chi) \) be a quadratic function of the random vector \( \chi \): \( Q(\chi) = C + B'\chi - \chi'A\chi \), where \( \chi \sim N(m, \Sigma) \), and \( A \) is a square, symmetric matrix whose dimension corresponds to that of \( \chi \). We then have

\[
E[\exp(Q(\chi))] = |\Sigma|^{-1/2} |2A + \Sigma^{-1}|^{-1/2} \times \\
\text{Exp}[C + B'm + m'Am + (1/2)(B' - 2m'\Sigma^{-1})(2A + \Sigma^{-1} - 1)(B - 2Am)]
\]

Let \( \varphi_{ij} \) and \( x_{ij} \) denote the information set and demand, respectively, of an agent \( i \) at date \( j \). The date 2 demand of the agent (from maximization of the mean-variance objective) is given by

\[
x_{i2} = \left[ E(F|\varphi_{i2}) - P2 \right] / \left[ \gamma_i \text{var}(F|\varphi_{i2}) \right]
\]

where \( \gamma_i \) denotes the risk aversion coefficient of agent \( i \). Let \( \mu_2 = E(F|\varphi_{i2}) \). Note that in period 1, the trader maximizes the derived expected utility of his time 2 wealth, which is given by

\[
E[\exp\{-\gamma_i[B_0 - x_{i1}P_1 + x_{i1}P_2 + [\mu_2 - P2]_2/(2\gamma_i \text{var}(F|\varphi_{i2}))]\}]|\varphi_{i1}.
\]

Let \( p \) and \( \mu \) denote the expectations of \( P_2 \) and \( \mu_2 \), and \( \Pi \) the variance-covariance matrix of \( P_2 \) and \( \mu_2 \), conditional on \( \varphi_{i1} \). Then, the expression within the exponential above (including terms from the normal density) can be written as

\[
- [{(1/2)y'Gy + h'y + w}],
\]

where

\[
y' = [\mu_2 - \mu, P_2 - p],
\]

\[
h = [-\gamma_i x_{i1} + (p - \mu)/\text{var}(F|\varphi_{i2}), (\mu - p)/\text{var}(F|\varphi_{i2})]
\]

\[
G = \left[ \Pi^{-1} + \begin{bmatrix} s^{-1} & -s^{-1} \\ -s^{-1} & s^{-1} \end{bmatrix} \right]^{-1},
\]

\[
w = \gamma_i x_{i1}(P_1 - p) + g,
\]

29
where \( s = \text{var}(F|\varphi_{i1}) \), and where \( g \) is an expression which does not involve \( x_{i1} \). From Lemma 1 and Bray (1981, Appendix), (2) is given by 
\[- \text{Det}(\Pi)^{-1/2} |\text{Det}(A)|^{-1/2} \exp[(1/2)h'G^{-1}h-w] \]
Thus, the optimal \( x_{i1} \) solves
\[
\left[ \frac{dh}{dx_{i1}} \right]'G^{-1}h - \frac{dw}{dx_{i1}} = 0.
\]
Substituting for \( h \) and \( w \), we have
\[
 x_{i1} = \frac{p - P_1}{\gamma_i G_1} + \frac{\mu - p}{\gamma_i \text{var}(F|\varphi_{i1})} \frac{G_1 - G_2}{G_1}
\]
where \( G_1 \) and \( G_2 \) are the elements in the first row of the matrix \( G^{-1} \). It follows that the demand \( x_1 \) is given by (1), with the \( S \) coefficients being the \( G_1 \) and \( G_2 \) coefficients above.
Table 1

Summary statistics

Quoted and effective spreads (QSPR and ESPR respectively) and order imbalances (OIB) are from the NYSE TAQ database. Spreads are averaged across the day for each stock and then value weighted across stocks using market capitalizations at the beginning of each calendar year. OIB# is the number of buyer-initiated less the number of seller-initiated trades divided by the total number of trades during a five-minute trading interval. OIB$ is the total dollars paid by buyer-initiators less the total dollars received by seller-initiators divided by the dollar volume of trading during a five-minute trading interval. N is the number of trading days and n is the number of five-minute trading intervals. The sample spans the years 1993 to 2002 inclusive, and consists of all stocks that traded every day on the NYSE during the sample period. The eighths regime spans January 1, 1993 to June 23, 1997, the sixteenths June 24, 1997 to January 28, 2001, and the decimal regime spans the remainder of the sample period.

<table>
<thead>
<tr>
<th></th>
<th>QSPR($)</th>
<th>ESPR($)</th>
<th>OIB#</th>
<th>OIB$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire sample</td>
<td>Mean</td>
<td>0.1350</td>
<td>0.0896</td>
<td>0.0633</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.1447</td>
<td>0.0916</td>
<td>0.0654</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.0464</td>
<td>0.0314</td>
<td>0.146</td>
</tr>
<tr>
<td>Eighths regime</td>
<td>Mean</td>
<td>0.1742</td>
<td>0.1176</td>
<td>0.0608</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.1710</td>
<td>0.1175</td>
<td>0.0625</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.0093</td>
<td>0.0031</td>
<td>0.164</td>
</tr>
<tr>
<td>Sixteenths regime</td>
<td>Mean</td>
<td>0.1305</td>
<td>0.0838</td>
<td>0.0642</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.1275</td>
<td>0.0842</td>
<td>0.0663</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.0130</td>
<td>0.0069</td>
<td>0.136</td>
</tr>
<tr>
<td>Decimal regime</td>
<td>Mean</td>
<td>0.0515</td>
<td>0.0349</td>
<td>0.0673</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.0489</td>
<td>0.0333</td>
<td>0.0685</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.0114</td>
<td>0.0073</td>
<td>0.115</td>
</tr>
</tbody>
</table>
Table 2

Contemporaneous correlations for five-minute trading intervals

Returns and order imbalances (OIB) are from the NYSE TAQ database. Five-minute returns are computed using the mid-points of the first and last quotes within each five-minute trading interval; n is the number of five-minute intervals. OIB# is the number of buyer-initiated less the number of seller-initiated trades divided by the total number of trades during a five-minute interval. OIB$ is the total dollars paid by buyer-initiators less the total dollars received by seller-initiators divided by the dollar volume of trading during a five-minute interval. The sample spans the years 1993 to 2002 inclusive, and consists of all stocks that traded every day on the NYSE during the sample period. The eighths regime spans January 1, 1993 to June 23, 1997, the sixteenths June 24, 1997 to January 28, 2001, and the decimal regime spans the remainder of the sample period.

<table>
<thead>
<tr>
<th></th>
<th>Return</th>
<th>OIB$</th>
<th>OIB#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire Sample</td>
<td>n=193844</td>
<td>0.538</td>
<td>0.510</td>
</tr>
<tr>
<td>Eighth Regime</td>
<td>n=86981</td>
<td>0.564</td>
<td>0.566</td>
</tr>
<tr>
<td>Sixteenth Regime</td>
<td>n=69886</td>
<td>0.646</td>
<td>0.617</td>
</tr>
<tr>
<td>Decimal Regime</td>
<td>n=36977</td>
<td>0.598</td>
<td>0.576</td>
</tr>
</tbody>
</table>
Table 3

Regression of five-minute returns on lagged order imbalance, 1993-2002

Returns and order imbalances (OIB) are from the NYSE TAQ database. Five-minute returns are computed using the mid-points of the first and last quotes within each five-minute trading interval. OIB#_t is the number of buyer-initiated less the number of seller-initiated trades divided by the total number of trades during five-minute interval t. OIB$_t is the total dollars paid by buyer-initiators less the total dollars received by seller-initiators divided by the dollar volume of trading during five-minute interval t. The sample spans the 1993 through 2002 inclusive, and consists of all stocks that traded every day on the NYSE during the sample period. The sample size is 191325. The first five-minute interval of each day is not used in the regressions, so the sample size is reduced by 2519 trading days from 193844, the total number of five-minute intervals from 1993 through 2002.

<table>
<thead>
<tr>
<th>Dependent Variable: Return_t</th>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.0036</td>
<td>-16.44</td>
</tr>
<tr>
<td>OIB#_{t-1}</td>
<td>0.0578</td>
<td>42.32</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.0093</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dependent Variable: Return_t</th>
<th>Coefficient</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.0020</td>
<td>-9.92</td>
</tr>
<tr>
<td>OIB$_{t-1}</td>
<td>0.0204</td>
<td>43.16</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.0096</td>
<td></td>
</tr>
</tbody>
</table>
Table 4

**Distribution of Liquid and Illiquid Days**

Effective spreads are obtained from the NYSE TAQ database. Spreads are averaged across the day for each stock and then value weighted across stocks using market capitalizations at the beginning of each calendar year. Illiquid days are those on which the daily effective spread is at least one standard deviation above the de-trended expected effective spread for the trading day. The de-trending is done separately for each tick size regime. The sample spans the years 1993 to 2002 inclusive, and consists of all stocks that traded every day on the NYSE during the sample period. The eighths regime spans January 1, 1993 to June 23, 1997, the sixteenths June 24, 1997 to January 28, 2001, and the decimal regime spans the remainder of the sample period.

<table>
<thead>
<tr>
<th>Regime</th>
<th>ESPR on liquid days</th>
<th>ESPR on illiquid days</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>0.1168</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.1170</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.0022</td>
</tr>
<tr>
<td></td>
<td>Number of days</td>
<td>976</td>
</tr>
<tr>
<td>Eighths regime</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.0819</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.0833</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.0072</td>
</tr>
<tr>
<td></td>
<td>Number of days</td>
<td>804</td>
</tr>
<tr>
<td>Sixteenths regime</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.0328</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>0.0321</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation</td>
<td>0.0049</td>
</tr>
<tr>
<td></td>
<td>Number of days</td>
<td>404</td>
</tr>
</tbody>
</table>
Table 5

Regression of five-minute returns on lagged order imbalances, and lagged order imbalances interacted with a dummy variable for low liquidity regimes

Returns and order imbalances (OIB) are from the NYSE TAQ database. The five-minute return, Return<sub>t</sub>, is the dependent variable in all regressions; it is computed using the mid-points of the first and last quotes within five-minute trading interval t. OIB<sub>#t-1</sub> is the number of buyer-initiated less the number of seller-initiated trades divided by the total number of trades during five-minute trading interval t-1. OIB$<sub>t-1</sub>$ is the total dollars paid by buyer-initiators less the total dollars received by seller-initiators divided by the dollar volume of trading during five-minute trading interval t-1. The illiquidity dummy, ILD, is 1.0 when the daily effective spread is at least one standard deviation above the de-trended expected effective spread for the trading day, otherwise zero. The sample spans the years 1993 to 2002 inclusive, and consists of all stocks that traded every day on the NYSE during the sample period. The eighths regime spans January 1, 1993 to June 23, 1997, the sixteenths June 24, 1997 to January 28, 2001, and the decimal regime spans the remainder of the sample period. n is the sample size. Each firm size category includes one-third of all sample firms. The sample size, n, is the number of available five-minute intervals in each sub-period, as reported in the table 1, less the number of days; (the first interval of each trading day is used only in the lagged regressors.)

<table>
<thead>
<tr>
<th></th>
<th>All Firms</th>
<th>Large Firms</th>
<th>Mid-Cap Firms</th>
<th>Small Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>t-statistic</td>
<td>Coefficient</td>
<td>t-statistic</td>
</tr>
<tr>
<td>Eighths</td>
<td>OIB$&lt;sub&gt;t-1&lt;/sub&gt;$</td>
<td>0.0359</td>
<td>58.00</td>
<td>0.0349</td>
</tr>
<tr>
<td></td>
<td>OIB$&lt;sub&gt;t-1&lt;/sub&gt;$*ILD</td>
<td>0.0264</td>
<td>15.63</td>
<td>0.0250</td>
</tr>
<tr>
<td></td>
<td>Intercept</td>
<td>-0.0017</td>
<td>-9.84</td>
<td>-0.0014</td>
</tr>
<tr>
<td></td>
<td>Adjusted R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.0539</td>
<td>0.0397</td>
<td>0.0557</td>
</tr>
<tr>
<td>Sixteenths</td>
<td>OIB$&lt;sub&gt;t-1&lt;/sub&gt;$</td>
<td>0.0624</td>
<td>11.52</td>
<td>0.0610</td>
</tr>
<tr>
<td></td>
<td>OIB$&lt;sub&gt;t-1&lt;/sub&gt;$*ILD</td>
<td>0.0624</td>
<td>11.52</td>
<td>0.0610</td>
</tr>
<tr>
<td></td>
<td>Intercept</td>
<td>-0.0023</td>
<td>-5.81</td>
<td>-0.0018</td>
</tr>
<tr>
<td></td>
<td>Adjusted R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.0066</td>
<td>0.0038</td>
<td>0.0212</td>
</tr>
<tr>
<td>Decimal</td>
<td>OIB$&lt;sub&gt;t-1&lt;/sub&gt;$</td>
<td>0.0285</td>
<td>3.34</td>
<td>0.0297</td>
</tr>
<tr>
<td></td>
<td>OIB$&lt;sub&gt;t-1&lt;/sub&gt;$*ILD</td>
<td>0.0285</td>
<td>3.34</td>
<td>0.0297</td>
</tr>
<tr>
<td></td>
<td>Intercept</td>
<td>-0.0010</td>
<td>-1.34</td>
<td>-0.0008</td>
</tr>
<tr>
<td></td>
<td>Adjusted R&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.0009</td>
<td>0.0006</td>
<td>0.0018</td>
</tr>
</tbody>
</table>
Table 6

Regression of five-minute returns on lagged order imbalances, and lagged order imbalances interacted with a dummy variable for low liquidity regimes, and with dummy variables for time of day

Returns and order imbalances (OIB) are from the NYSE TAQ database. The five-minute return, Return\(_t\), is the dependent variable in all regressions; it is computed using the mid-points of the first and last quotes within five-minute trading interval \(t\). OIB\(_{t-1}\) is the number of buyer-initiated less the number of seller-initiated trades divided by the total number of trades during five-minute trading interval \(t-1\). OIB$\(_{t-1}\) is the total dollars paid by buyer-initiators less the total dollars received by seller-initiators divided by the dollar volume of trading during five-minute trading interval \(t-1\). The illiquidity dummy, ILD, is 1.0 when the daily effective spread is at least one standard deviation above the de-trended expected effective spread for the trading day, otherwise zero. The dummy variables “morn” and “eve” respectively equal 1 if the five minute interval is between 9:30 and 12 noon or 2:00pm to 4:00pm, and zero otherwise. The sample spans the years 1993 to 2002 inclusive, and consists of all stocks that traded every day on the NYSE during the sample period. The eighths regime spans January 1, 1993 to June 23, 1997, the sixteenths June 24, 1997 to January 28, 2001, and the decimal regime spans the remainder of the sample period. Each firm size category includes one-third of all sample firms. The sample size, \(n\), is the number of available five-minute intervals in each sub-period, as reported in the table 1, less the number of days; (the first interval of each trading day is used only in the lagged regressors.)

(Table continued on next page.)
<table>
<thead>
<tr>
<th>Regime</th>
<th>Large Firms</th>
<th>Mid-Cap Firms</th>
<th>Small Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>Coefficient</td>
<td>Coefficient</td>
</tr>
<tr>
<td></td>
<td>t-statistic</td>
<td>t-statistic</td>
<td>t-statistic</td>
</tr>
<tr>
<td>Eighths Regime</td>
<td>0.0291</td>
<td>24.71</td>
<td>0.0201</td>
</tr>
<tr>
<td></td>
<td>0.0205</td>
<td>5.95</td>
<td>0.0092</td>
</tr>
<tr>
<td></td>
<td>0.0089</td>
<td>5.12</td>
<td>0.0102</td>
</tr>
<tr>
<td></td>
<td>0.0088</td>
<td>5.42</td>
<td>0.0093</td>
</tr>
<tr>
<td></td>
<td>0.0078</td>
<td>1.54</td>
<td>0.0021</td>
</tr>
<tr>
<td></td>
<td>0.0057</td>
<td>1.22</td>
<td>-0.0003</td>
</tr>
<tr>
<td></td>
<td>-0.0015</td>
<td>-7.34</td>
<td>-0.0017</td>
</tr>
<tr>
<td></td>
<td>0.0403</td>
<td>0.0573</td>
<td>0.0556</td>
</tr>
<tr>
<td>Sixteenths Regime</td>
<td>0.0570</td>
<td>6.54</td>
<td>0.0350</td>
</tr>
<tr>
<td></td>
<td>-0.0011</td>
<td>-0.22</td>
<td>0.0130</td>
</tr>
<tr>
<td></td>
<td>-0.0084</td>
<td>-1.84</td>
<td>0.0087</td>
</tr>
<tr>
<td></td>
<td>0.0024</td>
<td>0.16</td>
<td>0.0097</td>
</tr>
<tr>
<td></td>
<td>0.0089</td>
<td>0.61</td>
<td>0.0093</td>
</tr>
<tr>
<td></td>
<td>-0.0018</td>
<td>-4.09</td>
<td>-0.0034</td>
</tr>
<tr>
<td></td>
<td>0.0038</td>
<td>0.0216</td>
<td>0.0374</td>
</tr>
<tr>
<td>Decimal Regime</td>
<td>0.0061</td>
<td>0.95</td>
<td>0.0177</td>
</tr>
<tr>
<td></td>
<td>-0.0105</td>
<td>-0.67</td>
<td>-0.0153</td>
</tr>
<tr>
<td></td>
<td>-0.0003</td>
<td>-0.03</td>
<td>0.0089</td>
</tr>
<tr>
<td></td>
<td>0.0069</td>
<td>0.80</td>
<td>-0.0050</td>
</tr>
<tr>
<td></td>
<td>0.0362</td>
<td>1.60</td>
<td>0.0351</td>
</tr>
<tr>
<td></td>
<td>0.0821</td>
<td>3.67</td>
<td>0.0815</td>
</tr>
<tr>
<td></td>
<td>-0.0008</td>
<td>-1.09</td>
<td>-0.0019</td>
</tr>
<tr>
<td></td>
<td>0.0010</td>
<td>0.0025</td>
<td>0.0089</td>
</tr>
</tbody>
</table>
### Table 7- Daily Variance Ratios and Autocorrelations

Panel A presents (open-to-close)/ (close-to-open) per hour return variance ratios, based on mid-quote returns, across three tick size regimes (eighths, sixteenths, and decimals). Panel B presents daily first order return autocorrelations across these regimes. The sample spans the years 1993 to 2002 inclusive, and consists of all stocks that traded every day on the NYSE during the sample period. The eighths regime spans January 1, 1993 to June 23, 1997, the sixteenths June 24, 1997 to January 28, 2001, and the decimal regime spans the remainder of the sample period. All numbers in Panel A are significantly different from unity at the 5% level. In Panel B, P-values for the relevant numbers being statistically different from zero are provided in parentheses. A * in Panel A (Panel B) indicates that the relevant number is statistically higher than (different from) the corresponding number in the eighths regime at the 5% level. The statistical conclusions for Panel A derive from bootstrap analyses described in the text.

#### Panel A: Per Hour Open/Close Variance Ratios

<table>
<thead>
<tr>
<th></th>
<th>Eighths</th>
<th>Sixteenths</th>
<th>Decimals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large firms</td>
<td>8.78</td>
<td>10.6</td>
<td>11.1</td>
</tr>
<tr>
<td>Mid-cap firms</td>
<td>12.4</td>
<td>19.0*</td>
<td>17.3*</td>
</tr>
<tr>
<td>Small firms</td>
<td>15.0</td>
<td>24.5*</td>
<td>20.3</td>
</tr>
</tbody>
</table>

#### Panel B: First order autocorrelations of daily returns

<table>
<thead>
<tr>
<th></th>
<th>Eighths</th>
<th>Sixteenths</th>
<th>Decimals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large firms</td>
<td>0.0730</td>
<td>-0.0131 (0.696)</td>
<td>-0.0083 (0.858)</td>
</tr>
<tr>
<td>Mid-cap firms</td>
<td>0.1168 (0.000)</td>
<td>0.0842 (0.012)</td>
<td>0.0621 (0.177)</td>
</tr>
<tr>
<td>Small firms</td>
<td>0.1993 (0.000)</td>
<td>0.0961* (0.004)</td>
<td>0.0386* (0.402)</td>
</tr>
</tbody>
</table>
Figure 1. Market Inefficiency Trend, NYSE, 1993-2002
Five-Minute Return Predictions Using Lagged (by five minutes) Dollar Order Imbalance
Figure 2. Value-Weighted Daily Average Effective Spread, NYSE, 1993-2002