Rev Quant Finan Acc (2006) 27:343–364
DOI 10.1007/s11156-006-0042-3

Order preferencing, adverse-selection costs, and the probability of information-based trading

Kee H. Chung · Chairat Chuwonganant · D. Timothy McCormick

Abstract Although prior studies offer various conjectures on the causes and consequences of order preferencing, there is only limited empirical evidence. In this study, we show that the extent of order preferencing is significantly and negatively related to both the adverse-selection component of the spread and the probability of information-based trading. This result is consistent with the prediction of the clientele-pricing hypothesis that dealers (brokers) selectively purchase (internalize) orders based on information content. Our results suggest that order preferencing may not be as harmful as some researchers have suggested and offer some rationale for its prevalence in securities markets with heterogeneously informed traders.

Keywords Order preferencing · Internalization · Components of the spread · Adverse-selection costs · Information-based trading

JEL Classification G18 · G19

1 Introduction

One of the central features of the NASDAQ Stock Market is that a large portion of order flow from brokers to dealers is preferenced. Brokers route customer orders to

K. H. Chung (✉)
Department of Finance and Managerial Economics, State University of New York (SUNY) at Buffalo, Buffalo, NY 14260, USA
e-mail: keechung@buffalo.edu

C. Chuwonganant
Department of Accounting and Finance, Indiana University-Purdue University at Fort Wayne, Fort Wayne, IN 46805, USA

D. T. McCormick
The Securities and Exchange Commission, 100 F Street, NE Washington, DC 20549

specific dealers according to 'pre-arranged' preferencing agreements, irrespective of whether dealers are posting competitive quotes at the time of order submission.[1] As a result, many NASDAQ dealers receive a significant portion of order flow when they are *not* at the inside market.

Order preferencing makes it difficult for dealers to control their inventories as they are obligated to accept preferenced orders. It also changes the nature of competition from posting more aggressive quotes to treating and servicing customers in other ways, such as charging low commissions, offering size improvements, or providing other in-kind goods and services. Clearly, any dealer firm wants to get more order flow. Getting more order flow not only means more business, but it also helps dealers' price discovery. More importantly, dealers want to get more *uninformed* order flow as it generates greater revenues than informed order flow. In the present study, we examine whether brokers and dealers use order preferencing as a mechanism to differentiate between informed and uninformed order flows.

Because order preferencing reduces the incentive to post aggressive quotes, prior studies examine whether order preferencing exerts a negative impact on market quality.[2] Chung, Chuwonganant, and McCormick (2004) provide the first empirical evidence on the extent of order preferencing on NASDAQ and its impact on market quality for a large sample of stocks. They also show that although decimal pricing lowered order preferencing, the extent of order preferencing after decimalization is much higher than what prior studies have suggested.

Barclay, Hendershott, and McCormick (2003) show that trades executed on Electronic Communications Networks (ECNs) have greater information content than those executed by dealers. Order preferencing arrangements allow NASDAQ dealers to attract and retain less-informed retail orders. Because ECNs typically preserve the anonymity of the trading parties, trades are more likely to occur on ECNs when information asymmetry is higher. Barclay, Hendershott, and McCormick show that ECN trades exhibit greater permanent price impacts and more private information is revealed through ECN trades than though dealer trades. The authors also show that ECN trades incur greater execution costs than dealer trades because dealers can preference less-informed trades and offer them better executions.[3]

Despite its prevalence, possible causes of order preferencing and its ramification for market quality and investor welfare are not well understood. Although prior studies offer conjectures on why and how market participants use order preferencing, there is only limited empirical evidence on this issue, largely due to the lack of available data on order preferencing. Battalio, Greene, and Jennings (1997) and Peterson and Sirri (2003) analyze the effect of order preferencing on market quality for NYSE-listed

[1]Order preferencing generally entails either payments for order flow or internalization. Dealers offer direct monetary payments or in-kind goods or services to brokers in return for preferenced orders. Internalization 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.

[2]See Kandel and Marx (1997, 1999), Bloomfield and O'Hara (1998), Battalio and Holden (2001), and Kluger and Wyatt (2002).

[3]There have been no studies to date that have examined the impact of ECNs on order preferencing and internalization on NASDAQ. However, the rapid growth of ECN trading in NASDAQ stocks likely led to a reduction in overall preferencing and internalization. Our research finds a large degree of preferencing and internalization continued to persist four years after the introduction of ECN quotations on NASDAQ.

stocks. The results found on the NYSE may not be directly relevant for NASDAQ securities because market fragmentation and dealer competition on NASDAQ are quite different from those on the NYSE. Hansch, Naik, and Viswanathan (1999) analyze the effect of order preferencing on spreads and dealer profits for a sample of London Stock Exchange (LSE) stocks. However, they do not examine the order routing practice of brokers.

Our study provides further empirical evidence on order preferencing using proprietary data from NASDAQ. In particular, we examine how the extent of order preferencing at the level of individual securities is related to adverse-selection costs and the probability of information-based trading using a large cross-sectional dataset. This study helps us understand why order preferencing may not be as harmful as some researchers have suggested and thereby offers some rationale for its existence.

Benveniste, Marcus, and Wilhelm (1992) hold that long-term relationships between brokers and dealers can mitigate the effects of asymmetric information. They suggest that dealers that actively identify and sanction informed traders can provide low cost services to uninformed traders. Easley, Kiefer, and O'Hara (1996) find a significant difference in the information content of orders executed in New York and Cincinnati and interpret the result as that the preferencing arrangements are used to cream-skim uninformed liquidity traders.

Battalio, Jennings, and Selway (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. Battalio, Jennings, and Selway (2001b) examine order preferencing on NASDAQ using proprietary data from Knight Securities. The authors find that payment for order flow does not unambiguously harm traders. Chung, Chuwonganant, and McCormick (2004) compare the price impact of trades between preferenced and unpreferenced orders. However, they do not analyze the order routing behavior of brokers in reference to the information content of orders. In addition, their empirical analyses do not control for factors that may be related to both order preferencing and the price impact of trades. To the extent that both order preferencing and the price impact of trades are related to a common set of variables, it is imperative to control them to minimize spurious correlations between the two variables.

In our study, we perform an alternative test of the clientele-pricing hypothesis by analyzing the order routing behavior of brokers.[4] Trades that convey less private information have smaller adverse-selection costs/risk and thus are more profitable to execute. If brokers route only those orders with low adverse-selection costs/risks to affiliated dealers (i.e., internalization) or dealers with payment for order flow arrangements, the extent of order preferencing for a given stock would be lower when adverse-selection costs/risk are higher.[5] Based on these considerations, we predict a negative cross-sectional relation between the extent of order preferencing and both adverse-selection costs and the probability of information-based trading (PIN).

[4]See Parlour and Rajan (2003) for a dynamic model of price competition in broker and dealer markets.

[5]In contrast, brokers are likely to send informed orders to unaffiliated dealers or dealers with no payment for order flow arrangements. Informed orders are likely to receive poorer executions (e.g., smaller price or size improvements) than preferenced, uninformed orders. See Chung, Chuwonganant, and McCormick (2004) for empirical evidence.

We show that the proportion of internalized trades is significantly and negatively related to both adverse-selection costs and PIN, after controlling for both stock characteristics and dealer types. Similarly, we find that the proportion of orders received when dealers are not at the inside market is significantly and negatively related to adverse-selection costs and PIN. Overall, these results are consistent with the prediction of the clientele-pricing hypothesis advanced by Benveniste, Marcus, and Wilhelm (1992), Battalio and Holden (2001), and others that dealers (brokers) selectively purchase (internalize) orders based on their information content.

The paper is organized as follows. Section 2 describes data sources and sample selection procedures and presents descriptive statistics. Section 3 explains the variable measurement procedure. Section 4 analyzes the relation between order preferencing and adverse-selection costs/risks. Section 5 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 and June 2001. By conducting separate empirical analyses for each month, we check the robustness of our results for different study periods. We use proprietary data from NASDAQ to infer whether a trade is internalized or routed through an order flow agreement. A stock is included in our study sample if its data are available from the above two sources. The final study sample consists of 3,032 stocks in November 2000 and 2,983 stocks in June 2001.[6] We omit the following to minimize data errors: (1) quotes if either the ask or the bid price 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 or equal to 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) transaction 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, trade size by the average dollar transaction size, and firm size by the market value of equity at the beginning of our study period. The Herfindahl-index is used as a measure of dealer competition and trading concentration in each stock. The Herfindahl-index is calculated using the following formula:

$$\text{H-INDEX}(i) = \Sigma j[100V(i, j)/\Sigma j V(i, j)]^2, \quad (1)$$

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

[6] 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 1,158.

leading dealer increases. Thus, a high Herfindahl-index is associated with high concentration of trading.

We report select attributes of our study sample in Table 1. The average share price in November 2000 (June 2001) is \$13.81 (\$12.04), the average dollar trade size is \$9,819 (\$7,900), and the average number of transactions is 587.51 (568.75), respectively. The average standard deviation of daily returns in November 2000 (June 2001) is 0.0566 (0.0421). The average market capitalization in November 2000 (June 2001) is \$1,187 (\$672) million. The average Herfindahl-index based on dollar volume in November 2000 (June 2001) is 2,399 (2,402).

3 Measurement of the variables

In this section, we describe how we measure order preferencing, adverse-selection costs, and the probability of information-based trading (PIN) and present their descriptive statistics.

3.1 Measurement of order preferencing

Proprietary data from NASDAQ contain information on quotes and transactions of all market makers that allows us to determine whether a public trade is preferenced (i.e., either internalized or routed through an order flow agreement). We consider a trade internalized 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 (i.e., routed through an order flow agreement) if the quote is poorer than the prevailing inside market quote.[7] For instance, if the market maker bought 500 shares at the inside market bid price of \$20 while he was bidding at \$19.875, we consider the trade preferenced.

We measure the extent of preferencing for stock i, PREF(i), by the ratio of stock i's internalized volume plus any noninternalized volume executed by dealers not quoting at the inside market at the time of the trade to its total volume,[8] i.e.,

$$\mathrm{PREF}(i) = \Sigma j[\mathrm{VINT}(i, j) + \mathrm{VNINS}(i, j)]/\Sigma j V(i, j); \qquad (2)$$

where VINT(i, j) is stock i's internalized volume forwarded to dealer j, VNINS(i, j) is stock i's noninternalized volume executed by dealer j when the dealer is not at the inside market (non-inside volume), and $V(i, j)$ is stock i's total volume executed by dealer j. We measure trading volume both in dollars and number of shares. However,

[7] The rule of best execution requires dealers to execute customer orders at prices that are equal to or better than the inside market quotes even when their quotes are below the inside bid or above the inside ask. We consider an order preferenced only if the quote at the time of its execution is poorer than the prevailing inside market quote. If the dealer is at the inside market at the time of an order's execution, we do not consider the order preferenced.

[8] Note that stock i's volume executed by dealer j, $V(i, j)$, can divided into four components: $V(i, j) = \mathrm{VINT}(i, j) + \mathrm{VINS}(i, j) + \mathrm{VNINS}(i, j) + \mathrm{VE}(i, j)$; where VINT($i$, j) is stock i's internalized volume to dealer j, VINS(i, j) is stock i's noninternalized volume executed by dealer j when the dealer is at the inside market, VNINS(i, j) is stock i's noninternalized volume executed by dealer j when the dealer is not at the inside market, and VE(i, j) is stock i's volume on ECNs routed by dealer j.

Table 1 Descriptive statistics

This table shows the attributes of our study sample of NASDAQ stocks. 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

			Percentile						
Variable	Mean	Standard deviation	1	5	25	50	75	95	99
November 2000									
Share price ($)	13.81	23.00	0.18	0.71	2.94	8.06	16.90	44.04	87.38
Number of trades	587.51	3,314.56	1.09	2.52	12.00	47.45	192.88	1,607.48	10,248.70
Trade size ($)	9,819	10,718	609	996	2,719	6,163	12,982	30,839	47,729
Return volatility	0.0566	0.0474	0.0028	0.0082	0.0264	0.0479	0.0738	0.1303	0.2575
Market value of equity ($K)	1,186,507	12,198,622	2,908	6,750	27,065	86,888	336,267	2,679,240	18,517,100
H-INDEX	2,399	1,568	488	731	1,355	1,995	2,980	5,562	8,411
June 2001									
Share price ($)	12.04	13.34	0.19	0.66	2.56	7.67	16.61	37.02	63.51
Number of trades	568.75	2,932.50	0.52	1.38	8.00	42.48	206.38	1,615.14	11,702.62
Trade size ($)	7,900	8,395	436	778	2,100	5,221	10,877	22,427	37,175
Return volatility	0.0421	0.0345	0.0026	0.0068	0.0203	0.0354	0.0537	0.0975	0.1744
Market value of equity ($K)	671,709	5,344,392	2,501	5,514	24,046	74,089	282,206	1,749,775	9,843,899
H-INDEX	2,402	1,685	412	620	1,228	1,939	3,100	5,792	8,639

the results are qualitatively identical between the two measures. Hence, we report only the results based on dollar volume throughout the paper.

Our measure of order preferencing is imperfect and thus our results should be interpreted with some caution. Note that orders can be preferenced to market makers who are at the inside market alone or together with other market makers. Hence, Eq. (2) is likely to underestimate the actual level of preferencing. One way to correct this downward bias is to inflate VNINS(i, j) by the proportion of time during which the market maker is not at the inside. However, this measure is likely to overstate order preferencing because of a self-selection problem. Dealers quote at the inside when they want to attract unpreferenced order flow. Thus, the probability of attracting an unpreferenced order is much greater than the probability of attracting a preferenced order during this time. If preferenced volume is only a small proportion of total volume, the extent of overstatement would not be large because the above method will assign a low probability to attracting a preferenced order anyway. However, if preferenced volume is a large proportion of total volume, the measure is likely to assign too high a probability. This problem is then magnified if the dealer is at the inside a long time.[9] For these reasons, we do not make an upward adjustment of VNINS(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 non-inside volume to his total volume, i.e.,

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

Table 2 shows the mean and standard deviation of preferenced volumes, 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, and PREF(i) = INT(i) + NINS(i). Panel A shows that there is wide variation in the percentage of internalized volumes across stocks with a mean value of 26.43% (24.29%) in November 2000 (June 2001). The non-inside volume accounts for 36.86% and 37.19% of the total volume, respectively, in November 2000 and June 2001. On average, the preferenced volume accounts for 63.29% and 61.48% of the total volume, respectively, in November 2000 and June 2001.

Panel B shows the percentages of preferenced trades by dealers, where INT(j) is the ratio of dealer j's internalized volume to his total volume, NINS(j) is the ratio of dealer j's non-inside volume to his total volume, and PREF(j) = INT(j) + NINS(j). The results show that there is wide variation in the percentage of internalized volumes across dealers with a mean value of 35.33% (30.26%) in November 2000 (June 2001). The non-inside volume accounts for 24.39% and 25.17% of the total volume, respectively, in November 2000 and June 2001. On average, the preferenced volume accounts for 59.72% and 55.43% of the total volume, respectively, in November 2000 and June 2001.

3.2 Measurement of adverse-selection costs and risks

We use the spread component models developed by Glosten and Harris (1988) and Lin, Sanger, and Booth (1995) to measure the adverse-selection cost. We use the

[9] The authors thank Frank Hatheway for this point.

Table 2 Stock and dealer preferencing

Panel A shows the percentages of internalized (INT(i)), non-inside (NINS(i)), and preferenced volume (PREF(i)). 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, and PREF(i) = INT(i) + NINS(i). Panel B shows the percentages of internalized (INT(j)), non-inside (NINS(j)), and preferenced volume (PREF(j)). INT(j) is the ratio of dealer j's internalized volume to his total volume, NINS(j) is the ratio of dealer j's non-inside volume to his total volume, and PREF(j) = INT(j) + NINS(j)

Variable	Mean	Standard deviation	Percentile						
			1	5	25	50	75	95	99
A. Stock preferencing (in %)									
November 2000									
INT(i)	26.43	18.51	0	0.30	10.51	25.77	39.85	57.22	74.02
NINS(i)	36.86	13.72	8.40	16.13	26.18	36.59	46.98	58.72	68.18
PREF(i) (INT(i) + NINS(i))	63.29	10.75	31.39	45.34	58.01	63.48	68.82	81.02	91.28
June 2001									
INT(i)	24.29	18.48	0	0	7.71	23.25	38.29	54.58	73.93
NINS(i)	37.19	15.66	9.93	15.50	25.48	37.59	50.03	63.45	76.26
PREF(i) (INT(i) + NINS(i))	61.48	10.38	33.73	45.65	57.00	62.29	67.73	80.65	90.23
B. Dealer preferencing (in %)									
November 2000									
INT(j)	35.33	25.91	0	0	9.35	38.87	56.47	73.62	90.38
NINS(j)	24.39	17.86	0	4.11	10.96	20.73	32.86	60.87	87.63
PREF(j) (INT(j) + NINS(j))	59.72	19.15	0	22.68	49.70	63.10	70.14	88.06	98.33
June 2001									
INT(j)	30.26	24.15	0	0	3.81	31.52	50.39	68.04	88.83
NINS(j)	25.17	17.47	0	3.67	11.61	20.53	35.18	58.71	78.18
PREF(j) (INT(j) + NINS(j))	55.43	17.94	0	18.29	46.21	59.35	66.97	80.21	91.72

algorithm in Easley, Hvidkjaer, and O'Hara (2002) to estimate the adverse-selection risk.

3.2.1 Glosten and Harris (GH) model

The GH model uses the following ordinary-least-squares regression to estimate the adverse-selection component of the spread:

$$P_t - P_{t-1} = c_0(Q_t - Q_{t-1}) + c_1(Q_t V_t - Q_{t-1} V_{t-1}) + z_0 Q_t + z_1 Q_t V_t + \varepsilon_t; \tag{4}$$

where P_t is the transaction price at time t, V_t is the number of shares traded at time t, ε_t is the error term that captures both the rounding error and the arrival of public information, and Q_t equals 1 for buyer-initiated trades and –1 for seller-initiated trades. We use the Lee and Ready (1991) algorithm as modified by Bessembinder (2003) to classify a trade as a buy or sell.[10]

We use the estimates of c_0, c_1, z_0, and z_1 for each stock to calculate the adverse-selection and transitory components. We estimate the adverse-selection component by $Z_0 = 2(z_0 + z_1 V_t)$ and the transitory component by $C_0 = 2(c_0 + c_1 V_t)$. The bid-ask spread in the GH model is the sum of Z_0 and C_0. We use the average transaction volume for stock i, $\bar{V}_i$, to estimate the adverse-selection component, $2(z_{0,i} + z_{1,i}\bar{V}_i)$ and the total spread, $2(c_{0,i} + c_{1,i}\bar{V}_i) + 2(z_{0,i} + z_{1,i}\bar{V}_i)$. We measure the adverse-selection component (as a proportion of the spread) for stock i by the ratio of $2(z_{0,i} + z_{1,i}\bar{V}_i)$ to $2(c_{0,i} + c_{1,i}\bar{V}_i) + 2(z_{0,i} + z_{1,i}\bar{V}_i)$.

3.2.2 Lin, Sanger, and Booth (LSB) model

Lin, Sanger, and Booth (1995) develop a model which shows that quote revisions reflect the adverse information revealed by the trade at time t. We use the following regression model to estimate the adverse-selection component of the effective spread:

$$Quote_t - Quote_{t-1} = \lambda z_{t-1} + \varepsilon_t; \tag{5}$$

where $Quote_t$ is the quote midpoint at time t, z_t is the signed effective half spread defined as the transaction price at time t minus the quote midpoint at time t, and λ measures the adverse-selection component (as a proportion of the effective spread). We use logs of the transaction price and quote midpoint in the estimation.

3.2.3 Adverse-selection risks

Easley, Kiefer, and O'Hara (1996, 1997a, 1997b) and Easley, Hvidkjaer, and O'Hara (2002) employ a comprehensive empirical measure of the probability of information-based trading (PIN) to examine a variety of market microstructure issues. Heidle and Huang (2002) analyze how PIN changes when a stock moves from NASDAQ to the NYSE. We use the algorithm in Easley, Hvidkjaer, and O'Hara (2002) to

[10] Bessembinder (2003) shows that making no allowances for trade-reporting lags is optimal when assessing whether trades are buyer or seller initiated.

estimate the adverse-selection risk. In Easley, Hvidkjaer, and O'Hara (EHO)'s model, market makers observe trades, update their beliefs, and establish price quotes. Over time, the process of trading, and learning from trading, results in prices converging to full information values. The EHO model provides the structure necessary to extract information from the observable variables, i.e., the number of buys and sells. EHO show that the structural model can be estimated via the maximum likelihood method, providing a convenient method for determining the value of information parameters (and thus PIN) for a given stock.

The EHO model of the trade process for a trading day is represented by the following likelihood function:

$$L(\theta|B,S) = (1-\alpha)e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} + \alpha\delta e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-(\mu+\varepsilon_s)}\frac{(\mu+\varepsilon_s)^S}{S!} + \alpha(1-\delta)e^{-(\mu+\varepsilon_b)}\frac{(\mu+\varepsilon_b)^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!}; \tag{6}$$

where B is the number of buyer-initiated trades for the day, S is the number of seller-initiated trades for the day, α the probability that an event is information based, δ is the probability that an information event contains good news, $1-\delta$ the probability that an information event contains bad news, μ is the arrival rate of orders from informed traders, ε_b is the arrival rate of orders from uninformed buyers, ε_s is the arrival rate of orders from uninformed sellers, and $\theta = (\alpha, \mu, \varepsilon_b, \varepsilon_s, \delta)$ represents the parameter vector. Note that $\varepsilon_b = \varepsilon_s = \varepsilon$ in the case where the uninformed traders are equally likely to buy or sell.

The likelihood function for the entire study period for each stock is given by:

$$V = L(\theta|M) = \prod_{d=1}^{D} L(\theta|\, B_d\; S_d); \tag{7}$$

where B_d (S_d) is the number of buyer (seller)-initiated trades for day $d = 1, 2, \ldots,$ D, and M is the data set that contains $((B_1, S_1), \ldots, (B_d, S_d))$. The rate of information arrival for all trades for stock i is represented by $\alpha\mu + \varepsilon_b + \varepsilon_s$ and the rate of arrival for information-based trades for stock i is $\alpha\mu$. We obtain the probability of information-based trading for stock i (PIN(i)) by the ratio of $\alpha\mu$ to $\alpha\mu + \varepsilon_b + \varepsilon_s$.

Panel A of Table 3 shows the mean and standard deviation of the adverse-selection component of the spread, together with their percentile values. The mean value of the adverse- selection component is 0.1246 (0.1334) in November 2000 (June 2001) according to the spread component model of Lin, Sanger, and Booth (1995) and 0.1114 (0.1216) in November 2000 (June 2001) according to the spread component model of Glosten and Harris (1988). Panel B shows the estimates of PIN and other model parameters. It shows that the mean value of PIN for our sample of NASDAQ stocks is 0.3007 (0.2710) in November 2000 (June 2001) with a standard deviation of 0.1725 (0.1671).

4 Preferencing as a function of adverse-selection costs and PIN

In this section, we examine how order preferencing is related to both adverse-selection costs and the probability of information-based trading.

Table 3 Descriptive statistics of adverse-selection components and probability of information-based trading
Panel A shows the descriptive statistics for the adverse-selection components of our sample stocks calculated using the methods developed by Lin, Sanger, and Booth (LSB) and Glosten and Harris (GH). Panel B shows the descriptive statistics for the probability of information-based trading (PIN)

Variable	Mean	Standard deviation	Percentile 1	5	25	50	75	95	99
A. Adverse-selection component (as a percentage of the spread)									
November 2000									
LSB	0.1246	0.0684	0.0133	0.0435	0.0856	0.1134	0.1469	0.2442	0.3523
GH	0.1114	0.0731	0.0072	0.0301	0.0703	0.0989	0.1336	0.2365	0.3686
June 2001									
LSB	0.1334	0.0839	0.0156	0.0482	0.0877	0.1157	0.1561	0.2778	0.4548
GH	0.1216	0.0890	0.0083	0.0315	0.0721	0.1020	0.1448	0.2706	0.5194
B. Probability of information-based trading									
November 2000									
PIN	0.3007	0.1725	0	0	0.2000	0.2754	0.3922	0.6232	0.8492
δ	0.6675	0.2901	0	0.0004	0.5000	0.7193	0.9207	1	1
μ	26.6223	25.1318	0.0998	4.9816	10.0000	15.9268	35.7775	78.5240	118.1951
ε	16.0738	19.9121	0.5002	1.0568	3.1194	10.0000	19.2161	59.6241	94.1411
α	0.4417	0.2460	0	0	0.2848	0.4285	0.5639	0.9999	1
June 2001									
PIN	0.2710	0.1671	0	0	0.1839	0.2490	0.3557	0.5692	0.8265
δ	0.5985	0.3219	0	0	0.3887	0.6003	0.9017	1	1
μ	27.5021	28.2201	0.0805	4.6387	10.0000	15.0713	36.8829	86.0612	124.5184
ε	15.5001	19.2620	0.6846	1.0680	3.1166	10.0000	17.7805	57.6464	94.8573
α	0.4045	0.2564	0	0	0.2350	0.3835	0.5049	0.9999	1

4.1 Univariate analysis

We first classify our sample stocks into five portfolios according to the market value of equity (MVE). Portfolio 1 contains stocks with the smallest MVE and Portfolio 5 contains stocks with the largest MVE. We then classify stocks in each MVE portfolio into five portfolios based on the probability of information-based trading (PIN). Portfolio 1 contains stocks with the lowest PIN and Portfolio 5 contains stocks with the highest PIN. For each PIN portfolio within each MVE group, we calculate the mean values of INT(i), NINS(i), and PREF(i), where 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, and PREF(i) = INT(i) + NINS(i). For brevity, we report only the results for MVE portfolios 1, 3, and 5. Similarly, within each MVE portfolio, we report the values of INT(i), NINS(i), and PREF(i) for PIN portfolios 1, 3, and 5.

Table 4 shows that both internalized and non-inside volumes decrease with PIN. For stocks that belong to the smallest MVE portfolio, the mean value of INT for Portfolio 1 (PIN = 0.0845) is 29.1% in November 2000, whereas the corresponding value for Portfolio 5 (PIN = 0.5449) is 24%. The difference in INT between Portfolio 1 and Portfolio 5 is statistically significant at the 1% level. Similarly, the mean value (34.73%) of NINS for Portfolio 5 is significantly smaller than the corresponding value (39.25%) for Portfolio 1. As a result, the mean value (58.73%) of PREF for Portfolio 5 is significantly smaller than the corresponding value (68.35%) for Portfolio 1. The results are similar for other MVE portfolios. We obtain qualitatively identical results for June 2001. On the whole, these results indicate that order preferencing decreases with the extent of information-based trading.

Table 5 shows the results when five portfolios are first formed by MVE and then by adverse-selection costs. We classify stocks in each MVE portfolio into five portfolios according to the adverse-selection component based on Lin, Sanger, and Booth (LSB). Within each MVE portfolio, Portfolio 1 contains stocks with the smallest adverse selection component and Portfolio 5 contains stocks with the largest adverse selection component. For each portfolio, we calculate the mean values of INT(i), NINS(i), and PREF(i). Similarly, using stocks in each MVE portfolio, we form five portfolios according to the adverse-selection component calculated from the method developed by Glosten and Harris (GH) and calculate the mean values of INT(i), NINS(i), and PREF(i) for each portfolio.

The results show that both internalized and non-inside volumes decrease with the adverse-selection component, regardless of whether we estimate the adverse-selection component using the LSB or GH models. For stocks that belong to the smallest MVE portfolio, the mean value of INT for Portfolio 1 is 28.36% in November 2000, whereas the corresponding value for Portfolio 5 is 23.1% when portfolios are formed according to the adverse selection component estimated from the LSB model. The difference in INT between Portfolio 1 and Portfolio 5 is statistically significant at the 1% level. Similarly, the mean value of NINS for Portfolio 5 is significantly smaller than the corresponding value for Portfolio 1. As expected, the mean value of PREF for Portfolio 5 is significantly smaller than the corresponding value for Portfolio 1. The results are similar for other MVE portfolios. We obtain similar results when we form portfolios according to the adverse selection component estimated from the GH model. We obtain qualitatively identical results for June 2001. On the whole,

Table 4 Order preferencing and the probability of information-based trading
We first classify our sample stocks into five portfolios according to the market value of equity (MVE). Portfolio 1 contains stocks with the smallest MVE and Portfolio 5 contains stocks with the largest MVE. We then classify stocks in each MVE portfolio into five portfolios based on the probability of information-based trading (PIN). Portfolio 1 contains stocks with the lowest PIN and Portfolio 5 contains stocks with the highest PIN. For each PIN portfolio within each MVE group, we calculate the mean values of INT(i), NINS(i), and PREF(i), where 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, and PREF(i) = INT(i) + NINS(i). For brevity, we report only the results for MVE portfolios 1, 3, and 5. Similarly, within each MVE portfolio, we report the values of INT(i), NINS(i), and PREF(i) for PIN portfolios 1, 3, and 5. Numbers in parentheses are t-statistics testing the equality of mean values between Portfolio 5 and Portfolio 1

	Portfolios based on the probability of information-based trading (PIN)				
MVE Portfolio	PIN portfolio	PIN(i)	INT(i) (in %)	NINS(i) (in %)	PREF(i) (in %)
		November 2000			
Portfolio 1	Portfolio 1	0.0845	29.10	39.25	68.35
	Portfolio 3	0.2735	27.35	35.65	63.00
	Portfolio 5	0.5449	24.00	34.73	58.73
	Portfolio 5 – 1	0.4604**	−5.10**	−4.52**	−9.62**
		(36.89)	(−3.06)	(−3.70)	(−6.60)
Portfolio 3	Portfolio 1	0.0933	28.99	38.90	67.89
	Portfolio 3	0.2771	25.01	37.24	62.25
	Portfolio 5	0.5631	24.72	33.89	58.61
	Portfolio 5 – 1	0.4698**	−4.27**	−5.01**	−9.28**
		(35.84)	(−2.74)	(−3.79)	(−5.74)
Portfolio 5	Portfolio 1	0.0934	28.52	40.37	68.89
	Portfolio 3	0.2774	25.06	37.27	62.33
	Portfolio 5	0.5691	24.28	35.42	59.70
	Portfolio 5 – 1	0.4685**	−4.24**	−4.95**	−9.19**
		(31.05)	(−2.77)	(−3.62)	(−6.98)
		June 2001			
Portfolio 1	Portfolio 1	0.0686	27.24	39.16	66.40
	Portfolio 3	0.2453	24.52	37.81	62.33
	Portfolio 5	0.5366	22.60	35.50	58.10
	Portfolio 5 – 1	0.4680**	−4.64**	−3.66**	−8.30**
		(31.89)	(−2.99)	(−2.87)	(−6.52)
Portfolio 3	Portfolio 1	0.0585	26.79	39.63	66.42
	Portfolio 3	0.2598	25.78	37.56	63.34
	Portfolio 5	0.5139	22.44	35.61	58.05
	Portfolio 5 – 1	0.4554**	−4.35**	−4.02**	−8.37**
		(37.92)	(−2.83)	(−3.03)	(−6.50)
Portfolio 5	Portfolio 1	0.0693	24.65	40.90	65.55
	Portfolio 3	0.2513	23.29	38.55	61.84
	Portfolio 5	0.5188	20.73	36.09	56.82
	Portfolio 5 – 1	0.4495**	−3.92**	−4.81**	−8.73**
		(34.55)	(−2.76)	(−3.23)	(−6.76)

**Significant at the 1% level.

these results indicate that order preferencing decreases with the adverse-selection cost.

4.2 Regression results

In this section, we provide further evidence on the relation between order preferencing and the adverse-selection cost and risk using regression analysis. In particular, we

Table 5 Order preferencing and the adverse-selection component of the spread

We first classify our sample stocks into five portfolios according to the market value of equity (MVE). Portfolio 1 contains stocks with the smallest MVE and Portfolio 5 contains stocks with the largest MVE. We then classify stocks in each MVE portfolio into five portfolios based on the adverse-selection component calculated from the method developed by Lin, Sanger, and Booth (LSB(i)). Portfolio 1 contains stocks with the smallest adverse-selection component and Portfolio 5 contains stocks with the largest adverse-selection component. For each LSB portfolio within each MVE group, we calculate the mean values of INT(i), NINS(i), and PREF(i), where 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, and PREF(i) = INT(i) + NINS(i). We also use stocks in each MVE portfolio to form five portfolios according to the adverse-selection component calculated from the method developed by Glosten and Harris (GH(i)) and calculate the mean values of INT(i), NINS(i), and PREF(i) for each GH portfolio. For brevity, we report only the results for MVE portfolios 1, 3, and 5. Similarly, within each MVE portfolio, we report the values of INT(i), NINS(i), and PREF(i) for LSB and GH portfolios 1, 3, and 5. Numbers in parentheses are t-statistics testing the equality of mean values between Portfolio 5 and Portfolio 1

		Portfolios based on the LSB model				Portfolios based on the GH model			
MVE Portfolio	Adverse-selection cost portfolio	LSB(i)	INT(i) (in %)	NINS(i) (in %)	PREF(i) (in %)	GH(i)	INT(i) (in %)	NINS(i) (in%)	PREF(i) (in%)
				November 2000					
Portfolio 1	Portfolio 1	0.0551	28.36	39.76	68.12	0.0438	31.14	38.16	69.30
	Portfolio 3	0.1142	27.52	36.22	63.74	0.0100	26.25	35.82	62.07
	Portfolio 5	0.2225	23.10	35.20	58.30	0.2171	24.32	34.36	58.68
	Portfolio 5 – 1	0.1674**	−5.26**	−4.56**	−9.82**	0.1733**	−6.82**	−3.80*	−10.62**
		(25.54)	(−3.20)	(−3.66)	(−6.87)	(18.35)	(−3.93)	(−3.80)	(−6.91)
Portfolio 3	Portfolio 1	0.0582	27.56	39.68	67.24	0.0376	26.96	39.00	65.96
	Portfolio 3	0.1141	25.39	37.25	62.64	0.0987	26.07	37.27	63.34
	Portfolio 5	0.2363	22.63	35.03	57.66	0.2091	22.12	35.59	57.71
	Portfolio 5 – 1	0.1781**	−4.93**	−4.65**	−9.58**	0.1715**	−4.84**	−3.41**	−8.25**
		(21.49)	(−3.07)	(−3.60)	(−6.51)	(25.79)	(−2.99)	(−2.89)	(−5.33)
Portfolio 5	Portfolio 1	0.0530	27.18	41.02	68.20	0.0390	27.90	41.60	69.50
	Portfolio 3	0.1110	26.85	36.91	63.76	0.0972	26.36	36.55	62.91
	Portfolio 5	0.2198	22.42	34.10	56.52	0.2215	22.69	34.99	57.68
	Portfolio 5 – 1	0.1668**	−4.76**	−6.92**	−11.68**	0.1825**	−5.21**	−6.61**	−11.82**
		(18.07)	(−3.10)	(−3.95)	(−9.01)	(18.32)	(−3.29)	(−3.68)	(−8.31)

		June 2001							
Portfolio 1	Portfolio 1	0.0550	27.61	41.43	69.04	0.0384	28.26	39.97	68.23
	Portfolio 3	0.1141	25.28	38.10	63.38	0.1015	25.74	37.44	63.18
	Portfolio 5	0.2467	21.39	37.09	58.48	0.2665	22.41	35.37	57.78
	Portfolio 5 – 1	0.1967**	−6.22**	−4.34**	−10.56**	0.2281**	−5.85**	−4.60**	−10.45**
		(17.02)	(−3.66)	(−3.09)	(−7.59)	(22.39)	(−3.38)	(−3.20)	(−6.94)
Portfolio 3	Portfolio 1	0.0651	27.04	39.88	66.92	0.0474	27.12	39.77	66.89
	Portfolio 3	0.1167	25.74	37.41	63.15	0.1024	25.81	38.19	64.00
	Portfolio 5	0.2464	22.25	34.70	56.95	0.2429	22.91	33.82	56.73
	Portfolio 5 – 1	0.1813**	−4.79**	−5.18**	−9.97**	0.1955**	−4.21**	−5.95**	−10.16**
		(22.94)	(−3.14)	(−3.62)	(−8.13)	(16.96)	(−2.84)	(−4.03)	(−7.51)
Portfolio 5	Portfolio 1	0.0545	25.54	43.02	68.56	0.0469	26.01	42.21	68.22
	Portfolio 3	0.1156	23.52	38.33	61.85	0.1019	22.10	38.95	61.05
	Portfolio 5	0.2558	20.16	36.82	56.98	0.2441	21.71	36.76	58.47
	Portfolio 5 – 1	0.2013**	−5.38**	−6.20**	−11.58**	0.1972**	−4.30**	−5.45**	−9.75**
		(19.18)	(−3.35)	(−4.14)	(−9.77)	(17.74)	(−2.86)	(−3.66)	(−6.82)

**Significant at the 1% level.

examine whether the negative relation between order preferencing and the adverse-selection cost/risk shown in the previous section remains intact when we control for the effects of other variables that might influence order preferencing.

Chung, Chuwonganant, and McCormick (2004) show that the proportion of trades executed by dealers not quoting at the inside is higher for stocks with smaller trade sizes, higher share prices, lower trading volumes, and higher H-INDEX. The authors also show that the proportion of internalized trades is higher for stocks with larger trade sizes, larger trading volumes, lower share prices, larger spreads, and lower H-INDEX. They explained these results based on the fact that (1) common qualification for preferencing contracts includes small orders and orders on stocks with a certain minimum price; (2) institutional brokers have large internalized volumes and their trade sizes tend to be greater than those of wholesalers or wirehouses;[11] (3) institutional brokers are more likely to trade high volume stocks than low volume stocks; (4) stocks with concentrated market shares have lower quote-based competition; and (5) stocks with larger non-inside volumes have smaller internalized volumes and vice versa.

Chung, Chuwonganant, and McCormick (2004) also show that the extent of order preferencing depends on dealer types. Institutional brokers frequently act as both dealer and broker for their clients, who are primarily large institutions. Consequently, institutional brokers 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. Preferencing arrangements are more frequently made with wholesalers (than with institutional brokers and wirehouses) because wholesalers tend to specialize in small retail orders. Chung, Chuwonganant, and McCormick show that the proportion of internalized trades is greater for institutional brokers and wirehouses, but smaller for wholesalers. Conversely, the proportion of non-inside trades is smaller for institutional brokers and wirehouses, but greater for wholesalers.

Based on these considerations, we employ the following regression model to examine how order preferencing is related to the adverse-selection cost and the probability of information-based trading after controlling for the effects of stock characteristics/dealer types:

$$\begin{aligned}
\text{INT}(i, j),\ \ \text{NINS}(i, j), \text{or PREF}(i, j) &= \beta_0 + \beta_1 \text{LSB}(i)(\text{or GH}(i)) + \beta_2 \text{PIN}(i) \\
&+ \beta_3 \log(\text{PRICE}(i)) + \beta_4 \log(\text{NTRADE}(i)) + \beta_5 \log(\text{TSIZE}(i)) \\
&+ \beta_6 \log(\text{MVE}(i)) + \beta_7 \text{H-INDEX}(i) + \beta_8 \log(\text{SPRD}(i)) \\
&+ \beta_9 \text{DUMIB}(j) + \beta_{10} \text{DUMWH}(j) + \beta_{11} \text{DUMWS}(j) + \varepsilon(i, j); \qquad (8)
\end{aligned}$$

where $\text{INT}(i, j)$ is the ratio of stock i's internalized volume routed to dealer j to its total volume executed by dealer j, $\text{NINS}(i, j)$ is the ratio of stock i's non-inside volume executed by dealer j to its total volume executed by dealer j, $\text{PREF}(i, j) = \text{INT}(i, j) + \text{NINS}(i, j)$, $\text{LSB}(i)$ and $\text{GH}(i)$ denote the adverse-selection components of stock i calculated from the methods developed by Lin, Sanger, and

[11]Examples of institutional brokers are Donaldson, Lufkin and Jenrette, and Goldman Sachs. Examples of wholesalers are Kinght/Trimark Securities and Mayer & Schweitzer. Wirehouses are integrated retail and full discount brokers, such as Dean Witter Reynolds and Merrill Lynch.

Booth (1995) and Glosten and Harris (1988), respectively, PIN(i) is the probability of information-based trading for stock i, 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, MVE(i) is the market value of equity of stock i, H-INDEX(i) is the Herfindahl-index, SPRD(i) is the quoted spread of stock i, 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 6 shows the regression results. The results show that all three measures of order preferencing, INT(i, j), NINS(i, j), and PREF(i, j), are significantly and negatively related to both measures (i.e., LSB(i) and GH(i)) of the adverse-selection component of the spread. Similarly, the regression results show that all three measures of order preferencing decrease with the probability of information-based trading. These results are consistent with our conjecture and support the idea that brokers route orders with low adverse-selection costs/risks to affiliated dealers (i.e., internalization) or dealers with payment for order flow arrangements.

The results show that the NINS(i, j) is significantly and positively related to share price and the Herfindahl-index and negatively to trade size and number of trades. 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 find that both INT(i, j) and NINS(i, j) are positively related to the spread. 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, NINS(i, j) is negatively related to the dummy variables for institutional brokers and wirehouses, but positively related to the dummy variable for wholesalers.[12]

These results are all consistent with the findings of Chung, Chuwonganant, and McCormick (2004) and support the idea that the extent of order preferencing varies with stock characteristics and dealer types.[13] The positive relation between INT(i, j) and the spread is also consistent with the finding of Chung, Chuwonganant, and McCormick (2004) that there exists a positive and bi-directional relation between the spread and the proportion of internalized trades.[14]

[12]Note that while the regression model explains 25.6% (24.4%) of variation in NINS(i, j) (INT(i, j)), it explains only 2.7% of variation in PREF(i, j). The lower explanatory power for the PREF(i, j) model is largely due to the fact that our common explanatory variables have opposite effects on the two components (i.e., NINS(i, j) and INT(i, j)) of PREF(i, j).

[13]To the extent that INT increases with the adverse-selection component of the spread, the size of non-adverse selection components (order processing cost, inventory cost, and dealer rent) of the spread is likely to be positively related to INT. To confirm this, we first obtain the non-adverse selection component by subtracting the adverse selection component from the total spread. We then regress the non-adverse selection component on INT and stock characteristics that are believed to determine the spread. Consistent with our expectation, we find that the non-adverse selection component is significantly and positively related to INT.

[14]Chung, Chuwonganant, and McCormick (2004) hold that the positive relation between INT(i, j) and the spread may be driven by the fact that brokers have an incentive to route large-spread stocks to their affiliated dealers. They also note that investors sometimes value immediate liquidity more than price improvement in stocks with higher internalization. That is, internalization can be high in high-spread 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. The authors also predict a positive relation between NINS(i, j) and the spread.

Table 6 Order preferencing as a function of adverse-selection cost and the probability of information-based trading

We employ the following regression model to examine how order preferencing is related to the adverse-selection cost and the probability of information-based trading after controlling for the effects of stock attributes/dealer types:

$$\mathrm{INT}(i,j),\ \mathrm{NINS}(i,j),\text{ or }\mathrm{PREF}(i,j) = \beta_0 + \beta_1\mathrm{LSB}(i)(\text{or }\mathrm{GH}(i)) + \beta_2\mathrm{PIN}(i) + \beta_3\log(\mathrm{PRICE}(i)) + \beta_4\log(\mathrm{NTRADE}(i)) + \beta_5\log(\mathrm{TSIZE}(i))$$
$$+\beta_6\log(\mathrm{MVE}(i)) + \beta_7\text{H-INDEX}(i) + \beta_8\log(\mathrm{SPRD}(i)) + \beta_9\mathrm{DUMIB}(j) + \beta_{10}\mathrm{DUMWH}(j) + \beta_{11}\mathrm{DUMWS}(j) + \varepsilon(i,j);$$

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, NINS(i, j) is the ratio of stock i's non-inside volume executed by dealer j to its total volume executed by dealer j, PREF(i, j) = INT(i, j) + NINS(i, j), LSB(i) and GH(i) denote the adverse-selection components of stock i calculated from the methods developed by Lin, Sanger, and Booth (1995) and Glosten and Harris (1988), respectively, PIN(i) is the probability of information-based trading for stock i, 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, MVE(i) is the market value of equity of stock i, H-INDEX(i) is the Herfindahl-index of stock i, SPRD(i) is the spread of stock i, 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). Numbers in parentheses are t-statistics

	INT(i, j)	NINS(i, j)	PREF(i, j)	INT(i, j)	NINS(i, j)	PREF(i, j)
			November 2000			
Intercept	0.0674**	0.5937**	0.6611**	0.0718**	0.5798**	0.6516**
	(2.70)	(23.39)	(31.88)	(2.77)	(21.98)	(30.16)
LSB(i)	−0.1252**	−0.0938**	−0.2190**			
	(−5.08)	(−3.74)	(−8.02)			
GH(i)				−0.1302**	−0.1160**	−0.2462**
				(−4.51)	(−3.95)	(−7.59)
PIN(i)	−0.0232**	−0.0338**	−0.0570**	−0.0321**	−0.0335**	−0.0656**
	(−3.03)	(−4.33)	(−6.35)	(−4.19)	(−4.30)	(−7.43)

log(PRICE(i))	−0.0221**	0.0144**	−0.0077*	−0.0224**	0.0145**	−0.0079*
	(−4.79)	(3.06)	(−2.17)	(−4.83)	(3.08)	(−2.16)
log(NTRADE(i)))	0.0049**	−0.0074**	−0.0025**	0.0045**	−0.0063**	−0.0018*
	(2.97)	(−4.84)	(−2.59)	(2.93)	(−4.12)	(−2.19)
log(TSIZE(i))	0.0188**	−0.0128**	0.0060*	0.0181	−0.0120	0.0061*
	(5.11)	(−3.41)	(2.05)	(4.90)	(−3.19)	(2.07)
log(MVE(i))	0.0056**	−0.0085**	−0.0029*	0.0058**	−0.0088**	−0.0030*
	(3.06)	(−4.62)	(−2.38)	(3.17)	(−4.71)	(−2.26)
H-INDEX(i)/	−0.0700**	0.1005**	0.0305*	−0.0683**	0.0992**	0.0309*
10,000	(−3.55)	(5.01)	(2.16)	(−3.47)	(4.95)	(2.12)
log(SPRD(i))	0.0126**	0.0164**	0.0290**	0.0119**	0.0158**	0.0277**
	(2.80)	(3.58)	(5.09)	(2.77)	(3.44)	(4.92)
DUMIB(i)	0.1761**	−0.1542**	0.0219**	0.1760**	−0.1542**	0.0218**
	(39.91)	(−34.33)	(6.22)	(39.90)	(−34.34)	(6.20)
DUMWH(i)	0.2731**	−0.1933**	0.0798**	0.2729**	−0.1932**	0.0797**
	(49.33)	(−34.32)	(18.00)	(49.31)	(−34.31)	(18.00)
DUMWS(i)	−0.1244**	0.1525**	0.0281**	−0.1245**	0.1525**	0.0280**
	(−47.30)	(57.05)	(11.55)	(−47.33)	(57.06)	(11.59)
F-value	1,238.87**	873.77**	56.54**	1,238.76**	874.15**	56.42**
Adjusted R^2	0.281	0.216	0.017	0.281	0.217	0.017
			June 2001			
Intercept	−0.4295**	0.9769**	0.5474**	−0.4352**	0.9639**	0.5287**
	(−19.39)	(42.08)	(30.17)	(−19.21)	(40.60)	(28.48)
LSB(i)	−0.1217**	−0.0832	−0.2049**			
	(−6.41)	(−4.18)	(−9.01)			
GH(i)				−0.1146**	−0.0895**	−0.2041**
				(−5.53)	(−4.12)	(−8.14)

(*Continue on next page.*)

Table 6 (*continued.*)

	INT(i, j)	NINS(i, j)	PREF(i, j)	INT(i, j)	NINS(i, j)	PREF(i, j)
PIN(i)	−0.0245**	−0.0228**	−0.0474**	−0.0222**	−0.0229**	−0.0451**
	(−3.54)	(−3.15)	(−5.74)	(−3.21)	(−3.16)	(5.56)
log(PRICE(i))	−0.0320**	0.020688	−0.0114**	−0.0325**	0.0200**	−0.0125**
	(−7.85)	(4.81)	(−3.42)	(−7.95)	(4.67)	(−3.73)
log(NTRADE(i))	0.0044**	−0.0141**	−0.0097**	0.0047**	−0.0136**	−0.0089**
	(3.25)	(−9.62)	(−7.68)	(3.37)	(−9.29)	(−6.98)
log(TSIZE(i))	0.0843**	−0.0662**	0.0181**	0.0837**	−0.0654**	0.0183**
	(25.36)	(−19.00)	(6.64)	(25.05)	(−18.69)	(6.66)
log(MVE(i))	0.0043**	−0.0065**	−0.0022*	0.0047**	−0.0064**	−0.0017*
	(2.71)	(−3.85)	(−2.08)	(2.91)	(−3.75)	(−2.04)
H-INDEX(i)/	−0.0529**	0.0899**	0.0370**	−0.0502**	0.0883**	0.0381**
10,000	(−3.02)	(4.90)	(2.58)	(−2.87)	(4.81)	(2.66)
log(SPRD(i))	0.0158**	0.0108**	0.0266**	0.0157**	0.0112**	0.0269**
	(4.20)	(2.75)	(5.95)	(4.18)	(2.86)	(6.04)
DUMIB(i)	0.1614**	−0.1419**	0.0195**	0.1612**	−0.1418**	0.0194**
	(38.90)	(−32.64)	(5.72)	(38.86)	(−32.63)	(5.68)
DUMWH(i)	0.1983**	−0.1405**	0.0578**	0.1982**	−0.1404**	0.0578**
	(38.84)	(−26.24)	(13.83)	(38.81)	(−26.23)	(13.82)
DUMWS(i)	−0.1281**	0.1440**	0.0159**	−0.1282**	0.1441**	0.0159**
	(−52.66)	(56.46)	(7.95)	(−52.71)	(56.48)	(7.91)
F-value	1,166.19**	1,246.78**	100.62**	1,164.62**	1,247.41**	99.12**
Adjusted R^2	0.244	0.256	0.027	0.243	0.256	0.026

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

To determine whether the relation between order preferencing and the adverse selection cost/risk differs between high- and low-activity stocks, we cluster our study sample of stocks into three portfolios according to the number of trades. We then replicate Table 6 using only the low-activity stocks (i.e., Portfolio 1) and the high-activity stocks (i.e., Portfolio 3), respectively. We find that the results are qualitatively similar between the two groups, although the relation between order preferencing and the adverse-selection cost/risk is slightly stronger for the low-activity stocks.[15]

5 Summary and concluding remarks

Dealers expose themselves to informed traders when they offer firm quotes that any trader can take. Hence, dealers quote wider spreads than they would quote if they traded only with uninformed traders. Uninformed traders who trade at quoted prices therefore are indirectly hurt by the presence of informed traders because of the wider spreads that they must pay.

Payments for order flow can benefit uninformed traders when dealers can determine who they are. If dealers can skim the cream of the order flow through order preferencing arrangements, brokers are likely to charge uninformed traders lower commissions in compensation for the high spreads that they pay. That is, brokers are likely to route only those orders from uninformed traders to dealers with payments for order flow agreements and pass on the order flow payments to uninformed traders in the form of lower commissions. According to this theory, we expect that the extent of order preferencing decreases with the adverse-selection cost and the probability of information-based trading. Our empirical results are generally consistent with this prediction.

Although prior research shows that stocks with higher levels of preferenced orders have wider spreads, the ultimate effect of order preferencing on investor welfare is not so obvious. To the extent that execution costs of uninformed traders are reduced by lower commissions, the net effect of order preferencing on overall execution quality can be positive. In addition, preferencing may improve other dimensions of market quality, such as speed of execution and reliability. Although accurate quantification of these benefits is difficult and beyond the scope of the present study, the results of this study suggest that order preferencing could be an efficient way to deal with the problems associated with asymmetric information in the securities markets.

Acknowledgments The authors are solely responsible for the content. The authors thank two anonymous referees, Frank Hatheway, and seminar participants at the Securities and Exchange Commission (SEC) for many valuable comments. The authors are grateful to James Lastoskie for his assistance in collecting and understanding the data. The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This article expresses the authors' views and does not necessarily reflect those of the Commission, the Commissioners, or other members of the staff.

References

Barclay M, Hendershott T, McCormick DT (2003) Information and trading on Electronic Communications Networks. Journal of Finance 58:2637–2665

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

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, 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, 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 (2003) Issues in assessing trade execution costs. Journal of Financial Markets 6:233–257
Bloomfield R, O'Hara M (1998) Does order preferencing matter? Journal of Financial Economics 50:3–37
Chung KH, Chuwonganant C, McCormick DT (2004) Order preferencing and market quality on NASDAQ before and after decimalization. Journal of Financial Economics 71: 581–612
Easley D, Hvidkjaer S, O'Hara M (2002) Is information risk a determinant of asset returns? Journal of Finance 57:2185–2221
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
Easley D, Kiefer N, O'Hara M (1997a) One day in the life of a very common stock. Review of Financial Studies 10:805–835
Easley D, Kiefer N, O'Hara M (1997b) The information content of the trading process. Journal of Empirical Finance 4:159–186
Glosten L, Harris L (1988) Estimating the components of the bid-ask spread. Journal of Financial Economics 21:123–142
Heidle H, Huang R (2002) Information-based trading in dealer and auction markets: An analysis of exchange listings. Journal of Financial and Quantitative Analysis 37:391–424
Hansch O, Naik N, Viswanathan S (1999) Preferencing, internalization, best execution, and dealer profits. Journal of Finance 54:1799–1828
Huang R (2002) The quality of ECN and NASDAQ market maker quotes. Journal of Finance 57:1285–1319
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
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
Lin J, Sanger G, Booth G (1995) Trade size and components of the bid-ask spreads. Review of Financial Studies 8:1153–1183
Parlour C, Rajan U (2003) Payment for order flow. Journal of Financial Economics 68:379–411
Peterson M, Sirri E (2003) Order preferencing and market quality on US equity exchanges. Review of Financial Studies 16:385–415