Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Financial Economics 71 (2004) 581–612
JOURNAL OF Financial ECONOMICS
www.elsevier.com/locate/econbase

# Order preferencing and market quality on NASDAQ before and after decimalization ☆

Kee H. Chung[a,*], Chairat Chuwonganant[b], D. Timothy McCormick[c]

[a] *Department of Finance and Managerial Economics, State University of New York (SUNY) at Buffalo, Buffalo, NY 14260, USA*
[b] *Department of Accounting and Finance, Indiana University-Purdue University at Fort Wayne, Fort Wayne, IN 46805, USA*
[c] *The NASDAQ Stock Market, Inc., Rockville, MD 20850, USA*

Received 7 March 2003; accepted 28 March 2003

**Abstract**

Despite the widely held belief that order preferencing affects market quality, no hard evidence exists on the extent and determinants of order preferencing and its impact on dealer competition and execution quality. This study shows that the bid-ask spread (dealer quote aggressiveness) is positively (negatively) related to the proportion of internalized volume during both the pre- and post-decimalization periods. Although decimal pricing led to lower order preferencing on NASDAQ, the extent of order preferencing after decimalization is higher than what prior studies had predicted. The price impact of preferenced trades is smaller than that of unpreferenced trades and preferenced trades receive greater (smaller) size (price) improvements than unpreferenced trades.


☆The views expressed herein are not intended to represent the views of NASDAQ. The authors are solely responsible for the content. The paper benefited immensely from insightful comments of an anonymous referee. The authors thank Bill Schwert, Michael Barclay, Robert Battalio, Hendrik Bessembinder, Bidisha Chakrabarty, Frank Hatheway, Terry Hendershott, Kenneth Kim, Youngsoo Kim, Rob Nelson, Chul Park, Mike Piwowar, Jeff Smith, Robert Van Ness, Ingrid Werner, the NASDAQ ER group, and session participants at the 2002 NBER Market Microstructure Meeting, the WAFA Meeting, and the 2002 FMA Conference for many valuable comments and useful discussions. The authors are grateful to James Lastoskie for his assistance in collecting and understanding the data and Xin Zhao for her excellent research assistance.

*Corresponding author. Tel.: 716-645-3262; fax: 716-645-2131.
*E-mail addresses:* keechung@buffalo.edu (K.H. Chung), http://www.acsu.buffalo.edu/˜keechung/.

0304-405X/$ - see front matter 
doi:10.1016/S0304-405X(03)00174-0

*JEL classification:* G18; G19

*Keywords:* Decimalization; Internalization; Payment for order flow; Quote aggressiveness; Spreads; Execution quality

## 1. Introduction

Brokers and dealers on NASDAQ routinely direct or preference customer orders to any dealer who agrees to honor the best quoted price regardless of the price quoted by the dealer to whom the order is directed. In return for the routing of orders, dealers commonly offer either direct monetary payments or in-kind goods and services to brokers. Brokers and dealers also frequently internalize their order flow on NASDAQ. Internalization, a subset of preferencing, is the direction of order flow by a broker-dealer to an affiliated dealer or order flow executed by that broker-dealer as market maker.

Although prior studies offer both analytical predictions and experimental evidence regarding the effects of order preferencing on execution costs,[1] extant studies offer limited evidence on the extent and determinants of preferencing and its impact on market quality. In the present study, we address the following questions using a large sample of NASDAQ-traded stocks: (1) How extensive is preferencing on NASDAQ? Which stocks, which dealers, and which trades are more likely to be involved in preferencing? (2) How does decimal pricing affect preferencing? (3) How does preferencing affect spreads and dealer quote aggressiveness? (4) Does preferencing allow dealers to separate informed traders from uninformed traders? and (5) Do preferenced orders receive better price and size improvements?

Several studies analyze order preferencing for exchange-listed stocks. Hansch et al. (1999) investigate the effect of preferencing and internalization on spreads and dealer profits for a sample of London Stock Exchange (LSE) stocks. The authors find that while preferenced trades pay higher spreads, they do not generate higher dealer profits. In contrast, they find that internalized trades pay lower spreads. However, the study finds no evidence of significant relations between the spread and preferencing or internalization across stocks. Although the LSE, like NASDAQ, is largely a dealer market, the results found on the LSE may not hold on NASDAQ because dealer competition for order flow on the LSE is different from that on NASDAQ. On NASDAQ, there are more than 400 market makers competing for order flow. In contrast, there are only 21 market makers on the LSE and the majority of them compete for business primarily in the large (FTSE-100) stocks. Furthermore, it is illegal on the LSE to make

[1] See Christie and Schultz (1994), Battalio and Holden (1996), Godek (1996), Huang and Stoll (1996), Dutta and Madhavan (1997), Kandel and Marx (1997, 1999), Bloomfield and O'Hara (1998), Ackert and Church (1999), Bessembinder (1999), and Chung et al. (2001).

cash payments to purchase the order flow, whereas such payments are allowed on NASDAQ.

Chordia and Subrahmanyam (1995) suggest that brokers channel customer orders to non-NYSE market makers for order flow payments even if they are aware that NYSE orders execute at better than quoted prices. They find that trades executed on the NYSE have, on average, a higher price improvement rate than trades executed off the NYSE. Battalio et al. (1997) show that internalization of order flow on regional exchanges has little short-run effect on posted or effective bid-ask spreads at the national level. The authors attribute this result to the degree of market fragmentation being too small. Because the extent of market fragmentation and dealer competition for NASDAQ-traded stocks is greater than that for exchanged-listed stocks, preferencing may exert a stronger impact on NASDAQ.

Peterson and Sirri (2003) conclude that preferencing in exchange-listed stocks has not damaged market quality. The Securities and Exchange Commission (1997) also examined the issue of preferencing in detail in a 1997 report. The Commission concludes that while order preferencing reduces order interaction, it does not inhibit dealers from providing executions of customer orders at the best prices.

Battalio et al. (2001a,b) examine order preferencing on NASDAQ using proprietary data from Knight Securities. Battalio et al. (2001a) find that the dealer's gross market-making revenue varies substantially across brokers. The study concludes that payment for order flow can survive decimalization if dealers customize payment schedules by paying some brokers more than others based on the information content of order flow. Battalio et al. (2001b) examine the division of market-making revenue among dealer, broker, and trader. They find that the net trading cost of the broker refusing order-flow payments does not dominate the net trading cost of all brokers selling order flow to Knight. The study concludes that payment for order flow does not unambiguously harm traders and casts doubt on the conclusions of extant studies using only trade prices to assess market quality.

Although prior studies provide valuable insight into the impact of order preferencing on dealer competition and execution quality, there are still many unanswered questions. In this paper, we provide further evidence on the extent and effect of order preferencing on NASDAQ using data before and after decimalization. Unlike previous studies that utilize data from one market maker, we use quote and trade data for a large number of NASDAQ dealers. While prior research focuses mostly on order preferencing on regional exchanges and the nature of competition between NYSE and non-NYSE market makers, our focus is order preferencing on NASDAQ and competition among NASDAQ dealers. Because the nature and extent of competition among NASDAQ dealers are different from those between NYSE and non-NYSE market makers, our study sheds further light on the effect of order preferencing on market quality.

In addition, our study provides evidence regarding the impact of decimalization on order preferencing. Prior studies offer conflicting predictions regarding the effect of decimal pricing on order preferencing. Chordia and Subrahmanyam (1995), Kandel and Marx (1999), and Harris (1999) predict that decimal pricing could greatly reduce order preferencing. In contrast, Benveniste et al. (1992), Battalio and Holden (2001), and Battalio et al. (2001a) predict that decimalization has only a marginal effect on preferencing. We shed light on this debate by comparing the extent of preferencing immediately before and after decimalization. We also compare price impact and price/size improvements between preferenced and unpreferenced trades and thereby determine whether preferenced trades have smaller information contents and receive better executions as some researchers suggested.

Our empirical results show that order preferencing is prevalent on NASDAQ during both the pre- and post-decimalization periods. Although the implementation of decimal pricing lowered the extent of order preferencing, the proportion of preferenced trades after decimalization is much higher than what prior studies (e.g., Chordia and Subrahmanyam, 1995; Kandel and Marx, 1999; Harris, 1999) have suggested. We find a significant and positive (negative) relation between spreads (dealer quote aggressiveness) and the extent of internalization during both periods. The price impact of preferenced trades is significantly smaller than the price impact of unpreferenced trades, and preferenced trades receive greater (smaller) size (price) improvements than unpreferenced trades.

The paper is organized as follows. Section 2 describes data sources and sample selection procedures and presents descriptive statistics. Section 3 discusses the extent and determinants of order preferencing and presents evidence on how decimal pricing affects order preferencing and execution costs. Section 4 analyzes the effects of internalization on spreads and dealer quote aggressiveness. Section 5 compares price impact, price improvement, and size improvement between preferenced and unpreferenced trades. Section 6 provides a brief summary and concluding remarks.

## 2. Data sources and sample characteristics

We obtain data for this study from NASTRAQ® Trade and Quote Data. We use trade, inside quote, and dealer quote data for November 2000 (a pre-decimalization period) and June 2001 (a post-decimalization period), respectively. We use proprietary data from NASDAQ to determine whether each trade is preferenced and to calculate dealer market share (see Section 3.1 for a detailed description of the data). We include a stock in our study sample if its data are available for both periods. The final study sample consists of 3,242 NASDAQ-listed stocks. The total number of market makers in our study sample is 384 and, of those, 13 are institutional brokers, five are wirehouses, five are wholesalers, and 11 are Electronic Communication Networks (ECNs). The total number of order-entry firms is 1,228 (1,158) during our pre- (post-) decimalization study period.

We omit the following to minimize data errors: (1) quotes if either the ask or the bid is less than or equal to zero; (2) quotes if either the ask size or the bid size is less than or equal to zero; (3) quotes if the bid-ask spread is greater than $5 or less than zero; (4) before-the-open and after-the-close trades and quotes; (5) trades if the price or volume is less than or equal to zero; (6) trade price, $p_t$, if $|(p_t - p_{t-1})/p_{t-1}| > 0.5$; (7) ask quote, $a_t$, if $|(a_t - a_{t-1})/a_{t-1}| > 0.5$; and (8) bid quote, $b_t$, if $|(b_t - b_{t-1})/b_{t-1}| > 0.5$.

We measure share price by the mean daily closing quote midpoints and return volatility by the standard deviation of daily returns calculated from daily closing quote midpoints. We measure number of trades by the average daily number of transactions. We measure trade size by the average dollar transaction size. We measure firm size by the market value of equity at the beginning of each study period. We employ the Herfindahl-index as a measure of dealer competition and trading concentration in each stock. We calculate the Herfindahl-index for stock $i$ using the following formula:

$$\text{H-INDEX}(i) = \sum_j \left[100V(i,j) \Big/ \sum_j V(i,j)\right]^2, \tag{1}$$

where $V(i,j)$ is stock $i$'s dollar volume executed by dealer $j$. The Herfindahl-index increases as the number of dealers decreases or as the proportion of volume by the leading dealer increases. Thus, a high Herfindahl-index is associated with high concentration of trading.

Schultz (2000) holds that the Herfindahl-index is a better measure of dealer competition than the number of dealers because the latter tends to overstate the degree of competition. Using data from May 1997 through February 1998, Schultz finds that the average number of dealers exceeds ten, but the average Herfindahl-index is greater than 2,500. Because the Herfindahl-index of 2,500 would occur if four dealers share all the volume equally, the number of dealers tends to overstate the level of competition. In addition, Ellis et al. (2002) show that there are many dormant market makers for any given stock and entering market makers fail to capture a meaningful share of trading or profits. Thus, the number of market makers alone is unlikely to capture the competitive nature of the market.

We report select attributes of our study sample of 3,242 stocks in Table 1. The average share price is \$14.52 before decimalization and \$12.38 after decimalization. The average dollar trade size and the average number of transactions are \$10,752 and 635.24, respectively, before decimalization and \$7,888 and 575.28 after decimalization. The average standard deviation of daily returns is 0.0524 before decimalization and 0.0414 after decimalization. The average market capitalization is \$1,201 millions before decimalization \$685 millions after decimalization. The larger percentage change in market capitalization between the pre- and post-decimalization periods relative to the corresponding change in share price indicates that the firm with a greater number of shares outstanding experienced a larger reduction in share price.

The average Herfindahl-index based on dollar volume is 2,122 before decimalization and 2,110 after decimalization. Schultz (2000) shows that the average Herfindahl-index for NASDAQ stocks is 2,883 just prior to the 1997 order-handling rule change and 2,756 after the rule change. Our results suggest that there has been a significant decline in the Herfindahl-index since the 1997 rule change: from 2,756 just

Table 1

Descriptive statistics of 3,242 NASDAQ stocks before and after decimalization

This table shows the attributes of our study sample of 3,242 NASDAQ stocks before and after decimalization. We use November 2000 as the pre-decimalization study period and June 2001 as the post-decimalization study period. Share price is measured by the mean quote midpoint. Number of trades is the average daily number of transactions. Trade size is the average dollar transaction size. Return volatility is measured by the standard deviation of quote midpoint returns. Firm size is measured by the market value of equity. H-INDEX is the Herfindahl-index measured by the sum of squared dealer market share based on dollar volume.

| Variable | Decimalization | Mean | Standard deviation | Percentile | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Min | 5 | 25 | 50 | 75 | 95 | Max |
| Share price ($) | Before | 14.52 | 17.50 | 0.31 | 1.26 | 3.74 | 8.89 | 18.15 | 47.72 | 210.48 |
| | After | 12.38 | 13.24 | 0.13 | 0.79 | 2.71 | 7.99 | 17.48 | 37.86 | 123.60 |
| Number of trades | Before | 635.24 | 3,506.43 | 0.19 | 2.67 | 12.19 | 48.48 | 212.29 | 1,765.95 | 68,046.05 |
| | After | 575.28 | 2,897.61 | 0.14 | 1.71 | 9.67 | 48.55 | 219.52 | 1,615.14 | 51,792.76 |
| Trade size ($) | Before | 10,752 | 37,845 | 428 | 1,277 | 3,018 | 6,583 | 13,722 | 30,527 | 2,083,589 |
| | After | 7,888 | 9,001 | 228 | 757 | 2,094 | 5,189 | 10,916 | 22,411 | 225,104 |
| Return volatility | Before | 0.0524 | 0.0331 | 0.0001 | 0.0105 | 0.0276 | 0.0471 | 0.0704 | 0.1106 | 0.4208 |
| | After | 0.0414 | 0.0301 | 0.0002 | 0.0077 | 0.0217 | 0.0359 | 0.0534 | 0.0907 | 0.3593 |
| Market value of equity ($) (in thousands) | Before | 1,200,776 | 10,555,797 | 610 | 8,341 | 32,084 | 101,044 | 387,258 | 3,114,303 | 387,360,000 |
| | After | 684,600 | 5,226,588 | 413 | 5,894 | 25,773 | 80,547 | 298,393 | 1,807,287 | 194,270,000 |
| H-INDEX | Before | 2,122 | 1,294 | 345 | 672 | 1,258 | 1,808 | 2,632 | 4,574 | 10,000 |
| | After | 2,110 | 1,409 | 302 | 587 | 1,117 | 1,718 | 2,697 | 4,952 | 10,000 |

after the rule change to 2,110 after decimalization, a 23% drop.[2] On the whole, these results indicate that the level of market concentration on NASDAQ has declined considerably during the last several years.

## 3. Order preferencing and execution costs before and after decimalization

In this section, we first describe how order preferencing and execution costs are measured and compare order preferencing between the pre- and post-decimalization periods. We then examine how order preferencing is related to stock attributes and dealer types. Finally, we compare execution costs before and after decimal pricing.

### 3.1. Measurement of order preferencing and execution costs

Proprietary data from NASDAQ contain information on the quotes and transactions of all market makers that allows us to determine whether a public

[2] The average Herfindahl-index for our sample of NASDAQ stocks may not be directly comparable to the corresponding figure in Schultz (2000) unless both study samples include roughly identical stocks.

trade is preferenced. First, we consider a trade internalized (i.e., vertically preferenced) if the reporting market maker is also a contra-party in the trade. When the reporting market maker is not a contra-party in the trade, we trace the market maker's quote at the time of transaction and consider the trade preferenced if the quote is poorer than the prevailing inside market quote. For example, if a market maker bought 500 shares at the inside market bid price of \$20 while posting the bid price of \$19.875, we consider the trade preferenced. Blume and Goldstein (1997) and Hansch et al. (1999) employ the same approach to measure the trading volume captured through payment for order flow and other non-price means.

The Securities and Exchange Commission (SEC) notes that in its broadest sense, the term preferencing refers to the direction of order flow by a broker-dealer to a specific market maker or specialist, independent of whether or not some form of affiliation or inducement for the direction of order flow exists between the broker-dealer and the market maker or specialist.[3] Hence, our measure of preferencing is in line with the SEC's definition of preferencing.

Note that stock $i$'s volume executed by dealer $j$ when $j$ is not an ECN, $V(i,j)$, consists of four components including

$$V(i,j) = \mathrm{VINT}(i,j) + \mathrm{VINS}(i,j) + \mathrm{VNINS}(i,j) + \mathrm{VE}(i,j); \tag{2}$$

where $\mathrm{VINT}(i,j)$ is stock $i$'s internalized volume forwarded to dealer $j$, $\mathrm{VINS}(i,j)$ is stock $i$'s noninternalized volume executed by dealer $j$ when the dealer is at the inside market, $\mathrm{VNINS}(i,j)$ is stock $i$'s noninternalized volume executed by dealer $j$ when the dealer is not at the inside market, and $\mathrm{VE}(i,j)$ is stock $i$'s volume on ECNs routed by dealer $j$. Specifically, $\mathrm{VINS}(i,j)$ is defined as customer buy volume when the market maker is at the inside ask plus customer sell volume when the market maker is at the inside bid plus crosses when the market maker is at either the inside bid or inside ask, net of any internalized or ECN trades. When $j = \mathrm{ECN}$, Eq. (2) reduces to $V(i,j) = \mathrm{VINS}(i,j) + \mathrm{VNINS}(i,j)$ because $\mathrm{VINT}(i,j) = 0$ and $\mathrm{VE}(i,j) = 0$ for $j = \mathrm{ECN}$. Although $\mathrm{VNINS}(i,j)$ is likely to be zero or very small for ECNs, we calculate the variable the same way as we do for the market makers for consistency. We measure trading volume both in dollars and number of shares. However, the results are qualitatively identical between the two measures. Hence, we report only the results based on dollar volume throughout the paper.

From Eq. (2), we obtain $\sum_j V(i,j) = \sum_j \mathrm{VINT}(i,j) + \sum_j \mathrm{VINS}(i,j) + \sum_j \mathrm{VNINS}(i,j) + \sum_j \mathrm{VE}(i,j)$; where $\sum_j$ denotes summation over $j$, $\sum_j V(i,j)$ is stock $i$'s total volume, $\sum_j \mathrm{VINT}(i,j)$ is stock $i$'s internalized volume, $\sum_j \mathrm{VINS}(i,j)$ is stock $i$'s noninternalized volume at the inside market, $\sum_j \mathrm{VNINS}(i,j)$ is stock $i$'s noninternalized volume at the non-inside market, and $\sum_j \mathrm{VE}(i,j)$ is stock $i$'s volume on ECNs routed by dealers. We then measure the extent of preferencing for stock $i$,

[3] See Securities Exchange Act Release No. 37619A (September 6, 1996), 61 FR 48290 (September 12, 1996) at n.357 (File No. S7-30-95).

PREF($i$), by the ratio of stock $i$'s internalized volume plus any noninternalized volume at the non-inside market to its total volume, i.e.,

$$\text{PREF}(i) = \sum_j [\text{VINT}(i,j) + \text{VNINS}(i,j)] \Big/ \sum_j V(i,j). \tag{3}$$

A broker can send customer orders to a market maker with preferencing arrangement when the market maker is posting the inside market quotes. Hence, Eq. (3) is likely to underestimate the actual level of preferencing. If we assume that preferenced orders arrive at the same rate regardless of whether the market maker is at the inside, we can correct the downward bias by inflating VNINS($i,j$) by the proportion of time during which the market maker is not at the inside. Let $\Psi$ denote the magnitude of dealer $j$'s preferenced volume in stock $i$. If preferenced orders arrive randomly over time and the dealer is at the inside $100\lambda\%$ of the time, the expected value of non-inside volume (VNINS) would be $\Psi(1-\lambda)$. Thus, we can estimate the "unobservable" preferenced volume ($\Psi$) by dividing the non-inside volume by $(1-\lambda)$. Accordingly, we also calculate the adjusted measure of order preferencing (PREF$^{\text{A}}(i)$) using the following equation:

$$\text{PREF}^{\text{A}}(i) = \sum_j [\text{VINT}(i,j) + \text{VNINS}(i,j) \Big/ \{1 - \text{PTINS}(i,j)\}] / \sum_j V(i,j), \tag{4}$$

where PTINS($i,j$) is the proportion of time during which dealer $j$ is at the inside for stock $i$. Because the magnitude of dealer $j$'s preferenced volume in stock $i$ cannot be greater than $V(i,j) -$ VINT($i,j$)$-$VE($i,j$), we assume VNINS($i,j$)/$\{1-$PTINS($i,j$)$\}$ = $V(i,j) -$ VINT($i,j$) $-$ VE($i,j$) whenever VNINS($i,j$)/$\{1 -$ PTINS($i,j$)$\}$ turns out to be greater than $V(i,j) -$ VINT($i,j$) $-$ VE($i,j$).

Similarly, we measure the extent of preferencing for dealer $j$, PREF($j$), by the ratio of dealer $j$'s internalized volume plus any noninternalized volume at the non-inside market to his total volume, i.e.,

$$\text{PREF}(j) = \sum_i [\text{VINT}(i,j) + \text{VNINS}(i,j)] \Big/ \sum_i V(i,j). \tag{5}$$

As in Eq. (4), we also calculate the adjusted measure of dealer preferencing using the following equation:

$$\text{PREF}^{\text{A}}(j) = \sum_i [\text{VINT}(i,j) + \text{VNINS}(i,j)/\{1 - \text{PTINS}(i,j)\}] \Big/ \sum_i V(i,j). \tag{6}$$

We calculate both the time-weighted quoted spread and the trade-weighted effective spread during each period. The quoted spread is calculated as

$$\text{Quoted spread}_{it} = (A_{it} - B_{it})/M_{it}, \tag{7}$$

where $A_{it}$ is the posted ask price for stock $i$ at time $t$, $B_{it}$ is the posted bid price for stock $i$ at time $t$, and $M_{it}$ is the mean of $A_{it}$ and $B_{it}$.

To measure the cost of trading when it occurs at prices inside the posted bid and ask quotes, we calculate the effective spread using the following formula:

$$\text{Effective spread}_{it} = 2D_{it}(P_{it} - M_{it})/M_{it}; \tag{8}$$

where $P_{it}$ is the transaction price for security $i$ at time $t$, $M_{it}$ is the midpoint of the most recently posted bid and ask quotes for security $i$, and $D_{it}$ is a binary variable which equals one for customer buy orders and negative one for customer sell orders. We estimate $D_{it}$ using the method in Lee and Ready (1991). We use quotes that are at least one-second old. For each stock, we calculate the time-weighted quoted (QSPRD($i$)) and the trade-weighted effective spread (ESPRD($i$)) using all the time-series observations during each period.

### 3.2. Effect of decimal pricing on order preferencing

Panel A of Table 2 shows the mean and standard deviation of preferenced volumes before and after decimalization, together with their percentile values. INT($i$) is the ratio of stock $i$'s internalized volume to its total volume, NINS($i$) is the ratio of stock $i$'s non-inside volume to its total volume, $\text{NINS}^{\text{A}}(i)$ is the ratio of stock $i$'s adjusted non-inside volume to its total volume, $\text{PREF}(i) = \text{INT}(i) + \text{NINS}(i)$, and $\text{PREF}^{\text{A}}(i) = \text{INT}(i) + \text{NINS}^{\text{A}}(i)$. All else being equal, the probability that a quote is at the inside is likely to be lower with the smaller tick size. This tendency works against finding a decrease in our proxy for order preferencing (i.e., $\text{PREF}^{\text{A}}(i)$) after decimalization.

There is wide variation in the percentage of internalized volumes across stocks, ranging from zero to almost 100%, with a mean value of around 25%. The adjusted non-inside volume accounts for more than 50% of the total volume during both the pre- and post-decimalization periods. Mean percentages of both the internalized and adjusted non-inside volumes after decimalization are smaller than the corresponding figures before decimalization. On average, the adjusted preferencing volume declines from 79.92% pre-decimals to 75.64% post-decimals.

Although the observed change (−4.28%) in the percentage of preferenced volume is non-trivial and statistically significant ($t = -14.71$), it is not as dramatic as some prior studies had predicted. Chordia and Subrahmanyam (1995) develop a revenue-maximization model that suggests that the tick size reduction will greatly reduce preferencing. The intuition is that finite tick sizes could cause quoted prices to be greater than specialists' and off-floor market makers' reservation prices. By paying a small amount for order flow, off-floor market makers can still earn rents in excess of their reservation prices. Kandel and Marx (1999) develop a model that considers the institutional features of the NASDAQ market and show that a decrease in tick size can result in the elimination of preferencing and a reduction in internalized trades. Harris (1999) suggests that payment for order flow will decrease with decimal pricing and could cease entirely for stocks for which the decrease in average spreads is greater than the payments currently made for orders. The moderate reduction in preferencing shown in the present study is inconsistent with the predictions of these studies.

Table 2

Order preferencing before and after decimalization

Panel A shows the percentages of internalized (INT($i$)), non-inside (NINS($i$)), adjusted non-inside (NINS$^A$($i$)), preferenced (PREF($i$)), and adjusted preferenced trades (PREF$^A$($i$)). INT($i$) is the ratio of stock $i$'s internalized volume (i.e., $\sum_j$VINT($i,j$)) to its total volume (i.e., $\sum_j V(i,j)$), NINS($i$) is the ratio of stock $i$'s non-inside volume (i.e., $\sum_j$VNINS($i,j$)) to its total volume, NINS$^A$($i$) is the ratio of stock $i$'s adjusted non-inside volume (i.e., $\sum_j$VNINS($i,j$)$/1-$PTINS($i,j$)) to its total volume, PREF($i$) = INT($i$) + NINS($i$), and PREF$^A$($i$) = INT($i$) + NINS$^A$($i$). Panel B shows the percentages of internalized (INT($j$)), non-inside (NINS($j$)), adjusted non-inside (NINS$^A$($j$)), preferenced (PREF($j$)), and adjusted preferenced trades (PREF$^A$($j$)). INT($j$) is the ratio of dealer $j$'s internalized volume (i.e., $\sum_i$VINT($i,j$)) to its total volume (i.e., $\sum_i V(i,j)$), NINS($j$) is the ratio of dealer $j$'s non-inside volume (i.e., $\sum_i$VNINS($i,j$)) to its total volume, NINS$^A$($j$) is the ratio of dealer $j$'s adjusted non-inside volume (i.e., $\sum_i$VNINS($i,j$)$/1-$PTINS($i,j$)) to its total volume, PREF($j$) = INT($j$) + NINS($j$), and PREF$^A$($j$) = INT($j$) + NINS$^A$($j$). Numbers in parentheses are $t$-statistics testing the equality of the mean between the pre- and post-decimalization periods. Panel C shows how order-entry firms route their orders to different dealers. We measure the Herfindahl-index of each order-entry firm by H-INDEX($k$) = $\sum_j[100\sum_i N(i,j,k)/\sum_i\sum_j N(i,j,k)]^2$, where $N(i,j,k)$ is the number of stock $i$'s orders routed to dealer $j$ by order-entry firm $k$. To estimate the extent of order-flow concentration in a market with no preferenced order flow, we also calculate the Herfindahl-index based on simulated order flows for each order-entry firm. For this, we first measure the proportion of stock $i$'s order flow routed to dealer $j$ by $P(i,j) = \sum_k N(i,j,k)/\sum_k\sum_j N(i,j,k)$. We then randomly assign (i.e., simulate) each order-entry firm's orders to different dealers using $P(i,j)$ as the probability that dealer $j$ will get the trade in stock $i$. Finally, we calculate the Herfindahl-index of order-entry firm $j$ using the simulated values of $N(i,j,k)$. We repeat this procedure one thousand times and calculate the mean simulated Herfindahl-index for each order-entry firm.

| Variable | Decimalization | Mean | Standard deviation | Percentile | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Min | 5 | 25 | 50 | 75 | 95 | Max |
| *Panel A. Stock preferencing* | | | | | | | | | | |
| INT($i$) | Before | 26.09 | 18.82 | 0 | 0 | 9.84 | 24.93 | 39.61 | 57.56 | 99.98 |
| | After | 23.89 | 18.82 | 0 | 0 | 6.65 | 22.70 | 38.02 | 55.13 | 99.04 |
| | After–before | −2.20** (−4.69) | | | | | | | | |
| NINS($i$) | Before | 37.01 | 14.11 | 0 | 15.71 | 26.18 | 36.73 | 47.27 | 59.38 | 100 |
| | After | 38.36 | 16.15 | 0 | 14.92 | 25.47 | 37.64 | 50.34 | 64.87 | 100 |
| | After–before | 1.35** (3.59) | | | | | | | | |
| NINS$^A$($i$) | Before | 53.83 | 21.67 | 0 | 21.13 | 35.53 | 53.34 | 71.33 | 88.77 | 100 |
| | After | 51.75 | 22.42 | 0 | 18.82 | 32.90 | 51.04 | 70.22 | 88.74 | 100 |
| | After–before | −2.08** (−3.79) | | | | | | | | |
| PREF($i$) | Before | 63.10 | 11.55 | 0 | 43.90 | 57.66 | 63.39 | 68.96 | 81.68 | 100 |
| (INT($i$) + NINS($i$)) | After | 62.25 | 11.52 | 0 | 43.59 | 56.74 | 62.24 | 67.89 | 82.28 | 100 |
| | After–before | −0.85** (−2.94) | | | | | | | | |
| PREF$^A$($i$) | Before | 79.92 | 11.01 | 0 | 63.18 | 73.70 | 80.61 | 87.96 | 94.85 | 100 |
| (INT($i$) + NINS$^A$($i$)) | After | 75.64 | 12.32 | 0 | 58.20 | 68.03 | 75.78 | 83.95 | 94.55 | 100 |
| | After–before | −4.28** (−14.71) | | | | | | | | |
| *Panel B. Dealer preferencing* | | | | | | | | | | |
| INT($j$) | Before | 34.85 | 25.94 | 0 | 0 | 8.29 | 37.80 | 56.33 | 73.62 | 100 |
| | After | 29.99 | 24.18 | 0 | 0 | 3.80 | 30.18 | 50.33 | 68.04 | 100 |
| | After–before | −4.86* (−2.29) | | | | | | | | |
| NINS($j$) | Before | 24.17 | 17.83 | 0 | 3.37 | 10.85 | 20.41 | 32.35 | 60.87 | 100 |
| | After | 25.24 | 17.65 | 0 | 2.93 | 11.61 | 20.55 | 35.18 | 60.42 | 80.59 |
| | After–before | 1.07 (0.72) | | | | | | | | |
| NINS$^A$($j$) | Before | 40.98 | 24.62 | 0 | 8.23 | 21.84 | 36.04 | 57.04 | 88.44 | 100 |
| | After | 41.88 | 23.80 | 0 | 10.07 | 23.24 | 37.64 | 60.08 | 85.10 | 100 |
| | After–before | 0.90 (0.44) | | | | | | | | |

Table 2. (*Continued*)

| Variable | Decimalization | Mean | Standard deviation | Percentile | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Min | 5 | 25 | 50 | 75 | 95 | Max |
| PREF($j$) | Before | 59.02 | 19.64 | 0 | 21.57 | 48.17 | 62.38 | 70.10 | 88.06 | 100 |
| (INT($j$) + NINS($j$)) | After | 55.24 | 18.19 | 0 | 17.74 | 45.93 | 59.28 | 66.97 | 80.21 | 100 |
| | After–before | −3.78* (−2.36) | | | | | | | | |
| PREF$^A$($j$) | Before | 75.84 | 18.34 | 0 | 41.20 | 69.69 | 80.45 | 87.25 | 98.03 | 100 |
| (INT($j$) + NINS$^A$($j$)) | After | 71.87 | 18.01 | 0 | 38.50 | 64.71 | 74.97 | 83.25 | 96.44 | 100 |
| | After–before | −3.97** (−2.57) | | | | | | | | |
| *Panel C. Order-routing pattern of order-entry firms as measured by* H-INDEX($k$) | | | | | | | | | | |
| H-INDEX($k$) before | Actual | 5,568 | 3,276 | 340 | 1,018 | 2,634 | 5,018 | 9,661 | 10,000 | 10,000 |
| decimalization | Simulated | 997 | 869 | 331 | 398 | 525 | 720 | 1,095 | 2,500 | 10,000 |
| | Actual–simulated | 4,571** (53.70) | | | | | | | | |
| H-INDEX($k$) after | Actual | 5,251[a] | 3,315 | 352 | 938 | 2,323 | 4,608 | 9,193 | 10,000 | 10,000 |
| decimalization | Simulated | 939 | 1,118 | 260 | 292 | 396 | 587 | 1,017 | 2,653 | 10,000 |
| | Actual–simulated | 4,312** (48.18) | | | | | | | | |

[a] The mean value of the actual H-INDEX($k$) after decimalization is significantly smaller ($t = -2.24$) than the mean value of the actual H-INDEX($k$) before decimalization.

** Significant at the 1% level.

* Significant at the 5% level.

However, our findings are in line with the result of several previous studies that predicted the survival of preferencing after decimal pricing. Porter and Weaver (1997) find that the tick size change has a negligible impact on internalization using a sample of stocks listed on the Toronto Stock Exchange (TSE). Battalio and Holden (2001) show that order preferencing is viable with competitive dealers quoting in infinitesimal increments if they can use ex ante identifiable attributes to differentiate order flow profitability. Benveniste et al. (1992) argue that repeatedly dealing with the same brokers allows market makers to learn when brokers exploit private information, suggesting that broker identity may be one such attribute. Building on this insight, Battalio et al. (2001a) claim that order preferencing can survive decimalization if dealers customize payment schedules by paying some brokers more than others based on information content of their orders.

Another possible explanation for our result is that the effect of decimal pricing takes more time to develop since preferencing arrangements are like legal contracts that take time to renegotiate or end. The time lapse between the date (April 9, 2001) on which decimalization was completed on NASDAQ and our post-decimalization study period (June 2001) may not be long enough, and thus the full impact of decimal pricing on preferencing may yet be materialized. In addition, order preferencing could have declined even more with the implementation of SuperSOES in August 2001.

Panel B shows the percentages of preferenced trades by dealers, where INT($j$) is the ratio of dealer $j$'s internalized volume to its total volume, NINS($j$) is the ratio of dealer $j$'s non-inside volume to its total volume, NINS$^A$($j$) is the ratio of dealer $j$'s adjusted non-inside volume to its total volume, $\text{PREF}(j) = \text{INT}(j) + \text{NINS}(j)$, and $\text{PREF}^{\text{A}}(j) = \text{INT}(j) + \text{NINS}^{\text{A}}(j)$.

The results show that there is wide variation in the percentage of internalized volumes across dealers as well, ranging from zero to 100%, with a mean value of 35% (30%) during the pre- (post-) decimalization period. The adjusted non-inside volume accounts for more than 40% of the total volume during both periods. On average, adjusted preferencing volume declines from 75.84% pre-decimals to 71.87% post-decimals. The observed change (−3.97%) in the percentage of preferenced volume is nontrivial and statistically significant ($t = -2.57$).

### 3.3. Order preferencing from the perspective of the order-entry firm's routing decisions

To shed further light on order preferencing, this section analyzes how order-entry firms route their orders to different dealers. We measure the Herfindahl-index of each order-entry firm by $\text{H-INDEX}(k) = \sum_j [100\sum_i N(i,j,k)/\sum_i \sum_j N(i,j,k)]^2$, where $N(i,j,k)$ is the number of stock $i$'s trades executed by dealer $j$ for order-entry firm $k$.

To estimate the extent of order-flow concentration in a market with no preferenced order flow, we also calculate the Herfindahl-index based on simulated order flows for each order-entry firm. For this, we first measure the proportion of stock $i$'s order flow routed to dealer $j$ by $P(i,j) = \sum_k N(i,j,k)/\sum_k \sum_j N(i,j,k)$. We then randomly assign (i.e., simulate) each order-entry firm's orders to different dealers using $P(i,j)$ as the probability that dealer $j$ will get the trade in stock $i$. Finally, we calculate the Herfindahl-index of order-entry firm $j$ using the simulated values of $N(i,j,k)$. We repeat this procedure one thousand times and calculate the mean simulated Herfindahl-index for each order-entry firm.

Panel C of Table 2 compares the actual H-INDEX($k$) with the simulated H-INDEX($k$). The results show that the mean value (5,568) of the actual H-INDEX($k$) across all order-entry firms is significantly ($t = 53.70$) greater than the mean value (997) of the simulated H-INDEX($k$) during the pre-decimalization period. We find similar results from the post-decimalization period. These results indicate that the actual order routing of order-entry firms is much more concentrated on a small number of dealers than the extent of order-flow concentration in the absence of order preferencing.

Indeed, we find that 24% (293) of the 1,228 order-entry firms in our study sample routed their orders to a single dealer and 12% (152) of them routed orders to two dealers during the pre-decimalization period. Nearly one-half (605) of these order-entry firms routed orders to fewer than or equal to four dealers. Similarly, 23% and 11% of 1,158 order-entry firms routed orders to a single and two dealers, respectively, and 45% routed their orders to fewer than or equal to four dealers during the post-decimalization period. The actual number of dealers used by an order-entry firm is likely greater than these figures because our results are based on a limited window of one month before and after decimalization. In addition, most of the order-entry firms with one market maker in our database are nonregistered market makers who do not post quotes, but are required to report any trades they execute. Because they are acting as both the order-entry firm and market maker, there is no need for them to have an alternative connection.

Finally, we note that the mean value (5,251) of the actual H-INDEX($k$) after decimalization is significantly smaller ($t = -2.24$) than the mean value (5,568) of the actual H-INDEX($k$) before decimalization. This result is in line with our earlier finding that the extent of order preferencing is smaller after decimalization.

### 3.4. Preferencing as a function of stock attributes and dealer types

Preferencing contracts typically specify the nature of orders the dealer must accept. For example, according to the National Association of Securities Dealers (1991), common qualification includes small orders (usually less than 3,000 shares) and orders on stocks with a certain minimum price. Thus, the proportion of preferenced trades is likely to be higher for stocks with smaller trade sizes and higher share prices. Because trades occurred at the non-inside market are likely routed through preferencing arrangements, we expect that the proportion of trades executed at the non-inside ($\mathrm{NINS}^{\mathrm{A}}$) is higher for stocks with smaller trade sizes and higher share prices.

Internalized trades are likely to be larger than noninternalized trades because institutional brokers have large internalized volumes and their trade sizes tend to be greater than those of wholesalers or wirehouses. Hence, the proportion of internalized trades (INT) is likely to be higher for stocks with larger trade sizes. Similarly, INT is likely to be higher for stocks with larger trading volumes because institutional brokers are more likely to trade high volume stocks than low volume stocks. We expect stocks with concentrated market shares to exhibit lower quote-based competition and thus higher levels of preferenced volume. Hence, $\mathrm{NINS}^{\mathrm{A}}$ is likely to be higher for stocks with a larger H-INDEX.

Because a large portion of order flow on NASDAQ is either internalized or preferenced, stocks with larger preferenced (i.e., non-inside) volumes have smaller internalized volumes and vice versa. Hence, the expected relations between INT and stock attributes are likely opposite those between $\mathrm{NINS}^{\mathrm{A}}$ and stock attributes. Therefore, $\mathrm{NINS}^{\mathrm{A}}$ is likely negatively related to trading volume and INS is likely negatively related to share price and H-INDEX.

We note that stocks with wide spreads are more likely routed through preferencing arrangements than stocks with narrow spreads. We do not include the spread in the $\mathrm{NINS}^{\mathrm{A}}$ equation, however, because there is a spurious negative correlation between the spread and $\mathrm{NINS}^{\mathrm{A}}$ (see Section 4 for a detailed discussion of this issue).

The extent of order preferencing is likely to depend not only on stock attributes but also on dealer types. Institutional brokers frequently act as both dealer and broker for their clients, who are primarily large institutions. Consequently, institutional brokers are likely to have large internalized volumes. Integrated national firms (i.e., wirehouses) tend to have large retail brokerage forces. Thus, an integrated firm generates substantial order flows that are executed by the market-making arm of the firm. We expect that preferencing arrangements are more likely made with wholesalers (relative to institutional brokers and wirehouses) because wholesalers tend to specialize in small retail orders. These considerations suggest that INT is likely greater for institutional brokers and wirehouses, but smaller for wholesalers. Conversely, $\text{NINS}^{\text{A}}$ is likely smaller for institutional brokers and wirehouses, but greater for wholesalers.

We employ the following regression models to examine the relation between order preferencing and stock attributes/dealer types:

$$\begin{aligned}\text{INT}(i,j) = {} & \alpha_0 + \alpha_1 \log(\text{PRICE}(i)) + \alpha_2 \log(\text{NTRADE}(i)) + \alpha_3 \log(\text{TSIZE}(i)) \\ & + \alpha_4 \text{H-INDEX}(i) + \alpha_5 \text{DUMIB}(j) + \alpha_6 \text{DUMWH}(j) \\ & + \alpha_7 \text{DUMWS}(j) + \varepsilon_1(i,j); \end{aligned} \tag{9}$$

$$\begin{aligned}\text{NINS}^{\text{A}}(i,j) = {} & \beta_0 + \beta_1 \log(\text{PRICE}(i)) + \beta_2 \log(\text{NTRADE}(i)) + \beta_3 \log(\text{TSIZE}(i)) \\ & + \beta_4 \text{H-INDEX}(i) + \beta_5 \text{DUMIB}(j) + \beta_6 \text{DUMWH}(j) \\ & + \beta_7 \text{DUMWS}(j) + \varepsilon_2(i,j); \end{aligned} \tag{10}$$

$$\begin{aligned}\text{PREF}^{\text{A}}(i,j) = {} & \gamma_0 + \gamma_1 \log(\text{PRICE}(i)) + \gamma_2 \log(\text{NTRADE}(i)) + \gamma_3 \log(\text{TSIZE}(i)) \\ & + \gamma_4 \text{H-INDEX}(i) + \gamma_5 \text{DUMIB}(j) + \gamma_6 \text{DUMWH}(j) \\ & + \gamma_7 \text{DUMWS}(j) + \varepsilon_3(i,j); \end{aligned} \tag{11}$$

where INT($i,j$) is the ratio of stock $i$'s internalized volume routed to dealer $j$ to its total volume executed by dealer $j$, $\text{NINS}^{\text{A}}(i,j)$ is the ratio of stock $i$'s adjusted non-inside volume executed by dealer $j$ to its total volume executed by dealer $j$, $\text{PREF}^{\text{A}}(i,j) = \text{INT}(i,j) + \text{NINS}^{\text{A}}(i,j)$, PRICE($i$) is the average share price of stock $i$, NTRADE($i$) is the number of trades of stock $i$, TSIZE($i$) is the average dollar trade size of stock $i$, H-INDEX($i$) is the Herfindahl-index, DUMIB($j$) equals one for institutional brokers and zero otherwise, DUMWH($j$) equals one for wirehouses and zero otherwise, and DUMWS($j$) equals one for wholesalers and zero otherwise. We classify dealers into these types according to dealer categories provided in Huang (2002).

Table 3 shows the regression results. The results show that the $\text{NINS}^{\text{A}}(i,j)$ is significantly and positively related to share price and the Herfindahl-index and negatively to trade size and number of trades during both study periods. In contrast, INT($i,j$) is significantly and positively related to trade size and number of trades and negatively to share price and the Herfindahl-index. We also find that INT($i,j$) is significantly and positively related to the dummy variables for institutional brokers and wirehouses, but negatively related to the wholesaler dummy variable. Conversely, $\text{NINS}^{\text{A}}(i,j)$ is negatively related to the dummy variables for institutional brokers and wirehouses, but positively related to the dummy variable for wholesalers.[4] These results are consistent with our expectation

[4] Notice that the regression coefficients for the INT and NINS models have opposite signs because, by construction, INT($i,j$) and $\text{NINS}^{\text{A}}(i,j)$ are negatively correlated.

Table 3
Effects of dealer types and stock attributes on order preferencing
This table reports the results of the following regression model:

$$\mathrm{INT}(i,j),\ \mathrm{NINS}^{A}(i,j),\ \text{or}\ \mathrm{PREF}^{A}(i,j) = \beta_0 + \beta_1 \log(\mathrm{PRICE}(i)) + \beta_2 \log(\mathrm{NTRADE}(i)) + \beta_3 \log(\mathrm{TSIZE}(i)) + \beta_4\ \text{H-INDEX}(i) + \beta_5 \mathrm{DUMIB}(j) + \beta_6 \mathrm{DUMWH}(j) + \beta_7 \mathrm{DUMWS}(j) + \varepsilon(i);$$

where $\mathrm{INT}(i,j)$ is the ratio of stock $i$'s internalized volume routed to dealer $j$ to its total volume executed by dealer $j$, $\mathrm{NINS}^{A}(i,j)$ is the ratio of stock $i$'s adjusted non-inside volume executed by dealer $j$ to its total volume executed by dealer $j$, $\mathrm{PREF}^{A}(i,j) = \mathrm{INT}(i,j) + \mathrm{NINS}^{A}(i,j)$, PRICE($i$) is the average quote midpoint of stock $i$, NTRADE($i$) is the average daily number of trades of stock $i$, TSIZE($i$) is the average dollar trade size of stock $i$, H-INDEX($i$) is the Herfindahl-index, DUMIB($j$) equals one for institutional brokers and zero otherwise, DUMWH($j$) equals one for wirehouses and zero otherwise, and DUMWS($j$) equals one for wholesalers and zero otherwise. We report the results for both the pre- and post-decimalization periods. Numbers in parentheses are $t$-statistics.

| | Before decimalization | | | After decimalization | | |
|---|---|---|---|---|---|---|
| | $\mathrm{INT}(i,j)$ | $\mathrm{NINS}^{A}(i,j)$ | $\mathrm{PREF}^{A}(i,j)$ | $\mathrm{INT}(i,j)$ | $\mathrm{NINS}^{A}(i,j)$ | $\mathrm{PREF}^{A}(i,j)$ |
| Intercept | −0.4009** | 1.2096** | 0.8087** | −0.4564** | 1.1693** | 0.7129** |
| | (−20.94) | (56.15) | (50.46) | (−24.08) | (53.92) | (43.16) |
| log(PRICE(i)) | −0.0293** | 0.0106** | −0.0187** | −0.0286** | 0.0111** | −0.0175** |
| | (−12.70) | (4.10) | (−9.67) | (−12.53) | (4.25) | (−8.80) |
| log(NTRADE(i)) | 0.0012** | −0.0084** | −0.0072** | 0.0013** | −0.0165** | −0.0152** |
| | (4.46) | (−9.14) | (−6.54) | (4.62) | (−17.43) | (−14.02) |
| log(TSIZE(i)) | 0.0802** | −0.0847** | −0.0045* | 0.0837** | −0.0783** | 0.0054* |
| | (29.14) | (−27.36) | (−1.97) | (29.39) | (−24.02) | (2.19) |
| H-INDEX(i)/10,000 | −0.0525** | 0.1011** | 0.0486** | −0.0357** | 0.1158** | 0.0801** |
| | (−3.34) | (5.72) | (3.70) | (−3.32) | (6.58) | (5.97) |
| DUMIB(j) | 0.1513** | −0.1394** | 0.0119** | 0.1601** | −0.1526** | 0.0075* |
| | (43.36) | (−35.52) | (4.06) | (39.70) | (−33.06) | (2.15) |
| DUMWH(j) | 0.2420** | −0.1665** | 0.0755** | 0.1958** | −0.1470** | 0.0488** |
| | (51.76) | (−31.66) | (19.28) | (39.35) | (−25.83) | (11.24) |
| DUMWS(j) | −0.1654** | 0.2557** | 0.0903** | −0.1298** | 0.2033** | 0.0735** |
| | (−70.86) | (97.37) | (46.24) | (−54.81) | (75.02) | (35.60) |
| F−*value* | 2,729.71** | 3,930.90** | 766.01** | 1,927.13** | 3,073.87** | 780.76** |
| Adjusted $R^2$ | 0.304 | 0.386 | 0.109 | 0.241 | 0.336 | 0.114 |

** Significant at the 1% level.
* Significant at the 5% level.

and support the idea that the extent of order preferencing varies with stock attributes and dealer types.

Note that while the regression model explains 38.6% (30.4%) and 33.6% (24.1%) of variation in $\mathrm{NINS}^{A}(i,j)$ ($\mathrm{INT}(i,j)$) during the pre- and post-decimalization periods, respectively, it explains only 10.9% and 11.4% of variation in $\mathrm{PREF}^{A}$(i,j) during each period. The lower explanatory power for the $\mathrm{PREF}^{A}(i,j)$ model is largely due to the fact that our common explanatory variables have opposite effects on the two components (i.e., $\mathrm{NINS}^{A}(i,j)$ and $\mathrm{INT}(i,j)$) of $\mathrm{PREF}^{A}(i,j)$.

We find a high degree of intertemporal stability in INT and NINS$^A$ across dealers: 52.2% (40.9%) of inter-dealer variation in INT (NINS$^A$) during the post-decimalization period can be explained by the corresponding variation during the pre-decimalization period. These results should not come as a surprise because both INT and NINS$^A$ are strongly correlated to dealer types. Similarly, we find that 32.4% (41.7%) of inter-stock variation in INT (NINS$^A$) during the post-decimalization period can be explained by the corresponding variation during the pre-decimalization period.

### 3.5. Execution costs and order preferencing before and after decimalization

Table 4 shows the extent of order preferencing (PREF$^A(i)$) and bid-ask spreads during the pre- and post-decimalization periods, respectively. We find that decimal pricing has a significant impact on the bid-ask spreads of NASDAQ securities. During the pre-decimalization period, the average proportional quoted and effective spreads for our whole sample are 0.0330 and 0.0311, respectively. The corresponding figures are 0.0254 and 0.0224, respectively, after decimalization. Both the quoted and effective spread changes are statistically significant ($t = -10.12$ and $-12.75$). The mean and median values of the dollar quoted (effective) spread are \$0.2272 (\$0.2121) and \$0.1729 (\$0.1685) before decimalization and \$0.1474 (\$0.1282) and \$0.0944 (\$0.0844) after decimalization and the changes in both spread measures are statistically significant at the 1% level. These results are consistent with the findings of previous studies such as Chakravarty et al. (2001), NASDAQ (2001a, b), NYSE (2001), and Bessembinder (2003a).

To examine whether the impact of decimal pricing on order preferencing and spreads differs among stocks, we group our sample of stocks into quartiles based on the average trading volume during the pre-decimalization period. We then cluster stocks within each volume quartile into four groups according to the average share price during the pre-decimalization period. In Table 4 (Panel B), we show the pre- and post-decimalization values of preferencing and spreads for the four portfolios HVHP, HVLP, LVHP, and LVLP. HVHP includes stocks in the highest volume quartile and the highest price quartile, HVLP includes stocks in the highest volume quartile and the lowest price quartile, LVHP includes stocks in the lowest volume quartile and the highest price quartile, and LVLP includes stocks in the lowest volume quartile and the lowest price quartile.

Consistent with the prediction of Harris (1997, 1999), we find that the impact of the tick size change on spreads is greatest for high volume and low price stocks (i.e., HVLP). The mean quoted and effective spreads for this group declined by 0.0098 ($t = -9.46$) and 0.0107 ($t = -10.67$), respectively, after decimalization. This result suggests that the previous tick size was apparently too large for traders in these stocks. With the relaxation of the binding tick size, liquidity providers start using finer price grids, resulting in narrower spreads. We find that the impact of decimal pricing on preferencing is greatest for high volume and low price stocks as well. During the pre-decimalization period, the adjusted preferencing volume was about

Table 4

Comparisons of spreads and order preferencing during the pre- and post-decimalization periods

This table shows the extent of preferencing and bid-ask spreads during the pre- and post-decimalization periods. For each stock, we first calculate the proportion of adjusted preferenced trades ($\mathrm{PREF}^{A}(i)$), the time-weighted quoted spread ($\mathrm{QSPRD}(i)$), and the trade-weighted effective spread ($\mathrm{ESPRD}(i)$). Both $\mathrm{QSPRD}(i)$ and $\mathrm{ESPRD}(i)$ are measured as a proportion of share price. We then calculate the cross-sectional mean, median, and standard deviation of $\mathrm{PREF}^{A}(i)$, $\mathrm{QSPRD}(i)$ and $\mathrm{ESPRD}(i)$ for our study sample of 3,242 Nasdaq stocks. Panel A shows the pre- and post-decimalization values of preferencing and spreads for the whole sample. To examine whether the impact of decimal pricing on preferencing and spreads differs among stocks, we group our sample of stocks into quartiles based on the average trading volume during the pre-decimalization period. We then cluster stocks within each volume quartile into four groups according to the average share price during the pre-decimalization period. Panel B shows the pre- and post-decimalization values of preferencing and spreads for the four portfolios HVHP, HVLP, LVHP, and LVLP. HVHP includes stocks in the highest volume quartile and the highest price quartile, HVLP includes stocks in the highest volume quartile and the lowest price quartile, LVHP includes stocks in the lowest volume quartile and the highest price quartile, and LVLP includes stocks in the lowest volume quartile and the lowest price quartile.

| | | Before decimalization | | | After decimalization | | | Testing the difference in the mean between the two periods | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Median | Standard deviation | Mean | Median | Standard deviation | Difference (after–before) | $t$-value |
| *Panel A. Results from the whole sample* | | | | | | | | | |
| | $\mathrm{QSPRD}(i)$ | 0.0330 | 0.0229 | 0.0316 | 0.0254 | 0.0153 | 0.0289 | −0.0076** | −10.12 |
| | $\mathrm{ESPRD}(i)$ | 0.0311 | 0.0222 | 0.0288 | 0.0224 | 0.0135 | 0.0260 | −0.0087** | −12.75 |
| | $\mathrm{PREF}^{A}(i)$ | 0.7992 | 0.8061 | 0.1101 | 0.7564 | 0.7578 | 0.1232 | −0.0428** | −14.71 |
| *Panel B. Results from volume-price portfolios* | | | | | | | | | |
| HVHP | $\mathrm{QSPRD}(i)$ | 0.0040 | 0.0037 | 0.0025 | 0.0023 | 0.0014 | 0.0124 | −0.0017** | −4.40 |
| | $\mathrm{ESPRD}(i)$ | 0.0044 | 0.0041 | 0.0024 | 0.0024 | 0.0012 | 0.0061 | −0.0020** | −6.34 |
| | $\mathrm{PREF}^{A}(i)$ | 0.7123 | 0.7098 | 0.0347 | 0.6609 | 0.6573 | 0.0492 | −0.0514** | −12.16 |
| HVLP | $\mathrm{QSPRD}(i)$ | 0.0203 | 0.0174 | 0.0110 | 0.0105 | 0.0086 | 0.0098 | −0.0098** | −9.46 |
| | $\mathrm{ESPRD}(i)$ | 0.0214 | 0.0185 | 0.0107 | 0.0107 | 0.0087 | 0.0092 | −0.0107** | −10.67 |
| | $\mathrm{PREF}^{A}(i)$ | 0.8288 | 0.8258 | 0.0459 | 0.7409 | 0.7435 | 0.0624 | −0.0879** | −16.13 |
| LVHP | $\mathrm{QSPRD}(i)$ | 0.0295 | 0.0235 | 0.0233 | 0.0244 | 0.0201 | 0.0167 | −0.0051** | −2.54 |
| | $\mathrm{ESPRD}(i)$ | 0.0254 | 0.0205 | 0.0168 | 0.0209 | 0.0180 | 0.0147 | −0.0045** | −2.87 |
| | $\mathrm{PREF}^{A}(i)$ | 0.7804 | 0.8052 | 0.1639 | 0.7837 | 0.8376 | 0.1821 | 0.0033 | 0.19 |
| LVLP | $\mathrm{QSPRD}(i)$ | 0.0977 | 0.0897 | 0.0443 | 0.0909 | 0.0817 | 0.0431 | −0.0068** | −3.18 |
| | $\mathrm{ESPRD}(i)$ | 0.0906 | 0.0830 | 0.0406 | 0.0851 | 0.0726 | 0.0432 | −0.0055** | −3.71 |
| | $\mathrm{PREF}^{A}(i)$ | 0.7919 | 0.8358 | 0.1478 | 0.7852 | 0.8139 | 0.1591 | −0.0067 | −0.44 |

** Significant at the 1% level.

83% for these stocks whereas the corresponding figure is 74% during the post-decimalization period.

## 4. Effects of internalization on spreads and dealer quote aggressiveness

In the previous section, we show that decimalization led to a decline in both order preferencing and bid-ask spreads. In the present section, we examine the cross-sectional relation between bid-ask spreads and the proportion of internalized trades.

To examine the effect of internalization on execution costs after controlling for the effects of other variables, we estimate the following regression model:

$$\begin{aligned}\text{QSPRD}(i) \text{ or } \text{ESPRD}(i) = \beta_0 &+ \beta_1(1/\text{PRICE}(i)) + \beta_2 \log(\text{NTRADE}(i)) \\ &+ \beta_3 \log(\text{TSIZE}(i)) + \beta_4 \text{VOLATILITY}(i) \\ &+ \beta_5 \log(\text{MVE}(i)) + \beta_6 \text{H-INDEX}(i) \\ &+ \beta_7 \text{INT}(i) + \varepsilon(i);\end{aligned} \tag{12}$$

where QSPRD($i$) is the time-weighted quoted spread of stock $i$ (as a proportion of share price), ESPRD($i$) is the trade-weighted effective spread of stock $i$ (as a proportion of share price), PRICE($i$) is the average share price of stock $i$, NTRADE($i$) is the number of trades of stock $i$, TSIZE($i$) is the average dollar trade size of stock $i$, VOLATILITY($i$) is the standard deviation of stock $i$'s daily returns, MVE($i$) is the market value of equity of stock $i$, H-INDEX($i$) is the Herfindahl-index for stock $i$, INT($i$) is the proportion of internalized trades for stock $i$, and $\varepsilon(i)$ is the error term. To assess whether the effects of internalization and other variables on spreads differ between the pre- and post-decimalization periods, we estimate the above regression model using data for each period.

We use only the proportion of internalized trades as our measure of preferenced volume because the spread is likely to be spuriously correlated to both NINS($i$) and $\text{NINS}^{\text{A}}(i)$. As noted earlier, NINS($i$) is likely to underestimate the actual level of preferenced trades and the degree of underestimation increases with the dealer's time at the inside market, PTINS($i,j$). In so far as the bid-ask spread decreases with PTINS($i,j$), a spurious positive correlation is likely to exist between the spread and NINS($i$).

Note also that because we measure $\text{NINS}^{\text{A}}(i)$ by $\sum_j[\text{VNINS}(i,j)/\{1 - \text{PTINS}(i,j)\}]/\sum_j V(i,j)$, there exists a built-in positive correlation between $\text{NINS}^{\text{A}}(i,j)$ and PTINS($i,j$). Hence, to the extent that the bid-ask spread decreases with PTINS($i,j$), a negative correlation is likely to exist between the spread and $\text{NINS}^{\text{A}}(i)$, regardless of the true relation between the spread and the proportion of preferenced volume. Indeed, we find that the spread is positively related to NINS($i$) and negatively to $\text{NINS}^{\text{A}}(i)$ in our sample. These results are likely to be driven largely by measurement errors associated with the actual level of preferenced volume.

Table 5 reports the regression results. The results show that both quoted and effective spreads are positively related to return volatility, and negatively to share price, number of trades, and trade size during both the pre- and post-decimalization periods. These results are all consistent with the findings of prior studies including McInish and Wood (1992), Schultz (2000), Stoll (2000), and Chung et al. (2003). The spreads are positively and significantly related to the Herfindahl-index. This indicates that spreads are wider for stocks with greater inequality in dealer market shares or fewer dealers. This result is consistent with the finding of Klock and McCormick (1999), Schultz (2000), and Ellis et al. (2002).

The quoted and effective spreads are positively and significantly related to the extent of internalization during both the pre- and post-decimalization periods

Table 5
Effect of internalization on spreads
This table reports the results of the following regression model:

$$\text{QSPRD}(i), \text{ESPRD}(i), \ \text{ESPRD}^{\text{S}}(i) = \beta_0 + \beta_1(1/\text{PRICE}(i)) + \beta_2 \log(\text{NTRADE}(i)) + \beta_3 \log(\text{TSIZE}(i)) + \beta_4\text{VOLATILITY}(i) + \beta_5 \log(\text{MVE}(i)) + \beta_6\text{H-INDEX}(i) + \beta_7\text{INT}(i) + \varepsilon(i); \ \text{ or ESPRD}^{\text{L}}(i)$$

where QSPRD($i$) is the quoted spread as a proportion of share price of stock $i$, ESPRD($i$) is the effective spread as a proportion of share price of stock $i$, ESPRD$^{\text{S}}$($i$) is the effective spread as a proportion of share price of stock $i$ using only trades that are smaller than or equal to the quoted depth at the time of transaction, ESPRD$^{\text{L}}$($i$) is the effective spread as a proportion of share price of stock $i$ using only trades that are larger than the quoted depth at the time of transaction, PRICE($i$) is the average quote midpoint of stock $i$, NTRADE($i$) is the average daily number of transactions of stock $i$, TSIZE($i$) is the average dollar trade size of stock $i$, VOLATILITY($i$) is the standard deviation of stock $i$'s quote midpoint returns, MVE($i$) is the market value of equity of stock $i$, H-INDEX($i$) is the Herfindahl-index for stock $i$, INT($i$) is the proportion of internalized trades of stock $i$, and $\varepsilon(i)$ is the error term. We report the results for both the pre- and post-decimalization periods. Numbers in parentheses are $t$-statistics.

| | Before decimalization | | | | After decimalization | | | |
|---|---|---|---|---|---|---|---|---|
| | QSPRD($i$) | ESPRD($i$) | ESPRD$^{\text{S}}$($i$) | ESPRD$^{\text{L}}$($i$) | QSPRD($i$) | ESPRD($i$) | ESPRD$^{\text{S}}$($i$) | ESPRD$^{\text{L}}$($i$) |
| Intercept | 0.0768** | 0.0807** | 0.0381** | 0.0023 | 0.0850** | 0.0777** | 0.0572** | 0.0070 |
| | (12.87) | (15.63) | (7.73) | (0.61) | (17.48) | (18.13) | (23.00) | (1.34) |
| 1/PRICE($i$) | 0.0499** | 0.0372** | 0.0348** | 0.0023** | 0.0103** | 0.0077** | 0.0043** | 0.0024** |
| | (27.03) | (23.30) | (22.82) | (4.11) | (14.09) | (11.98) | (11.44) | (4.47) |
| log(NTRADE($i$)) | −0.0069** | −0.0058** | −0.0040** | −0.0010** | −0.0056** | −0.0046** | −0.0027** | −0.0010** |
| | (−17.13) | (−16.62) | (−12.12) | (−4.76) | (−14.43) | (−13.49) | (−13.76) | (−5.00) |
| log(TSIZE($i$)) | −0.0073** | −0.0077** | −0.0050** | −0.0024** | −0.0094** | −0.0089** | −0.0062** | −0.0017** |
| | (−9.94) | (−12.20) | (−8.27) | (−5.87) | (−14.55) | (−15.71) | (−18.70) | (−4.46) |
| VOLATILITY($i$) | 0.2874** | 0.2319** | 0.1813** | 0.0192** | 0.2150** | 0.1797** | 0.1045** | 0.0033** |
| | (22.40) | (20.90) | (17.11) | (3.31) | (14.70) | (13.96) | (13.97) | (3.26) |
| log(MVE($i$)) | 0.0012** | 0.0010** | 0.0012** | 0.0009** | 0.0017** | 0.0015** | 0.0006** | 0.0017** |
| | (2.67) | (2.53) | (3.19) | (2.96) | (3.57) | (3.66) | (2.63) | (3.11) |
| H-INDEX($i$)/ 10,000 | 0.0458** | 0.0337** | 0.0340** | 0.0321** | 0.0469** | 0.0406** | 0.0351** | 0.0428** |
| | (13.19) | (11.22) | (11.86) | (7.16) | (13.46) | (13.23) | (12.46) | (5.53) |
| INT($i$) | 0.0069** | 0.0083** | 0.0095** | 0.0076** | 0.0070** | 0.0092** | 0.0097** | 0.0077** |
| | (2.98) | (3.97) | (5.77) | (3.88) | (2.94) | (4.37) | (6.91) | (3.80) |
| $F$-value | 1,257.87** | 1,152.62** | 727.83** | 264.01** | 820.13** | 745.59** | 837.30** | 251.15** |
| Adjusted $R^2$ | 0.731 | 0.713 | 0.611 | 0.353 | 0.639 | 0.617 | 0.644 | 0.340 |

** Significant at the 1% level.

when all trades are considered. Before decimal pricing, the estimated coefficients for INT($i$) in the quoted and effective spread equations are 0.0069 ($t = 2.98$) and 0.0083 ($t = 3.97$), respectively. The corresponding figures after decimalization are 0.007 ($t = 2.94$) and 0.0092 ($t = 4.37$), respectively. Hence, during the pre-decimalization period, an increase in the degree of internalization from 10% to 40% is associated with a corresponding increase of 0.00207 (i.e., $0.0069 \times (0.4 - 0.1)$) in the quoted spread and 0.00249 (i.e., $0.0083 \times (0.4 - 0.1)$) in the effective spread. After decimalization, the corresponding figures are 0.0021 (i.e., $0.007 \times (0.4 - 0.1)$) for the quoted spread and 0.00276 (i.e., $0.0092 \times (0.4 - 0.1)$) for the effective spread.[5] These results suggest that internalization on NASDAQ may have a significant detrimental effect on execution costs.

Huang and Stoll (1996) show that larger trades receive greater price improvements on NASDAQ, resulting in smaller effective spreads for larger trades. Furthermore, internalization of institutional trades may be characteristically different from internalization of noninstitutional trades. Hence, we also estimate Eq. (12) using two different measures of the effective spread. We cluster trades into two groups according to whether the actual trade size is greater or smaller than the quoted depth at the time of transaction. We then calculate the average effective spread ($\mathrm{ESPRD}^{\mathrm{S}}(i)$) using only those trades that are smaller than or equal to the quoted depth. Similarly, we calculate the average effective spread ($\mathrm{ESPRD}^{\mathrm{L}}(i)$) using only those trades that are larger than the quoted depth. The regression results show that internalization has a slightly greater effect on $\mathrm{ESPRD}^{\mathrm{S}}(i)$ than on $\mathrm{ESPRD}^{\mathrm{L}}(i)$ during both periods.[6]

To examine whether the effect of internalization on spreads varies across stocks with different volumes, we estimate the above regression model using only the most active stocks (top quartile based on trading frequency) in our study sample. We find that the estimated coefficients (*t*-statistics) for INT(*i*) in the quoted and effective spread equations are 0.0049 (3.26) and 0.0066 (4.93), respectively, before decimal pricing and 0.001 (2.73) and 0.0016 (2.88), respectively, after decimal pricing. Hence, the effect of internalized trades on spreads for high volume stocks is qualitatively similar to that of our whole sample.

The positive relation between the spread and the proportion of internalized trades shown in Table 5 could be driven by other reasons than the one suggested in this study. For example, it could reflect the fact that stocks with larger spreads have more internalized trades because brokers have an incentive to route large-spread stocks to their affiliated dealers. To examine this possibility, we estimate a structural model in which both the spread and the proportion of internalized trades are specified as endogenous variables. As exogenous variables, we include 1/PRICE(*i*), log(NTRADE(*i*)), log(TSIZE(*i*)), VOLATILITY(*i*), log(MVE(*i*)), and H-INDEX in the spread equation, and log(PRICE(*i*)), log(NTRADE(*i*)), log(TSIZE(*i*)), and H-INDEX(*i*) in the INT equation. The result shows that the estimated coefficient for INT(*i*) in the spread equation and the estimated coefficient for the spread in the INT equation are both positive and significant, indicating a positive and bi-directional relation between the spread and the proportion of internalized trades.

[5] The spread cost differences could be overwhelmed by differences in other costs of trading. For example, a 0.3% difference in effective spread for a \$5,000 trade equals a cost of \$15. Commission costs differences between full-service and discount brokers can easily exceed \$15. Likewise, opportunity cost differences between faster executions of preferenced orders and slower executions of unpreferenced orders can also exceed the spread costs found.

[6] These results could be driven by investor behavior rather than a difference in market quality. For example, investors sometimes value immediate liquidity more than price improvement in stocks with higher internalization. That is, internalization can be high in these stocks because demand for immediate liquidity is high. Investors can also choose to trade "net" to a greater degree in stocks with high internalization, which implies that the effective spread includes an implicit commission.

To further examine the effect of internalization on market quality, we analyze how dealer quote aggressiveness varies with the extent of internalization. We measure dealer $j$'s quote aggressiveness for stock $i$ by: (1) the proportion of time during which dealer $j$ is at the inside (PTINS($i,j$)); (2) the proportion of time during which dealer $j$ is at the inside alone (PTINSA($i,j$)); and (3) the ratio of dealer $j$'s spread to the average spread across all dealers for stock $i$ (RELSPRD($i,j$)). We then estimate the following regression model:

$$\mathrm{QA}(i,j) = \beta_0 + \beta_1\ \mathrm{INT}(i,j) + \text{Control variables} + \varepsilon(i,j); \tag{13}$$

where QA($i,j$) is a measure of dealer $j$'s quote aggressiveness for stock $i$ (PTINS($i,j$), PTINSA($i,j$), or RELSPRD($i,j$)), INT($i,j$) is the proportion of internalized trades for stock $i$ by dealer $j$, and $\varepsilon(i,j)$ is the error term. To control for the effect of stock attributes on quote aggressiveness, we include log(PRICE($i$)), log(NTRADE($i$)), log(TSIZE($i$)), VOLATILITY($i$), and log(MVE($i$)) as control variables.

We estimate Eq. (13) in three different ways. First, we estimate the model using the panel data of entire stock-dealer quotes. To assess the sensitivity of our results to different estimation methods, we also estimate the model for each stock using individual dealer quote data and calculate the mean $\beta_1$ coefficient across stocks and the $z$-statistic. In this case, we do not include control variables in the regression model. We obtain the $z$-statistic by adding individual $t$-statistics across stocks and then dividing the sum by the square root of the number of coefficients. We also show the panel data regression results when we include a dummy variable for each stock (fixed effects) instead of control variables. For the panel data regression, we report the estimated $\beta_1$ coefficient, its $t$-statistic, $F$-value, and adjusted $R^2$ from regression model (13). For the stock-by-stock regressions, we report the mean $\beta_1$ coefficient and its $z$-statistic.

Table 6 shows the regression results. Panel A shows the results for the pre-decimalization period and Panel B shows the results for the post-decimalization period. In each panel, the first three columns show the results ($\beta_1$ estimate with its $t$-statistic, $F$-value, and adjusted $R^2$) from the panel data regressions with control variables, the next three columns show the results from the panel data regressions with fixed effects, and the last column shows the results (mean $\beta_1$ coefficient and its $z$-statistic) from stock-by-stock regressions, respectively.

The results show that internalization has a significant effect on dealer quote aggressiveness. We find that both PTINS($i,j$) and PTINSA($i,j$) are significantly and negatively related to INT($i,j$), and RELSPRD($i,j$) is significantly and positively related to INT($i,j$). For the pre-decimalization period, estimates of $\beta_1$ from the panel data regressions with control variables are $-0.0156$ ($t = -19.46$), $-0.0040$ ($t = -17.02$), and 0.0434 ($t = 12.09$) when we measure dealer quote aggressiveness by PTINS($i,j$), PTINSA($i,j$), and RELSPRD($i,j$), respectively. We find similar results from the panel data regressions with fixed effects. The corresponding figures from stock-by-stock regressions are $-0.0166$ ($z = -16.81$), $-0.0037$ ($z = -41.25$), and 0.0583 ($z = 10.58$). The regression results (see Panel B) for the post-decimalization period are qualitatively similar to those for the pre-decimalization period.

Table 6

Effect of internalization on dealer quote aggressiveness

This table reports the results of the following regression model: $QA(i,j) = \beta_0 + \beta_1 \log(INT(i,j))$ + Control variables + $\varepsilon(i,j)$; where QA($i,j$) is a measure of dealer $j$'s quote aggressiveness for stock $i$ (PTINS($i,j$), PTINSA($i,j$), or RELSPRD($i,j$)), INT($i,j$) is the proportion of internalized trades for stock $i$ by dealer $j$, and $\varepsilon(i,j)$ is the error term. To control for the effect of stock attributes on dealer quote aggressiveness, we include log(PRICE($i$)), log(NTRADE($i$)), log(TSIZE($i$)), VOLATILITY($i$), and log(MVE($i$)) as control variables. PRICE($i$) is the average quote midpoint of stock $i$, NTRADE($i$) is the average daily number of transactions of stock $i$, TSIZE($i$) is the average dollar trade size of stock $i$, VOLATILITY($i$) is the standard deviation of stock $i$'s quote midpoint returns, and MVE($i$) is the market value of equity of stock $i$. We also show the regression results when we include a dummy variable for each stock (fixed effects) in lieu of control variables. To assess the sensitivity of our results to a different estimation method, we estimate the model for each stock using individual dealer quote data and calculate the mean $\beta_1$ coefficient across stocks and the $z$-statistic. We obtain the $z$-statistic by adding individual regression $t$-statistics across stocks and then dividing the sum by the square root of the number of regression coefficients. For the regression with control variables and fixed effects regression, we report the estimated $\beta_1$ coefficient, its $t$-statistic, $F$-value, and adjusted $R^2$. For the stock-by-stock regression, we report the mean $\beta_1$coefficient and its $z$-statistic. Numbers in parentheses are $t$-statistics or $z$-statistics.

| | Regression with control variable results | | | Fixed effects regression results | | | Stock-by-stock regression results |
|---|---|---|---|---|---|---|---|
| | $\beta_1$ estimate ($t$-statistic) | $F$-value | Adj. $R^2$ | $\beta_1$ estimate ($t$-statistic) | $F$-value | Adj. $R^2$ | Mean $\beta_1$ estimate ($z$-statistic) |
| *Panel A. Before decimalization* | | | | | | | |
| PTINS($i,j$) | −0.0156** | 609.58** | 0.152 | −0.0129** | 275.13** | 0.013 | −0.0166** |
| | (−19.46) | | | (−16.59) | | | (−16.81) |
| PTINSA($i,j$) | −0.0040** | 1,240.87** | 0.268 | −0.0041** | 494.55** | 0.024 | −0.0037** |
| | (−17.02) | | | (−22.24) | | | (−41.25) |
| RELSPRD($i,j$) | 0.0434** | 464.45** | 0.120 | 0.0466** | 182.34** | 0.009 | 0.0583** |
| | (12.09) | | | (13.50) | | | (10.58) |
| *Panel B. After decimalization* | | | | | | | |
| PTINS($i,j$) | −0.0145** | 675.67** | 0.174 | −0.0093** | 142.56** | 0.007 | −0.0086** |
| | (−17.98) | | | (−11.94) | | | (−7.45) |
| PTINSA($i,j$) | −0.0039** | 1,194.19** | 0.271 | −0.0030** | 219.25** | 0.011 | −0.0017** |
| | (−15.03) | | | (−14.81) | | | (−19.47) |
| RELSPRD($i,j$) | 0.0521** | 396.33** | 0.110 | 0.0600** | 110.61** | 0.006 | 0.0610** |
| | (8.92) | | | (10.52) | | | (6.20) |

** Significant at the 1% level.

These results indicate that if a high portion of a dealer's volume on a stock is internalized, the dealer is less likely to post the inside market quote on that stock. Similarly, the dealer is likely to quote wider spreads, relative to other dealers' spread quotes on the stock. The results are quite robust and similar between the pre- and post-decimalization periods. Our findings are consistent with the observation of Bessembinder (1999, p. 406) that "If, in contrast, a large portion of order flow is subject to preferencing agreements, then posting a better quote may not attract order flow, leaving little incentive to improve quotes." The results are also consistent with Kluger and Wyatt's (2002) recent experimental finding that the average dealer and inside spread are wider when dealers have the opportunity to internalize their order flow.

Although the above results are consistent with the notion that internalization discourages dealer quote competition, it is difficult to determine the causality of the observed relation. As suggested above, the relative spread could be wide for stocks with large internalized order flow because the dealer's need (or incentive) to be at the inside is low for such stocks. Conversely, the relation could have been driven by a reverse causality. For example, with a large relative spread, the dealer is seldom at the inside and thus receives only internalized orders. In this case, it is the large dealer spread that causes greater internalization. Hence, our results should be interpreted with some caution.

## 5. Preferencing, price impact, and execution quality

In this section, we examine whether the price impact of trades and execution quality differ between preferenced and unpreferenced trades.

### *5.1. Price impact of preferenced and unpreferenced trades*

Benveniste et al. (1992) hold that long-term relationships between brokers and dealers can mitigate the effects of asymmetric information. The authors suggest that dealers who actively identify and sanction informed traders can provide low cost services to uninformed traders more so than dealers who do not make such efforts. Easley et al. (1996) find a significant difference in the information content of orders executed in New York and Cincinnati and interpret the result as evidence that the preferencing arrangements are used to cream-skim uninformed liquidity traders. Battalio et al. (2001a) conjecture that dealers utilize broker identity to distinguish between profitable and unprofitable order flow and show that NASDAQ dealers' trading gross revenues vary substantially among routing brokers after controlling for order size.

We provide an alternative and direct test of the clientele-pricing hypothesis by comparing the information content of preferenced and unpreferenced trades. Trades that convey less private information lead to smaller post-trade price movements and thus are more profitable to execute. If brokers route only those orders with low adverse-selection risks to affiliated dealers (i.e., internalization) or dealers with preferencing arrangements, the price impact of preferenced trades would be smaller than the price impact of unpreferenced trades.

We measure the price impact of trades by $\text{IMPACT}(t) = 100D(t)[\{M(t+5) - M(t)\}/M(t)]$, where $M(t)$ and $M(t+5)$ are quote midpoints at time $t$ and $t+5$ min, respectively, and $D(t)$ is a trade direction indicator that equals $+1(-1)$ for buyer (seller) initiated trades. Next, for each stock, we calculate the mean value of IMPACT($t$) for each of the six trade groups (i.e., INS, ECN, INS + ECN, INT, NINS, and INT + NINS) by weighting each trade equally within each trade-size group. Following SEC Rule11Ac1-5 Report, we classify trades into the following four size groups: 100–499 shares, 500–1,999 shares, 2,000–4,999 shares, and 5,000+ shares. Finally, we calculate the mean value of IMPACT($t$) across stocks.

Panel A of Table 7 shows the mean price impact for each trade-size group during the pre- and post-decimalization periods. To determine whether the average price impact differs significantly between preferenced and unpreferenced trades, we compare the average price impact of preferenced trades with the average price impact of unpreferenced trades. In each cell, we show the average price impact of preferenced trades (INT, NINS, or INT + NINS), the average price impact of unpreferenced trades (INS, ECN, or INS + ECN), and the difference between the two groups. Numbers in parentheses are paired-comparison $t$-statistics for the equality of mean between preferenced trades and unpreferenced trades (e.g., INT − INS, INT − ECN, NINS − ECN, etc.).

The results show that the price impact of internalized trades (INT) is significantly smaller than the price impact of unpreferenced trades (INS, ECN, or (INS + ECN)) across all trade-size categories during both the pre- and post-decimalization period. Similarly, we also find that the price impact of non-inside trades (NINS) is significantly smaller than the price impact of unpreferenced trades in two large trade-size categories during both periods. The price impact of preferenced trades as a whole (INT + NINS) is significantly smaller than the price impact of unpreferenced trades as a whole (INS + ECN) in all trade-size categories during both periods. Overall, these results are consistent with the prediction of the clientele-pricing hypothesis advanced by Battalio and Holden (2001) and others that dealers (brokers) selectively purchase (internalize) orders based on their information content.

To assess the sensitivity of our results, we also employ an alternative empirical method. We regress the price impact of trades on dummy variables for internalized and non-inside trades, dummy variables for trade-size categories two, three, and four, and four stock attributes (share price, number of trades, return volatility, and market capitalization). The regression results show that the estimated coefficients for dummy variables for internalized and non-inside trades are negative and significant at the 1% level during both the pre- and post-decimalization periods, indicating that the price impact of preferenced trades is smaller than the price impact of unpreferenced trades.[7] Hence, our findings are quite robust and not sensitive to different empirical methods.

### 5.2. *Preferencing and price/size improvement*

Quoted bid and ask prices are not necessarily the prices at which trades take place because it is possible to trade at inside the quoted prices. Quoted prices are the starting point for a negotiation, not the price at which trades take place. Huang and Stoll (1996) find that the proportion of trades inside the quotes is 0.267 on NASDAQ and 0.379 on the NYSE from a matching sample of NASDAQ and NYSE stocks in 1991.

[7] The results are available from the authors upon request.

Table 7

Comparisons of price impact and price/size improvements between preferenced and unpreferenced trades

We measure the price impact of trades by $\text{IMPACT}(t) = 100D(t)[M(t+5) - M(t)/M(t)]$, where $M(t)$ and $M(t+5)$ are quote midpoints at time $t$ and $t+5$ min, respectively, and $D(t)$ is a trade direction indicator that equals $+1(-1)$ for buyer (seller) initiated trades. Next, for each stock, we calculate the mean value of $\text{IMPACT}(t)$ for each of the six trade groups (i.e., INS, ECN, INS + ECN, INT, NINS, and INT + NINS) by weighting each trade equally within each trade-size group. Finally, we calculate the mean value of $\text{IMPACT}(t)$ across stocks. Panel A shows the mean price impact for each trade-size group during the pre- and post-decimalization periods. To determine whether the average price impact differs significantly between preferenced and unpreferenced trades, we compare the average price impact of preferenced trades with the average price impact of unpreferenced trades. In each cell, we show the average price impact of preferenced trades (INT, NINS, or INT + NINS), the average price impact of unpreferenced trades (INS, ECN, or INS + ECN), and the difference between the two groups. Numbers in parentheses are paired-comparison $t$-statistics for the equality of mean between preferenced trades and unpreferenced trades (e.g., INT–INS, INT– ECN, NINS–ECN, etc.). Panel B reports the price improvement rates for pre- and post-decimalization periods. We measure the price improvement rate for each trade by $\text{PI}(t) = 100[\text{IAP}(t) - P(t)/\text{IAP}(t)]$ if $D(t) = 1$ and $\text{PI}(t) = 100[P(t) - \text{IBP}(t)/\text{IBP}(t)]$ if $D(t) = -1$, where $P(t) =$ trade price at time $t$, $\text{IAP}(t) =$ the inside ask price at time $t$, and $\text{IBP}(t) =$ the inside bid price *at time* $t$. Panel C presents the size improvement rates for pre- and post-decimalization periods. We measure the size improvement rate by $\text{SI}(t) = 100\,\text{Max}[S(t) - \text{IAS}(t)/\ \text{IAS}(t), 0]$ if $D(t) = 1$ and $\text{SI}(t) = 100\,\text{Max}[S(t) - \text{IBS}(t)/\text{IBS}(t), 0]$ if $D(t) = -1$, where $S(t) =$ trade size at time $t$, $\text{IAS}(t) =$ the inside ask size at time $t$, and $\text{IBS}(t) =$ the inside bid size at time $t$.

| Trade size | | Before decimalization | | | After decimalization | | |
|---|---|---|---|---|---|---|---|
| | | INS | ECN | INS + ECN | INS | ECN | INS + ECN |
| *Panel A. Price impact* | | | | | | | |
| 100–499 | INT | 0.1976–0.5116 = −0.3140** (−11.31) | 0.1976–0.3416 = −0.1440** (−5.70) | 0.1976–0.4109 = −0.2133** (−8.68) | 0.0920–0.3222 = −0.2302** (−18.20) | 0.0920–0.1946 = −0.1026** (−5.27) | 0.0920–0.2409 = −0.1489** (−9.99) |
| | NINS | 0.3835–0.5116 = −0.1281** (−9.82) | 0.3835–0.3416 = 0.0419 (1.22) | 0.3835–0.4109 = −0.0274** (−4.08) | 0.2249–0.3222 = −0.0973** (−9.35) | 0.2249–0.1946 = 0.0303 (1.06) | 0.2249–0.2409 = −0.0160** (−4.74) |
| | INT + NINS | 0.3357–0.5116 = −0.1759** (−12.46) | 0.3357–0.3416 = −0.0059 (−1.30) | 0.3357–0.4109 = −0.0752** (−6.31) | 0.2001–0.3222 = −0.1221** (−14.02) | 0.2001–0.1946 = 0.0055 (0.56) | 0.2001–0.2409 = −0.0408** (−7.28) |
| 500–1,999 | INT | 0.2296–0.4706 = −0.2410** (−9.38) | 0.2296–0.3709 = −0.1413** (−6.72) | 0.2296–0.4193 = −0.1897** (−8.84) | 0.1332–0.3015 = −0.1683** (−10.29) | 0.1332–0.2320 = −0.0988** (−4.57) | 0.1332–0.2569 = −0.1237** (−8.33) |
| | NINS | 0.3740–0.4706 = −0.0966** (−5.43) | 0.3740–0.3709 = 0.0031 (0.13) | 0.3740–0.4193 = −0.0453** (−4.83) | 0.2574–0.3015 = −0.0441** (−4.87) | 0.2574–0.2320 = 0.0254 (1.46) | 0.2574–0.2569 = 0.0005 (0.05) |
| | INT + NINS | 0.3306–0.4706 = −0.1400** (−7.82) | 0.3306–0.3709 = −0.0403** (−4.70) | 0.3306–0.4193 = −0.0887** (−7.56) | 0.2234–0.3015 = −0.0781** (−8.92) | 0.2234–0.2320 = −0.0086 (−0.50) | 0.2234–0.2569 = −0.0335** (−5.25) |
| 2,000–4,999 | INT | 0.1063–0.4483 = −0.3420** (−12.31) | 0.1063–0.3900 = −0.2837** (−9.04) | 0.1063–0.4218 = −0.3155** (−12.48) | 0.0707–0.2872 = −0.2165** (−9.43) | 0.0707–0.2919 = −0.2212** (−7.27) | 0.0707–0.2863 = −0.2156** (−8.29) |

Table 7. (*Continued*)

| Trade size | | Before decimalization | | | After decimalization | | |
|---|---|---|---|---|---|---|---|
| | | INS | ECN | INS + ECN | INS | ECN | INS + ECN |
| | NINS | 0.2432–0.4483 = −0.2051** (−9.74) | 0.2432–0.3900 = −0.1468** (−5.53) | 0.2432–0.4218 = −0.1786** (−9.88) | 0.1202–0.2872 = −0.1670** (−9.38) | 0.1202–0.2919 = −0.1717** (−5.31) | 0.1202–0.2863 = −0.1661** (−7.68) |
| | INT + NINS | 0.1903–0.4483 = −0.2580** (−12.68) | 0.1903–0.3900 = −0.1997** (−7.72) | 0.1903–0.4218 = −0.2315** (−13.45) | 0.0980–0.2872 = −0.1892** (−10.90) | 0.0980–0.2919 = −0.1939** (−7.04) | 0.0980–0.2863 = −0.1883** (−8.85) |
| 5,000+ | INT | 0.0233–0.4312 = −0.4079** (−14.79) | 0.0233–0.4651 = −0.4418** (−7.54) | 0.0233–0.4317 = −0.4084** (−10.96) | 0.0098–0.2835 = −0.2737** (−11.62) | 0.0098–0.3471 = −0.3373** (−8.30) | 0.0098–0.3163 = −0.3065** (−11.01) |
| | NINS | 0.0857–0.4312 = −0.3455** (−12.50) | 0.0857–0.4651 = −0.3794** (−6.81) | 0.0857–0.4317 = −0.3460** (−9.51) | 0.0416–0.2835 = −0.2419** (−10.60) | 0.0416–0.3471 = −0.3055** (−7.76) | 0.0416–0.3163 = −0.2747** (−10.18) |
| | INT + NINS | 0.0576–0.4312 = −0.3736** (−15.18) | 0.0576–0.4651 = −0.4075** (−7.21) | 0.0576–0.4317 = −0.3741** (−10.85) | 0.0253–0.2835 = −0.2582** (−13.17) | 0.0253–0.3471 = −0.3218** (−8.53) | 0.0253–0.3163 = −0.2910** (−11.98) |
| *Panel B. Price improvement* | | | | | | | |
| 100–499 | INT | 0.1260–0.2029 = −0.0769** (−6.65) | 0.1260–0.3133 = 0.1873** (−10.89) | 0.1260–0.2266 = −0.1006** (−7.05) | 0.0481–0.0891 = −0.0410** (−5.58) | 0.0481–0.1453 = −0.0972** (−9.48) | 0.0481–0.1146 = −0.0665** (−8.44) |
| | NINS | 0.2175–0.2029 = 0.0146 (1.45) | 0.2175–0.3133 = −0.0958** (−7.91) | 0.2175–0.2266 = −0.0091** (−3.34) | 0.0855–0.0891 = −0.0036 (−0.44) | 0.0855–0.1453 = −0.0598** (−6.18) | 0.0855–0.1146 = −0.0291** (−3.67) |
| | INT + NINS | 0.1859–0.2029 = −0.0170** (−3.77) | 0.1859–0.3133 = −0.1274** (−9.49) | 0.1859–0.2266 = −0.0407** (−5.95) | 0.0800–0.0891 = −0.0091** (−3.82) | 0.0800–0.1453 = −0.0653** (−7.14) | 0.0800–0.1146 = −0.0346** (−5.88) |
| 500–1,999 | INT | 0.0888–0.1537 = −0.0649** (−5.20) | 0.0888–0.2518 = −0.1630** (−12.39) | 0.0888–0.1841 = −0.0953** (−9.92) | 0.0456–0.0761 = −0.0305** (−5.70) | 0.0456–0.1348 = −0.0892** (−8.38) | 0.0456–0.1012 = −0.0556** (−7.48) |
| | NINS | 0.1556–0.1537 = 0.0019 (0.95) | 0.1556–0.2518 = −0.0962** (−9.62) | 0.1556–0.1841 = −0.0285** (−5.42) | 0.0708–0.0761 = −0.0053 (−1.04) | 0.0708–0.1348 = −0.0640** (−7.64) | 0.0708–0.1012 = −0.0304** (−5.50) |
| | INT + NINS | 0.1359–0.1537 = −0.0178** (−4.18) | 0.1359–0.2518 = −0.1159** (−11.70) | 0.1359–0.1841 = −0.0482** (−8.14) | 0.0652–0.0761 = −0.0109** (−4.20) | 0.0652–0.1348 = −0.0696** (−8.40) | 0.0652–0.1012 = −0.0360** (−6.69) |
| 2,000–4,999 | INT | 0.0791–0.1306 = −0.0515** (−6.48) | 0.0791–0.2057 = −0.1266** (−10.61) | 0.0791–0.1560 = −0.0769** (−7.69) | −0.0147–0.0611 = −0.0758** (−7.68) | −0.0147–0.0701 = −0.0848** (−8.03) | −0.0147–0.0632 = −0.0779** (−7.96) |
| | NINS | 0.1456–0.1306 = 0.0150 (1.38) | 0.1456–0.2057 = −0.0601** (−7.24) | 0.1456–0.1560 = −0.0104** (−4.38) | 0.0171–0.0611 = −0.0440** (−5.09) | 0.0171–0.0701 = −0.0530** (−5.06) | 0.0171–0.0632 = −0.0461** (−5.29) |
| | INT + NINS | 0.1227–0.1306 = −0.0079** (−4.08) | 0.1227–0.2057 = −0.0830** (−8.40) | 0.1227–0.1560 = −0.0333** (−5.77) | 0.0124–0.0611 = −0.0487** (−6.07) | 0.0124–0.0701 = −0.0577** (−6.86) | 0.0124–0.0632 = −0.0508** (−6.38) |
| 5,000+ | INT | −0.1530–0.0861 = −0.2391** (−13.70) | −0.1530–0.1525 = −0.3055** (−13.70) | −0.1530–0.1132 = −0.2662** (−14.55) | −0.2559–0.0557 = −0.3116** (−14.03) | −0.2559–0.0691 = −0.3250** (−9.57) | −0.2559–0.0673 = −0.3232** (−12.07) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | NINS | 0.0277–0.0861 = −0.0584** (−5.49) | 0.0277–0.1525 = −0.1248** (−7.68) | 0.0277–0.1132 = −0.0855** (−7.43) | −0.1453–0.0557 = −0.2010** (−11.06) | −0.1453–0.0691 = −0.2144** (−6.83) | −0.1453–0.0673 = −0.2126** (−9.10) |
| | INT + NINS | −0.0470–0.0861 = −0.1331** (−13.95) | −0.0470–0.1525 = −0.1995** (−12.48) | −0.0470–0.1132 = −0.1602** (−15.04) | −0.1750–0.0557 = −0.2307** (−17.17) | −0.1750–0.0691 = −0.2441** (−8.46) | −0.1750–0.0673 = −0.2423** (−12.17) |
| *Panel C. Size improvement* | | | | | | | |
| 100–499 | INT | 9,286–7,577 = 1,709** (18.32) | 9,286–8,643 = 643** (6.18) | 9,286–8,166 = 1,120** (12.09) | 11,677–8,213 = 3,464** (28.44) | 11,677–8,993 = 2,684** (21.73) | 11,677–8,798 = 2,879** (24.06) |
| | NINS | 8,603–7,577 = 1,026** (21.05) | 8,603–8,643 = −40 (−0.31) | 8,603–8,166 = 437** (8.39) | 8,954–8,213 = 741** (13.55) | 8,954–8,993 = −39 (−0.43) | 8,954–8,798 = 156** (5.15) |
| | INT + NINS | 8,803–7,577 = 1,226** (27.94) | 8,803–8,643 = 160** (4.45) | 8,803–8,166 = 637** (14.32) | 9,471–8,213 = 1,258** (28.33) | 9,471–8,993 = 478** (8.67) | 9,471–8,798 = 673** (16.07) |
| 500–1,999 | INT | 41,261–36,647 = 4,614** (10.76) | 41,261–34,248 = 7,013** (15.23) | 41,261–34,667 = 6,594** (16.12) | 48,159–41,666 = 6,493** (14.10) | 48,159–38,417 = 9,742** (20.59) | 48,159–39,419 = 8,740** (19.73) |
| | NINS | 38,231–36,647 = 1,584** (7.21) | 38,231–34,248 = 3,983** (13.54) | 38,231–34,667 = 3,564** (18.52) | 41,295–41,666 = −371 (−1.12) | 41,295–38,417 = 2,878** (10.42) | 41,295–39,419 = 1,876** (9.22) |
| | INT + NINS | 39,272–36,647 = 2,625** (11.58) | 39,272–34,248 = 5,024** (17.08) | 39,272–34,667 = 4,605** (23.44) | 43,399–41,666 = 1,733** (7.43) | 43,399–38,417 = 4,982** (18.10) | 43,399–39,419 = 3,980** (19.31) |
| 2,000–4,999 | INT | 121,758–109,204 = 12,554** (8.69) | 121,758–103,074 = 18,684** (12.51) | 121,758–104,041 = 17,717** (14.20) | 146,298–135,824 = 10,474** (6.51) | 146,298–126,757 = 19,541** (12.54) | 146,298–129,607 = 16,691** (12.24) |
| | NINS | 116,081–109,204 = 6,877** (5.95) | 116,081–103,074 = 13,007** (10.12) | 116,081–104,041 = 12,040** (12.90) | 139,286–135,824 = 3,462** (3.57) | 139,286–126,757 = 12,529** (9.43) | 139,286–129,607 = 9,679** (9.16) |
| | INT + NINS | 118,480–109,204 = 9,276** (8.49) | 118,480–103,074 = 15,406** (12.76) | 118,480–104,041 = 14,439** (17.21) | 141,849–135,824 = 6,025** (5.89) | 141,849–126,757 = 15,092** (12.86) | 141,849–129,607 = 12,242** (13.86) |
| 5,000+ | INT | 528,624–293,574 = 235,050** (9.33) | 528,624–294,852 = 233,772** (9.24) | 528,624–290,930 = 237,694** (10.72) | 698,913–381,808 = 317,105** (14.60) | 698,913–351,421 = 347,492** (14.91) | 698,913–361,602 = 337,311** (15.42) |
| | NINS | 397,020–293,574 = 103,446** (7.91) | 397,020–294,852 = 102,168** (7.09) | 397,020–290,930 = 106,090** (8.25) | 459,705–381,808 = 77,897** (7.30) | 459,705–351,421 = 108,284** (8.27) | 459,705–361,602 = 98,103** (9.19) |
| | INT + NINS | 452,970–293,574 = 159,396** (11.50) | 452,970–294,852 = 158,118** (10.54) | 452,970–290,930 = 162,040** (13.44) | 562,644–381,808 = 180,836** (16.53) | 562,644–351,421 = 211,223** (15.66) | 562,644–361,602 = 201,042** (18.20) |

** Significant at the 1% level.

Hansch et al. (1999) hold that if posting the best price indicates a greater willingness to transact, then dealers who post the best price offer a better price improvement than dealers who do not post the best price. Because preferenced trades are routed to dealers who may or may not be at the inside (whereas the unpreferenced trades are always routed to dealers who are at the inside market), price improvement is likely lower for preferenced trades than unpreferenced trades.

In contrast, Seppi (1990) holds that the transaction is not anonymous with negotiated price improvements and thus a penalty technology could force repeat customers to trade at the quotes when they are informed and ask for an improvement only when they are uninformed. Barclay and Warner (1993) and Rhodes-Kropf (2001) hold that negotiation allows dealers to assess the customer's information, and they suggest that informed customers remain anonymous and trade at the quotes while uninformed customers negotiate for price improvements. In so far as preferenced orders have lower information content than unpreferenced orders, these studies suggest that preferenced orders receive greater price improvements than unpreferenced orders.

We measure the price improvement rate for each trade by $\mathrm{PI}(t) = 100[\{\mathrm{IAP}(t) - P(t)\}/\mathrm{IAP}(t)]$ if $D(t) = 1$ and $\mathrm{PI}(t) = 100[\{P(t) - \mathrm{IBP}(t)\}/\mathrm{IBP}(t)]$ if $D(t) = -1$, where $P(t) =$ trade price at time $t$, $\mathrm{IAP}(t) =$ the inside ask price at time $t$, $\mathrm{IBP}(t) =$ the inside bid price at time $t$, and $D(t) =$ a trade direction indicator that equals $+1(-1)$ for buyer (seller) initiated trades. Next, for each stock, we calculate the mean value of $\mathrm{PI}(t)$ for each of the four trade groups by weighting each trade equally within each trade-size group. Finally, we calculate the mean value of PI across stocks.

Panel B of Table 7 shows the average price improvement rate for each trade-size group during the pre- and post-decimalization periods. Consistent with the result in Bacidore et al. (2001), the price improvement rates after decimal pricing are smaller than the corresponding figures before decimalization. More importantly, preferenced trades generally receive smaller price improvements than unpreferenced trades. The average price improvement rates for internalized trades are significantly smaller than the corresponding figures for unpreferenced trades in all four trade-size categories during both periods. The average price improvement rates for non-inside trades (NINS) are significantly smaller than the corresponding figures for unpreferenced trades in most trade-size categories. The average price improvement rates for preferenced trades as a whole (INT + NINS) are significantly smaller than the corresponding figures for unpreferenced trades as a whole (INS + ECN) in all four trade-size categories during both periods.

Overall, these results are consistent with Hansch et al. (1999) conjecture that preferenced trades receive smaller price improvements than unpreferenced trades. The results are in line with Bessembinder's (2003b) finding that trades executed at a market displaying inferior quotes receive poorer execution prices, as compared to matched trades in the same stock that are executed on the same market when quotes are competitive. The results are also in line with the finding of Chordia and

Subrahmanyam (1995) that trades executed off the NYSE (which are largely preferenced) receive smaller price improvements than trades executed on the NYSE.

As in the case of the price impact of trades, we also examine the robustness of our results by regressing the price improvement rate against the same set of explanatory variables. We find that the estimated coefficients for dummy variables for internalized and non-inside trades are both negative and significant during the post-decimalization period, indicating smaller price improvements for preferenced trades.

Oftentimes, order preferencing is accompanied by execution size guarantees that are larger than the typical size at the inside quotes. As a result, preferenced trades are likely to receive greater size improvement. We measure the size improvement rate by $\mathrm{SI}(t) = 100\,\mathrm{Max}[\{S(t) - \mathrm{IAS}(t)\}/\mathrm{IAS}(t), 0]$ if $D(t) = 1$ and $\mathrm{SI}(t) = 100\,\mathrm{Max}[\{S(t) - \mathrm{IBS}(t)\}/\mathrm{IBS}(t), 0]$ if $D(t) = -1$, where $S(t) =$ trade size at time $t$, $\mathrm{IAS}(t) =$ the inside ask size at time $t$, and $\mathrm{IBS}(t) =$ the inside bid size at time $t$. As noted by Bacidore et al. (2002), the accurate measurement of depth (size) improvement (the order-weighted depth improvement) requires data on order size to obtain the number of shares that are eligible for depth improvement. Our measures of depth improvement are imperfect because we do not use order size information due to our data limitation. For each stock, we calculate the mean value of $\mathrm{SI}(t)$ for each of the four trade groups by weighting each trade equally and then obtain the mean value of SI across stocks.

Panel C shows that preferenced trades generally receive greater size improvements than unpreferenced trades across all trade-size categories during both periods, with few exceptions. The differences are all statistically significant at the 1% level. Not surprisingly, the magnitude of size improvements increases with trade size for both preferenced and unpreferenced trades. When we employ the regression approach, we find that the estimated coefficients for dummy variables for internalized and non-inside trades are positive and significant at the 1% level. Hence, market makers offer greater size improvements selectively for trades routed by their favored brokers with preferencing arrangements. Market makers also offer greater size improvements for internalized trades. These results are consistent with the clientele-pricing hypothesis addressed in Benveniste et al. (1992), Huang and Stoll (1996), Battalio and Holden (2001), and Battalio et al. (2001a). In addition to direct payments for order flow to brokers based on information content, market makers also offer size improvements.

## 6. Summary and concluding remarks

Despite the widely held notion that order preferencing has detrimental effects on market quality and may have led to wider bid-ask spreads on NASDAQ, there has been no prior direct evidence on the cross-sectional relation between NASDAQ execution costs and the extent of order preferencing. In the present study, we provide new evidence on these issues using a large sample of NASDAQ stocks before and after decimalization.

Our empirical results suggest that order preferencing is prevalent on NASDAQ during both the pre- and post-decimalization periods. Although decimal pricing results in lower preferencing, the post-decimal proportion of preferenced volume is still much higher than what some people had predicted. Dealer quote aggressiveness (the spread) is significantly and negatively (positively) related to the extent of internalization during both the pre- and post-decimalization periods. Consistent with the prediction of the clientele-pricing hypothesis advanced in several recent studies, we find that the price impact of preferenced trades is smaller than that of unpreferenced trades. Market makers help affiliated brokers and brokers with preferencing agreements by offering greater size improvements.

Although the present study provides some new evidence regarding the effects of order preferencing on dealer quote competition and execution costs, the net effect of order preferencing on investor welfare is not obvious because order preferencing is likely to have broad and diverse ramifications for investor welfare. For example, order preferencing can reduce broker search costs, allowing the savings to be passed along to customers in the form of reduced commissions. In addition, there are dimensions of market quality other than price competition, such as speed of execution and reliability. Preferencing can also improve these areas of market quality. The accurate quantification of these benefits is likely difficult and is well beyond the scope of the present study. Further investigations into these issues would be a fruitful area for future research.

## References

Ackert, L., Church, B., 1999. Bid-ask spreads in multiple dealer settings: some experimental evidence. Financial Management 28, 75–88.

Bacidore, J., Battalio, R., Jennings, R., Farkas, S., 2001. Changes in order characteristics, displayed liquidity, and execution quality on the New York Stock Exchange around the switch to decimal pricing. Unpublished working paper. The New York Stock Exchange.

Bacidore, J., Battalio, R., Jennings, R., 2002. Depth improvement and adjusted price improvement on the New York Stock Exchange. Journal of Financial Markets 5, 169–195.

Barclay, M., Warner, J., 1993. Stealth trading and volatility: which trades move prices? Journal of Financial Economics 34, 281–306.

Battalio, R., Holden, C., 1996. Does increased competition in securities markets always lower trading costs? Unpublished working paper. Indiana University.

Battalio, R., Holden, C., 2001. A simple model of payment for order flow, internalization, and total trading costs. Journal of Financial Markets 4, 33–71.

Battalio, R., Greene, J., Jennings, R., 1997. Do competing specialists and preferencing dealers affect market quality? Review of Financial Studies 10, 969–993.

Battalio, R., Jennings, R., Selway, J., 2001a. The potential for clientele pricing when making markets in financial securities. Journal of Financial Markets 4, 85–112.

Battalio, R., Jennings, R., Selway, J., 2001b. The relationship among market-making revenue, payment for order flow, and trading costs for market orders. Journal of Financial Services Research 19, 39–56.

Benveniste, L., Marcus, A., Wilhelm, W., 1992. What's special about the specialist? Journal of Financial Economics 32, 61–86.

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., 2003a. Trade execution costs and market quality after decimalization. Journal of Financial and Quantitative Analysis, forthcoming.

Bessembinder, H., 2003b. Quote-based competition and trade execution costs in NYSE-listed stocks. Journal of Financial Economics, Article in Press.

Bloomfield, R., O'Hara, M., 1998. Does order preferencing matter? Journal of Financial Economics 50, 3–37.

Blume, M., Goldstein, M., 1997. Quotes, order flow, and price discovery. Journal of Finance 52, 221–244.

Chakravarty, S., Harris, S., Wood, R., 2001. Decimal trading and market impact: the NASDAQ experience. Unpublished working paper. University of Memphis.

Chordia, T., Subrahmanyam, A., 1995. Market making, the tick size, and payment-for-order flow: theory and evidence. Journal of Business 68, 543–575.

Christie, W., Schultz, P., 1994. Why do NASDAQ market makers avoid odd-8th quotes? Journal of Finance 49, 1813–1840.

Chung, K.H., Van Ness, B., Van Ness, R., 2001. Can the treatment of limit orders reconcile the differences in trading costs between NYSE and NASDAQ issues? Journal of Financial and Quantitative Analysis 36, 267–286.

Chung, K.H., Van Ness, B., Van Ness, R., 2003. Trading costs and quote clustering on the NYSE and NASDAQ after decimalization. Journal of Financial Research, forthcoming.

Dutta, P., Madhavan, A., 1997. Competition and collusion in dealer markets. Journal of Finance 52, 245–276.

Easley, D., Kiefer, N., O'Hara, M., 1996. Cream-skimming or profit sharing? The curious role of purchased order flow. Journal of Finance 51, 811–834.

Ellis, K., Michaely, R., O'Hara, M., 2002. The making of a dealer market: from entry to equilibrium in the trading of NASDAQ stocks. Journal of Finance 57, 2289–2316.

Godek, P., 1996. Why NASDAQ market makers avoid odd-eighth quotes. Journal of Financial Economics 41, 465–474.

Hansch, O., Naik, N., Viswanathan, S., 1999. Preferencing, internalization, best execution, and dealer profits. Journal of Finance 54, 1799–1828.

Harris, L., 1997. Decimalization: a review of the arguments and evidence. Unpublished working paper. University of Southern California.

Harris, L., 1999. Trading in pennies: a survey of the issues. Unpublished working paper. University of Southern California.

Huang, R., 2002. The quality of ECN and NASDAQ market maker quotes. Journal of Finance 57, 1285–1319.

Huang, R., Stoll, H., 1996. Dealer versus auction markets: a paired comparison of execution costs on NASDAQ and the NYSE. Journal of Financial Economics 41, 313–357.

Kandel, E., Marx, L., 1997. NASDAQ market structure and spread patterns. Journal of Financial Economics 45, 61–90.

Kandel, E., Marx, L., 1999. Payments for order flow on NASDAQ. Journal of Finance 54, 35–66.

Klock, M., McCormick, D., 1999. The impact of market maker competition on NASDAQ spreads. Financial Review 34, 55–74.

Kluger, B., Wyatt, S., 2002. Preferencing, internalization of order flow, and tacit collusion: evidence from experiments. Journal of Financial and Quantitative Analysis 37, 449–469.

Lee, C., Ready, M., 1991. Inferring trade direction from intraday data. Journal of Finance 46, 733–746.

McInish, T., Wood, R., 1992. An analysis of intraday patterns in bid/ask spreads for NYSE stocks. Journal of Finance 47, 753–764.

National Association of Securities Dealers (NASD), 1991. Inducement for order flow. A Report of the Order Flow Committee to the Board of Governors. NASD, Washington, D.C.

NASDAQ Stock Market Inc., 2001a. Impact of decimalization. Press release by NASDAQ Economic Research, June 12, 2001.

NASDAQ Stock Market Inc., 2001b. The impact of decimalization on the NASDAQ stock market. Final report to the Securities and Exchange Commission. Manuscript prepared by the NASDAQ Economic Research, June 11, 2001.
New York Stock Exchange (NYSE), 2001. Comparing bid-ask spreads on the New York Stock Exchange and NASDAQ immediately following NASDAQ decimalization. Manuscript prepared by NYSE Research, July 26, 2001.
Peterson, M., Sirri, E., 2003. Order preferencing and market quality on US equity exchanges. Review of Financial Studies 16, 385–415.
Porter, D., Weaver, D., 1997. Tick size and market quality. Financial Management 26, 5–26.
Rhodes-Kropf, M., 2001. Price improvement in dealership markets. Unpublished working paper. Columbia University.
Schultz, P., 2000. The market for market-making. Unpublished working paper. University of Notre Dame.
Securities and Exchange Commission (SEC), 1997. Report on the Practice of Preferencing. SEC, Washington, DC.
Seppi, D., 1990. Equilibrium block trading and asymmetric information. Journal of Finance 45, 73–94.
Stoll, H., 2000. Friction. Journal of Finance 55, 1479–1514.