An analysis of trade-size clustering and its relation to stealth trading

Article *in* Journal of Financial Economics · February 2007

DOI: 10.1016/j.jfineco.2006.02.005

2 authors:

Gordon J. Alexander
University of Minnesota Twin Cities

Mark A. Peterson
Southern Illinois University Carbondale

Available online at www.sciencedirect.com
ScienceDirect
Journal of Financial Economics 84 (2007) 435–471
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/jfec

An analysis of trade-size clustering and its relation to stealth trading ☆

Gordon J. Alexander[a,*], Mark A. Peterson[b]

[a]Carlson School of Management, University of Minnesota, Minneapolis, MN 55455, USA
[b]College of Business and Administration, Southern Illinois University, Carbondale, IL 62901, USA

Received 16 June 2004; received in revised form 8 November 2005; accepted 24 February 2006
Available online 17 January 2007

Abstract

NYSE and Nasdaq trades increasingly cluster on multiples of 500, 1,000, and 5,000 shares. Such clustering varies over time and across stocks, and tends to increase with the level of trading activity. Furthermore, rounded trades tend to have more persistence both in occurrence and in trade initiation. Finally, medium-sized rounded trades tend to have greater relative price impact than large rounded trades. From these observations we surmise that trade-size clustering is consistent, at least in part, with the actions of stealth traders who tend to use medium-sized rounded transactions in an attempt to disguise their trades.

JEL classification: D82; G12; G14

Keywords: Trade-size clustering; Stealth trading; Informed traders; Price impact

☆The authors thank an anonymous referee, seminar participants at the Office of Economic Analysis at the US Securities and Exchange Commission, UCLA, and the University of Missouri-Columbia, and Adam Y. C. Lei for helpful comments. This paper, previously entitled "Trade-size Clustering on the NYSE: An Alternative Test of the Negotiations Hypothesis," was presented at the 2003 Annual Meeting of the Financial Management Association, where it received the Nasdaq Award for Best Paper in Market Microstructure.
*Corresponding author. Tel.: +1 612 624 8598.
E-mail address: galexander@csom.umn.edu (G.J. Alexander).

0304-405X/$ - see front matter
doi:10.1016/j.jfineco.2006.02.005

1. Introduction

In this paper we examine the relatively unexplored area of trade-size clustering in financial markets. While several studies examine the relations between trade size and things such as prices, information, and spreads, our focus is different in that we seek to determine whether trade-size clustering exists for the typical NYSE or Nasdaq stock, and if so, its economic consequences.[1]

Price clustering has been found in US stock markets and many other settings such as initial public offerings, seasoned equity offerings, foreign exchange markets, gold markets, derivatives markets, and bank deposit rates.[2] Various hypotheses have been proposed to explain such clustering. For instance, Ball, Torous, and Tschoegl (1985), Harris (1991), and Grossman, Miller, Cone, Fischel, and Ross (1997) develop the negotiations hypothesis. This hypothesis argues that while a more "precise" price that is mutually acceptable to both the buyer and the seller can be reached by continuing negotiations, the incremental benefit to each side decreases and the exposure of each side to reporting and price risk increases.[3] Thus, the negotiations hypothesis suggests that a trade will take place when each party views the incremental risks from continuing negotiations as offsetting the incremental benefits. In this scenario, price clustering occurs naturally as agents try to simplify the negotiation process.

In a second line of reasoning, Wyckoff (1963, pp. 106–107) notes that "we think in round numbers and try to sell at round numbers." Similarly, Niederhoffer and Osborne (1966, p. 914) state that "the tendency of traders to prefer integers seems to be a fundamental and stable principle of stock market psychology," and Ikenberry and Weston (2003, p. 3), who cite previous studies of people having a preference for even numbers or numbers ending in a zero or five, argue that "price clustering may be a collective preference by investors to voluntarily trade at particular price levels in order to … minimize cognitive processing costs." We refer to this line of reasoning for price clustering as the behavioral hypothesis.

Third, Christie and Schultz (1994) hypothesize that the relative absence of odd-eighth quotes in the Nasdaq market was the result of implicit collusion among dealers for the purpose of increasing market-making revenues; the authors provide supporting evidence. Thus, the collusion hypothesis is a third hypothesis for price clustering.

Of these three explanations of price clustering, collusion is not likely to be applicable to trade-size clustering, as the gains from colluding are not apparent. However, the negotiations hypothesis is plausible, as using a rounded size may reduce reporting risk and speed up negotiations. The behavioral hypothesis is also plausible. Institutional investors, who have greater involvement relative to individuals as trade size increases (see, for example, Chakravarty, 2001, p. 300), may tend to favor rounded sizes because doing so is

[1]See, for example, Easley and O'Hara (1987) and Brennan and Subrahmanyam (1998).

[2]See, for example, Ball, Torous, and Tschoegl (1985), Harris (1991), Christie and Schultz (1994), Grossman, Miller, Cone, Fischel, and Ross (1997), Gwilym, Clare, and Thomas (1998), Kahn, Pennachi, and Sopranzetti (1999), Chen and Ritter (2000), Yeoman (2001), Sopranzetti and Datar (2002), Ikenberry and Weston (2003), Mola and Loughran (2004), and Moulton (2005).

[3]*Reporting risk* refers to an error in reporting the agreed-upon price that subsequently leads to costs being incurred in resolving the dispute; it is more likely to occur when the agreed-upon price is more "precise." *Price risk* refers to the chance that the price will move away from the reservation price of one of the traders, and is increasingly likely to occur as negotiations continue.

increasingly inconsequential in terms of portfolio management. For example, rounding an order for 14,900 shares to 15,000 is less consequential than rounding an order from 4,900 shares to 5,000, as the effect in the aggregate purchase price and the security's portfolio weight is proportionally much smaller.

Turning to the issue of trade-size clustering, Hodrick and Moulton (2005) recently analyzed such clustering in a rational expectations framework. In their model, when many heterogeneous uninformed investors are present an asset will trade at an increasing number of distinct sizes as the desire by investors for trading exact quantities increases. Accordingly, Moulton (2005) hypothesizes that there is less trade-size clustering shortly before the end of calendar quarters, perhaps because of the desire of portfolio managers to have their portfolios more fully aligned with their given objectives. Note that this conjecture is consistent with what Moulton observes in her empirical analysis of the foreign exchange market. We refer to this hypothesis as the end-of-quarter hypothesis.

Several other papers are also relevant to our work. First, Admati and Pfleiderer (1988) show that informed traders will attempt to disguise their trades by placing them during times of abnormally heavy trading. Second, Barclay and Warner (1993), Hasbrouck (1995), and Chakravarty (2001) document the presence of stealth trading by institutional investors by showing that medium-sized trades tend to have a disproportionately greater aggregate price impact that is attributable to informed traders disguising at least some of their trades by using such sizes (here a medium-sized trade is defined to be a trade between 500 and 9,900 shares). Third, Keim and Madhavan (1995) and others provide empirical evidence that institutions often break up their orders and fill them over time. Fourth, Chordia and Subrahmanyam (2004) present a model in which traders find it optimal under certain conditions to break up their orders, and find empirical evidence that trades in one direction tend to be followed by trades in the same direction, supporting their model.[4] We refer to the attempt to disguise large informed trades by using smaller sizes as the stealth trading hypothesis.

Our analysis of NYSE and Nasdaq stocks indicates that orders and trades tend to cluster at three levels. In increasing intensity, these levels are multiples of 500, 1,000, and 5,000 shares. For example, there are significantly more trades for 3,500 shares than for either 3,400 or 3,600 shares. There are significantly more trades for 4,000 shares than for 3,500 or 4,500 shares, and there are significantly more trades for 5,000 shares than at 4,000 or 6,000 shares. Additionally, there is substantial variation in the amount of trade-size clustering over time for the typical stock and across stocks on a typical day, with some diminution in the overall amount after decimalization. Surprisingly, trade-size clustering tends to occur simultaneously with price clustering, suggesting the possibility that the negotiations hypothesis is behind both forms of clustering.

In subsequent analysis we make the following observations. First, rounded trades tend to have a greater price impact than unrounded trades. Second, rounded trades of medium size tend have a greater relative price impact than those of large size. Third, rounded trades are more likely to occur if the previous trade was rounded. Fourth, trade initiation is more persistent for rounded trades, i.e., buyer- (seller-) initiated trades are more likely to follow buyer- (seller-) initiated trades when the trades are rounded. Fifth, trade-size rounding tends to increase when trading activity is abnormally heavy. From these observations we

[4] Bertsimas and Lo (1998) derive optimal trading strategies for breaking up a large block of shares over time while Seppi (1990) presents a model for why an institutional investor would not want to break up a large block.

surmise that trade-size clustering is consistent, at least in part, with the actions of stealth traders who tend to use rounded medium-sized transactions in an attempt to disguise their trades, particularly during periods when there is an unusually large amount of trading taking place.

Other observations are that trade-size clustering is more likely to occur for large orders and orders that exceed the quoted depth, but is less likely to occur for high-priced stocks and stocks that are members of the S&P 1500. Interestingly, rounded trades are less likely to occur if the trade occurs shortly before quarter-end, consistent with the end-of-quarter hypothesis.

These observations might be of value to a stealth trader in that using a sequence of unrounded medium sizes appears to be less costly than a sequence of rounded medium sizes because of the smaller price impact associated with the unrounded medium sizes.[5] However, without order data this implication is more speculative than conclusive. The results also suggest that specialists and dealers are not fooled by the actions of stealth traders, as they appear to adjust the spread midpoint within five minutes of the execution of a rounded medium-sized trade.

This paper is organized as follows. Section 2 describes the distribution of order and trade size and presents evidence of trade-size clustering. Section 3 explores the relation between trade-size and price clustering. Section 4 presents a bivariate probit model to examine variables that might influence whether a rounded or unrounded trade size is used when a trade is executed. Sections 5 and 6 compare effective spreads and price impacts, respectively, for rounded and unrounded trade sizes. Section 7 examines alternative explanations for trade-size clustering. Lastly, Section 8 presents the conclusions.

2. Presence of size clustering

Our initial analysis for the presence of size clustering involves the sample of NYSE stocks contained in the TORQ (Trades, Orders, Reports, and Quotes) database, the only publicly available database that contains NYSE information on both orders and trades.

2.1. *TORQ sample*

The TORQ database consists of 144 randomly chosen NYSE stocks during 63 trading days from November 1990 through January 1991. More specifically, TORQ contains SuperDot system order data in the SOD file and audit trail data in the CAUD file, as well as trade and quote data from the Securities Industry Automation Corporation (SIAC). Because more recent data are available only for NYSE trades, we examine TORQ to see if there are any notable differences in the extent of size clustering between orders and trades. If none are observed, then trade-size results using samples from the larger and more recent TAQ (Trades and Quotes) database of all NYSE trades are likely to be applicable to orders.[6] We also analyze samples of Nasdaq trades in order to determine whether the NYSE results are attributable to its different market structure.

[5] Offsetting the benefit of smaller price impacts associated with medium-sized unrounded trades is the observation that under penny pricing, such trades tend to have wider effective spreads.

[6] Chakravarty (2001) notes there is noise in the reported trade size on the NYSE, as specialists have flexibility in how they report trades when there are multiple parties on at least one side of the trade. For example, a single

2.2. TAQ samples

The second, third, and fourth samples involve trades and quotes of three separate sets of 200 randomly chosen NYSE common stocks during 1995, 1998, and 2002. We choose these in order to analyze trade-size clustering for a full calendar year under three distinct pricing regimes. On June 24, 1997, the minimum tick size changed from \$1/8 to \$1/16 (1995 versus 1998) and on January 29, 2001, trades began being executed and reported using decimals (1998 versus 2002). Comparing samples from the three different sample years allows us to analyze tradeoffs between coarseness in price regimes versus coarseness in trade sizes.

The fifth and sixth samples involve trades and quotes of two separate sets of 200 randomly chosen Nasdaq common stocks during 1998 and 2002. While a NYSE sample was taken for 1995, we omit a Nasdaq sample for this year because the implementation of the Order Handling Rules of 1997 substantially changed the way stocks are traded on Nasdaq, requiring, for example, the display of limit orders.[7]

In sampling the NYSE and Nasdaq universe of stocks, we exclude certain stocks in order to reduce the chance of admitting potentially confounding effects. Specifically, we use the CRSP database to identify and exclude any stock that (1) was other than an ordinary common share, (2) changed its ticker symbol over the year, (3) split during the year, or (4) was an ADR, closed-end fund, REIT, or share of beneficial interest, unit, or certificate. Many NYSE stocks trade elsewhere such as on regional stock exchanges or the third market. In our analysis of the NYSE stock samples, only those trades executed on the NYSE are examined, except as indicated.[8] In addition, opening trades on the NYSE and trades prior to 9:45 am on Nasdaq are excluded.

Table 1 provides a cross-sectional summary of the sample stocks and their trading statistics. Not surprisingly, the Nasdaq samples consist of stocks that have, on average, lower market capitalizations, share prices, and membership in the S&P 1500 relative to the NYSE stock samples.

2.3. Distribution of order and trade size

We begin by using the TORQ and TAQ databases to estimate the probability distributions of order and trade size in increments of 100 shares for each stock, and then we calculate the average probability for each order or trade size across stocks. While only the round lot portions of *trades* that include partial odd lots are reported in the TORQ and TAQ files, the TORQ *order* file includes partial odd lots.[9] To be consistent with the trade files, order sizes are rounded down to the lower 100-share multiple unless the resulting size is a multiple of 500 shares, in which case the size is rounded up because any order that

(*footnote continued*)
buyer of 8,000 shares could be matched with two 4,000-share sellers, in which case the specialist could report one trade of 8,000 shares or two trades of 4,000 shares. By comparing orders and trades with TORQ data, we can determine the extent to which such reporting introduces bias in trade sizes. The results reported in our tables indicate no substantive differences.

[7]See *Federal Register* (1996).

[8]Our analysis of off-exchange trades in NYSE stocks reveals a similar pattern of trade-size clustering.

[9]Partial odd lots are involved in 10.2% of the orders on TORQ.

Table 1
Cross-sectional daily trade summary and stock characteristics

	NYSE TORQ 1990–1991		NYSE TAQ trades			Nasdaq TAQ trades	
	Orders	Trades	1995	1998	2002	1998	2002
A. Number of days	63	63	252	252	252	252	252
B. Sample							
Number of stocks	144	144	200	200	200	200	200
Number in S&P 1500	40	40	105	112	134	17	28
Market capitalization (in $billions)							
Mean	2.8	2.8	2.0	4.4	5.1	0.2	0.4
Median	0.3	0.3	0.6	0.6	1.2	0.1	0.1
Standard deviation	9.8	9.8	4.2	14.6	16.8	0.6	1.1
C. Orders and trades							
Average price							
Mean	NA	$19.40	$26.91	$29.62	$25.05	$13.24	$11.19
Median	NA	$15.94	$22.12	$25.35	$22.10	$10.49	$8.23
Standard deviation	NA	$19.35	$17.92	$25.00	$16.06	$11.41	$9.80
Average number of trades/orders per day							
Overall							
Mean	91.8	49.0	61.8	127.3	571.5	79.9	629.7
Median	28.2	16.4	35.7	52.1	358.8	30.4	85.1
Standard deviation	197.9	105.6	80.3	225.9	645.1	159.2	2,137.2
Small							
Mean	44.9	20.6	22.8	46.0	326.5	28.2	437.7
Median	11.1	6.5	13.8	18.2	245.8	8.5	46.5
Standard deviation	103.1	44.4	25.6	79.5	301.9	58.4	1,461.0
Medium							
Mean	44.8	25.9	35.0	75.7	231.5	50.2	188.4
Median	14.0	8.0	19.2	30.6	95.5	20.1	24.3
Standard deviation	94.5	60.4	49.5	137.7	344.2	100.6	684.6
Large							
Mean	2.1	2.5	3.9	5.6	13.5	1.5	3.8
Median	0.5	0.3	1.4	1.1	2.0	0.4	0.5
Standard deviation	5.1	6.2	7.2	11.2	33.3	3.5	11.3

This table reports cross-sectional statistics for the (1) daily average number of trades (or orders) per stock, and (2) average trade price for all of the stocks in TORQ and for 200 randomly selected NYSE and Nasdaq common stocks in TAQ. TORQ orders include all regular-way buy and sell orders. Trades from the TAQ database and from the NYSE's CT file are filtered as in Bessembinder (1997). Small-size trades are for 100–400 shares (100–499 for TORQ orders). Medium-sized trades are for 500–9,900 shares (500–9,999 for TORQ orders). Large-sized trades are for 10,000 or more shares. Information regarding the S&P 1500 comes from *S&P Reports*. Market capitalization in $billions, is based on the previous year's ending price and shares outstanding. Trades executed on the regional exchanges, opening trades on the NYSE, and trades executed on Nasdaq prior to 9:45 am are excluded.

involves an odd lot is, by our definition, not rounded.[10] Note that this procedure is conservative in its treatment of order size rounding, as it biases downward the reported frequencies of sizes involving multiples of 500 shares.

[10]Thus, an order for 490 shares is treated as an order for 400 shares and an order for 510 shares is treated as an order for 600 shares, whereas orders for either 710 or 790 shares are rounded down to 700 shares.

Fig. 1 displays the average distributions on a log scale (up to 6,000 shares to conserve space) and suggests that there are three levels of size clustering in both orders and trades. In increasing intensity, these levels are multiples of 500, 1,000, and 5,000 shares. For example, there are more trades at 3,500 shares than either 3,400 or 3,600 shares. Furthermore, there are more trades at 4,000 shares than at 3,500 or 4,500 shares. Lastly, there are more trades at 5,000 shares than at 4,000 or 6,000 shares. Interestingly, there appears to be little difference (1) in the pattern of clustering between orders and trades in 1990–1991, (2) in trades over the four sample periods, and (3) between NYSE and Nasdaq trades.[11]

To test whether size clustering is significant, we estimate the following regression equation separately for orders and trades:

$$LnFreq_i = \alpha + \beta_{500}\, D500_i + \beta_{1000}\, D1000_i + \beta_{5000}\, D5000_i + \beta_{LnSize}\, LnSize_i + \varepsilon_i, \tag{1}$$

where ε_i is a random error term, and

$LnFreq_i$ = natural log of the percentage of orders or trades in the average stock that are of size i,

$D500_i = 1$ if order or trade size i is a multiple of 500 shares, 0 otherwise,

$D1000_i = 1$ if order or trade size i is a multiple of 1,000 shares, 0 otherwise,

$D5000_i = 1$ if order or trade size i is a multiple of 5,000 shares, 0 otherwise, and

$LnSize_i$ = natural log of order or trade size i measured in number of shares.

Table 2 reports the regression results, correcting for autocorrelation in the error terms, for TORQ in Panel A and TAQ in Panels B and C. The results indicate that all the estimates of β_{500} are positive and significant for orders and trades. This observation is consistent with clustering occurring on multiples of 500 shares. Furthermore, all the estimates of β_{1000} and β_{5000} are also positive and significant, indicating increasing levels of clustering associated with multiples of 1,000 and 5,000 shares. The adjusted-R^2 values are similar, all being in excess of 0.92 and attesting to the good fit of the model. Importantly, the results are very similar across orders and trades as well as over time and across markets.

Fig. 1 and the significant negative estimates of β_{LnSize} in Table 2 indicate that the frequency of both orders and trades decrease with size. To test for clustering while controlling for size levels, we conduct one-tailed binomial z-tests of proportions of the average order-size and trade-size clustering frequencies by aggregating the data across stocks. This aggregation is done for each sample year in individual and pooled trade-size levels. Consider the individual tests using, for example, 1995 NYSE trades. Initially, the sum (N) of the number of trades at 400, 500, and 600 shares across all stocks is 525,688, with the 500-share proportion p being $295{,}226/525{,}688 = 0.5616$ with $q = 1 - p = 1 - 0.5616 = 0.4384$. Accordingly, the square root (s) of pq/N is $(0.5616 \times 0.4384/525{,}688)^{1/2} = 0.0007$ and the resulting z-statistic is $(p - 0.3333)/s = (0.5616 - 0.3333)/0.0007 = 333.55$; the p-value from the unit normal cumulative probability density function is 0.0000. We repeat this process for each multiple of 500 shares, ending with 20,000 shares, which results in 40

[11] Our analysis of time-weighted quoted depths reveals similar levels of size clustering, where the time-weighted depth is based on the aggregate time a particular depth existed during the sample year across all stocks. For the NYSE stocks, we define the depth as the maximum contemporaneous quoted depth across all markets at the inside quotes. For the Nasdaq stocks we use the reported depths at the inside quotes disseminated by Nasdaq.

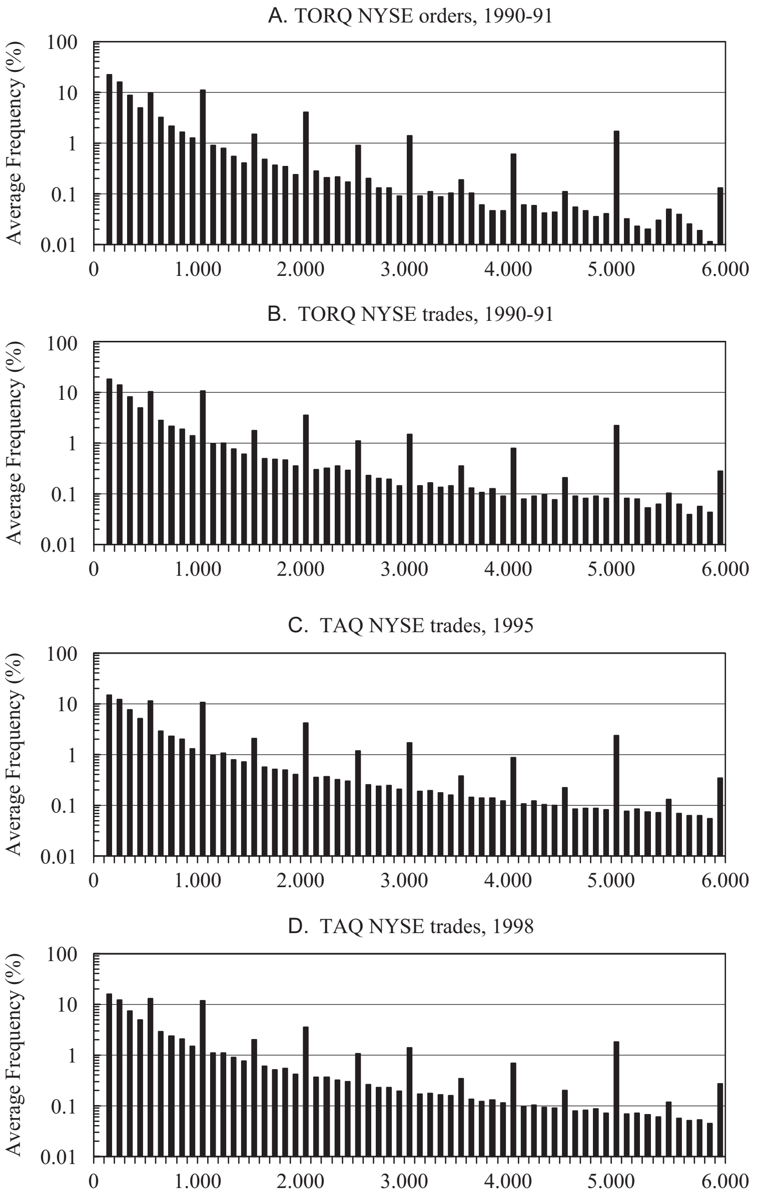

Fig. 1. Distribution of order and trade sizes. These figures display the cross-sectional average probability distribution of order and trade size of stocks in the TORQ and TAQ databases as described in Table 1. For TORQ, regular-way buy and sell orders are taken from the SOD file. SOD orders with an odd lot component are rounded down to the lower size in hundreds, except for sizes that are multiples of 500. For example, an order for 510 shares would be rounded to 600 shares, but an order for 790 shares would be rounded to 700 shares. The figures are truncated to 6,000 shares to conserve space. (A) TORQ NYSE orders, 1990–1991; (B) TORQ NYSE trades, 1990–1991; (C) TAQ NYSE trades, 1995; (D) TAQ NYSE trades, 1998; (E) TAQ NYSE trades, 2002; (F) TAQ Nasdaq trades, 1998; and (G) TAQ Nasdaq trades, 2002.

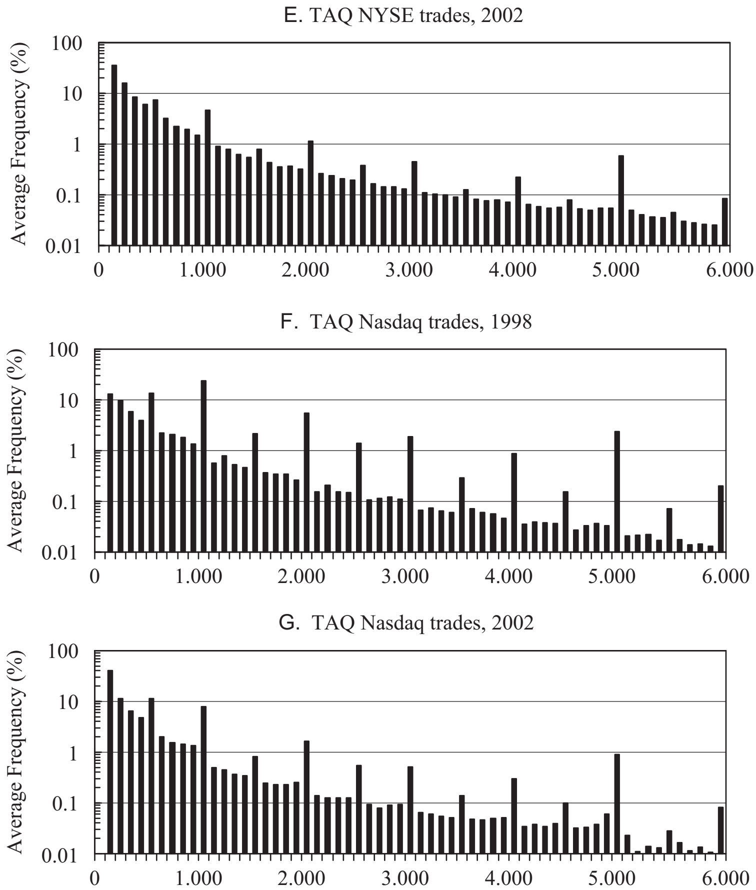

Fig. 1. (*Continued*)

individual test p-values. We determine the percentage of these p-values that are less than or equal to 0.05.

For pooled tests, we sum all trades that are multiples of 500 shares, along with all trades that are either 100 shares more or less than multiples of 500 shares. Thus, we sum the number of trades at 400, 500, 600, 900, 1,000, 1,100, and so on ending with 20,100 shares, along with the number of trades at multiples of 500 shares (i.e., 500, 1,000, 1,500, and so on, ending with 20,000 shares). These two sums are used to determine a pooled proportion p and ultimately a pooled z-statistic and p-value in the same manner as for individual trade sizes.

These tests have a null hypothesis that the probability of the interior size, relative to the interior and two adjacent exterior sizes, is less than or equal to 1/3. Given the results in Table 2, where frequency is a linear function of $1/size^2$, this is a conservative test in that it is biased toward not rejecting the null. That is, if the null is that rounding is not present and trade frequency is convex in size, then the relative frequency of a rounded size is strictly less than 1/3.

Table 2
Regression tests for the presence of size clustering

	Coefficient estimates					
	Intercept	β_{500}	β_{1000}	β_{5000}	β_{LnSize}	Adjusted R^2
A. NYSE 1990–1991						
Orders	13.030**	0.929**	1.192**	2.418**	−2.519**	0.924
	(1.098)	(0.118)	(0.182)	(0.289)	(0.121)	
Trades	9.377**	0.942**	0.900**	2.164**	−1.984**	0.970
	(0.655)	(0.084)	(0.117)	(0.186)	(0.072)	
B. NYSE trades						
1995	8.497**	0.856**	0.908**	2.164**	−1.855**	0.982
	(0.590)	(0.081)	(0.121)	(0.189)	(0.064)	
1998	9.562**	0.954**	0.884**	2.011**	−2.005**	0.981
	(0.739)	(0.067)	(0.094)	(0.184)	(0.080)	
2002	10.291**	0.424**	0.819**	2.062**	−2.143**	0.990
	(0.503)	(0.052)	(0.065)	(0.202)	(0.055)	
C. Nasdaq trades						
1998	12.059**	1.573**	1.094**	2.405**	−2.425**	0.960
	(1.055)	(0.156)	(0.204)	(0.185)	(0.115)	
2002	12.242**	1.042**	1.118**	2.189**	−2.470**	0.969
	(0.739)	(0.132)	(0.142)	(0.163)	(0.081)	

We estimate a regression using the cross-sectional average size probability distributions (see Fig. 1) with the following model:

$$LnFreq_i = \alpha + \beta_{500}\,D500_i + \beta_{1000}\,D1000_i + \beta_{5000}\,D5000_i + \beta_{LnSize}\,LnSize_i + \varepsilon_i,$$

where $LnFreq_i$ is the natural log of the percentage of orders or trades in the average stock that are of size i, $D500_i$ is a dummy variable set to one if size i is a multiple of 500 shares, zero otherwise, $D1000_i$ is a dummy variable set to one if size i is a multiple of 1,000 shares, zero otherwise, $D5000_i$ is a dummy variable set to one if size i is a multiple of 5,000 shares, zero otherwise, and $LnSize_i$ is the natural log of the order or trade size i in shares. **/* indicates significance at the 1%/5% level using a t-test. Standard errors of the coefficient estimates, corrected for autocorrelation, are in parentheses.

Looking first at the pooled tests, the top part of Panel A of Table 3 indicates that over 70% of the orders or trades are at a rounded size except for 2002, when it is 51.6% for NYSE and 61.5% for Nasdaq. Hence, it is not surprising that the binomial z-tests reject the null at the 1% significance level for every sample. Moving on to the individual tests, it can be seen that 100% of the 40 multiples of 500 shares in each sample have significant rounding in size with one minor exception (TORQ orders), where 80%, or 32 of the 40 clusters, report significant rounding. Thus, there appears to be size clustering at multiples of 500 shares.

Orders and trades are similarly examined for clustering at multiples of 1,000 shares. Here, for each sample we contrast the frequencies at each multiple with trade sizes that are *exactly* 500 shares more or less than the multiple, beginning with a trade size of 1,000 shares. Thus, for individual tests the number of trades at 1,000 shares is compared to the summed number of trades at 500, 1,000, and 1,500 shares, with similar additional tests for other multiples of 1,000 shares. As before, we also conduct a pooled test resulting in a

Table 3
Variation (%) in trade-size clustering

	NYSE 1990–1991		NYSE trades			Nasdaq trades	
	Orders	Trades	1995	1998	2002	1998	2002
A. Percent of trades at rounded sizes							
% Rounded, 500X	71.5**	71.6**	72.1**	70.1**	51.6**	83.3**	61.5**
	(80.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)
% Rounded, 1000X	55.3**	55.5**	56.0**	52.8**	45.1**	66.6%**	43.7**
	(89.5%)	(100.0)	(100.0)	(100.0)	(100.0)	(89.5)	(78.9)
% Rounded, 5000X	75.3**	76.5**	75.2**	72.9**	72.1**	75.8**	77.4**
	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)	(100.0)
B. Distribution of daily cross-sectional percent rounded							
Mean	72.1	67.9	69.2	69.5	50.4	81.4	65.2
	(2.0)	(2.4)	(1.8)	(1.7)	(2.1)	(2.0)	(2.7)
Standard deviation	22.4	24.3	19.6	18.3	15.6	19.4	21.2
	(1.9)	(2.0)	(1.6)	(1.5)	(1.5)	(2.0)	(1.5)
C. Distribution of individual stock time-series percent rounded							
Mean	72.4	68.6	69.4	69.5	50.6	81.0	65.8
	(10.7)	(9.9)	(7.5)	(6.5)	(7.3)	(5.8)	(8.3)
Standard deviation	19.1	22.7	17.0	15.8	12.1	18.7	19.5
	(9.5)	(11.4)	(8.4)	(9.0)	(8.4)	(10.0)	(10.8)

Panel A reports the percent of trades at rounded sizes. The row labeled % Rounded, 500X is based on rounded sizes at 500X, $X = 1$–40 and unrounded sizes plus or minus 100 shares of rounded size. The row labeled % Rounded, 1,000X is based on rounded sizes at 1,000X, $X = 1$–19 and unrounded sizes plus or minus *exactly* 500 shares of rounded size. The row labeled % Rounded, 5,000X is based on rounded sizes at 5,000X, $X = 1$–3 and unrounded sizes plus or minus *exactly* 1,000 shares of rounded size. The first figure in each cell reports the pooled percent rounded, and the figure in parentheses is the percent of clusters at 500X, 1,000X, or 5,000X for which significant rounding in size exists based on a binomial z-test of proportions. **/* indicates significance at the 1%/5% level. Panel B reports measures of the cross-sectional distribution of size clustering around multiples of 500 shares plus or minus 100 shares. For each day in the sample period, the percent of trades that involve a multiple of 500 shares is determined for each stock, ignoring all trades not within plus or minus 100 shares of a 500-share multiple. The cross-sectional mean and standard deviation of this distribution are calculated daily. Each cell contains the mean and, in parentheses, the standard deviation of the daily means and standard deviations over the sample period. Panel C reports measures of the time-series distribution of the percent rounded for each stock. The time-series mean and standard deviation are calculated for each stock. Each cell similarly contains the mean and standard deviation of the time-series means and standard deviations across stocks.

single p-value for each sample where all trades at multiples of 1,000 shares are summed and compared to the sum of all trades that are either multiples of 1,000 shares or exactly 500 shares more or less than these multiples. Hence, both types of binomial z-tests are designed to determine whether there is a second level of clustering on multiples of 1,000 shares in comparison to trade sizes that are exactly 500 shares more or less than the multiples. As the middle part of Panel A shows, the pooled test results indicate that over 50% of orders or trades involve a multiple of 1,000 shares except for 2002, when it is 45.1% for NYSE and 43.7% for Nasdaq. Again, it is not surprising that the null is rejected at the 1% significance level for every one of these pooled tests. Furthermore, the individual tests reject the null at the 5% level for each one of the 19 multiples of 1,000 shares for each sample except for two multiples with TORQ orders, two with 1998 Nasdaq trades, and four with 2002 Nasdaq trades. Thus, there appears to be a second level of size clustering at multiples of 1,000 shares.

Lastly, we examine trades for clustering at multiples of 5,000 shares by comparing the trade-size frequencies at each multiple of 5,000 shares with trade sizes that are *exactly* 1,000 shares more or less than the multiple, beginning with a trade size of 5,000 shares. Thus, for individual tests the number of trades at 5,000 shares is compared to the summed number of trades at 4,000, 5,000, and 6,000 shares, with similar additional tests for other multiples of 5,000 shares. Pooled tests in which the data are summed over all 5,000-share multiples, plus or minus 1,000 shares, are also conducted. As the bottom part of Panel A shows, the pooled test results support at the 1% significance level the notion of size clustering at multiples of 5,000 shares, as over 70% of orders or trades involve a multiple of 5,000 shares. Because the individual test results similarly reject the null for all of the three multiples of 5,000 shares in each sample, there appears to be a third level of size clustering at 5,000 shares.

2.4. *Cross-sectional and time-series variation in size clustering*

Next, we seek to determine if size clustering occurs with consistency. This is done by calculating the percent of size clustering for each stock on each day as the number of trades at a multiple of 500 shares divided by the total number of trades. Note that all orders or trades that are neither a multiple of 500 shares nor within 100 shares of a 500-share multiple are excluded. Using these percentages, we conduct cross-sectional and time-series analyses of size clustering.

First, for each day in the sample period we construct a cross-sectional distribution of size clustering. Panel B of Table 3 reports the distribution of these daily cross-sectional means and standard deviations. The typical daily mean on the NYSE is around 70% until 2002, when it drops to 50.4%. On Nasdaq the typical daily mean is even higher at 81.4% in 1998, and then, as for the NYSE, drops to 65.2%. Thus, every day more than one-half of all orders and trades within 100 shares of a multiple of 500 shares involve a rounded size. However, there is substantial variation across stocks in the amount of clustering, as the typical standard deviation in the daily cross-sectional distribution is between 15.6% and 24.3%.[12] Finally, there is little variation in the mean and standard deviation from day to day, as the standard deviations of the daily means and standard deviations are less than 3% for each sample.

Second, we use the percent clustering for a given stock for every day in a given sample year to construct a time-series distribution of its size clustering. Panel C of Table 3 reports the distribution of the means of these time-series distributions, which are nearly the same as the means of the cross-sectional distributions shown in Panel B because both panels involve using the same data to take an average of averages, with the slight differences being attributable to a few stocks without any orders or trades on some days. More interesting is the range in the standard deviations of 5.8–10.7%, indicating substantial cross-sectional variability in the mean amount of size clustering.

Panel C also shows the means of the standard deviations of the time-series distributions, which fall steadily on the NYSE over time from 19.1% for orders and 22.7% for trades in

[12] Medians, while not reported, are nearly identical to the means.

1990–1991 to 12.1% in 2002. This finding indicates substantial but declining variability over time in the degree of size clustering for the typical stock. However, no such decline is noted on Nasdaq, where the percentage increases slightly from 18.7% in 1998 to 19.5% in 2002.

In summary, Table 3 indicates that (1) there are three levels of size clustering on the NYSE and Nasdaq, (2) there is substantial variation in size clustering each day across stocks, and (3) the amount of clustering in a given stock varies substantially over time. However, the overall amount of clustering across stocks seems relatively stable over time until decimalization, after which both NYSE and Nasdaq report a decline. Lastly, there are similar amounts of clustering in TORQ orders and trades in 1990–1991.

3. Clustering in price and size

We next seek to determine whether trade-size and price clustering occur with similar frequency and whether they occur independently of each other. In doing so, only trades that are either a multiple of 500 shares or 100 shares above or below a multiple of 500 shares are considered. On a given day, we calculate the proportion of these trades that involve a multiple of 500 shares for each stock and then average these proportions across stocks. In turn, we average these daily averages across the sample time period. The results, shown in the top line of Panel A, Table 4, represent the average daily percentage of trades that are rounded for the average stock in each sample.[13] Binomial z-tests of proportions reveal that these averages are all significantly greater than the null of 33.33%, indicating an abnormally large degree of trade-size clustering.[14]

The second line of Panel A, Table 4, presents the test results for price rounding. A trade is considered rounded in price if its price is a multiple of \$0.25 for 1990–1991 and 1995, a multiple of \$0.125 for 1998, and a multiple of \$0.05 for 2002. Hence, the null is that 50% of trades in 1990–1991, 1995, and 1998 have rounded prices, while in 2002 the null is 20%. Again, only those trades that involve a multiple of 500 shares, plus or minus 100 shares, are examined on a given day with the proportion rounded in price determined for each stock and, in turn, averaged across stocks. Finally, these daily averages are themselves averaged over the sample time period. Thus, the table reports the average daily percentage of prices that are rounded for the average stock in each sample. Binomial z-tests of proportions reveal that these averages are all significantly greater than their respective nulls. Thus, there is an abnormally large degree of price clustering in each sample year.[15]

We next conduct Wilcoxon signed rank tests in order to see if there is a different relative incidence of clustering in size than in price. In conducting these tests, the null is that the difference in the amount of clustering in price and size is equal to -16.67% ($=33.3\%-50.0\%$), except for 2002 when it is equal to 13.33% ($=33.3\%-20.0\%$). Because the test requires that the two samples be independent, the degree of size rounding

[13] Note that these results are similar to those shown in the top part of Panel A of Table 3, where the data are pooled.

[14] Only trade-size clustering is analyzed with the TORQ 1990–1991 data because of the unavailability of trade prices for orders that do not execute.

[15] Ikenberry and Weston (2003) also find evidence of price clustering at multiples of \$0.05 after decimalization.

Table 4
Frequency and independence of clustering in price and size

	NYSE				Nasdaq	
	1990–1991	1995	1998	2002	1998	2002
A. Percent rounded in size and pric-e						
% Rounded in size	68.3**	69.0**	69.5**	50.4**	81.3**	65.2**
% Rounded in price	53.3**	52.7**	60.0**	42.2**	73.8**	43.4**
Difference (%)	15.0**	16.3**	9.5**	8.2**	7.5**	21.8**
Null (%)	–16.7	–16.7	–16.7	13.3	–16.7	13.3
B. Contingency chi-square test statistics for trades at 500-share multiples$\pm$100 shares						
	6**	78**	2,153**	20,293**	514**	12,801**

This table reports the frequency and independence of clustering in size and price. Only trades at $500X$ shares plus or minus 100 shares are included. Even-numbered days are used in estimating trade-size clustering and odd-numbered days are used in estimating price clustering. On a given day, the percentage of trades that are rounded in price and the percentage of trades that are rounded in size are computed for each stock. A trade is considered rounded in size if it is at a multiple of 500 shares and a trade is considered rounded in price if it occurs on even eighths in 1990–1991 and 1995, even sixteenths in 1998, and even twentieths in 2002; unrounded in price are all other prices. The average percent rounded is calculated across stocks each day and the time-series average is reported in Panel A. For the row labeled "% Rounded in size," the null is 33.3% of trades are rounded in size. For the row labeled "% Rounded in price," the null is 50.0% of the trades are rounded in price for all years except 2002, when it is 20.0%. The % Rounded in size and % Rounded in price are tested using a sign test. The row labeled "Null" is computed as $33.3\% - 50.0\% = -16.7\%$ except for 2002, when it is $33.3\% - 20\% = 13.3\%$, and is tested using a Wilcoxon test. In Panel B the chi-square test statistic is used for the analysis of independence in contingency tables between trade-size and price rounding based on all trades in the sample period. **/* indicates significance at the 1%/5% level.

is based on even-numbered days while the degree of price rounding is based on odd-numbered days.[16] The results, shown in the bottom part of Panel A in Table 4, indicate that the null is rejected in every sample, suggesting there is significantly more rounding in size than in price, except in 2002 when there is more rounding in price than in size on the NYSE.

In order to investigate possible simultaneous clustering of price and size, we form 2×2 contingency tables based on all days in the sample period. Here, we classify each trade for each stock as having (1) a rounded or unrounded size, and (2) a rounded or unrounded price, again focusing on trades that are multiples of 500 shares, plus or minus 100 shares. The associated chi-square values, found in Panel B of Table 4, indicate that the number of observations involving unrounded sizes and prices as well as rounded sizes and prices are significantly greater than expected. Thus, it appears that, in general, size and price clustering are not independent.

In sum, our tests verify that significant trade-size and price clustering exists on both the NYSE and Nasdaq. Furthermore, clustering in size and price are not independent events. Importantly, the TORQ trade-size and order-size results are very similar, and the

[16]Whether a day is even or odd is determined by whether the integer representing the SAS (ver. 9.1) date is even or odd. For consistency, the results of size and price rounding shown in the top and middle lines of Panel A are based on these even- and odd-day subsamples.

trade-size results for TORQ and TAQ are also very similar. These two observations strengthen our confidence that our results using TAQ trade data can be generalized to orders.

4. Modeling the rounding decision

We now turn to determining what market, firm, and trade factors are associated with the use of a rounded trade size. Because the analysis is conducted in an intraday setting, we estimate a bivariate probit regression model of trades.

4.1. Bivariate probit model

A bivariate probit model is preferred over a univariate probit model because of the dependence of price and trade-size clustering observed in the chi-square contingency table tests as shown in Table 4. Such dependence suggests that the error terms of trade-size and price univariate probit equations are likely to be correlated, given the two equations have similar independent variables. This means that the equations' estimated coefficients are likely to be inconsistent. However, the following bivariate probit model produces consistent estimates:

$$\Pr\{\mathit{Round\ size}\} = \Phi\left(X_1'\beta_1 + \varepsilon_1\right), \tag{2}$$

$$\Pr\{\mathit{Round\ price}\} = \Phi\left(X_2'\beta_2 + \varepsilon_2\right), \tag{3}$$

$$\mathit{Cov}(\varepsilon_1, \varepsilon_2) = \rho, \tag{4}$$

where X_1 and X_2 are matrices of independent variables, β_1 and β_2 are vectors of coefficients to be estimated, and ε_1 and ε_2 are vectors of error terms.[17]

In the first equation of the bivariate probit model, we set the dependent variable equal to one if the trade size is a multiple of 500 shares and zero if it is plus or minus 100 shares of a multiple of 500 shares; all other trades are excluded from analysis. In the second equation of the bivariate probit model, we set the dependent variable equal to one if a trade was executed at a rounded price and zero otherwise. Again, a price is considered rounded if it is a multiple of \$0.25 in 1990–1991 and 1995, \$0.125 in 1998, and \$0.05 in 2002.

4.2. Independent variables

Three of the independent variables utilized in our trade-by-trade analysis are closely linked to uncertainty in a stock's value, and thus are similar to those used by Harris (1991) in his study of price rounding. The negotiations hypothesis posits that greater uncertainty leads to more rounding in order to speed up the time to execution. To test this hypothesis, we measure the demeaned and standardized number of trades in the 15 minutes prior to each trade's execution; the resulting value is denoted $DmTrd_{it}$ for a trade at time t of stock i. This is done by first dividing a given sample year into even and odd trading days with the bivariate probit analysis being performed on trades that occur on the even-numbered days. Then, for a given trade on an even-numbered day, the mean number of trades during the corresponding previous 15 minutes on odd-numbered days is determined, along with the

[17]See Greene (2003, pp. 710–719) for an introduction to bivariate probit analysis.

standard deviation. Next, we subtract this mean from the actual number of trades during the previous 15 minutes on the even-numbered day in order to demean the number. Lastly, we divide the demeaned number by the standard deviation in order to standardize it.[18] Thus, the negotiations hypothesis implies that there will be more rounding in price when $DmTrd_{it}$ is high, as this will correspond to greater-than-normal uncertainty in value.[19]

As for price and firm size, we employ two more of Harris's independent variables. Specifically, we calculate the inverse of the quote midpoint at the time of the trade ($InvPrice_{it}$) and the natural log of the market capitalization measured at the end of the previous year ($LnMktCap_{it}$). The negotiations hypothesis suggests that there should be less price rounding for firms with either lower market prices, as such firms have less uncertainty in value as measured in dollars, or greater market capitalization, as more information is available about such firms.

When it comes to trade-size rounding, a similar relation is expected for the variable $DmTrd_{it}$, as it is faster and less risky in terms of reporting errors when communicating the size of a rounded trade than a unrounded one during times of abnormally heavy trading. However, the opposite relation should exist when it comes to market price. The reason is that trade-size rounding is more consequential (in terms of aggregate purchase price) for high-priced stocks. Furthermore, increasing or decreasing the trade size by a given number of shares for a high-priced stock so as to have it rounded will lead to a greater deviation from its desired weight in the investor's portfolio. Consequently, a trade is hypothesized to be less likely to be rounded in size if the stock's price is relatively high. Lastly, because there is no clear relation between firm size and the likelihood of trade-size rounding, the variable $LnMktCap_{it}$ is not used as an independent variable in the trade-size probit.

We add six additional independent variables to the analysis. First, we include EOQ_{it}, a dummy variable that equals one if the trade occurs in the last two weeks of a calendar quarter, and zero otherwise in the trade-size probit. Moulton (2005) hypothesizes that the number of distinct trade sizes used in the two weeks before a quarter ends increases, perhaps because portfolio managers seek to more fully align their portfolios with their given objectives. Hence, less rounding is expected to occur in the last two weeks of a calendar quarter. Second, we include $Large_{it}$, a dummy variable that equals one if trade size is 10,000 shares or more, and zero otherwise, in both the trade-size and price probits. Barclay and Warner (1993), Hasbrouck (1995), and Chakravarty (2001) indicate that large trades individually (but not collectively) tend to have a greater effect on prices than medium-sized trades. Hence, we hypothesize that large trades are more likely to be rounded in both size and price than smaller trades due to a greater desire to have them executed quickly. Third, we include $GTDepth_{it}$, a dummy variable that equals one if trade size is greater than the quoted depth, and zero otherwise.[20] In deciding whether to compare the trade size to the inside depth at the quoted bid or ask, the trade must be classified as being initiated by either a seller or a buyer, respectively. This is accomplished by using the Ellis, Michaely, and O'Hara (2000) method with the exception that there is no time lag between trades and quotes (Bessembinder, 2003; Peterson and Sirri, 2003). Trades that

[18]Similarly, Werner (2003) uses demeaned measures of trading activity and price volatility.

[19]It might seem that this is inconsistent with Harris, who finds less price rounding for stocks with larger numbers of trades. However, he uses the annual number of trades while we use trade-by-trade data along with a measure of the abnormal number of trades associated with each trade.

[20]Bessembinder (1999, p. 390) notes that the depth shown on TAQ for Nasdaq does not reflect the depths of all dealers matching the inside quote. Hence, it is a noisy measure for the 1998 and 2002 Nasdaq samples.

exceed the quoted depth are more likely to involve coordination with other traders, and hence are hypothesized to be more likely to be rounded in both price and trade size. Fourth, we include $Persist_{it}$, a dummy variable that equals one if the previous trade involved a rounded size, and zero otherwise. It is included only in the trade-size probit because, as we show later, a rounded (unrounded) trade size is more likely to occur if the previous trade involved a rounded (unrounded) trade size. Fifth, we include $SP1500_{it}$, a third dummy variable that equals one if the stock belongs to the S&P 1500, and zero otherwise. It is included only in the trade-size probit equation because index funds and index arbitragers tend to place orders for specific sizes to reduce tracking error. We therefore hypothesize that such trades are less likely to be rounded in size. Finally, $QSpread_{it}$ is the quoted spread at the time of the trade. It is included only in the price probit equation because wider quoted spreads indicate more room for negotiating a price within the spread. Hence, we hypothesize that trades associated with wider quoted spreads are more likely to be rounded in price.

4.3. Results

Table 5 presents the bivariate probit model results.[21] In Panel A the results generally indicate that a trade is more likely to be rounded in size when there is heavy trading volume, the stock's price is low, the trade is large, the trade follows a rounded trade, and the trade involves a stock that is not a member of the S&P 1500. Furthermore, we find mild but consistent evidence that there is less trade-size rounding in the last two weeks of a quarter, as all of the coefficients are negative, with two of them being statistically significant at the 0.01 level and another one (Nasdaq 1998) at the 0.10 level. All of these results are as expected; the latter result is consistent with the end-of-quarter hypothesis.

The price-rounding results, displayed in Panel B, are largely consistent with the negotiations hypothesis in that a trade made during a time of heavy volume involving a stock with a relatively higher price and smaller market capitalization is more likely to have been executed at a rounded price. In addition, a trade is more likely to be rounded in price if it is large in size and is executed when the quoted spread is relatively wide.[22]

Panel C shows that the parameter *Rho*, which estimates the covariance of the error terms, is significant for both samples in 1998 and in 2002, indicating that the error terms are correlated and that the bivariate probit model is appropriate. Panel D shows that the likelihood ratio test for each year strongly rejects the null that the coefficients of the independent variables are jointly zero, indicating the models have significant explanatory power.

[21]We analyze a random sample of approximately 100,000 trades in 1990–1991, 1995, 1998, and 2002 in order to assure convergence in a reasonable amount of time. Note that TORQ orders for 1990–1991 are not analyzed in Table 5 or any of the subsequent tables, as trade prices are not provided for those orders that did not execute.

[22]We also measure volatility in the stock's price by dividing the high by the low during the 15 minutes prior to execution, and taking the natural log of the resulting number. Like $DmTrd_{it}$, we demean and standardize this number by using trades on odd-numbered days and denote it $DmHiLo_{it}$. The hypothesis that there will tend to be more rounding in size and price when $DmHiLo_{it}$ is high, as this will also be a time of greater-than-normal uncertainty in value, is supported when $DmHiLo_{it}$ is used in lieu of $DmTrd_{it}$. These two variables are not used jointly since they are highly correlated and, as noted by Jones, Kaul, and Lipson (1994, p. 645), "any information in the trading behavior of agents is almost entirely contained in the frequency of trades during a particular interval."

Table 5
Pooled bivariate probit estimates of choice to round in size and in price

Coefficient estimates	NYSE				Nasdaq	
	1990–1991	1995	1998	2002	1998	2002
A. Rounding in size						
Intercept	0.445**	0.424**	0.471**	−0.008	0.684**	0.233**
	(0.010)	(0.013)	(0.015)	(0.013)	(0.009)	(0.010)
DmTrd	0.023**	0.037**	0.031**	0.020**	−0.000	0.012**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.000)	(0.001)
InvPrice	0.055**	1.112**	0.918**	0.647**	0.075**	0.070**
	(0.017)	(0.102)	(0.123)	(0.049)	(0.020)	(0.010)
EOQ	−0.021	−0.001	−0.017	−0.030**	−0.022	−0.039**
	(0.013)	(0.012)	(0.012)	(0.011)	(0.013)	(0.012)
Large	0.629**	0.641**	0.694**	0.748**	0.690**	1.003**
	(0.019)	(0.018)	(0.021)	(0.021)	(0.044)	(0.046)
GTDepth	−0.073**	−0.053**	−0.019*	−0.002	0.051**	0.051**
	(0.013)	(0.011)	(0.010)	(0.008)	(0.011)	(0.008)
Persist	0.269**	0.247**	0.244**	0.316**	0.486**	0.385**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)
SP1500	−0.027**	−0.048**	−0.125**	−0.102**	−0.008**	−0.131**
	(0.010)	(0.011)	(0.014)	(0.012)	(0.011)	(0.009)
B. Rounding in price						
Intercept	0.104**	0.386**	0.340**	−0.352**	0.706**	1.065**
	(0.038)	(0.076)	(0.064)	(0.065)	(0.065)	(0.060)
DmTrd	0.013**	0.012**	0.016**	0.012**	0.001**	0.004**
	(0.003)	(0.002)	(0.003)	(0.004)	(0.000)	(0.001)
InvPrice	−0.234**	−0.429**	−0.624**	−0.614**	−0.606**	−0.249**
	(0.035)	(0.101)	(0.120)	(0.055)	(0.020)	(0.014)
LnMktCap	−0.009**	−0.019**	−0.016**	−0.009**	−0.030**	−0.073**
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Large	0.111**	0.158**	0.312**	0.738**	0.154**	0.315**
	(0.015)	(0.014)	(0.017)	(0.019)	(0.027)	(0.033)
GTDepth	0.104**	0.023*	0.118**	0.033**	−0.035**	0.048**
	(0.012)	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)
QSpread	0.534**	0.428**	1.395**	4.256**	2.811**	0.233**
	(0.051)	(0.060)	(0.055)	(0.101)	(0.037)	(0.039)
C. *Rho*	0.005	0.002	0.019**	0.062**	0.039**	0.044**
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
D. *Likelihood ratio*	2,898**	4,380**	4,158**	7,188**	12,412**	4,206**

This table reports parameter estimates (standard errors) of the simultaneous probability of rounding in size and price. The unrounded trades are sizes plus or minus 100 from the rounded sizes, which are at multiples of 500 shares; all other trades are excluded. For prices, the model has rounded prices at even eighths (1990–1991; 1995), even sixteenths (1998), and even twentieths (2002); all other prices are considered unrounded. The independent variables include: *DmTrd* (demeaned standardized number of trades in previous 15 minutes); *InvPrice* (inverse of the quote midpoint for the current trade); *EOQ* (one if trade occurs within 10 trading days of quarter-end, zero otherwise); *LnMktCap* (natural log of the market cap at the end of the previous year); *Large* (one if trade size is greater than or equal to 10,000 shares, zero otherwise); *GTDepth* (one if trade size is greater than depth, zero otherwise); *Persist* (one if previous trade was rounded, zero otherwise); *SP1500* (one if stock is member of S&P 1500, zero otherwise); and *QSpread* (quoted spread). Opening trades on the NYSE and trades prior to 9:45 am on Nasdaq are excluded. **/* indicates significance at the 1%/5% level. Random samples of 100,000 trades are selected for all models. *Rho* is the estimated covariance of the error terms. The likelihood ratio test is described by Greene (2003, pp. 484–486).

5. Analyzing trade-execution costs

We now proceed to examine a commonly used measure of trade-execution costs, namely proportional effective half-spreads, in order to determine whether rounded trades receive comparable execution relative to unrounded trades.[23]

5.1. Research procedure

In measuring proportional effective half-spreads (hereafter "spreads") we use the following procedure. Consider analyzing spreads for rounded and unrounded trades during a given sample year by using the following standard equation for trade t of stock i:

$$Proportional\ Effective\ Half\text{-}Spread_{it} = \frac{D_{it}\ (Price_{it} - Midpoint_{it})}{Midpoint_{it}}, \tag{5}$$

where D_{it} is the direction of the trade using the Ellis, Michaely, and O'Hara (2000) method that has $D_{it} = +1$ if the trade is a buy and -1 if it is a sell, with the exception that there is no time lag between trades and quotes, and $Midpoint_{it}$ is the average of the inside bid and ask quoted prices at the time of execution.[24] This spread measure takes into account both quoted and actual prices, and thus is a more meaningful measure of a trade's execution cost than the proportional quoted half-spread.[25]

First, we calculate the spread for each trade of each stock during a given sample period. Second, for a given stock-day we calculate the average spread at each trade size up to 20,100 shares, resulting in a maximum of 201 spread estimates per stock-day. Third, we discard those average spreads for trade sizes that were neither a multiple of 500 shares nor within 100 shares of a 500-share multiple, resulting in a maximum of 120 average spreads per stock-day. Fourth, if on a given stock-day any multiple of 500 shares had no trades at either that multiple or at both of the adjacent 100-share levels, then we discard that "three-tuple" for that day. For example, if there are two trades at 7,000 shares but none at either 6,900 or 7,100 shares on a given stock-day, then the 6,900–7,000–7,100 three-tuple is discarded for that stock-day. Similarly, the three-tuple would be discarded if there were three trades at 6,900 shares and two at 7,100 shares but no trades at 7,000 shares. Fifth, we average the remaining average spreads at each multiple of 500 shares for each stock-day; similarly, the average of all the adjacent trade-size average spreads are averaged.[26] Thus, for each stock-day there is an overall average spread for rounded trade sizes and another overall average spread for unrounded trade sizes. Sixth, we average these average spreads across stocks on each day, resulting in an average rounded spread and an average unrounded spread for each day of the sample period. Finally, the difference between these two sets of daily average spreads is calculated and tested using a t-test and Wilcoxon signed-rank test. We also use a binomial sign test to determine whether the percentage of stocks with wider spreads for rounded trades is significantly different from 50%.

[23]For more on this and other measures of trade-execution costs, see Bessembinder (2003). For a comparison of NYSE and Nasdaq trade-execution costs using Dash-5 reports, see Boehmer (2005).

[24]Similar results also obtain using the Lee and Ready (1991) method for classifying trades.

[25]Jones and Lipson (2003, p. 6) note that effective spreads "are not perfect measures of trading costs for many reasons" but are "simple to measure, [and] readily available."

[26]If there are trades at only one of the two adjacent trade sizes (e.g., three trades at 6,900 shares but none at 7,100 shares), then we use the average of that adjacent size as the average of the other adjacent size.

Table 6
Comparison of proportional effective half-spreads for rounded and unrounded sizes

Sample	Rounded	Unrounded	Difference	% Rounded > % unrounded
A. NYSE				
1990–1991	35.67	35.44	0.23	52.4
1995	30.57	30.52	0.05	47.6
1998	22.54	22.51	0.04	52.0
2002	12.58	13.29	–0.71**,††	30.2§§
B. Nasdaq				
1998	58.12	57.50	0.62	54.4
2002	36.16	36.13	0.03	52.8

This table reports the mean proportional effective half-spread (= D_{it} *[Price*$_{it}$*−Midpoint*$_{it}$*]/Midpoint*$_{it}$) for the sample described in Table 1, where D_{it} is the trade side (1 or −1) using the Ellis, Michaely, and O'Hara (2000) method without a lag, *Price*$_{it}$ is the price observed at time *t*, and *Midpoint*$_{it}$ is the quote midpoint observed at time *t*. For each stock-day, the average effective half-spread is calculated at each trade size. For each rounded size (multiple of 500 shares), if there are trades at the rounded size and either or both of the rounded size plus or minus 100 shares, then the spread of the rounded size and the average spread of the unrounded size are output. The average spread of the unrounded size is the average of the spread at the rounded size minus 100 shares and the spread at the rounded size plus 100 shares. For each stock-day the average rounded and unrounded spread is computed. Next, the equally weighted average spread is calculated across stocks each day. The table reports the average across days. The sample size for 1990–1991 is 63 days and 252 days for all other years. **/*, ††/†, and §§/§ indicate significance at the 1%/5% level using a *t*-test, Wilcoxon signed-rank test, and binomial sign test, respectively. Half-spreads are reported in basis points.

5.2. Univariate analysis

In comparing the four sample periods, Table 6 shows that proportional effective half-spreads declined consistently from 1990 to 1991 through 2002 for both rounded and unrounded trades. This is as expected, given the reduction in the minimum tick size from eighths to sixteenths to pennies and the continual increase in trading volume over this time period as shown in Table 1. Furthermore, the results indicate an insignificant difference between spreads of rounded and unrounded trade sizes except in 2002 for NYSE when rounded trades have narrower spreads.

5.3. Ordered probit analysis

The proportional effective half-spreads reported in Table 6 are likely to vary subject to factors such as trading intensity. In order to see if these spreads vary after accounting for such factors we begin with the following regression model:

$$\mathit{Effective\ Half\text{-}Spread}_{it} = \alpha + \beta_1\, \mathit{LnSize}_{it} + \beta_2\, \mathit{GTDSize}_{it} + \beta_3\, \mathit{DmTrd}_{it} + \beta_4\, \mathit{Persist}_{it} + \varepsilon_{it}, \tag{6a}$$

where *Effective Half-Spread*$_{it}$ for trade *t* of stock *i* is the proportional effective half-spread given in Eq. (5) multiplied by *Midpoint*$_{it}$, which is equivalent to using the half-spreads that are not measured proportional to price.[27] The independent variables $LnSize_{it}$, $DmTrd_{it}$, and $Persist_{it}$ are described above with regard to Table 5. The interactive term $GTDSize_{it}$ is defined to be equal to $GTDepth_{it} \times (LnSize_{it} - LnDepth_{it})$, where $GTDepth_{it}$ and $LnSize_{it}$ are as described above and $LnDepth_{it}$ is the log of the quoted depth at the inside bid if the trade is initiated by a seller and at the inside ask if it is initiated by a buyer. In this context, β_1 represents the marginal effect of trade size on the effective spread when the trade size is less than or equal to the inside quoted depth, while the sum $\beta_1+\beta_2$ represents the marginal effect when the trade is greater than the inside depth. We hypothesize that effective spreads increase with trade size, but more rapidly when the trade size is greater than the quoted depth, i.e., we hypothesize that $\beta_1>0$ and $\beta_2>0$. The dark solid kinked line in Panel A of Fig. 2 illustrates the equation with regard to trade size and quoted depth. Note how the slope of the line to the left of the kink is β_1 and the slope to the right of the link is $\beta_1+\beta_2$.

Given our interest in the effect of trade-size rounding on effective spreads, Eq. (6a) is expanded in the following manner:

$$\begin{aligned} \mathit{Effective\ Half\text{-}Spread}_{it} = {} & \alpha + \alpha' \, \mathit{Round}_{it} + \beta_1 \, \mathit{LnSize}_{it} + \beta_1' \, \mathit{Round}_{it} \times \mathit{LnSize}_{it} \\ & + \beta_2 \, \mathit{GTDSize}_{it} + \beta_2' \, \mathit{Round}_{it} \times \mathit{GTDSize}_{it} + \beta_3 \, \mathit{DmTrd}_{it} \\ & + \beta_4 \, \mathit{Persist}_{it} + \varepsilon_{it}. \end{aligned} \tag{6b}$$

This equation can be thought of as consisting of two kinked lines like the ones shown in Panel B of Fig. 2 – one for unrounded trades (the solid line) and the other for rounded trades (the dashed line). The differences in the two lines appear in the estimates of α', β_1', and β_2'. We separately estimate Eq. (6b) as a fixed-effects ordered probit regression equation for each even-numbered day during the sample period, as odd-numbered days are used to estimate $DmTrd_{it}$. We then average and analyze the time series of the estimated coefficients, accounting for any time-series autocorrelation in the errors.

Note that the dependent variable *Effective Half-Spread*$_{it}$ can assume only a limited number of values. As an example, in 1990–1991 and 1995 it can be either 0, 1/16, 1/8, 3/16, and so on. Let *Effective Half-Spread*$_{it}$ equal zero if the half-spread is 0, one if it is 1/16, two if it is 1/8, three if it is 3/16, or four if it is 1/4, at which point 95% of the trades are accounted for. Thus, *Effective Half-Spread*$_{it} = 4$ if the trade has an effective half-spread of 1/4 or greater. Since the number of intercept terms α in Eq. (6) is equal to one less than the number of values that *Effective Half-Spread*$_{it}$ can assume, there are four intercept terms for 1990–1991.[28]

The results are presented in Panel A of Table 7. Note that the coefficients on the variable $Round_{it}$ are of different signs depending on the sample, being significantly positive in the early years and significantly negative in the later years. According to Greene (2003, pp. 736–740), interpreting the marginal effect of the estimated coefficients of an ordered probit model must be done carefully. After running the fixed-effects ordered probit equation on a given even-numbered day, the mean and standard deviation for each independent variable are determined for that day and used to simulate the marginal effect of the estimated coefficients for that day's ordered probit equation. However, the average trade size relative to the average quoted depth for that day affects how certain variables are treated.

[27] Hausman, Lo, and MacKinlay (1992) are the first to use ordered probit analysis to examine microstructure issues.

[28] The number of values that the dependent variable assumes in 1990–1991, 1995, 1998, and 2002 was 5, 5, 7, and 15, respectively.

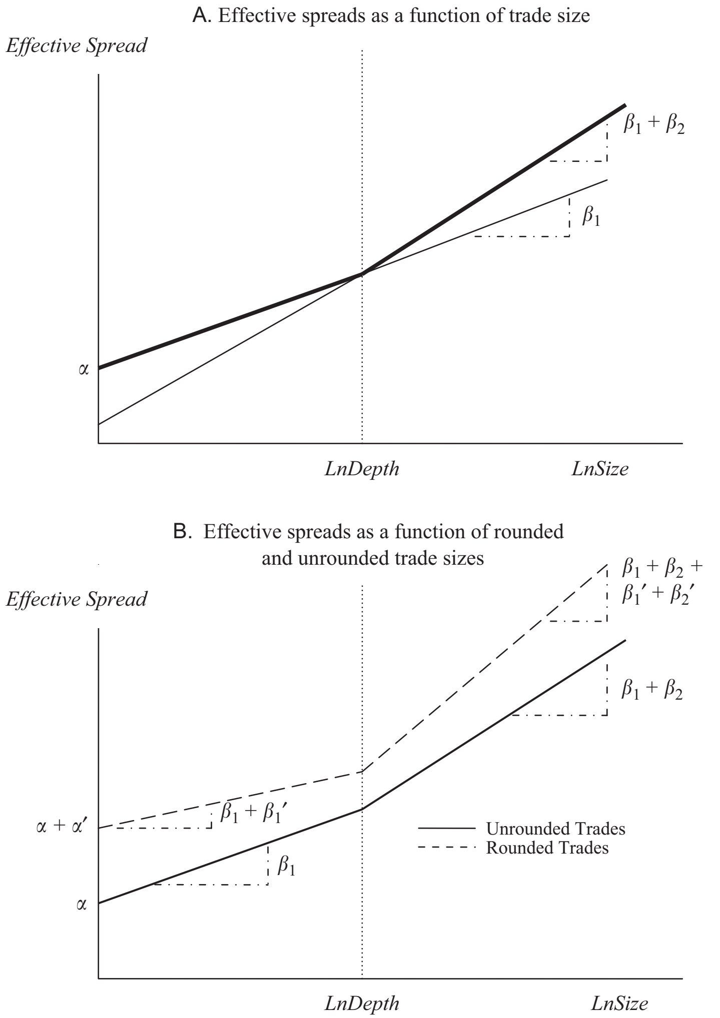

Fig. 2. The impact of trade size on effective spreads. In Panel A the dark solid kinked line represents a graphical depiction of the following regression equation for the effective half-spread associated with a trade of stock *i* at time *t*:

$$EHS_{it} = \alpha + \beta_1\, LnSize_{it} + \beta_2\, GTDSize + \varepsilon_{it},$$

where *LnSize* is the log of the trade size, and *GTDSize* equals *LnSize* minus the log of the quoted depth if positive and zero otherwise where the quoted depth is the size of the inside bid for trades initiated by sellers and the inside ask for those initiated by buyers. Panel B depicts the following regression equation where $Round_{it}$ equals one if the trade involves a round size and zero otherwise:

$$EHS_{it} = \alpha + \alpha'\, Round_{it} + \beta_1 LnSize_{it} + \beta_1'\, Round_{it} LnSize_{it} + \beta_2 GTDSize_{it} + \beta_2'\, Round'\, GTDSize_{it} + \varepsilon_{it}.$$

The equation reduces to the solid kinked line for unrounded trades and the solid line for rounded trades.

For example, consider the marginal effect of the variable of greatest importance to our analysis, $Round_{it}$. In the case of rounded trades (i.e., when $Round_{it} = 1$), the variables $GTDSize_{it}$ and $Round_{it} \times GTDSize_{it}$ are set equal to zero if the average trade size is less

Table 7
Ordered probit analysis of effective half-spreads

	NYSE				Nasdaq	
	1990–1991	1995	1998	2002	1998	2002
A. Time-series averages of selected coefficient estimates						
Round	0.130**,††	0.048*,†	0.070**,††	–0.048**,††	–0.640**,††	–0.667**,††
LnSize	–0.050**,††	0.015**,††	0.052**,††	0.020**,††	–0.076**,††	0.030**,††
Round × *LnSize*	–0.013**	0.001	–0.005*,††	0.004††	0.099**,††	0.095**,††
GTDSize	0.033*	0.226**,††	0.216**,††	0.159**,††	0.023**,††	0.041**,††
Round × *GTDSize*	0.032	–0.032**,††	–0.052**,††	–0.013**,††	0.002††	0.016**,††
DmTrd	–.007	–0.021**,††	0.027**,††	0.048**,††	0.019**,††	0.039**,††
Persist	–0.055**,††	–0.037**,††	0.030**,††	0.015**,††	0.003	–0.049**,††
B. Simulated effects of rounding on effective half-spreads (in ¢)						
Rounded	6.59	6.28	4.97	2.03	7.65	2.42
Unrounded	6.27	6.13	4.80	2.09	8.97	3.74
Difference	0.32**,††	0.14**,††	0.17**,††	–0.06**,††	–1.32**,††	–1.32**,††
C. Simulated effects of other variables on effective half-spreads (in ¢)						
DmTrd + σ	6.29	6.10	4.93	2.20	9.40	3.91
DmTrd–σ	6.34	6.27	4.70	1.97	8.89	3.56
Difference	–0.05	–0.17**,††	0.23**,††	0.23**,††	0.51**,††	0.36**,††
Persist = 1	6.21	6.12	4.87	2.11	9.15	3.62
Persist = 0	6.38	6.22	4.78	2.07	9.13	3.76
Difference	–0.17**,††	–0.10**,††	0.10**,††	0.03**,††	0.02	–0.16**,††
Rounded						
LnSize + σ	6.36	6.34	5.19	2.07	7.25	2.50
LnSize–σ	6.82	6.21	4.76	1.99	8.06	2.35
Difference	–0.45**,††	0.13**,††	0.44**,††	0.08**,††	–0.81**,††	0.15**,††
Unrounded						
LnSize + σ	6.05	6.20	5.02	2.14	8.54	3.83
LnSize–σ	6.49	6.07	4.59	2.05	9.41	3.65
Difference	–0.45**,††	0.13**,††	0.43**,††	0.09**,††	–0.87**,††	0.19**,††

This table reports results from an analysis of effective half-spreads ($= D_{it}[Price_{it} - Midpoint_{it}]$), where D_{it} is the trade side (1 or −1) using the Ellis, Michaely, and O'Hara (2000) method without a lag and $Midpoint_{it}$ is the quote midpoint observed at time t. A fixed-effects ordered probit model of the following form: $\Pr\{ES_t \leqslant k \mid z_t, X_t\} = \Phi(\Sigma\alpha_k + z'_t\omega + \beta' X_t + \varepsilon_t)$ is estimated for each even day in the sample period with $k = 5/5/7/15$ intercepts for 1991/1995/1998/2002 samples. z is an indicator variable for $n-1$ stocks, where n is the number of stocks. The independent variables include: *Round* (one if trade was for a multiple of 500 shares, zero otherwise); *LnSize* (log of trade size); *Round* × *LnSize*; *GTDSize* (*LnSize* – log of depth if trade size is greater than depth, zero otherwise); *Round* × *GTDSize* (*LnSize* – log of depth if trade size is rounded and greater than depth, zero otherwise); *DmTrd* (the demeaned standardized number of trades in the previous 15 minutes); and *Persist* (one if previous trade was rounded, zero otherwise). Opening trades on the NYSE and trades prior to 9:45 am on Nasdaq are excluded. For each regression the slopes as well as the means and standard deviations of the independent variables are retained. In Panel A the time-series averages are reported and tested after correcting for autocorrelation. Panels B and C report the simulated effects of rounding on half-spreads. On each even day in the sample period the estimated half-spread is computed for the trade type indicated, setting the variables to levels that would occur for the average trade on a given day. The time-series average differences are tested after correcting for autocorrelation. **/* and ††/† indicate significance at the 1%/5% level using a t-test and a binomial sign test, respectively.

than the average depth; the mean value is used if the average trade size is more than the average depth. All other variables are set to their means. In estimating the marginal effect of not rounding (i.e., when $Round_{it} = 0$), the same procedure is followed except that $Round_{it}$, $Round_{it} \times LnSize_{it}$, and $Round_{it} \times GTDSize_{it}$ are set equal to zero regardless of the size of the trade relative to the depth. We employ analogous treatment when the other variables are examined in Panel C. Once this analysis is complete, we separately average the daily values for $Round_{it} = 1$ and 0, which we refer to as *Simulated Effects* in Panel B of the table.

The results indicate that for the six sample periods, rounded trades realize significantly higher effective half-spreads during three of them.[29] This suggests that the negotiations hypothesis is not a complete explanation for rounding. In Panel C, we obtain inconsistent results for all of the explanatory variables, which is not surprising given the results in Panel A.

6. Analyzing the price impact of trades

The bivariate probit analysis of Table 5 suggests that rounded trade sizes tend to be used when there is abnormally heavy trading, a period during which informed trading is especially likely to be present (see Admati and Pfleiderer, 1988). We now seek to determine whether informed traders tend to make relatively more or less use of rounded sizes than noise or liquidity traders.

6.1. Estimating and comparing proportional price impacts

The price impact of trade t of stock i is defined as

$$Price\ Impact_{it} = D_{it}(Midpoint_{it+5\,\text{minutes}} - Midpoint_{it}), \tag{7}$$

where $Price\ Impact_{it}$, the movement of the quote midpoint from the time of trade t to 5 minutes later, reflects the influence of the information contained in the trade on the price of the stock and hence can be viewed as a measure of the trade's adverse selection cost. In order to standardize this measure, the price impact of a trade is divided by the quote midpoint, resulting in the proportional price impact of the trade:

$$Proportional\ Price\ Impact_{it} = \frac{D_{it}(Midpoint_{it+5\,\text{minutes}} - Midpoint_{it})}{Midpoint_{it}}. \tag{8}$$

Note that trades resulting in greater proportional price impacts are more likely, on average, to have been made by informed traders than noise or liquidity traders.

We calculate and analyze the proportional price impact for each trade in our sample in a manner identical to the procedure used to analyze effective half-spreads. Thus, for each day there is an overall average proportional price impact for rounded trade sizes and a second overall average for unrounded trade sizes; the difference between these two sets of daily average price impacts is examined using a t-test, a Wilcoxon signed-rank test, and a binomial sign test. Table 8 shows that the proportional price impact is greater for rounded than unrounded trade sizes in each sample year, and is significantly so except for 1990–1991, leaving open the possibility that informed traders are more likely than noise or liquidity traders to use rounded sizes.

[29] As a robustness check, we replace the right-hand side of Eq. (6b) with $\alpha+\alpha'\,Round_{it}+\beta_1\,LnSize_{it}+\beta_2\,GTDepth_{it}+\beta_3\,DmTrd_{it}+\beta_4\,Persist_{it}+\varepsilon_{it}$ and obtain similar results regarding rounded trades.

Table 8
Comparison of proportional price impact for rounded and unrounded sizes

Sample	Rounded	Unrounded	Difference	% Rounded > % unrounded
A. NYSE				
1990–1991	11.47	11.03	0.45	55.6
1995	14.65	13.30	1.35**,††	74.2§§
1998	18.24	15.64	2.60**,††	84.1§§
2002	10.43	9.76	0.66**,††	61.9§§
B. Nasdaq				
1998	15.14	11.64	3.50**,††	78.6§§
2002	11.89	10.13	1.76**,††	67.1§§

This table reports the mean proportional price impact ($= D_{it}[Midpoint_{it+5\text{minutes}} - Midpoint_{it}]/Midpoint_{it}$) for the sample described in Table 1, where D_{it} is the trade side (1 or −1) using the Ellis, Michaely, and O'Hara (2000) method without a lag, $Midpoint_{it}$ is the quote midpoint observed at time t, and $Midpoint_{it+5\text{ minutes}}$ is the quote midpoint observed 5 minutes after time t. For each stock-day, the average price impact is calculated at each trade size. For each rounded size (multiple of 500 shares), if there are trades at the rounded size and either or both of the rounded size plus or minus 100 shares, then the average price impacts of the rounded size and unrounded size are determined. The average price impact of the unrounded size is the average of the price impact at the rounded size minus 100 shares and the average price impact at the rounded size plus 100 shares. For each stock-day the average rounded and unrounded price impact is computed. Next, the equally weighted average price impact is calculated across stocks each day. The table reports the average across days. The sample size for 1990–1991 is 63 days and 252 days for all other years. **/*, ††/†, and §§/§ indicate significance at the 1%/5% using a t-test, Wilcoxon signed-rank test, and binomial sign test, respectively. Price impacts are reported in basis points.

The possibility remains that the price impact results above are due to the specialist moving the midpoint more after a rounded trade than after an unrounded trade. However, analysis of those trades that were not followed by another trade within the next 5 minutes provide price impact results similar to those we obtain when analyzing those trades that were followed by another trade in the next 5 minutes. Furthermore, similar results obtain with the two Nasdaq samples, in which there are dealers, not specialists. Thus, it appears that the midpoint moves by a similar amount 5 minutes after a rounded trade regardless of whether or not a trade takes place in this 5-minute time period. In sum, market-makers, be they specialists or dealers, appear to react quickly in adjusting their quotes to a rounded trade.

6.2. Ordered probit analysis

Like proportional effective half-spreads, the proportional price impacts reported in Table 8 are likely to vary subject to a number of factors. Hence, we again utilize a fixed-effects ordered probit regression equation similar to Eq. (6b) in order to see if the price impact of a trade differs based on whether it involves a rounded size. Consistent with Hasbrouck (1991), who notes that the price impact of a trade is an increasing concave function of the trade's size, we use the equation:

$$\begin{aligned} Price\ Impact_{it} = \alpha + \alpha'\, Round_{it} + \beta_1\, LnSize_{it} + \beta_1'\, Round_{it} \times LnSize_{it} \\ + \beta_2\, GTDSize_{it} + \beta_2'\, Round_{it} \times GTDSize_{it} + \beta_3\, DmTrd_{it} \\ + \beta_4\, Persist_{it} + \varepsilon_{it}, \end{aligned} \tag{9}$$

which is identical to Eq. (6b) that we use for effective spreads except that here $Price\ Impact_{it}$ for trade t of stock i, defined in Eq. (7), is used as the dependent variable.[30] Similar to the analysis of effective half-spreads, we estimate Eq. (9) for each even-numbered day in the sample period (odd-numbered days are used to estimate $DmTrd_{it}$) and then average the coefficients, accounting for any time-series autocorrelation in the errors. The results are displayed in Table 9.

Keeping in mind that care must be taken in interpreting the coefficients, the results in Panel B reveal that rounded trades have significantly greater price impacts than unrounded trades in each sample.[31] Thus, the results in Table 8 continue to hold when allowing for control variables. Furthermore, both tables show that price impacts are notably smaller since decimalization. Interestingly, the table also shows in Panel C that price impacts are likely to be greater in each sample when (1) there is below-normal trading, and (2) the previous trade was rounded. Also, larger trade sizes have greater price impacts in all NYSE samples. However, they have smaller price impacts on Nasdaq, which can be attributed to larger informed trades of Nasdaq stocks taking place relatively more frequently on electronic communication networks (ECNs), where effective and realized spreads are smaller than those provided by market makers, implying smaller price impacts (Barclay, Hendershott, and McCormick, 2003, particularly Table 9).

6.3. Persistence analysis

The observation that rounded trades have larger price impacts (or stated differently, greater information content) raises the possibility that such trades are used by "stealth" traders, as initially defined by Barclay and Warner (1993). Such traders do not want to place large orders, as doing so would be more likely to indicate that they are trading on the basis of information that is not widely available. Because trading in small sizes may involve substantial trading costs, these traders tend to use medium-sized trades, which are defined to involve 500–9,900 shares. As mentioned earlier, Chakravarty (2001) finds that such trades tend to be made by institutions. Because Keim and Madhavan (1995) and others observe that institutions break up their orders into smaller amounts that are subsequently executed over time, it seems possible that they, as stealth traders, tend to break up their orders more frequently into rounded sizes than unrounded ones. If so, two types of persistence should exist as stealth traders attempt to minimize the probability that their information becomes public before the component orders have been executed.[32]

First, there should be an increased probability that a trade will involve a rounded size if the previous trade involved a rounded size. Evidence that this happens is provided in the bivariate probit analysis of Table 5, where the estimated coefficient on the independent variable $Persist_{it}$ is found to be significantly positive each year. As a robustness test, we examine transition matrices (Bhat, 1972) in order to see if rounded (unrounded) trades tend to be followed by rounded (unrounded) trades more than expected by chance. For every year the chi-square test statistic is highly significant, indicating persistence.

[30] The dependent variable assumes 8, 8, 21, and 29 values in 1990–1991, 1995, 1998, and 2002, respectively.

[31] A robustness check in which Eq. (9) is modified as in footnote 29 produces similar results for rounded trades.

[32] See Barclay and Warner (1993). Also, "follow-on" traders, who successfully mimic the actions of informed traders, will cause persistence; see Chakravarty and Sarkar (2002).

Table 9
Ordered probit analysis of price impact

	NYSE				Nasdaq	
	1990–1991	1995	1998	2002	1998	2002
A. Time-series averages of selected coefficient estimates						
Round	0.002	0.136**,††	0.210**,††	0.099**,††	0.661**,††	0.172**,††
LnSize	0.065**,††	0.091**,††	0.055**,††	0.011**,††	−0.001	−0.013**,††
Round × *LnSize*	0.008	−0.008**,††	−0.022**,††	−0.011**,††	−0.082**,††	−0.023**,††
GTDSize	0.079**,††	0.127**,††	0.077**,††	0.035**,††	0.061**,††	0.043**,††
Round × *GTDSize*	−0.032	−0.036**,††	−0.020**,††	−0.013**,††	−0.026**,††	−0.007**,††
DmTrd	−0.014**,††	−0.006**,††	−0.006**,††	−0.007**,††	−0.013**,††	−0.008**,††
Persist	0.029**,††	0.042**,††	0.018**,††	0.010**,††	0.015**,††	0.001
B. Simulated effects of rounding on price impact (in ¢)						
Rounded	2.88	4.09	5.19	1.83	5.00	2.12
Unrounded	2.72	3.47	3.60	1.35	1.57	1.16
Difference	0.15	0.62**,††	1.59**,††	0.48**,††	3.43**,††	0.96**,††
C. Simulated effects of other variables on price impact (in ¢)						
DmTrd + σ	2.73	3.57	3.71	1.31	1.53	1.08
DmTrd − σ	2.91	3.66	3.89	1.40	2.26	1.27
Difference	−0.17**,††	−0.09**,††	−0.18**,††	−0.09**,††	−0.73**,††	−0.19**,††
Persist = 1	2.91	3.75	3.91	1.41	1.96	1.18
Persist = 0	2.77	3.51	3.73	1.35	1.82	1.17
Difference	0.14**,††	0.24**,††	0.18**,††	0.06**,††	0.14**,††	0.01
Rounded						
LnSize + σ	3.35	4.88	5.98	1.90	4.98	2.02
LnSize − σ	2.43	3.34	4.41	1.75	5.03	2.21
Difference	0.92**,††	1.54**,††	1.57**,††	0.16**,††	−0.05	−0.19**,††
Unrounded						
LnSize + σ	3.18	4.22	4.37	1.43	1.55	1.07
LnSize − σ	2.29	2.77	2.84	1.27	1.60	1.26
Difference	0.90**,††	1.46**,††	1.53**,††	0.16**,††	−0.05	−0.19**,††

This table reports results from an analysis of price impact ($= D_{it}[Midpoint_{it+5\text{minutes}} - Midpoint_{it}]$), where D_{it} is the trade side (1 or −1) using the Ellis, Michaely, and O'Hara (2000) method without a lag, $Midpoint_{it}$ is the quote midpoint observed at time t, and $Midpoint_{it+5\text{minutes}}$ is the quote midpoint observed five minutes after time t. A fixed-effects ordered probit model of the form: $\Pr\{PI_t \leqslant k \mid z_t, X_t\} = \Phi(\Sigma\alpha_k + z'_t\omega + \beta' X_t + \varepsilon_t)$ is estimated for each even day in the sample period with $k = 8/8/21/29$ intercepts for the 1991/1995/1998/2002 samples. z is an indicator variable for $n-1$ stocks, where n is the sample size. The independent variables include: *Round* (one if trade was for a multiple of 500 shares, zero otherwise); *LnSize* (log of trade size); *Round* × *LnSize*; *GTDSize* (*LnSize* – log of depth if trade size is greater than depth, zero otherwise); *Round* × *GTDSize* (*LnSize* – log of depth if trade size is rounded and greater than depth, zero otherwise); *DmTrd* (the demeaned standardized number of trades in the previous 15 minutes); and *Persist* (one if previous trade was rounded, zero otherwise). Opening trades on the NYSE and trades prior to 9:45 am on Nasdaq are excluded. For each regression the slopes as well as the means and standard deviations of the independent variables are retained. In Panel A the time-series averages are reported and tested after correcting for autocorrelation. Panels B and C report the simulated effects of rounding on price impact. On each even day in the sample period the estimated price impact is computed for the trade type indicated, setting the variables to levels that would occur for the average trade on a given day. The time-series average differences are tested after correcting for autocorrelation. **/* and ††/† indicate significance at the 1%/5% level using a t-test and binomial sign test, respectively.

The second type of persistence involves trade initiation. The Lin-Sanger-Booth (1995) spread decomposition model can be used to estimate the probability that a trade will be initiated by a buy (sell) order if the previous trade was initiated by a buy (sell) order.[33] This is accomplished by utilizing the following equation for stock i:

$$z_{it+1} = \theta_i z_{it} + \eta_{it+1}, \tag{10}$$

where z_{it} is the difference between the log of trade t's price and the log of the quoted spread midpoint when trade t is executed, θ_i is a measure of persistence, and η_{it+1} is a random error term. In this model the probability of a buy (sell) following a buy (sell) is $\delta_i = (\theta_i + 1)/2$.

In order to test for the existence of two regimes, one for rounded trade sizes and the other for unrounded trade sizes, the following version of Eq. (10) is used:

$$z_{it+1} = \theta_{i1} D_{i1t} z_{it} + \theta_{i2} D_{i2t} z_{it} + \theta_{i3} D_{i1t} D_{i2t} z_{it} + \theta_{i4} z_{it} + \eta_{it+1}, \tag{11}$$

where D_{i1t} equals one if trade t is for a multiple of 500 shares, zero otherwise, and D_{i2t} equals one if trade t+1 is for a multiple of 500 shares, zero otherwise. The sum $\theta_{i1}+\theta_{i2}+\theta_{i3}+\theta_{i4}$ is the persistence measure for rounded trade sizes while θ_{i4} is the measure for unrounded sizes. Accordingly, $\theta_{i1}+\theta_{i2}+\theta_{i3}$ is the difference in persistence between rounded and unrounded trade sizes.

Once Eq. (11) is estimated for each stock in a given sample year, we calculate cross-sectional means of the estimated coefficients. The Rounded–Rounded and Unrounded–Unrounded values are shown in Table 10. Also shown in the table are the mean estimated values of the difference $\theta_{i1}+\theta_{i2}+\theta_{i3}$, along with the proportion of stocks with a positive difference and the results from t-tests, Wilcoxon signed-rank tests, and binomial sign tests.

The results from all three tests indicate that rounded trades have significantly higher persistence than unrounded trades in each year except 1990–1991.[34] Thus, there tends to be more persistence in trade initiation for rounded than unrounded trades.[35] Interestingly, persistence for both rounded and unrounded trades seems to have become more frequent as the tick size is reduced, perhaps attributable to the reduced liquidity displayed at the inside quotes associated with smaller ticks.

6.4. Rounding and price changes

In order to explore the possibility of relations existing among informed trading, rounding, and trade size, each year every trade for each stock is classified as small (less than 500 shares), medium (500–9,900 shares), or large (greater than 9,900 shares) as in Barclay and Warner (1993) and Chakravarty (2001). For this analysis the proportional price impact for each trade between 100 and 20,000 shares is estimated using Eq. (8) and then averaged each day for each size from 100 to 20,100 shares for each stock. On a given stock-day, the averages for 100, 200, 300, and 400 are averaged to determine the average

[33] Using different models, others such as Hasbrouck (1991) observe this type of persistence.

[34] We calculate the number of small (S), medium (M), and large (L) trades for each sample year. The ratios M/L and $(S+M)/L$ are both significantly greater in 1998 and 2002, suggesting relatively more block trades were executed and less stealth trading took place under eighth pricing.

[35] We also use the Lin-Sanger-Booth model to estimate the price impact (i.e., adverse selection costs) of rounded and unrounded trades in a manner similar to Eq. (11). The results show that in each year rounded trades have a significantly higher price impact than unrounded trades, consistent with Table 9.

Table 10
Persistence analysis of trades

Coefficient estimates	NYSE				Nasdaq	
	1990–1991	1995	1998	2002	1998	2002
Rounded–Rounded						
$\theta_{i1}+\theta_{i2}+\theta_{i3}+\theta_{i4}$	0.340	0.452	0.516	0.774	0.547	0.608
Unrounded–Unrounded						
θ_{i4}	0.384	0.360	0.413	0.727	0.519	0.559
Test of difference	$-0.041^{*,\dagger}$	$0.092^{**,\dagger\dagger}$	$0.103^{**,\dagger\dagger}$	$0.047^{**,\dagger\dagger}$	$0.028^{**,\dagger\dagger}$	$0.049^{**,\dagger\dagger}$
$\theta_{i1}+\theta_{i2}+\theta_{i3}$	[0.430]	$[0.825]^{\S\S}$	$[0.849]^{\S\S}$	$[0.750]^{\S\S}$	[0.551]	$[0.640]^{\S\S}$

For each stock in the sample described in Table 1, persistence measures based on Lin, Sanger, and Booth (1995) are calculated for rounded and unrounded trade sizes using the following time-series regression equation for stock i:

$$z_{it+1} = \theta_{i1}D_{i1t}z_{it} + \theta_{i2}D_{i2t}z_{it} + \theta_{i3}D_{i1t}D_{i2t}z_{it} + \theta_{i4}z_{it} + \eta_{it+1},$$

where Q_{it} is the log of the quote midpoint at time t; P_{it} is the log of the trade price at time t; D_{i1t} equals one if trade t is for a multiple of 500 shares, zero otherwise; D_{i2t} equals one if trade t+1 is for a multiple of 500 shares, zero otherwise; and $z_{it} = P_{it} - Q_{it}$. The sum $\theta_{i1}+\theta_{i2}+\theta_{i3}$ is the estimated difference in the persistence measures for rounded versus unrounded trade sizes. Cross-sectional means of the coefficients are calculated and shown in the row labeled "rounded–rounded" and "unrounded–unrounded." In the row labeled "test of difference" the first element is the mean difference and the second is the proportion of stocks with a greater average for rounded trades than unrounded trades. $^{**}/^{*}$, $^{\dagger\dagger}/^{\dagger}$, and $^{\S\S}/^{\S}$ indicate significance at the 1%/5% level using a t-test, Wilcoxon signed-rank test, and binomial sign test, respectively. Any stock with an implied negative spread for rounded or unrounded trades is excluded. Results are similar when such stocks are included.

price impact for small trades on that day. These small trade averages are averaged across stocks on that day and then over the sample year, and are reported in Panel A of Table 11. Interestingly, Nasdaq small trades realize a smaller price impact than NYSE in 1998, but a greater one in 2002.

For medium and large trades, we consider only the average price impacts of trades involving sizes that are multiples of 500 shares or 100 shares larger or smaller than a multiple using a procedure similar to the one used to construct Panel A of Table 8. Thus, if on a given stock-day a multiple of 500 shares does not have a trade, or if neither of the adjacent 100-share sizes have a trade, then all three sizes are discarded for that day. The resulting average price impacts for multiples of 500 shares ranging from 500 to 9,500 shares are averaged for each stock and then across stocks on each day. We use the average of these daily averages as a measure of the price impact of medium-sized rounded trades. Similarly, the averages for 500-share multiples ranging from 10,000 to 20,000 shares are averaged for each stock and then across stocks for each day, with the average of the daily averages used to determine the price impact of large rounded trades. For medium-sized and large unrounded sizes, the averages of the 100-shares larger or smaller than a multiple of 500 shares are similarly averaged each day. Panel A also displays the average of these daily averages.

As expected, medium-sized trades have a greater price impact than small trades. Surprisingly, medium-sized trades have greater price impacts than large trades regardless of rounding except for NYSE in 1990–1991 and 1995. Interestingly, the price impact is greater for medium-sized rounded trades than unrounded trades in every sample period, with the difference being significant in five of the six years. In contrast, the difference for

Table 11
Stealth trading analysis

	NYSE				Nasdaq	
	1990–1991	1995	1998	2002	1998	2002
A. Average proportional price impact, assuming time-series independence						
Small	7.32	7.89	10.44	8.70	9.37	9.61
Medium						
Rounded	11.40	14.57	18.20	10.38	15.15	11.93
Unrounded	10.90	13.23	15.59	9.72	11.77	10.12
Difference	0.50	1.34**,††	2.61**,††	0.66**,††	3.38**,††	1.81**,††
Large						
Rounded	9.67	16.37	12.49	7.86	11.99	2.52
Unrounded	11.96	14.72	11.48	7.33	−4.50	6.97
Difference	−2.28	1.65	1.00	0.53	16.50**,††	−4.45**,††
B. Ordered probit analysis of price impact attributable to stealth trading						
Round	−0.247	0.161†	0.180**,††	0.144**,††	−0.049	−0.012
Round × *Medium*	0.226	−0.044	−0.013	−0.035	0.570**,††	0.131*,††
LnSize	0.064**,††	0.091**,††	0.055**,††	0.011**,††	−0.001	−0.013**,††
Round × *LnSize*	0.036	−0.006	−0.013†	−0.012*,††	−0.022	−0.005
Round × *Medium* × *LnSize*	−0.025	−0.001	−0.002	−0.001	−0.040**,††	−0.010†
GTDSize	0.080**,††	0.127**,††	0.077**,††	0.035**,††	0.061**,††	0.043**,††
Round × *GTDSize*	−0.098**,††	−0.096**,††	−0.063**,††	−0.023**,††	−0.024**,††	−0.023**,††
Round × *Medium* × *GTDSize*	0.138**,††	0.114**,††	0.064**,††	0.013**,††	0.007	0.018**,††
DmTrd	−0.014**,††	−0.006**,††	−0.006**,††	−0.007**,††	−0.013**,††	−0.008**,††
Persist	0.028**,††	0.041**,††	0.017**,††	0.010**,††	0.015**,††	0.001
C. Simulated effects of rounding on price impact (in ¢)						
Medium rounded	0.13	4.32	6.61	1.87	4.39	1.52
Medium unrounded	0.97	3.80	5.77	1.49	3.21	1.17
Medium difference	*0.16*	*0.52***,††	*0.84***,††	*0.38***,††	*1.18***,††	*0.35***,††
Large rounded	0.01	6.26	7.74	2.22	2.15	0.90
Large unrounded	0.11	5.97	7.69	2.11	4.20	1.63
Large difference	*−0.10*	*0.29**	*0.05*	*0.11*	*−2.05***,††	*−0.73***,††
Difference of differences	**0.26**	**0.23**†	**0.79****,††	**0.27****,††	**3.23****,††	**1.09****,††

D. Simulated effects of other variables on price impact (in ¢)						
DmTrd + σ	2.73	3.56	5.54	1.31	3.16	1.08
DmTrd − σ	2.90	3.65	5.68	1.40	3.69	1.27
Difference	−0.17**,††	−0.09**,††	−0.16**,††	−0.09**,††	−0.53**,††	−0.19**,††
Persist = 1	2.90	3.74	5.69	1.41	3.47	1.18
Persist = 0	2.76	3.51	5.56	1.35	3.37	1.17
Difference	0.14**,††	0.23**,††	0.18**,††	0.06**,††	0.10**,††	0.01
Medium rounded:						
LnSize + σ	3.41	4.76	6.96	1.92	4.39	1.46
LnSize − σ	2.87	3.89	6.26	1.83	4.40	1.57
Difference	0.54**,††	0.87**,††	0.70**,††	0.09**,††	−0.01	−0.11**,††
Large rounded						
LnSize + σ	4.23	6.66	8.01	2.26	2.14	0.84
LnSize − σ	3.80	5.86	7.47	2.19	2.16	0.95
Difference	0.43**,††	0.80**,††	0.54**,††	0.07**,††	−0.02	−0.11**,††
Unrounded						
LnSize + σ	3.17	4.21	6.04	1.43	3.17	1.07
LnSize − σ	2.28	2.76	4.88	1.27	3.20	1.25
Difference	0.90**,††	1.46**,††	1.16**,††	0.16**,††	−0.03	−0.18**,††

Panel A reports the mean proportional price impact ($= D_{it}[Midpoint_{it+5\,\text{minutes}} - Midpoint_{it}]/Midpoint_{it}$) for the sample described in Table 1, where D_{it} is the trade side (1 or −1) using the Ellis, Michaely, and O'Hara (2000) method without a lag, $Midpoint_{it}$ is the quote midpoint observed at time t, and $Midpoint_{it+5\,\text{minutes}}$ is the quote midpoint observed five minutes after time t. For each stock-day, the average price impact is calculated at each trade size. For each rounded size (multiple of 500 shares), if there are trades at the round size and either or both of the rounded size plus or minus 100 shares, then the average price impact of the rounded size and the average price impact of the unrounded size are determined. The average price impact of the unrounded size is the average of the average price impact at the rounded size minus 100 shares and the average price impact at the rounded size plus 100 shares. For each stock-day the average rounded and unrounded price impact is computed for medium-sized (500–9,900 shares) and large trades (10,000 or more shares). Similarly, for each stock-day the average price impact for small trades (100–400 shares) is computed. Next, the equally weighted average price impact is calculated across stocks each day. The table reports the average across days. In Panel B a fixed-effects ordered probit model analyzing price impact ($= D_{it}[Midpoint_{it+5\,\text{minutes}} - Midpoint_{it}]$) is estimated. This model is similar to that found in Table 9, except that the additional term *Medium* that equals one if the trade size is between 500 and 9,900 shares and zero otherwise is included. In Panels C and D simulated effects of rounding on price impacts are computed similarly as in Panels B and C in Table 9 with the exception that daily means of medium-sized and large trades are computed separately. Significance is tested as described in Table 9. **/* and ††/† indicate significance at the 1%/5% level using a t-test (accounting for time-series correlation in Panel B) and binomial sign test, respectively. Values are in basis points in Panel A.

large trades is significantly positive for only one of the six years. These results suggest the possibility that informed traders have a greater tendency to use rounded medium trade sizes than rounded large trade sizes. However, caution is in order due to the small number of observations that are available in determining the price impact of large trades in this panel.[36]

In order to examine this possibility more thoroughly, we use the following fixed-effects ordered probit regression equation similar to Eq. (9) in order to detect stealth trading:

$$\begin{aligned} Price\ Impact_{it} = {} & \alpha + \alpha'\, Round_{it} + \alpha''\, Round_{it} \times Medium_{it} + \beta_1\, LnSize_{it} \\ & + \beta_1'\, Round_{it} \times LnSize_{it} + \beta_1''\, Round_{it} \times Medium_{it} \times LnSize_{it} + \beta_2\, GTDSize_{it} \\ & + \beta_2'\, Round_{it} \times GTDSize_{it} + \beta_2''\, Round_{it} \times Medium_{it} \times GTDSize_{it} \\ & + \beta_3\, DmTrd_{it} + \beta_4\, Persist_{it} + \varepsilon_{it}, \end{aligned} \tag{12}$$

where $Medium_{it}$ equals 1 if trade t of stock i is between 500 and 9,900 shares, zero otherwise, and the other variables are as defined earlier. This equation can be thought of as being represented in Panel B of Fig. 2 except that there are now three kinked lines, which represent unrounded trades, medium-sized rounded trades, and large rounded trades. Again, since $DmTrd_{it}$ is estimated using odd-numbered days, only the even-numbered days are used in the daily estimation of Eq. (12). The resulting estimated coefficients are averaged and analyzed for significance, accounting for any time-series autocorrelation in the errors.

Note that this equation allows for the price impact of a medium-sized rounded trade to be an increasing convex function of its size that is different from an increasing convex function for a large rounded trade. Furthermore, it allows unrounded trades to have a different increasing convex function from those of either medium-sized or large rounded trades. Thus, the intercept coefficient α'' on $Round_{it} \times Medium_{it}$, and the slope coefficients β_1'' and β_2'' on $Round_{it} \times Medium_{it} \times LnSize_{it}$ and $Round_{it} \times Medium_{it} \times GTDSize_{it}$, respectively, will suggest the presence of stealth traders if α'' and β_1'' are not significantly negative and β_2'' is significantly positive. In such a situation, medium-sized rounded trades can be viewed as being represented by the dashed kinked line in Panel B of Fig. 2, while large rounded trades are represented by the solid line. Interestingly, this is what is observed for the most part in Panel B of Table 11. First, in all six samples α'' is not significantly negative, being insignificantly negative in three samples and positive in the other three. Second, β_1'' is negative in all six samples but only once is it clearly significant. Third, β_2'' is positive in all six samples, and significant in five of them.

Panel C of Table 11 displays the simulated effects of rounding on medium-sized and large trades based on a methodology similar to the one used in Tables 7 and 9. Of note is that in every sample period rounded medium-sized trades have a greater price impact than unrounded medium-sized trades, with the difference being significant in all but one of the samples. In contrast, rounded large trades have a greater price impact for only half of the samples, with the difference being significant in only one of the samples. Most important is that the row entitled "Difference of Differences," which indicates that in every sample the price impact of rounding is greater for medium-sized trades than large ones, with the

[36]Specifically, there are fewer trades at sizes adjacent to larger multiples of 500 shares for each stock. Also, the seemingly anomalous negative price impact for large unrounded trades in 1998 on Nasdaq is consistent with Barclay, Hendershott, and McCormick (2003), who note that prices often move in the opposite direction of a large Nasdaq trade (p. 2639).

difference being significant in all but one sample.[37] These results are largely consistent with those shown in Panel A.

Panel D of Table 11 reports results for $DmTrd_{it}$ and $Persist_{it}$ that are similar in sign and magnitude to those shown in Table 9. Thus, price impacts of trades are likely to be greater when trading is relatively slow and when the previous trade was rounded. Regarding $LnSize_{it}$, larger trades lead to greater price impacts except for in the two Nasdaq samples.[38]

These results are consistent with informed traders using medium-sized rounded trades, and thus potentially support the stealth trading hypothesis. They also are consistent with the behavioral hypothesis in that rounding is more prevalent for large trades, yet the difference in price impact associated with rounding for such trades is generally insignificant, suggesting a similar degree of informed trading for large rounded and unrounded orders.[39]

7. Alternative explanations for trade-size clustering

Although our analysis indicates trade-size clustering is consistent with the stealth trading hypothesis, there may be other reasons for trades to cluster in size. These alternative hypotheses concern (1) trade commission schedules, (2) market mechanisms such as NYSE's Direct +, (3) trading rules based on size, such as block trading exceptions, and (4) the negotiations hypothesis. We address each of these issues next.

7.1. Trade commission schedules

If trade commission breakpoints are set at rounded sizes, this would encourage trades at such sizes due to their lower cost. First, consider the commissions paid by institutional traders. In a study on institutional trading patterns, Goldstein, Irvine, Kandel, and Wiener (2004) suggest that there is relatively little variation in the distribution of per-share commissions. They claim that the most important determinant of the per-share commission of an order is the prior-period commission paid by an institutional client to that same broker, stating that "factors potentially affecting the execution cost of an order, such as the order size or trade difficulty, are relatively unimportant in determining the per-share commission" (p. 3). Furthermore, in a survey of equity traders through TraderForum of the *Institutional Investor*, Economides and Schwartz (1995) report that commissions are viewed as an unimportant cost (see Table 9, p. 12). Therefore, the lack of evidence in support of commission breakpoints for institutional traders suggests that commission schedules of such traders are not driving our results.

Second, consider the commissions paid by retail investors. If commission schedules cause such investors to choose rounded sizes, then one would expect to see more non-informed

[37] When we replace the right-hand side of Eq. (12) with $\alpha+\alpha' \, Round_{it}+\alpha'' \, Round_{it} \times Medium_{it}+\beta_1 \, LnSize_{it}+\beta_2 \, GTDepth_{it}+\beta_3 \, DmTrd_{it}+\beta_4 \, Persist_{it}+\varepsilon_{it}$ as a robustness check, similar results regarding the relative impact of rounding for medium-sized and large trades obtain.

[38] As mentioned earlier, this observation is consistent with Barclay, Hendershott, and McCormick (2003).

[39] Note that any large order that is rounded to an even multiple of 500 shares will, if divided equally into two orders, result in two rounded medium-sized trades upon execution; large orders for odd multiples of 500 shares (e.g., 13,500) can be divided so as to result in two nearly equal rounded medium-sized trades (e.g., 6,500 and 7,000). Hence, it is possible that the observed price impact results can be attributable, at least in part, to such orders being divided by stealth traders into smaller medium-sized rounded orders that are subsequently executed.

trades at rounded sizes. However, the price impact results shown in Table 8 indicate that there are proportionately more informed trades at rounded sizes. Therefore, trade commission schedules of retail traders do not explain our results.

7.2. NYSE Direct+

The NYSE's Direct + is a facility that automatically executes orders up to 1,099 shares. Such a market trading mechanism may cause TAQ users to observe an undue number of trades at 1,000 shares. One way to determine whether Direct + has had a material effect on trade size would be to compare the NYSE results before and after Direct + was fully implemented in April 2001. Because the NYSE results of 1995 and 1998 are similar to those in 2002, the idea that Direct + explains our results can be rejected. Further, the results for Nasdaq stocks are similar in most aspects to the results at the NYSE. Given that Nasdaq does not use Direct + further supports the notion that Direct + has not influenced our results.

7.3. Block trading rules

Block trades (orders of at least 10,000 shares) are often subject to certain more favorable trading rules.[40] For example, block trades are exempt from ITS trade-through rules. This and related provisions might encourage some traders to increase their order size to exactly 10,000 shares in order to qualify as a block trade. At most, this explanation can only address large orders. Accordingly, we repeat our analysis excluding trade sizes equal to or greater than 9,900 shares. The results are unchanged, and thus block trading cannot entirely explain our results.

7.4. Efficiency and the negotiations hypothesis

One empirically supported argument for the existence of price clustering (see Table 5 and Harris, 1991) is that it occurs to simplify the negotiation process, thereby reducing reporting and price risk to both the buying and the selling parties. That is, the use of rounded prices benefits both the buyer and the seller as it reduces the chance of both misreporting and the market price moving away from the reservation price of one of the parties. Perhaps something similar happens with respect to trade size. That is, it may be more efficient for a floor broker working an order in pieces to shout or signal "5,000" shares rather than "4,900" when speed of execution is of particular importance and the possibility of reporting errors is of heightened concern. However, Panel B of Table 7 shows that the evidence is mixed in this regard. Half of the sample periods indicate that rounding is associated with wider effective spreads, an observation that appears to be inconsistent with the negotiations hypothesis since it suggests that rounding often does not result in lower execution costs. Nevertheless, the evidence that price and trade-size rounding tend to occur simultaneously and that trade-size rounding tends to occur when trading is abnormally heavy is consistent with the negotiations hypothesis applying to floor traders for some trades.

[40] See, for example, Keim and Madhavan (1996) for more on block trading.

8. Conclusion

We present evidence that trade-size clustering increasingly occurs on the NYSE and Nasdaq in three levels–multiples of 500, 1,000, and 5,000 shares, and in greater relative frequency than price clustering. The data also show that there is substantial variation in the amount of trade-size clustering over time for the typical stock and across stocks on a typical day, with some diminution in the overall amount after decimalization. Furthermore, trades involving a rounded size are also likely to involve a rounded price.

Bivariate probit model analysis reveals that traders are more likely to use rounded sizes when trading activity is relatively high, which are periods when informed trading is especially likely to be present (see Admati and Pfleiderer, 1988). In addition, trade-size rounding is less likely to occur for high-priced stocks, as it is more costly to round such orders, and for stocks that are members of the S&P 1500, as they are frequently purchased by index funds and index arbitragers in specific amounts in order to reduce tracking error. Also, there is mild evidence that less rounding occurs shortly before the end of calendar quarters, which is consistent with Moulton's (2005) end-of-quarter hypothesis. Lastly, rounding is more likely if the trade involves 10,000 or more shares, as rounding large orders involves a smaller percentage alteration in the amount invested.

The negotiations hypothesis, initially developed to explain price clustering, suggests that rounding benefits both the buyer and the seller as it reduces the chance of both misreporting and the market price moving away from the reservation price of one of the parties. Analysis of trade-execution costs indicates that trade-size rounding has no clear bearing on a trade's effective spread but is associated with a significantly greater price impact. These results continue to hold when an ordered probit model with control variables is used, and appear to be inconsistent with the negotiations hypothesis applying to all trades, as they suggest the use of round sizes often does not result in lower execution costs. However, we find other evidence that is consistent with this hypothesis. In particular, the evidence that price and trade-size rounding tend to occur simultaneously and that trade-size rounding tends to occur when trading is abnormally heavy is consistent with the negotiations hypothesis applying to floor traders for some trades.

The observation that rounded trades have greater price impact than unrounded trades suggests that such trades tend to be placed by informed traders. Furthermore, these traders appear to have a tendency to use medium-sized trades, as we also observe that rounding tends to lead to a greater relative price impact for medium-sized trades than large trades. Tests of persistence reveal two additional interesting results. First, rounded trades tend to follow rounded trades more than would be expected by chance. Second, trades initiated by buyers (sellers) have a greater tendency to follow trades initiated by buyers (sellers) if they are rounded than if they are not rounded.

The stealth trading hypothesis of Barclay and Warner (1993), according to which informed traders use rounded medium-sized orders, appears to link several of these results together. When breaking up a large order such traders (and follow-on traders who mimic their actions) are likely to persist in initiating a number of similar-sized trades in the same direction that tend to have a significantly positive large price impact. In sum, trade-size clustering is possibly attributable, at least in part, to the actions of informed traders who strategically seek to disguise their trades by placing multiple medium-sized rounded orders during times of abnormally heavy trading.

At this point there remain two unanswered questions. First, why do stealth traders tend to round their medium-sized trades? Second, why does clustering occur increasingly at multiples of 500, 1,000, and 5,000 shares? A possible answer to both questions lies in the behavioral hypothesis and the observation that large orders and trades are generally rounded. The large *orders* that stealth traders break up are often rounded. This might lead to the use of medium-sized rounded *trades*, e.g., breaking up an order involving 12,000 shares is likely to produce several medium-sized rounded trades. Since large orders are likely to involve informed institutions, given the analysis of Hasbrouck (1991) and Chakravarty (2001), the subsequent medium-sized rounded trades are more likely to be information-based. While there may be other explanations, the abnormal use of rounded sizes for large orders and trades is consistent with the behavioral hypothesis. However, it should be noted that rounding such orders and trades is relatively inconsequential, as the effect of rounding on the aggregate purchase price and the security's portfolio weight is proportionally much smaller for a large size than a smaller one. In response to the second question, the three levels of trade-size clustering that we observe are consistent with the behavioral use of focal points (Schelling, 1960).

Additional research needs to be done to more fully understand trade-size clustering. A good starting point would be to analyze recent order data (which are currently unavailable to the public). In particular, institutions may release an order in stages to brokers or floor traders, who may in turn break them up into smaller pieces. Having data that indicate what happens at each stage would be of great value in further understanding the rounding phenomenon described in this paper.

References

Admati, A.R., Pfleiderer, P., 1988. A theory of intraday patterns: volume and price variability. Review of Financial Studies 1, 3–40.

Ball, C., Torous, W., Tschoegl, A., 1985. The degree of price resolution: the case of the gold market. Journal of Futures Markets 5, 29–43.

Barclay, M.J., Hendershott, T., McCormick, D.T., 2003. Competition among trading venues: information and trading on electronic communications networks. Journal of Finance 58, 2637–2665.

Barclay, M., Warner, J., 1993. Stealth trading and volatility: which trades move prices. Journal of Financial Economics 34, 281–305.

Bertsimas, D., Lo, A.W., 1998. Optimal control of execution costs. Journal of Financial Markets 1, 1–50.

Bessembinder, H., 1997. The degree of price resolution and equity trading costs. Journal of Financial Economics 45, 9–34.

Bessembinder, H., 1999. Trade execution costs on Nasdaq and the NYSE: a post-reform comparison. Journal of Financial and Quantitative Analysis 34, 387–407.

Bessembinder, H., 2003. Issues in assessing trade execution costs. Journal of Financial Markets 6, 233–257.

Bhat, U., 1972. Elements of Applied Stochastic Processes. Wiley, New York.

Boehmer, E., 2005. Dimensions of execution quality: recent evidence for US equity markets. Journal of Financial Economics 78, 553–582.

Brennan, M., Subrahmanyam, A., 1998. The determinants of average trade size. Journal of Business 71, 1–25.

Chakravarty, S., 2001. Stealth trading: which traders' trades move stock prices? Journal of Financial Economics 61, 289–307.

Chakravarty, S., Sarkar, A., 2002. A model of broker's trading, with application to order flow internalization. Review of Financial Economics 11, 19–36.

Chen, H., Ritter, J., 2000. The seven percent solution. Journal of Finance 55, 1105–1131.

Chordia, T., Subrahmanyam, A., 2004. Order imbalance and individual stocks returns: theory and evidence. Journal of Financial Economics 72, 485–518.

Christie, W., Schultz, P., 1994. Why do NASDAQ market makers avoid odd-eighth quotes? Journal of Finance 49, 1813–1840.
Easley, D., O'Hara, M., 1987. Price, trade size, and information in securities markets. Journal of Financial Economics 19, 69–90.
Economides, N., Schwartz, R.A., 1995. Equity trading practices and market structure: assessing asset managers' demand for immediacy. Financial Markets, Institutions and Instruments 4, 1–47.
Ellis, K., Michaely, R., O'Hara, M., 2000. The accuracy of trade classification rules: evidence from Nasdaq. Journal of Financial and Quantitative Analysis 35, 529–551.
Federal Register, 1996. Order execution obligations. September 12, 48290–48332.
Goldstein, M., Irvine, P., Kandel, E., Wiener, Z., 2004. Brokerage commissions and institutional trading patterns. Unpublished working paper, Babson College.
Greene, W., 2003. Econometric Analysis. Prentice-Hall, Upper Saddle River, New Jersey.
Grossman, S., Miller, M., Cone, K., Fischel, D., Ross, D., 1997. Clustering and competition in asset markets. Journal of Law and Economics 40, 23–60.
Gwilym, O., Clare, A., Thomas, S., 1998. Extreme price clustering in the London equity index futures and options markets. Journal of Banking and Finance 22, 1193–1206.
Harris, L., 1991. Stock price clustering and discreteness. Review of Financial Studies 4, 389–415.
Hasbrouck, J., 1991. Measuring the information content of stock trades. Journal of Finance 46, 179–207.
Hasbrouck, J., 1995. One security, many markets: determining the contributions to price discovery. Journal of Finance 50, 1175–1199.
Hausman, J., Lo, A.W., MacKinlay, A.C., 1992. An ordered probit analysis of transaction stock prices. Journal of Financial Economics 31, 319–379.
Hodrick, L.S., Moulton, P.C., 2005. Liquidity. Unpublished working paper, Columbia University.
Ikenberry, D., Weston, J.P., 2003. Clustering in US stock prices after decimalization. Unpublished working paper, Rice University.
Jones, C.M., Kaul, G., Lipson, M.L., 1994. Transactions, volume, and volatility. Review of Financial Studies 7, 631–651.
Jones, C., Lipson, M., 2003. Are retail orders different? Unpublished working paper, University of Virginia.
Kahn, C., Pennachi, G., Sopranzetti, B., 1999. Bank deposit clustering: theory and empirical evidence. Journal of Finance 54, 2185–2214.
Keim, D., Madhavan, A., 1995. Anatomy of the trading process: empirical evidence on the behavior of institutional traders. Journal of Financial Economics 37, 371–399.
Keim, D., Madhavan, A., 1996. The upstairs market for large-block transactions: analysis and measurement of price effects. Review of Financial Studies 9, 1–36.
Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. Journal of Finance 46, 733–746.
Lin, J., Sanger, G., Booth, G., 1995. Trade size and components of the bid-ask spread. Review of Financial Studies 8, 1153–1183.
Mola, S., Loughran, T., 2004. Discounting and clustering in seasoned equity offering prices. Journal of Financial and Quantitative Analysis 39, 1–23.
Moulton, P., 2005. You can't always get what you want: trade-size clustering and quantity choice in liquidity. Journal of Financial Economics 78, 89–119.
Niederhoffer, V., Osborne, M.F.M., 1966. Market making and reversal on the stock exchange. Journal of the American Statistical Association 61, 897–916.
Peterson, M., Sirri, E., 2003. Evaluation of the biases in execution cost estimation using trade and quote data. Journal of Financial Markets 6, 259–280.
Schelling, T.C., 1960. The Strategy of Conflict. Harvard University Press, Cambridge, MA.
Seppi, D.J., 1990. Equilibrium block trading and asymmetric information. Journal of Finance 45, 73–94.
Sopranzetti, B.J., Datar, V., 2002. Price clustering in foreign exchange spot markets. Journal of Financial Markets 5, 411–417.
Werner, I.M., 2003. NYSE order flow, spreads, and information. Journal of Financial Markets 6, 309–335.
Wyckoff, P., 1963. The Psychology of Stock Market Timing. Prentice-Hall, Englewood Cliffs, NJ.
Yeoman, J.C., 2001. The optimal spread and offering price for underwritten securities. Journal of Financial Economics 62, 169–198.