ELSEVIER

Journal of Financial Markets 4 (2001) 85–112

Journal of FINANCIAL MARKETS

www.elsevier.nl/locate/econbase

The potential for clientele pricing when making markets in financial securities☆

Robert Battalio[a], Robert Jennings[b,*], Jamie Selway[c]

[a]*University of Notre Dame, USA*

[b]*New York Stock Exchange, International and Research, 17th Floor, 11 Wall Street, New York, NY 10005, USA*

[c]*Archipelago, Inc., USA*

Received 1 May 2000; received in revised form 1 July 2000

Abstract

Benveniste et al. J. Financial Econom. 32 (1992) 61–86 argue that repeatedly dealing with the same brokers allows market makers to know when brokers exploit private information. This suggests that broker identity may allow market makers to differentiate between customers when pricing market-making services even when market makers can provide separate quoted prices for each order size. Estimates of a major Nasdaq dealer's gross trading revenue vary substantively among routing brokers after controlling for order size. Furthermore, these differences exhibit a degree of stability over time. This suggests that market makers may effectively enforce clientele pricing schedules in a world where security prices are quoted without a minimum price variation and the limit order

☆We thank Knight-Trimark Group, Inc. for providing data and Jeff Bacidore, Shane Corwin, Jenny Drake, Larry Harris, Leslie Marx, David C. Smith, Jeff Smith, H. Douglas Witte, and participants at the 1999 Financial Management Association meeting, the 1999 Nasdaq-Notre Dame Microstructure Conference, and the 2000 Western Finance Association meeting for comments on earlier versions of the paper. At the initiation of this project, Battalio was Visiting Economist and Selway was Associate Director of Research at the National Association of Security Dealers, Inc.. Opinions expressed herein are those of the authors and not necessarily those of Knight-Trimark Group, Inc., the National Association of Security Dealers, Inc., the member, officers, or directors of the New York Stock Exchange, Inc., or Archipelago, Inc.

* Corresponding author. Tel.: + 212-656-5656; fax: + 1-212-656-2401.

E-mail address: bjennings@nyse.com (R. Jennings).

PII: S1386-4181(00)00017-3

book is available to all market participants.

Keywords: Microstructure; Payment for order flow; Decimalization; Internalization; Price discrimination

1. Introduction

The decimalization of domestic equity security prices is scheduled to begin in earnest on August 28, 2000 and be fully implemented by April 2001.[1] In addition, major markets may publicly display their limit order books (e.g., see NYSE (2000) for a discussion of the NYSE's proposal). Should these two events occur, market makers will be able to post prices as a function of order size, i.e., market makers could post nearly continuous supply and demand schedules. Consistently with financial theory (e.g., Easley and O'Hara, 1987), firms such as Bernard L. Madoff Investment Securities and Knight Securities currently use order size to partition orders into those less likely to be motivated by information and, therefore, more profitable to the market maker. Proponents of the display of the limit order book and decimalization argue that these changes will consolidate order flow by eliminating a market maker's ability to use order size to selectively execute orders. As noted in Harris (1993), however, this claim depends crucially on the assumption that order size is the only observable security characteristic related to order flow profitability in an economically significant way.

Extant research suggests there may be opportunities for clientele pricing even when quotes are posted as continuous functions of order size. Angel (1994), Knez and Ready (1996), Ready (1999), and Harris and Panchapagesan (1999) find evidence consistent with the claim that New York Stock Exchange specialists exploit advantages in location/information to strategically interact with order flow. Harris and Schultz (1998) present empirical evidence suggesting orders placed by day traders through Nasdaq's Small Order Execution System (SOES) generate less market making revenue than other comparably-sized orders. Finally, Affleck-Graves et al. (1994), Easley et al. (1996), Lin et al. (1998), and Bessembinder (1999) provide evidence that there may be order characteristics other than order size useful in distinguishing between profitable and unprofitable order flow. Theoretically, Fishman and Longstaff (1992) and Benveniste et al. (1992) suggest broker identity may be one characteristic related to order flow profitability and Massimbe and Phelps (1994) and Franke and Hess

[1] See SEC (2000) and Bloomberg News (2000). Throughout the paper, we use decimalization to refer to the elimination of the minimum price variation. Datek Online began allowing customers to trade in decimals on July 3, 2000.

(1995) provide empirical support for their claims. We use proprietary data from a major Nasdaq market maker to show that, after controlling for order size and the security's price level, price volatility, and trading volume, the routing broker's identity helps explain cross-sectional differences in estimates of market making revenue.[2]

Revenue differences must be predictable and economically significant for dealers to charge differentially for their services. We examine the accuracy of a naive prediction model extrapolating a broker's past levels of gross revenue into the future. Although there is randomness in relative levels of our estimates of gross trading revenue, rank-order correlation coefficients generally are positive. In addition, we find order flow associated with extremely high or extremely low revenue levels tends to remain at that relative level over an extended period of time. Inter-broker differences in dealer revenue are as large as four cents per share. These results suggest there will be opportunities for clientele pricing even when prices are expressed as a function of order size.

In the following section, we discuss our data and describe orders sent to a major Nasdaq market maker, Knight Securities. Section 3 presents estimated gross market maker revenue from different brokers' order flow. During late-1996, gross revenue varied by more than \$0.04 per share across brokers. With an average order size of 377 shares, this is an estimated difference in revenue of \$15 per order. Section 4 presents an econometric model of the revenue from Knight's order flow and shows that conditioning on broker identity helps explain revenue levels after controlling for order and stock characteristics. In Section 5, we examine the time-series stability of our trading-revenue metric. Monthly rank-order correlation coefficients of revenue estimates are less than one, but generally are positive. Furthermore, statistically significant differences in estimated revenue in one month typically are maintained in the following month. We find that this persistence extends over a long period of time, implying that market makers can predict order-flow profitability based on at least one dimension on which they do not quote prices (i.e., broker identity). This allows the possibility of clientele pricing of dealer services. The final section concludes by discussing some implications of our findings.

2. Data description

To estimate the gross trading revenue associated with order flow, we require detailed order, trade, and quote data. We obtain proprietary audit-trail order

[2] The idea that information asymmetries may affect competitive equilibria is not new. In the context of insurance markets, Rothschild and Stiglitz (1976) state, that "the single price equilibrium of conventional competitive analysis was shown to be no longer viable" with even a "small amount of imperfect information" introduction.

data from Knight Securities and quote data from the National Association of Security Dealers, Inc. (NASD).

2.1. Knight Securities, L.P

Knight was founded on July 24, 1995 as a market maker in Nasdaq and other non-listed equity securities. It traded 93 million shares (a 0.9% market share) its first full month of existence. During February 2000, Knight traded 11.2 billion shares (a 21.7% market share), the most of all NASD dealers. Knight began as a consortium of 25 retail brokerage firms trying to recognize scale economies as they vertically integrated into market making.[3] The members in September, 1996, the beginning of our first sample period, and March 1998, the end of our second sample period, appear in Table 1.

Four of these firms (E*Trade, Waterhouse, Ameritrade, and Discover) are among the eight on-line brokers having the largest trading volume, with a combined market share of 32% of on-line trading (see Wall Street Journal, 1998). A particular correspondent may route Knight order flow from several retail brokers, possibly including its own retail brokerage unit (e.g., Ameritrade routes order flow from Accutrade, Ameritrade, Aufhauser, Ceres, and Ebroker during our sample period).

2.2. Data

We obtain order audit-trail data for each Knight order in the fourth quarter of 1996 (4Q96) and the first quarter of 1998 (1Q98).[4] We examine only orders routed directly to Knight, excluding orders Knight receives via Nasdaq's SOES and SelectNet and Knight's proprietary trades. Each record contains security name, order type (e.g., market or limit), an indication of which party initiates the order, order quantity, execution price and quantity, to-the-second order receipt time, to-the-minute execution time, and a code identifying the firm routing the order to Knight. Quotation data for the same periods come from the NASD. We initially examine the 1996 period.

We estimate the market-making revenue Knight earns from providing liquidity. The value of this liquidity (and, hence, the revenue associated with providing it) depends on the stock's trading characteristics. For example, if trading is infrequent, then the value of Knight's liquidity is greater than if natural buyers and sellers are plentiful. To control for cross-sectional differences in liquidity's value, we limit our sample to the stocks included in the Nasdaq-100 Index. These are the 100 largest Nasdaq stocks based on market capitalization.

[3] For details, see *Traders* (1996) and Knight's S-1 filing in June 1998. Brokerage firms routing Knight orders are referred to as corresponding brokers.

[4] We are unable to obtain data for December 11, 1996 because of technical problems.

Table 1
An alphabetically-arranged list of the brokerage-firm owners of Knight Securities.

Corresponding Broker[a]	Original Owner[b]	Owner as of 3/31/98[c]
Ameritrade[d]	Yes	Yes
BHC Securities Inc.	Yes	Yes
BHF Securities	Yes	Yes
Bidwell & Co.	Yes	Yes
Brown & Co.[e]	Yes	Yes
Burke, Christensen & Lewis Securities	Yes	Yes
Cowles Sabol & Co., Inc.	Yes	Yes
David A. Noyes & Co.	No	Yes
Direct Access Brokerage Services, Inc.	Yes	Yes
E*Trade Securities	Yes	Yes
Gruntal Financial Corp.	Yes	Yes
Hanifen Imhoff Clearing Corp.[f]	Yes	Yes
Howe Barnes Investments	Yes	Yes
International Correspondents Trading, Inc.	No	Yes
J. W. Charles Securities, Inc.	Yes	Yes
Josephthal & Co. Inc.	No	Yes
Lombard Institutional Brokerage[g]	Yes	Yes
Nathan & Lewis Securities, Inc.	No	Yes
Primevest Financial Services, Inc.	No	Yes
R. J. Forbes Group	Yes	Yes
R. P. Assignee Corp.[h]	Yes	Yes
R. P. R. Clearing Services	Yes	Yes
Richardson Greenshields	Yes	No
Sanders Morris Mundy	Yes	Yes
Scottsdale Securities	Yes	Yes
Southwest Securities	Yes	Yes
Stockcross	Yes	Yes
Thomas F. White & Co.	Yes	Yes
Van Kasper & Co.	Yes	Yes
Waterhouse Securities	Yes	Yes

[a]Corresponding brokers are brokers routing orders to Knight Securities, Inc.

[b]See Trader's Magazine, September 1996.

[c]See S – 1 filing for Knight Trimark Group, Inc.

[d]Formerly operated under the names of Accutrade, Aufhauser, Ceres, Ebroker.

[e]Brown & Co. was a subordinated note holder. It received payments for its order flow under a schedule similar to the original owners of Knight. Brown was granted the right to purchase an equity stake in the company in April 1996. This right was exercised at the closing of the initial public offering of Knight Trimark Group, Inc.

[f]Hanifen Imhoff Clearing Corp. currently operates as Fiserv Correspondent Services, Inc.

[g]Lombard Securities currently operates as Discover Brokerage.

[h]R. P. Assignee Corp. and R. P. R. Clearing Services currently operate as Dain Rauscher, Inc.

During 4Q96, Knight receives ten or more orders from 226 corresponding brokers, but the 25 initial consortium firms account for 91% of the order flow. In these months, Knight receives 289,809 orders in Nasdaq-100 stocks. To be included in our study, an order must: (1) not be stop order, (2) be in a stock priced at $10.00 per share or greater, (3) arrive between 9:30 a.m. and 4:00 p.m., (4) arrive when the bid-ask spread is positive (i.e., the market is not locked or crossed), and (5) be a market order. The first filter eliminates 14,246 orders that are effective only if the stock's price reaches a specified value. Because stocks priced at $10 per share and greater can be quoted only in $0.125 increments and stocks priced below $10 can be quoted in $0.03125 increments during this period, comparing trading profits between these groups may be misleading. Our data contain 36,148 orders in stocks priced below $10 per share. We eliminate the 15,405 orders sent outside of trading hours because we require benchmark quotes. We also eliminate 198 orders arriving when the market is locked or crossed. Market orders demand liquidity, which Knight provides at a price. Limit orders supply liquidity and earn the price of liquidity for the investor submitting the order, denying Knight the opportunity to interact with an order demanding liquidity. Because differences in order aggressiveness suggest differences in dealer revenue, we focus on market orders in the remainder of the paper. This does not imply that Knight does not value limit orders, only that measuring this value requires a different approach than measuring a market order's value. There are 129,699 limit orders, leaving 94,113 Nasdaq-100 market orders in our sample.

Table 2 describes Knight's Nasdaq-100 market orders by correspondent. The ten brokers sending Knight the most market orders are given identifying codes (B1 through B10).

Knight receives 47% of its orders from 3 correspondents and 62% from five. The volume-weighted average price (VWAP) is the market value of all sample trades divided by the number of shares traded. The fact the VWAP varies from under $40 per share to almost $72 among brokers suggests that different brokers route Knight orders in different stocks. Average order size also varies considerably by broker, from 203 shares to 811 shares. Overall, 19% of Knight's market orders are odd lots (less than 100 shares), 68% are round lots or partial round lots for fewer than 1000 shares, 8% are for exactly 1000 shares, and 5% are orders for more than 1000 shares. During this period, 1000 shares is the minimum-sized quote allowed for Nasdaq-100 securities. As such, it represents the largest order size that should receive immediate execution. The percentage of market orders for exactly 1000 shares varies across routing broker from less than 3% to over 16%. Differences among broker's order flow in traits such as order size and security suggest across-broker differences in market-making revenue. In the following section, we demonstrate that this is true.

3. Descriptive statistics of estimated gross trading revenue

We use the realized liquidity premium (RLP) as a proxy for Knight's per share gross trading revenue. The RLP is the distance between the trade price and the

Table 2
Descriptive statistics on price and size of market orders in Nasdaq-100 securities received by Knight between October and December 1996[a]

Broker[b]	Number of Market Orders	VWAP[c]	Mean order size (shares)	1–99 Shares	100–499 Shares	500–999 Shares	1000 Shares	> 1000 Shares
B1	22,402	$47.39	209.00	22.62%	66.87%	6.71%	2.66%	1.13%
B2	11,733	52.18	447.00	14.81	50.92	14.48	15.01	4.78
B3	10,549	49.07	414.00	14.77	56.24	13.73	10.45	4.82
B4	6,626	42.55	402.00	17.07	56.01	12.48	8.99	5.45
B5	7,148	71.84	203.00	43.06	45.73	6.44	2.95	1.82
B6	5,602	39.68	353.00	17.86	62.66	9.68	5.57	4.23
B7	5,061	65.22	811.00	5.87	43.76	13.75	16.04	20.58
B8	4,707	53.21	257.00	22.75	59.27	10.69	5.93	1.36
B9	2,620	50.58	350.00	21.15	54.89	11.30	8.44	4.24
B10	2,264	51.85	483.00	21.82	50.75	12.63	7.11	7.69
All firms	94,113	52.05	377.00	19.39	56.73	11.20	7.97	4.72

[a] Columns 4 through 8 may not total to 1.00 on each row because of rounding.
[b] Brokers are coded B1–B10 based on the number of market orders routed to Knight in the fourth quarter of 1996.
[c] VWAP = volume weighted average price per share = total dollar volume divided by total share volume.

inside quoted bid–ask spread's midpoint five minutes after the trade time.[5] Assuming that the spread midpoint is a good estimate of the stock's value and that the liquidity provider unwinds trades (or computes unrecognized gains/losses) at the spread's midpoint an average of five minutes after executing customer orders, the RLP proxies the liquidity providers's revenue. A buy order's RLP is the difference between the trade price and the midpoint of the inside bid-ask spread prevailing five minutes after the order executes. For sell orders, the RLP is the difference between the inside spread's midpoint five minutes after the trade time and the order's execution price. The formal definition of the realized liquidity premium is,

$$\text{RLP} = \text{I} \times (\text{transaction price} - \text{the bid–ask spread's midpoint five minutes after the trade time}), \tag{1}$$

where I is $+1$ (-1) if the trade is initiated by a buyer (seller). For example, suppose that the submission-time bid and ask prices are \$20.00 and \$20.125, respectively, and that orders execute at quoted prices. If quotes are unchanged five minutes after the execution time, then the benchmark spread's midpoint is \$20.0625 and the buy-order's (sell-order's) RLP is \$0.0625 = \$20.125-\$20.0625 (= \$20.0625-\$20.00). The bid/ask price may change between the time the order executes and five minutes later. Suppose that each side of the quote changes by \$0.125. If the quote increases, then the bid moves to \$20.125 and the ask to \$20.25. The RLP for the buy (sell) order is -\$0.0625 = \$20.125-\$20.1875 (\$0.1875 = \$20.1875-\$20.00). In the case where the quoted prices decrease, the quote becomes \$19.875 to \$20.00 and the buy (sell) order's RLP is \$0.1875 = \$20.125-\$19.9375 (− \$0.0625 = \$19.9375-\$20.00). If the bid remains the same five minutes after the execution time but the ask increases by \$0.125 during this interval, then the buy (sell) order's RLP is \$0.0000 (\$0.125).[6]

Suppose, for example, that odd-lot trades are uninformed and that Knight executes these orders at the submission-time posted quotes. We anticipate that an odd-lot order's RLP equals the submission-time quoted half-spread because we expect no systematic movement in security prices after uninformed traders place orders. If Knight's informed customers successfully anticipate short-term

---

[5] The inside spread is the highest bid price and lowest ask price across all market makers quoting a stock.

[6] Lee(1993) provides a more complete discussion of the liquidity premium concept. Huang and Stoll (1996) discuss the realized half-spread. If we had to-the-second execution time, then we would simply look at the quote 300 seconds after execution. As in Huang and Stoll (1996), we find that RLP are not particularly sensitive to the time we allow to lapse before obtaining a benchmark quote, times of from three to 15 minutes give similar results. Sofianos (1995) suggests a more accurate methodology for estimating dealer gross trading revenues. Our data do not include the inventory positions necessary to replicate his calculations.

price movements, however, then Knight buys (sells) just prior to a price decrease (increase). This suggests that Knight cannot realize the quoted half-spread as trading revenue. For small spread changes in the customer's favor, Knight's revenue is less than the quoted half-spread. For substantial changes, Knight may unwind the position at a loss (generating a negative RLP). Thus, when Knight deals with "informed" customers, we anticipate systematically lower RLP than when trading with uninformed investors. Given findings in Easley and O'Hara (1987) and Harris and Schultz (1998), larger orders and 1000-share orders may provide Knight less revenue than other orders.

Table 3 reports the RLP and the number of shares included in its calculation for sample orders received in \$0.125-spread (Panel A) and \$0.25-spread (Panel B) markets conditional on order size from the ten correspondents sending the most market orders to Knight during our sample period.

Table 3
Realized liquidity premiums for market orders in Nasdaq-100 securities routed to Knight during the fourth quarter of 1996.[a]

	Order size (shares)						
Broker[b]	< 100	100–499	500–999	1000	1001–2499	2500–5000	All orders
Panel A: Realized liquidity premiums for market orders submitted in \$0.125-spread markets							
B1	\$0.0672*	**\$0.0472***	**\$0.0300***	\$0.0288	\$0.0333	\$0.0748*	**\$0.0428***
	134	1,492	518	397	225	140	2,906
B2	\$0.0598*	**\$0.0394***	**\$0.0348***	**−\$0.0081**	**\$0.0187**	\$0.0383	**\$0.0185***
	42	700	601	1,239	499	340	3,421
B3	\$0.0643*	\$0.0595*	\$0.0665*	\$0.0499*	**\$0.0240**	\$0.0437	\$0.0514*
	39	660	489	672	335	382	2,577
B4	\$0.0793*	**\$0.0484***	**\$0.0663**	**\$0.0530**	**\$0.0412**	**\$0.0928**	**\$0.0586***
	23	377	283	369	256	227	1,535
B5	\$0.0746*	\$0.0696*	\$0.0606*	**−\$0.0355**	\$0.0158	\$0.0376	\$0.0434
	82	380	180	146	129	86	1,003
B6	\$0.0548*	\$0.0600*	\$0.0578*	\$0.0407*	\$0.0541	\$0.0710*	\$0.0564*
	24	357	181	185	180	122	1,049
B7	\$0.0908	\$0.0555	\$0.0496	\$0.0110	\$0.0478	\$0.0481	\$0.0412
	8	277	243	525	893	507	2,453
B8	\$0.0771*	\$0.0742*	\$0.0461*	**−\$0.0333**	**−\$0.0073**	\$0.0091	\$0.0333
	28	310	162	192	65	24	781
B9	\$0.0520*	\$0.0560*	\$0.0762*	\$0.0091	− \$0.0025	\$0.1208*	\$0.0408
	12	159	90	130	89	28	508
B10	\$0.0602*	\$0.0590*	\$0.0717*	\$0.0335	\$0.1135*	\$0.0908	\$0.0747*
	14	139	105	110	128	139	635
ALL	\$0.0668*	**\$0.0546***	**\$0.0454***	**\$0.0167***	**\$0.0316***	\$0.0620*	**\$0.0412***
	461	5,775	3,610	4,857	3,528	2,356	20,587
χ^2 Test for RLP	536						
differences over	40						
order size[c]	0.001						

Table 3 (*continued*)

Broker[b]	Order size (shares) < 100	100–499	500–999	1000	1001–2499	2500–5000	All orders
Panel B: Realized liquidity premiums for market orders submitted in $0.25-spread markets							
B1	$0.1168*	**$0.1132***	$0.1162*	$0.0827*	$0.1052*	$0.0714*	**$0.1102***
	63	717	241	148	71	45	1,285
B2	$0.1230*	**$0.0949***	**$0.0837***	**$0.0468***	**$0.0644***	$0.1574*	**$0.0786***
	24	339	267	389	198	92	1,309
B3	$0.1144*	**$0.1112***	$0.0945*	$0.0931*	$0.0453	$0.1269*	**$0.0968***
	21	323	250	309	147	145	1,195
B4	$0.1208*	$0.1110*	$0.1060*	$0.1113*	**$0.0667***	**$0.0312***	$0.0934*
	18	233	134	161	163	78	787
B5	$0.1213*	**$0.1004***	$0.0920*	$0.1019*	$0.1214	$0.2109	**$0.1050***
	21	148	61	45	26	8	309
B6	$0.1066*	$0.1191*	$0.1090*	$0.0848*	**$0.0827***	$0.0560	$0.0980*
	14	184	105	98	74	65	540
B7	$0.0725	**$0.0970***	$0.0961*	$0.1014*	**$0.0960***	$0.1254	$0.1016*
	3	112	113	210	500	172	1,110
B8	$0.1207*	$0.1251*	**$0.0533**	$0.0458	− $0.019	no	$0.0832
	11	139	75	55	19	orders	299
B9	$0.1311*	$0.1215*	$0.045*	$0.0471	$0.0057	$0.1924*	$0.1009*
	9	84	62	73	44	25	297
B10	$0.1195*	$0.1363*	$0.1048*	$0.1339*	$0.1185*	$0.1628*	$0.1345*
	5	59	40	40	43	66	253
ALL	**$0.1174***	**$0.1107***	**$0.0953***	**$0.0764***	**$0.0710***	$0.1000*	**$0.0927***
	221	2,840	1,725	1,926	1,663	850	9,225
χ^2 Test RLP	279						
differences over	40						
order sizes	0.001						

[a]The realized liquidity premium (RLP) is the difference between an order's execution price and the bid–ask spread's midpoint five minutes after order execution time. Calculations exclude orders in securities priced below $10, orders received outside normal hours or in locked or crossed markets, and orders lacking a valid quote when received or five minutes after the minute of execution. The entries in each cell are: (1) the share-weighted RLP and (2) the number of shares ('000 omitted). Bold numbers indicate that the mean RLP differs statistically from the quoted half-spread at the 0.01 significance level. An asterisk indicates that the mean RLP differs from zero at the 0.01 significance level. Two-tailed *t*-tests are used to examine both null hypothesis.

[b]Brokers are coded B1–B10 based on the number of market orders routed to Knight in the fourth quarter of 1996.

[c]The three numbers are: (1) the Chi-square test statistic; (2) the degrees of freedom; and, (3) the *p*-value.

The first entry in each cell of Table 3 is the share-weighted mean RLP and the second is the number of shares (in thousands) included in that broker:order-size category. For small order sizes, RLP generally are significantly positive (in a two-tailed test at the 0.01 level). Although it is uncommon for an individual broker's RLP to be significantly less than one-half the quoted spread (except for B2 and B4), the aggregate RLP typically is less than the submission-time quoted half-spread. This indicates that the quoted spread is a poor proxy for dealer revenue.

Average RLP vary by order size. Odd-lots' RLP equal the quoted half-spread in \$0.125-spread markets, suggesting that these orders are informationless. Consistently with the notion that larger orders are more likely to be motivated by information, RLP generally decrease with order size. RLP in the 2500–5000 share order-size category are large because orders of this size more frequently exceed quoted size allowing Knight to execute them at prices outside the order-receipt-time quoted spread. Consistently with the notion that they are used extensively by aggressive traders (see Harris and Schultz, 1998), 1000-share orders generate less revenue than do slightly smaller and slightly larger orders in \$0.125-spread markets. We also find that RLP vary considerably across brokers. For example, the RLP associated with B2's market orders in \$0.125-spread markets are between 40% and 80% of B7's for most order-size categories, and negative for 1000-share orders. We strongly reject the hypothesis that the mean RLP are equal across brokers (χ^2-statistic $= 536$). These results suggest that the dealer revenue associated with order flow varies across brokers holding order size constant.

4. An empirical model of trading revenue

To determine whether order characteristics other than size are systematically related to dealer revenue, we estimate a statistical RLP model. As motivated by Easley and O'Hara (1987), Huang and Stoll (1996), and Bessembinder and Kaufman (1997), respectively, we use order size, stock price, and trading volume as control variables. In addition, stock-price volatility may influence dealer revenue because traditional dealers favor liquid stocks with stable prices. Without an information event, this provides dealers with many orders alternating between buyers lifting the offer and sellers hitting the bid. Volatility and trading frequency are measured during the month prior to the order date. Finally, because the RLP distribution depends on the quoted spread, we use the quoted spread at order-submission time as a control variable by estimating our model separately with trades occurring in markets where the submission-time spread is \$0.125 and \$0.25.

4.1. The model

Most of Knight's executions occur at prices evenly divisible by \$0.125, implying that RLP assume a limited number of discrete values, each evenly divisible by \$0.0625.[7] The discrete nature of the dependent variable makes estimation via ordinary least squares inappropriate. Thus, we use an ordered-response model.[8] Actual RLP are converted to multinomial-choice variables. Let RLP_i denote the multinomial-choice variable representing the observed RLP for the ith order, where

$$RLP_i \in \{\$0.4375\ \textit{and less}, \$0.375, \$0.3125, \ldots, \$0.4375, \$0.50, \$0.5625\ \textit{and more}\}$$

represent the categories for orders submitted in \$0.125-spread markets and

$$RLP_i \in \{\$0.375\ \textit{and less}, \$0.3125, \$0.25, \ldots, \$0.50, \$0.5625, \$0.625\ \textit{and more}\}$$

represent the categories for orders in \$0.25-spread markets.[9] Consider two examples. If RLP is $-\$0.4375$, then it is assigned a categorical value of zero because it is an element of the lowest category. An RLP of \$0.0625 is assigned a categorical value of eight in \$0.125-spread markets and seven in \$0.25-spread markets. That is, $\$0.0625 < -\$0.4375 \times (\$0.0625 \times 8)$ for \$0.125-spread markets and $\$0.0625 < -\$0.375 \times (\$0.0625 \times 7)$ in \$0.25-spread markets.

RLP_i is related to the realization of an unobserved response variable, RLP_i^*, whose mean is a linear function of several variables. More formally, we estimate the latent regression model,

$$\mathrm{RLP}_i^* = \beta' x_i + \varepsilon_i, \tag{2}$$

where β is a coefficient vector, x_i a vector of explanatory variables, and ε_i a zero-mean error vector. Although RLP_i^* is unobservable, we observe,

$$\begin{aligned} &\mathrm{RLP}_i = 0 \quad \text{if} \quad \mathrm{RLP}_i^* \leqslant \gamma_0, \\ &\mathrm{RLP}_i = j \quad \text{if} \quad \gamma_{j-1} < \mathrm{RLP}_i{}^* \leqslant \gamma_j \quad \text{for} \quad j = 1, 2, 3, \ldots, 15 \end{aligned} \tag{3}$$

and

$$\mathrm{RLP}_i = 16 \quad \text{if} \quad \gamma_{15} \leqslant \mathrm{RLP}_i^*.$$

The γ are unknown parameters to be estimated with β. From Eq. (3), we can show that

$$\begin{aligned} &\mathrm{Prob}(\mathrm{RLP}_i = 0) = F(\gamma_0 - \beta' x), \\ &\mathrm{Prob}(\mathrm{RLP}_i = j) = F(\gamma_j - \beta' x) - F(\gamma_{j-1} - \beta' x) \quad \text{for}\, j = 1, 2, 3, \ldots, 15 \end{aligned} \tag{4}$$

---

[7] Nasdaq permits dealers to execute orders in decimal (\$0.01) prices. Less than 1% Knight's executions are executed on a price that is not evenly divisible by \$0.125.

[8] Hausman et al. (1992) and Keim and Madhavan (1995) among others use ordered response models to deal with discreteness in financial modeling.

[9] Categories are chosen so that no more than 2% of the RLP observations fall into the extreme categories.

and

$$\text{Prob}(\text{RLP}_i = 16) = 1 - F(\gamma_{15} - \beta' x),$$

where F is the error term's cumulative distribution function. In order for the estimated probabilities to be positive, we require that $\gamma_0 < \gamma_1 < \ldots < \gamma_{15}$. The probability of observing a particular RLP category depends on the location of the conditional mean of the underlying response variable, $\beta' x_i$, relative to the partition γ_j. Keim and Madhavan (1995) argue that the usual choices for F (i.e., the logistic and the normal) produce similar results, so (following them) we choose the more commonly-used normal distribution and estimate (4) with ordered probit.

With 16 partitions ($j = 0, 1, 2, \ldots, 15$), we calculate the parameter estimates by maximizing the following likelihood function, L,

$$L = \prod_{i=1}^{n} \prod_{j=0}^{16} [\Phi(\gamma_j - \beta' x_i) - \Phi(\gamma_{j-1} - \beta' x_i)]^{y_{ij}}, \tag{5}$$

where Φ represents the cumulative standard normal distribution, $\Phi(\gamma_{-1} - \beta' x) \equiv 0, \Phi(\gamma_{16} - \beta' x) \equiv 1, y_{ij} = 1$ if $\text{RLP}_i = j$ and 0 otherwise, and n is the number of observations. Estimating β with respect to the underlying continuous response variable RLP_i^* and not the discrete RLP_i values, means we need data on the partitions' frequencies to interpret the economic significance of the estimated coefficients.

We model the mean of the response variable RLP_i^* as

$$\beta' x_i = \beta_1 S_i + \beta_2 S_i^2 + \beta_3 D_i^{1000} + \beta_4 \ln(P_i) + \beta_5 \ln(N_i) + \beta_6 \ln(V_i), \tag{6}$$

where, for ith order, S_i is order size in shares, D_i^{1000} is a binary variable taking a value of one if the order is for 1000 shares and zero otherwise, P_i is the stock's price, N_i is the number of trades in the prior month for the stock of interest, and V_i is the standard deviation of log daily returns for the stock of interest over the prior 20 trading days. Our model's specification of the effect of order size on trading revenue is allowed to be non-linear with a discontinuity at the 1000-share order size. The variable N_i proxies for the stock's liquidity.[10] We take logarithms of the indicated variables because the anticipated relationship with RLP_i^* is relative, not absolute. As previously indicated, we separately estimate the model for orders received in \$0.125-spread and \$0.25-spread markets. Finally, because the functional relationship between the independent variables and RLP_i^* may not be constant among brokers, we estimate the model

[10] Jones et al. (1994) suggest we do not use trading volume in addition to number of trades.

separately for each of the five brokers sending the most orders to Knight.[11] We estimate 22 (= 17 response categories − 1 + 6 slope coefficients) parameters for each broker.

Holding price constant, theory suggests that informativeness increases and trading revenue decreases in order size. If RLP is a good proxy for dealer revenue, then RLP decreases in order size. The dealer, however, can execute orders exceeding quoted size at prices other than those quoted. This may offset the information advantage of large orders. Also, the RLP for 1000-share orders may be less than the RLP for other orders because 1000-share orders are profitable, on average, for traders (see Harris and Shultz, 1998). Thus, we expect β_1 and β_3 to be negative and β_2 to be positive.

The other control variables account for across-stock differences in dealer revenue. Gross revenue is expected to be greater in more liquid and less volatile stocks, thus, the RLP is greater in those stocks. This suggests that β_5 is positive and β_6 negative. Finally, conditional on the order's submission-time spread, we expect RLP to increase in the stock's price because the discrete tick size presents more of a constraint on price moves for low-priced stocks. Hence β_4 should be positive.

4.2. The results

Panels A of Table 4 ($0.125-spread market) and 5 ($0.25-spread market) present estimates of the six slope coefficients and their asymptotic standard errors for the probit model estimated using maximum likelihood (partition boundary estimates are available from the authors). Panels B contain frequency counts for the seventeen ordered response categories.

Consistently with our prediction that the explanatory variables affect brokers' RLP differently, a likelihood ratio test (not reported) rejects, at the 0.01 level, the null hypothesis that the vector $\hat{\beta}$ is the same for the five order-flow sources displayed in Tables 4 and 5 in both spread-width markets. This suggests the functional relationship between RLP and order flow varies across brokers.

Table 4 shows a significant order-size:RLP relationship in the predicted direction in $0.125-spread markets for B1, B2, and B5. In $0.25-spread markets, this relationship is significant and has the predicted sign only for B2. These results suggest that dealer revenue from B2's orders initially is inversely related to order size, but becomes positively related to size for large orders reflecting Knight's ability to re-price orders exceeding quoted size. The 1000-share binary

[11] Using total order flow allows us to include brokers with very different limit-market order ratios, order sizes, and frequencies of 1000-share orders. Each of these variables acts as a proxy for the sophistication of the broker's clientele. Including additional brokers means we estimate the model with fewer observations, add little to the sample's stratification along the indicated variables, and increases the reporting burden without additional insight.

variable's coefficient is significantly negative (positive) for orders from B2 (B7) in \$0.125-spread markets indicating these orders generate significantly less (more) revenue for Knight than orders of nearly identical size. The 1000-share dummy's coefficient is not significant for any of the five sources of order flow in \$0.25-spread markets. The other control variables' coefficient estimates are not consistently significant.[12]

The relationship between RLP and order size varies across broker. To illustrate the differences, we present a specific example. We estimate brokers' RLP conditional on order size using estimates from our ordered probit analysis and the median values for the control variables: 4.0279 for ln(P), 12.8047 for ln(N), and $-$ 3.4588 for ln(V).[13] Fig. 1a (1b) contains a plot of these estimates for orders submitted in \$0.125- (\$0.25-) spread markets.

For small orders, there is little difference in RLP across brokers; Knight earns approximately the half-spread (\$0.0625 in Fig. 1a and \$0.125 in Fig. 1b) on each share traded. As order size increases, however, dealer revenue varies widely across order sizes and among brokers. Consider Fig. 1a. Trading revenue associated with orders from B1, B3, and B5 is relatively uncorrelated with order size; Knight earns over \$0.05 per share regardless of order size. A nonlinear relationship appears to exist for B2 and B7; estimated RLP initially fall and then rise. The revenue generated by B2's order flow falls from \$0.054 per share for 100-share orders to \$0.013 for 2160-share orders, but then increases to over \$0.02 for 3000-share orders because Knight can execute large orders at prices outside of the receipt-time quote. The revenue associated with 1000-share trades is less than that of the surrounding order sizes for B1 and B2, but greater for B7. The order flow from B2 uniformly provides Knight less revenue than that from other brokers. In the case of a 1500-share order in a \$0.125-spread market, the difference in revenue is about \$0.04 per share. Fig. 1b suggests that orders from B2 are the least profitable in \$0.25-spread markets also.

To determine whether the order-size:RLP relationship differs statistically among brokers, we estimate a model with binary variables to isolate each broker's orders. We define the coefficients relative to B2's and estimate the differences in dealer revenue between order flow from B2 (the lowest-revenue broker) and the order flow from each of the remaining top-five brokers. Single-broker regressions suggest that the RLP:order-size relationship differs among brokers, so we include the binary variables in interactions with the size

[12] To be accurate, we note that the only unambiguous claims we can make in interpreting the coefficients regards the probabilities of observing a RLP in one of the extreme categories. Thus, a positive coefficient indicates that an increase in the independent variable increases (decreases) the likelihood the RLP is in the largest (smallest) category.

[13] We estimate RLP only for those order sizes well represented in our sample. For this reason, we estimate RLP for orders of 3000 (2000) shares and less in \$0.125- (\$0.25-) spread markets.

Table 4
Estimates of an ordered response model for realized liquidity premia (RLP) in \$0.125-spread markets

The table reports estimates of the ordered-probit model, $\text{Prob}[y_{i,j} = 1|x_i] = \Phi(\gamma_j - \beta' x_i) - \Phi(\gamma_{j-1} - \beta' x_i)$, where $y_{i,j}$ is 1 if the ith order results in the jth RLP category, γ_j is an unknown partition, β is a vector of unknown coefficients, x_i is a vector of independent variables, and Φ is the cumulative normal distribution. The ith order falls in category 1 if the order's RLP is less than or equal to – \$0.4375, category k if the order's RLP exceeds – \$0.5625 + [k(\$0.0625)] but is less than or equal to – \$0.50 + [k(\$0.0625)] for $k \in \{2, 3, \ldots, 16\}$ and category 17 if the order's RLP is greater than or equal to \$0.5625. The linear combination is

$$\beta' x_i = \beta_1 S_i + \beta_2 (S_i)^2 + \beta_3 D_i^{1000} + \beta_4 \ln(P_i) + \beta_5 \ln(N_i) + \beta_6 \ln(V_i),$$

where, for the ith order, S_i = the order's size, D_i^{1000} = a binary variable taking the value of one if size is 1000 shares and zero otherwise, P_i = midpoint of the bid–ask spread at the time the ith order is submitted, N_i is the number of trades in the stock in the prior month, and V_i is the standard deviation of log daily returns for the stock over the prior twenty days. The maximum likelihood estimates of the slope coefficients and their asymptotic standard errors are reported in Panel A below. Frequency counts for the ordered-response categories are reported in Panel B.

	B1		B2		B3		B5		B7	
	Point est.	Std. error	Point est.	Std. error	Point est.	Std. error	Point est.	Std. error	Point est.	Std. error
Panel A: Point estimates (point est.) and standard errors (std. error)										
β_1[a]	**−19.18**	6.30	**−22.77**[c]	5.70	– 7.051	5.90	**−32.95**	8.80	– 0.930	5.30
β_2[b]	**5.28**	2.06	**5.33**	1.72	1.04	1.56	**7.38**	2.85	– 0.089	1.39
β_3	– 0.0528	0.0636	**−0.1188**	0.0424	– 0.0017	0.0521	– 0.1621	0.1067	**0.1274**	0.0531
β_4	0.0176	0.0190	– 0.0461	0.0301	0.0580	0.0298	– 0.0142	0.0399	– 0.0582	0.0578
β_5	**0.0425**	0.0175	0.0025	0.0247	– 0.0148	0.0250	0.0043	0.0321	0.0461	0.0396
β_6	– 0.0070	0.0122	**−0.0630**	0.0223	– 0.0259	0.0241	0.0289	0.0251	0.0074	0.0327

Realized Liquidity Premium	Broker B1	B2	B3	B5	B7
Panel B: Frequency counts for the ordered-response categories					
< − \$0.3750	409	288	160	125	201
− 0.3750	55	34	28	12	11
− 0.3125	260	190	111	126	108
− 0.2500	103	61	51	26	21
− 0.1875	548	405	297	301	205
− 0.1250	210	194	105	66	58
− 0.0625	1,494	841	661	674	367
0.0000	850	539	424	258	158
0.0625	6,019	2,723	2,698	1,910	965
0.1250	833	485	422	218	175
0.1875	1,421	641	654	651	386
0.2500	252	146	121	64	48
0.3125	609	282	238	278	205
0.3750	85	48	54	26	19
0.4375	235	132	127	123	105
0.5000	34	29	26	7	9
> 0.5000	276	149	135	141	120
Total	13,693	7,187	6,312	5,006	3,161

[a]Reported numbers should be multiplied by 10^{-5}
[b]Reported number should be multiplied by 10^{-8}.
[c]Bold indicates significance at the 5% level in a two-tailed test.

Table 5
Estimates of an ordered response model for realized liquidity premiums (RLP) in \$0.25-spread markets

The table reports estimates of the ordered-probit model, $\mathrm{Prob}[y_{ij}=1|x_i]=\Phi(\gamma_j-\beta' x_i)-\Phi(\gamma_{j-1}-\beta' x_i)$, where $y_{i,j}=1$ if the ith order is in the jth RLP category, γ_j is an unknown partition, β is a vector of unknown coefficients, x_j is a vector of independent variables, and Φ is the cumulative normal distribution. The ith order is in category $j=1$ if the order's RLP is less than or equal to – \$0.375, category $j=k$ if the order's RLP is greater than – \$0.50 + [k(\$0.0625)] and less than or equal to – \$0.4375 + [k(\$0.0625)] for $k\in\{2,\ldots,16\}$, and category $j=17$ if the order's RLP is greater than or equal to \$0.625. The linear combination $\beta' x_i$ is given by

$$\beta' x_i=\beta_1 S_i+\beta_2(S_i)^2+\beta_3 D_i^{1000}+\beta_4\ln(P_i)+\beta_5\ln(N_i)+\beta_6\ln(V_i),$$

where, for the ith order, S_i = the order's size, D_i^{1000} = a binary variable taking the value of one if size equals 1000 shares and zero otherwise, P_i equals the midpoint of the bid/ask spread at the time of the ith order's submission, N_i is the number of trades in the stock in the prior month, and V_i is the standard deviation of log daily returns for the stock over the prior twenty days. The maximum likelihood estimates of the partition boundaries and slope coefficients and their asymptotic standard errors are reported in Panel A below. Frequency counts for the ordered-response categories are reported in Panel B.

	B1		B2		B3		B5		B7	
	Point est.	Std. error	Point est.	Std. error	Point est.	Std. error	Point est.	Std. error	Point est.	Std. error
Panel A: Point estimates (point est.) and standard errors (std. error)										
β_1[a]	2.26[c]	9.40	**–40.28**	8.30	**–18.13**	8.60	– 37.14	22.50	– 9.42	8.00
β_2[b]	– 2.83	3.18	**10.70**	2.56	3.58	2.19	0.196	0.118	3.95	2.31
β_3	– 0.1398	0.1014	– 0.07911	0.0696	0.0087	0.0769	0.0736	0.1850	0.0414	0.0816
β_4	0.0201	0.0299	– **0.0815**	0.0403	0.0583	0.04327	0.0259	0.0607	0.0013	0.0823
β_5	0.0253	0.0188	– 0.0356	0.0285	– 0.0117	0.0280	– 0.0079	0.0378	– 0.0357	0.0414
β_6	– 0.0028	0.0209	– **0.0906**	0.0356	– 0.0163	0.0362	– 0.0864	0.0462	– 0.0575	0.0483

Realized Liquidity Premium	Broker B1	B2	B3	B5	B7
Panel B: Frequency counts for the order-response categories					
⩽ − 0.375	145	114	90	23	47
− $0.3125	59	54	31	17	25
− 0.3125	58	31	21	9	14
− 0.2500	134	101	65	34	43
− 0.1875	153	97	82	36	44
− 0.1250	320	195	158	97	79
− 0.0625	420	230	210	85	89
0.0000	788	432	403	227	168
0.0625	2,705	1,232	1,183	644	475
0.1250	763	366	342	190	141
0.1875	283	141	128	59	54
0.2500	287	129	119	78	73
0.3125	130	64	66	25	33
0.3750	101	52	63	35	21
0.4375	41	21	26	8	10
0.5000	50	24	23	11	10
⩾ 0.5625	118	63	60	11	28
Total	6,555	3,346	3,070	1,589	1,354

[a] Reported numbers should be multiplied by 10^{-5}.
[b] Reported number should be multiplied by 10^{-8}.
[c] Bold indicates significance at the 5% level.

variables. Because most orders arrive when the spread is \$0.125, we focus on these market settings. Specifically, we propose to estimate the following model,

$$\beta' x_i = \sum_{k=1}^{4} \alpha_k D_k + \beta_1 S_i + \sum_{k=1}^{4} \delta_k D_k S_i + \beta_2 S_i^2 + \sum_{k=1}^{4} \zeta_k D_k S_i^2 + \beta_3 D_i^{1000}$$

$$+ \sum_{k=1}^{4} \lambda_k D_k D_i^{1000} + \beta_4 \ln(P_i) + \beta_5 \ln(N_i) + \beta_6 \ln(V_i), \quad (7)$$

where D_k equals 1 if the order is from broker k ($k = 1, 3, 5$, or 7) and 0 otherwise. In Eq. (7), the α's represent the average additional revenue of the indicated broker relative to B2. The δ's and ζ' s represent the differential order-size effects and the λ's the differential effect of a 1000-share trade. Our results are reported on Table 6.

None of the intercept terms are significant at traditional levels, suggesting that the average revenue associated with order flow is not significantly different across brokers for the smallest order size. This is consistent with the fact that the intercepts in Fig. 1a do not differ substantially. As before, the overall order-size coefficient (β_1) is significantly negative, implying that RLP initially decline with order size. The interaction terms for order size are consistently positive, with B5's coefficient being statistically significant at the 0.01 level and B3's and B7's coefficients at (about) the 0.06 level.[14] This suggests that the revenue associated with B5's order flow is significantly less sensitive to order size than that for B2. Fig. 1a shows that B5's RLP have little sensitivity to order size (relative to B2's). A significant coefficient estimate on size-squared (β_2) suggests that the RLP:order size relationship is non-linear, with RLP eventually increasing in order size. B5's interaction term is significantly positive at approximately the .01 level and B7's at the 0.07 level, suggesting that the RLP:order-size relationship for order flow from B5 is significantly less non-linear than is that relationship for B2. Finally, the 1000-share order dummy coefficient estimate (β_3) suggests that, on average, these orders generate less revenue than orders of nearly equal size. The only significant interaction terms in this case are for B1 (at the 0.02 level) and B7 (at the 0.07 level). As illustrated in Fig. 1a, although 1000-share trades from B2 generate less revenue for Knight than similarly-sized orders, they generate more revenue when from B7. The other control variables' coefficient estimates have the expected sign, with the coefficient on price volatility significant at just below the 95% confidence level. We can reject the model of Eq. (6), imposing a common RLP:order-size relationship across brokers, in favor of the model of Eq. (7) allowing the coefficients related to order size (β_1, β_2, and β_3) to

[14] To be conservative, we report two-tailed significance levels although theory supports a one-tailed test.

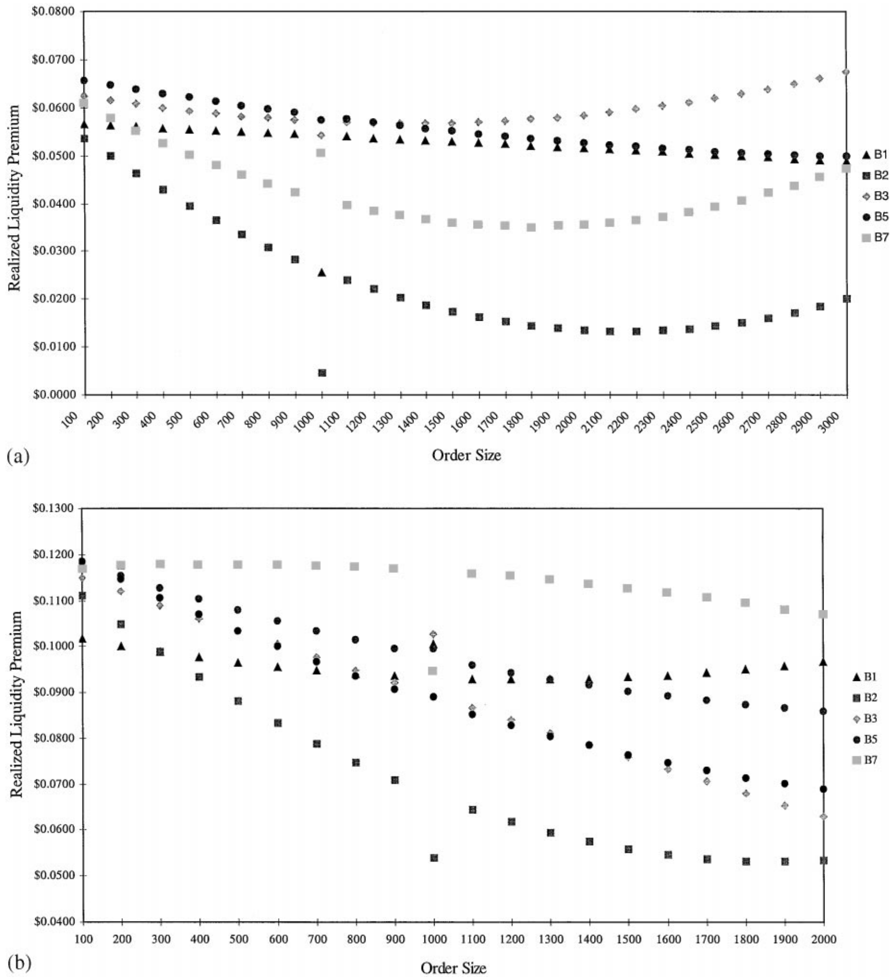

Fig. 1. Predicted values are obtained from the ordered response model for realized liquidity premiums described by Eq. (6) in the text. Predicted values are conditioned on the median values of the control variables (other than order size) across all orders. The values in $0.125-spread markets are 4.0828, 13.2489, and − 3.5302 for ln (price), ln (# trades), and ln (volatility) respectively. In $0.25-spread markets, these values are 4.0279, 12.8047, and − 3.4588. These values roughly correspond to the median values for the order flow routed by each broker.

vary across the brokers sending Knight the most market orders at a 0.0001 significance level.[15]

[15] We obtain qualitatively similar results estimating Eq. (7) with data from individual stocks. There are so few stocks with sufficient data to provide statistically reliable estimates that we do not present those results.

Table 6
Estimates of an ordered-response model for realized liquidity premium in \$0.125-spread markets with broker-specific interaction variables

The table reports estimates of the ordered-probit model, $Prob[y_{i,j} = 1|x_i] = \Phi(\gamma_j - \beta' x_i) - \Phi(\gamma_{j-1} - \beta' x_i)$, where $y_{i,j}$ is 1 if the ith order results in the jth RLP category, γ_j is an unknown partition, β is a vector of unknown coefficients, x_i is a vector of independent variables, and Φ is the cumulative normal distribution. The ith order falls in category $j = 1$ if the order's RLP is less than or equal to $-$\$0.4375, category $j = k$ if the order's RLP exceeds $-$\$0.5625 + [k(\$0.0625)] but is less than or equal to $-$\$0.50 + [k(\$0.0625)] for $k \in \{2, 3, \ldots, 16\}$ and category $j = 17$ if the order's RLP is greater than or equal to \$0.5625. The linear combination is

$$\beta' x_i = \sum_{k=1}^{4} \alpha_k D_k + \beta_1 S_i + \sum_{k=1}^{4} \delta_k D_k S_i + \beta_2 S_i^2 + \sum_{k=1}^{4} \zeta_k D_k S_i^2 + \beta_3 D_i^{1000}$$
$$+ \sum_{k=1}^{4} \lambda_k D_k D_i^{1000} + \beta_4 \ln(P_i) + \beta_5 \ln(N_i) + \beta_6 \ln(V_i), \tag{8}$$

where, for the ith order, S_i = the order's size, D_i^{1000} = a binary variable taking the value of one if size is 1000 shares and zero otherwise, P_i = midpoint of the bid–ask spread at the time the ith order is submitted, N_i is the number of trades in the stock in the prior month, and V_i is the standard deviation of log daily returns for the stock over the prior twenty days. The maximum likelihood estimates of the coefficients, their asymptotic standard errors, and the p-levels are reported.

Coefficient	Parameter Estimate	Standard error	p-value
α_1	0.0307	0.0235	0.1928
α_3	0.0241	0.0326	0.4604
α_5	– 0.0194	0.0354	0.5832
α_7	0.0331	0.2750	0.2282
β_1	– 0.00024	0.000057	0.0001
δ_1	0.00003	0.000084	0.7417
δ_3	0.00018	0.000094	0.0608
δ_5	0.00023	0.000078	0.0025
δ_7	0.00016	0.000081	0.0552
β_2	5.55E-08	1.72E-08	0.0013
ζ_1	7.15E-14	2.65E-08	0.9785
ζ_3	– 3.32E-08	2.59E-08	0.1995
ζ_5	– 5.60E-08	2.20E-08	0.0109
ζ_7	– 4.31E-08	2.31-E-08	0.0625
β_3	– 0.1279	0.0425	0.0026
λ_1	0.1845	0.0759	0.0151
λ_3	0.1052	0.0803	0.1902
λ_5	– 0.0303	0.0682	0.6564
λ_7	0.1253	0.0671	0.0619
β_4	0.0129	0.0127	0.3070
β_5	0.0186	0.0108	0.0851
β_6	– 0.0174	0.0090	0.0519

5. Stability of gross trading revenue estimates

Although the previous section's analysis finds significant across-broker RLP differences and differences in the RLP:order-size relationship, these differences might be random. For Knight to use broker identity to implement clientele pricing, RLP differences must be predictable. One could posit many forecasting models with varying levels of complexity. Each model is likely to include variables intended to estimate the sophistication of the broker's customers. The market-to-limit order ratio, the frequency of large orders, and the broker's commission level are examples of variables that might capture the relevant investor characteristics. Here we examine an extremely simple forecasting model. Specifically, we analyze the time-series stability of RLP differences in a univariate setting. That is, we forecast that a broker's future RLP equals its past RLP (at least relative to other brokers' RLP). We examine both short- and longer-term stability. Short-term stability is assessed by examining monthly changes in brokers' relative RLP. To examine longer-term stability, we employ our 1Q98 data to compare the across-broker differences in relative trading revenue in 1996 to those in 1998. Finding that the revenue rankings (at least in the extremes) do not change across time suggests that Knight might be able to use broker identity to predict RLP and to charge brokers differentially for its services.[16]

We begin by examining mean RLP conditional on correspondent. Table 7 provides the share-weighted RLP for the seven corresponding brokers that are among the ten brokers sending Knight the most market orders in both 4Q96 and 1Q98. We limit the sample to trades occurring when the quoted spread is no more than $0.25.

The decreased minimum price variation between 4Q96 and 1Q98 reduces estimated dealer revenue. Estimated revenue in 1Q98 is about 60% of its 4Q96 level. To determine if relative revenue levels among brokers are stable, we convert the RLP figures to rankings each month and compute Person rank-order correlation coefficients between adjacent periods.[17] For example, in October, 1996, B3's order flow is associated with the largest RLP and B2's with the smallest. In November, the order flow from B3 is the fifth most profitable and that from B2 the second least profitable. Although there is monthly

---

[16] At best, serial correlation is but a necessary condition for Knight to discriminate between brokers.

[17] Rank order correlation coefficients are used because (as noted) the change in the minimum price variation has a substantial effect on RLP levels. In addition, it seems that Knight's first-order interest is to determine whether it can identify brokers with consistently high or low RLP relative to the others. Correlation coefficients computed on levels are higher if the two months being compared have the same minimum price variation and lower if the months being compared are from different years.

Table 7
Share-weighted mean realized liquidity premium (SW–RLP) for orders received by Knight Securities in Nasdaq-100 stocks by corresponding broker when the quoted bid–ask spread is less than or equal to $0.25 during the fourth quarter of 1996 and the first quarter of 1998 for the seven brokers sending Knight the most market-order order flow in each period.[a]

Panel A. First quarter 1996.

	October, 1996		November, 1996		December, 1996		Overall	
Broker[b]	S–W RLP	# Trades	S–W RLP	# Trades	S–W RLP	# Trades	S–W RLP	# Shares ('000)
B1	$0.0697	6,612	$0.0408	6,089	$0.0751	7,564	$0.0630	4,192.1
B2	0.0388	4,666	0.0467	3,332	0.0148	2,550	0.0351	4,731.7
B3	0.0734	2,650	0.0505	3,141	0.0732	3,596	0.0658	3,772.3
B4	0.0661	2,331	0.0740	1,509	0.0733	1,862	0.0704	2,322.7
B5	0.0492	2,183	0.0595	2,129	0.0621	2,279	0.0566	1,313.3
B6	0.0723	1,706	0.0662	1,807	0.0737	1,512	0.0705	1,589.7
B7	0.0491	1,137	0.0644	1,407	0.0633	1,972	0.0600	3,563.2

Panel B. First quarter 1998.

	January, 1998		February, 1998		March, 1998		Overall	
Broker	S–W RLP	# Trades	S–W RLP	# Trades	S–W RLP	# Trades	S–W RLP	# Shares ('000)
B1	$0.0365	14,497	$0.0361	21,155	$0.0331	26,130	$0.0350	16,788.2
B2	0.0303	14,469	0.0277	20,451	0.0285	25,542	0.0287	19,792.4
B3	0.0381	8,106	0.0321	9,463	0.0292	10,504	0.0331	10,685.7
B4	0.0384	6,082	0.0338	8,856	0.0391	13,871	0.0374	8,190.9
B5	0.0254	2,988	0.0214	4,073	0.0114	4,438	0.0186	3,000.6
B6	0.0313	4,696	0.0375	5,198	0.0309	5,210	0.0333	5,762.6
B7	0.0233	4,949	0.0215	6,350	0.0154	9,993	0.0193	16,892.7

Panel C. Rank-order correlations between adjacent time periods.

Time periods	Rank-order correlation coefficient	*p*-level
October–November 1996	0.0714	0.8790
November–December 1996	0.1071	0.8191
December 1996–January 1998	0.5714	0.1802
January–February 1998	0.6429	0.1194
February–March 1998	0.8571	0.0137
All 4Q96–all 1Q98	0.7143	0.0737

[a]RLP = the difference between a buy (sell) order's execution price and the bid–ask spread's midpoint five minutes after order execution time (times -1).
[b]Brokers are coded B1–B7 based on the number of market orders submitted to Knight in the fourth quarter of 1996.

variation in revenue levels and ranking among brokers, the correlations generally are positive; ranging from 0.07 between October and November 1996 to 0.86 between February and March 1998. For the latter, we reject the hypothesis that it equals zero at the 0.02 level despite having only seven brokers. The rank-order correlation between broker revenue rankings from Overall 4Q96 and Overall 1Q98 is 0.71, which is statistically positive at the 0.08 level. Thus, over a period as long as one year, the rank ordering of dealer revenue among brokers is relatively stable. Furthermore, the brokers providing Knight order flow generating extremely high (B4 and B6) and low (B2, B5, and B7) levels of revenue in 4Q96 are the same brokers at the extremes in 1Q98.

We repeat the analysis after grouping orders into size categories to investigate if broker-specific dealer revenue is predictable within order size category. Results appear in Table 8.

We present the rank-order correlation coefficient of RLP between selected time periods and the frequency with which statistically significant RLP differences from one time period are reversed in the rankings during the immediately subsequent period. For example, between January 1998 and February 1998, the rank-order correlation coefficient of the RLP associated with odd-lot orders among the seven correspondents is 0.15. In January, there are five statistically significantly different pairwise gross revenue comparisons (from a total of 21 unique broker pairs). The relative rankings for only one of those differences is maintained in February. Overall, there is little evidence that Knight can predict the RLP associated with odd lot orders. There appears to be more potential for success in predicting the RLP associated with orders from 100 to 1000 shares. With the exception of the November–December 1996 periods, the rank-order correlation coefficients are positive and almost all of the significant RLP differences are maintained in the immediately following period. Despite the mixed success in predicting RLP for some order sizes and time periods, it is rare that statistically significant differences in RLP are reversed. Overall, we interpret the evidence to suggest that dealer revenue varies by broker in a potentially predictable manner for the most popular order sizes.

6. Conclusions

Benveniste et al. (1992) posit that market makers may know when brokers route them informed order flow by repeatedly dealing with them in the financial markets. If this is true, it suggests that market makers may not charge each dealer the same price their services. Market makers may price their services to attract a particular brokerage clientele or charge different clienteles different prices. To examine this issue, we obtain order audit-trail data from a major Nasdaq market maker and estimate the trading revenue associated with order flow from different brokers. We find significant differences in realized liquidity

Table 8
An analysis of the time-series stability of the relative levels of across-broker realized liquidity premiums (RLP) associated with order flow to Knight Securities, L.P. in one-tick markets during the fourth quarter of 1996 and the first quarter of 1998 by order size.[a]

	Oct.–Nov. 1996		Nov.–Dec. 1996		Jan.–Feb. 1998		Feb.–Mar. 1998		4Q96–1Q98	
Order size (shares)	Rank-order cor. Coef.[b]	Maintained relationships[c]	Rank-order cor. Coef.	Maintained relationships	Rank-order cor. Coef.	Maintained relationships	Rank-order cor. Coef.	Maintained relationships	Rank-order cor. Coef.	Maintained relationships
< 100	– 0.1181	0/0	0.5363	0/0	0.1545	1/5	– 0.0454	5/7	0.3214	0/0
100–499	0.1909	8/17	0.4454	8/10	0.4363	10/10	0.8363***	16/16	0.2500	4/9
500–999	0.7455***	12/12	– 0.1636	2/6	0.5455*	24/25	0.7818***	17/17	0.2142	3/4
1000	0.0727	9/12	0.6636**	2/2	0.8545***	3/3	0.7636***	14/15	0.7857**	10/10
1001–2499	– 0.3273	2/4	0.4182	0/0	0.4091	20/24	0.5091*	8/8	– 0.0714	0/0
2500–5000	– 0.0363	0/0	– 0.1771	0/0	– 0.0090	8/14	0.4091	10/10	0.1785	0/0

[a] RLP is the difference between a buy (sell) order's execution price and the bid–ask spread's midpoint five minutes after order execution time (times – 1).

[b] Rankings of the brokers' mean RLP for each order size category are taken in each of the two indicated time periods, e.g., October and November 1996. The number in the cell is the Person correlation coefficient computed between the two rankings. One asterisk indicates that the correlation coefficient is significantly different than zero at the 0.10 level, two asterisks indicate a 0.05 significance level, and three asterisks indicate a 0.01 significance level.

[c] In the denominator is the number of statistically significantly different between-broker RLP pairwise comparisons in the earlier time period indicated in the column heading. In the numerator is the number of these pairwise comparative RLP rankings maintained in the latter period. For example, the entry of 8/17 in the Oct.–Nov. 1996 column and the 100–499 row indicates that of the 17 statistically significant pairwise differences in RLP existing in October 1996 for orders with 100–499 shares, 8 maintain the same rank order in November 1996.

premiums (our proxy for market-maker revenue) among routing brokers. When the order-submission-time quoted spread is \$0.125, the across- broker difference in mean RLP exceeds four cents per share in our early sample period. This across-broker difference in dealer revenue persists in a multivariate setting controlling for order size and other factors that might affect market-making revenue. Furthermore, we find that RLP differences tend to persist across time. If dealers can predict these differences, then they may wish to differentiate among brokers in pricing their services. Although, it is unclear whether "informed" or "uninformed" order flow is more valuable, traditional market makers appear to value "uninformed" flow. Uninformed order flow, however, may be worthless to dealers generating profits through proprietary trading.

Currently, with the exception of primary market specialists, it is common to find market makers attracting order flow believed to be desirable by providing certain clienteles with better-than-quoted prices. With the inverse relationship between an order's size and its estimated market-making revenue, order size is frequently used to separate clienteles (i.e., market makers take only orders with sizes less than a stated maximum). For the clientele selected, the market maker provides enhanced services, e.g., frequent trading within the bid–ask spread (high price improvement rates) and/or making a direct payment to the broker (payment for order flow). Proponents of decimalization and the public display of the limit order book suggest that these changes will eliminate the ability of market makers to continue clientele pricing and lead to a consolidation of order flow. We expect that the incentive to differentially price market-making services will persist in the face of these changes in market structure, despite the fact that the changes will allow market makers to provide a different quoted price for each order size. As noted in Harris (1993) and Battalio and Holden (2000), these market-structure changes will eliminate the incentives for clientele pricing only if order size is the only order characteristic correlated with market-making revenue. Our paper suggests there is at least one other factor (broker identity) that must be priced in quotes before order flow will naturally consolidate in any one market.

Our objective is to demonstrate that the value a dealer places on order flow may vary along a dimension other than order size and to note the implication of this variance for the distribution of order flow across trading venues. We do not attempt to determine why the value of the order flow may vary across brokers, nor do we consider limit orders. Clearly, understanding the value-drivers for all types of order flow would be an important topic for future research.

References

Angel, J., 1994. Who gets price improvement on the NYSE? Georgetown University, unpublished paper.

Affleck-Graves, J., Hedge, S., Miller, R., 1994. Trading mechanisms and components of the bid–ask spread. Journal of Finance 49, 1471–1488.

Battalio, R., Holden, C., 2001. A simple model of payment for order flow, internalization, and total trading cost. Journal of Financial Markets 4(1), this issue.
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., Kaufman, H., 1997. A comparison of trade execution costs for NYSE and NASDAQ-listed stocks. Journal of Financial and Quantitative Analysis 32, 287–310.
Bloomberg News, 2000. FedEx and 6 others picked by big board for decimal trading. July 6.
Easley, D., O'Hara, M., 1987. Price, trade size and information in security markets. Journal of Financial Economics 19, 69–90.
Easley, D., Keifer, N., O'Hara, M., 1996. Cream-skimming or profit sharing? The curious role of purchased order flow. Journal of Finance 51, 811–834.
Fishman, M., Longstaff, F., 1992. Dual trading in futures markets. Journal of Finance 47, 643–671.
Franke, G., Hess, D., 1995. Anonymous electronic trading versus floor trading. University of Konstanz, unpublished paper.
Harris, L., 1993. Consolidation, fragmentation, segmentation, and regulation. Financial Markets, Institutions, and Instruments 2.
Harris, L., Panchapagesan, V., 1999. The information-content of the limit order book: evidence from NYSE specialist actions. University of Southern California, unpublished manuscript.
Harris, J., Schultz, P., 1998. The trading profits of SOES bandits. Journal of Financial Economics 50, 39–62.
Hausman, J., Lo, A., McKinley, C., 1992. An ordered probit analysis of transaction stock prices. Journal of Financial Economics 31, 319–379.
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, 313–351.
Jones, C., Kaul, G., Lipson, M., 1994. Transactions, volume, and volatility. Review of Financial Studies 7, 631–651.
Keim, D., Madhavan, A., 1995. Anatomy of the trading process: empirical evidence on the behavior of institutional traders. Journal of Financial Economics 37, 371–398.
Knez, P., Ready, M., 1996. Estimating the profits from trading strategies. Review of Financial Studies 9, 1121–1164.
Lee, C., 1993. Market integration and price execution for NYSE-listed securities. Journal of Finance 48, 1009–1038.
Lin, J-C., Sanger, G., Booth, G., 1998. External information costs and the adverse selection problem: a comparison of Nasdaq and NYSE stocks. International Review of Financial Analysis 7, 113–136.
Massimbe, M., Phelps, B., 1994. Electronic trading, market structure, and liquidity. Financial Analysts' Journal 50, 39–50.
New York Stock Exchange, 2000. nyse.com to go real-time – live market data will support new products in network NYSE. NYSE Press Release, July 19.
Ready, M., 1999. The specialist's discretion: stopped orders and price improvement. Review of Financial Studies 12, 1075–1112.
Rothschild, M., Stiglitz, J., 1976. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. Quarterly Journal of Economics 90, 629–649.
Securities and Exchange Commission, 2000. Order directing exchanges and National Association of Securities Dealer's, Inc. to submit a phase-in plan to implement decimal pricing on equity securities and options. Release No. 34-42914, June 8.
Sofianos, G., 1995. Specialist gross trading revenues at the New York stock exchange. NYSE Working Paper No. 95–01.
Traders, 1996. Jousting the giants for order flow. September.
Wall Street Journal, 1998. Wall street firms in on-line battle. June 2, p. C1.