Brokerage Commissions and Institutional Trading Patterns

Michael A. Goldstein
Babson College

Paul Irvine
University of Georgia

Eugene Kandel
Hebrew University and CEPR

Zvi Wiener
Hebrew University

The institutional brokerage industry faces an ever-increasing pressure to lower trading costs, which has already driven down average commissions and shifted volume toward low-cost execution venues. However, traditional full-service brokers that bundle execution with services remain a force and their commissions are still considerably higher than the marginal cost of trade execution. We hypothesize that commissions constitute a convenient way of charging a prearranged fixed fee for long-term access to a broker's premium services. We derive testable predictions based on this hypothesis and test them on a large sample of institutional trades from 1999 to 2003. We find that institutions negotiate commissions infrequently, and thus commissions vary little with trade characteristics. Institutions also concentrate their order flow with a relatively small set of brokers, with smaller institutions concentrating their trading more than large institutions and paying higher per-share commissions. These results are stable over time, are consistent with our predictions, and cannot be explained by cost-minimization alone. Finally, we discuss the evolution of the institutional brokerage market within the proposed framework and make informal predictions about future developments in the industry. (*JEL*: G23, G24)

We would like to thank Abel/Noser, Greenwich Associates, and the Institutional Broker Estimate Service for providing the data. We also thank Ekkehart Boehmer, Chitru Fernando, Terence Lim, Marc Lipson, Maureen O'Hara, Christine Parlour, Chester Spatt, George Sofianos, Daniel Weaver, seminar participants at the Hebrew University, HEC (France), Tel Aviv University, Texas A&M, the NASDAQ Economic Advisory Board, and participants at the New York Stock Exchange conference, the Yale-Nasdaq conference, and the FIRS Capri conference for their comments. We also thank Granit San for her assistance with the CDA/Spectrum data, David Hunter for his help with Thompson mutual funds data, and Michael Borns for his expert editorial assistance. Goldstein gratefully acknowledges financial support from the Babson College Board of Research; Kandel and Wiener are grateful to the Krueger Center for Financial Research at the Hebrew University for financial support. We apologize for any errors remaining in the article. Send correspondence to Paul Irvine, 444 Brooks Hall, Terry College of Business, Athens, GA 30602; telephone: 706-542-3661; fax: 706-542-9434. E-mail: pirvine@uga.edu.

doi:10.1093/rfs/hhp083 Advance Access publication October 24, 2009

The dominant business model in the institutional brokerage industry is under attack from discount brokers, crossing networks, and ECNs, which provide trade execution at less than half the cost of full-service commissions. In contrast, full-service brokers produce research; provide capital, time, and expertise to facilitate trade execution; and allocate initial public offerings. These valuable but costly services are bundled with trade execution and paid for with a high per-share commission. In this article, we examine the economics of the bundled commission market and find associated patterns of institutional trading that are not generally recognized. We also use this framework to describe the structure of the institutional brokerage industry.

Although commissions were deregulated on May 1, 1975, the continued reliance on high per-share commissions is puzzling. In the other industries deregulated in the 1970s, such as trucking, banking, and airlines, the bundling of services and inducement in kind disappeared quickly after the onset of competition. In addition, high per-share commissions do not appear to be the most natural way for brokers to charge for trade execution, since, as with any transaction cost, commissions should significantly reduce turnover, as argued in Constantinides (1986) and Vayanos (1998).[1] Yet the practice persists over thirty years after deregulation.[2]

The main reason for such a prolonged survival of bundling of execution and services after deregulation is most likely the safe-harbor provision of Section 28(e) of the 1975 Amendments to the Securities Act. Section 28(e) permits institutions to pay for various investment-related services out of brokerage commissions, rather than out of the management fee. While this exception facilitates the continuance of bundling, it is not a sufficient condition, as payment for these services can take other forms. The underlying economics of per-share commissions and their impact remain largely unexplored. Similar to airlines and restaurants, brokers provide many services that they either cannot, or do not wish to, sell outright. Instead, they allocate them to their best clients as a reward for past business and an inducement for future business. We contend that brokers and their institutional clients enter into long-term agreements specifying a level of service (premium or standard) and the overall payment for it. The payment for these services is rendered through the appropriate allocation of order flow to brokers, as institutional per-share commissions are already set in the contract. These contracts may be informal, yet well understood by the parties involved. Cumulative institutional commissions therefore represent a metering device that determines the allocation of commission dollars, making

[1] Since the level of trading volume remains one of the more puzzling problems in finance, any market feature that impedes trading makes it even more puzzling. We argue later that full-service brokerage commissions do not constitute marginal cost and thus do not significantly impede trading.

[2] Total commission revenues have been steadily increasing over time: from $1.74 billion from all sources in 1974 to $13.2 billion paid by institutional investors alone in 2005. Despite the growth of electronic trading, full-service commission payments still dominate U.S. institutional execution. Sofianos (2001) notes that institutional commission rates remain considerably higher than the marginal cost of trade execution.

it simple for a broker to keep a detailed profit-and-loss account for each client, as noted in Kelly and Hechinger (2004).

In this article, we suggest that per-share commissions constitute a convenient and legally safe-harbored way of charging a prearranged fixed payment for a broker's premium services. If one takes this view of commissions, then the predictions of the *quid pro quo* theoretical models of commissions and trading, such as Brennan and Hughes (1991) and Brennan and Chordia (1993), may change, and empirical estimates of marginal trading costs need to be reexamined. Our framework can explain the continued existence of high per-share commissions despite notable competition from discount brokers and ECNs, and also severs the link between the characteristics of a trade (such as price and size) and the commission applied to it. On the other hand, it links commissions to the value of the premium services supplied by full-service brokers. Finally, this framework helps predict institutions' allocation of trading volume across brokers.

We use a proprietary database of institutional trades in 1999–2003 from Abel/Noser, which allows us to identify over 25 million trades in NYSE-traded stocks submitted by over six hundred institutions (identified by an ID number, which we can follow) to over one thousand brokers, whose identities we know. The data identify the security, the trade size, the average execution price, and the commission. The evidence from these data supports our hypothesis. First, we show that there is relatively little variation in per-share institutional commissions across transactions, regardless of the institution or broker involved. In fact, the majority of institutional client-broker pairs use only one or two different per-share commissions for all their transactions, which indicates that the characteristics of a trade are not driving commissions. Indeed, we find that the most important determinant of the per-share commission on any trade is the prior-period commission paid by that institutional client to that same broker. These results are stable through time and are consistent with commissions being a metering device used to pay for a broker's premium services, which implies that full-service commissions are an average and not a marginal cost of trading.

Second, if institutions pay for premium services through commissions, this should affect their order flow allocation across brokers. Gargantuan institutions, such as Fidelity Investments, can allocate small proportions of their volume and still obtain the premium status from most brokers. Most institutions, however, face a trade-off between the need to hide their trading strategy by dispersing their trades and the benefits of concentrating their order flow with a small set of brokers, for whom they become important clients and receive premium services. Consistent with this hypothesized trade-off, we find that institutions indeed concentrate their volume with a few brokers, and smaller institutions concentrate significantly more. These findings are consistent over time. Third, we find that smaller institutions also pay higher per-share commissions and tend to have higher turnover, two facts consistent with their

desire to increase their total payment so as to receive premium services from at least some brokers. Fourth, we contrast our hypothesis with a simple alternative model wherein institutions allocate their order flow to minimize their total execution costs. While some results on the patterns of institutional order flow are consistent with both hypotheses (as these are not mutually exclusive), others are inconsistent with the cost-minimization model, yet corroborate our hypotheses.

Finally, we show that the proportion of volume executed at discount commission rates increases over time at the expense of full-service commissions placing downward pressure on full-service commissions rates. Even the practice of bundling services with execution is under competitive pressure. We show that a stronger emphasis on buy-side execution costs forces many full-service brokers to provide low-cost execution in-house. Recent regulatory measures appear to have reduced the profitability of premium broker services (Kadan et al. 2006). In addition, Hintz and Tang (2003) document brokers' increasing reliance on hedge funds for commission revenue, but these clients demand liquidity rather than proprietary research. Together, these factors imply that the value of brokers' premium services has declined for many institutional clients, while the importance of liquidity has increased. Consequently, it is reasonable to presume that the process of the unbundling of research from execution, which has already begun, will accelerate in the future.

The article is organized as follows. Section 1 puts the commissions in the perspective of the extant literature. Section 2 presents commissions in the context of a long-term contract and presents supporting evidence. Section 3 examines the market for a broker's premium services. Section 4 generates and tests hypotheses regarding the allocation of trading volume across brokers. Section 5 concludes.

1. History of Commissions and the Literature

Prior to 1975, commissions were tightly regulated by the SEC, and essentially fixed. Copeland (1979) reports that prior to 1975, institutional commissions on the NYSE were a direct function of both price and shares traded and calculated as:

$$Commission\ per\ share = A + B \times Price. \tag{1}$$

The coefficients A and B could vary with trade size and commissions on trades above \$300,000 could be negotiated. As with many other industries under price regulation, such as airlines, banking, and trucking, brokers who were prohibited from competing for clients with lower-priced commissions reverted to offering auxiliary services. Thus, prior to 1975, the bundling of services with execution was the norm.

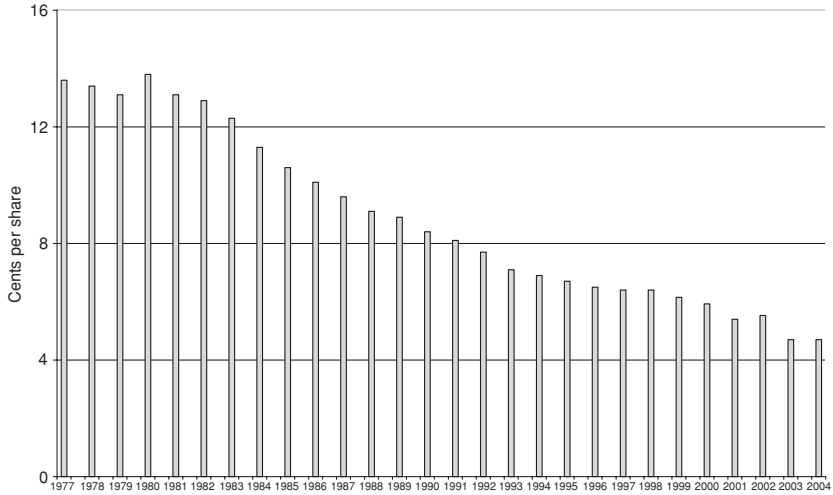

Figure 1
Historical commissions per share
Institutional investor average cents-per-share commission, 1977–2004. Data are from Greenwich Associates and represent the unweighted average commission that is calculated from proprietary survey data.

The May 1975 deregulation abolished fixed commissions, resulting in two major impacts on securities trading. First, commissions fell rapidly, though not uniformly, across all trade sizes (Tinic and West 1980). Figure 1 presents a 1977–2004 time series of average institutional commission rates reconstructed from Greenwich Associates survey data. The figure shows how the average institutional commissions fell from the mid-teens (in cents per share) in the late 1970s to just under 5 cents per share in 2004. The decline in real terms is much more dramatic.

The second major impact of deregulation was that discount brokers began to trade NYSE-listed stocks. For the first time, institutional investors were able to unbundle trade execution from the provision of ancillary services. Initially, discount brokers captured little institutional trading volume: the discount market share was only 6% in 1980 (Jarrell 1984). By 2003, over 40% of institutional volume in our sample was executed at discount prices. While this is a significant change, it is still small relative to other industries that underwent deregulation.

Early research modeled the post-deregulation commissions as a negotiated *marginal* cost of trade execution, but the evidence was mixed. While Ofer and Melnick (1978) claimed that commission rates represent the costs of executing various trades, Jarrell (1984) finds that commissions per share were relatively invariant to their estimated per-share cost, with the profits from large trades subsidizing losses from executing small trades.

Starting with Edmister (1978) and Edmister and Subramanian (1982), the focus shifts to measuring commissions as a percentage of price. Reflecting this view, Figure 2 presents percentage institutional commissions on the NYSE in

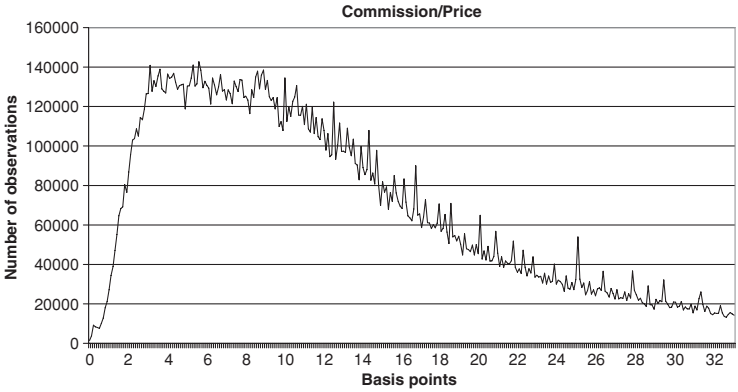

Figure 2
Institutional percentage commission costs on the NYSE
The distribution of commissions in Abel/Noser's NYSE-listed institutional trading data 1999–2003, as a percentage of stock price. Commissions per share are divided by the reported execution price to calculate percentage commission transactions cost. Zero cents per share commissions are not analyzed and the distribution is truncated at 33 bps.

our 1999–2003 sample. In this graph, institutional commissions appear to be a continuously distributed transaction cost, which seems consistent with the extant literature. The largest frequencies are between 5 and 15 basis points of the stock price, and there is a long right tail, which gradually dies out (we truncate it at 33 bps for ease of presentation). However, the representation of commissions in basis points is misleading for U.S. stocks.[3] In fact, the variation in commissions in Figure 2 comes primarily from price variation rather than from commission variation. To illustrate this point, Figure 3 presents commissions in cents per share in 1999 and in 2003. For clarity, we round commissions to the nearest tenth of a cent, making one hundred different price points available to institutional brokers. Ignoring most available prices, brokers in the United States price commissions primarily in exact cents per share. Commissions of 5 and 6 cents constitute the majority of observations in the 1999 sample, with the bulk of the rest executed at 2, 3, or 1 cent per share, respectively. While Figure 3 shows the distribution of institutional commissions in 2003 to be similar to that in 1999, the increased competition from ECNs has significantly reduced average commissions. This trend is mainly reflected in the paucity of commissions per share above 5 cents in 2003; commissions of 6 cents per share have almost disappeared.

Figure 4 presents the frequency of commissions by five trade-size categories. For trades under five hundred shares, low commissions are somewhat more prevalent: 25% of all small trades are executed at 2 cents per share, while 52% are executed at 5 or 6 cents per share. For large trades of over 10,000 shares, only 14% are executed at 2 cents per share, while 67% are executed at 5 or

[3] On the other hand, it is common for commissions for European or Japanese stocks to be quoted in basis points. As we demonstrate later, our results are not dependent on the type of quoting mechanism.

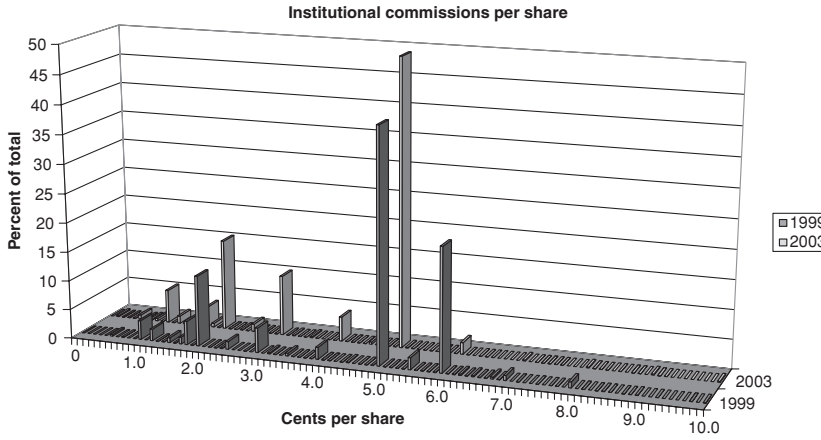

Figure 3
Per share institutional commissions for the NYSE-listed stocks in 1999 and 2003
All commissions per share are rounded to the nearest one-tenth of one cent. Zero cents per share commissions are not analyzed in this distribution, and the distribution is truncated above ten cents per share, where only a few observations reside. The resulting distribution of commissions is presented below. Few of the possible pricing nodes are actively used; institutions rely on whole number pricing, primarily at 2 and 5 cents per share.

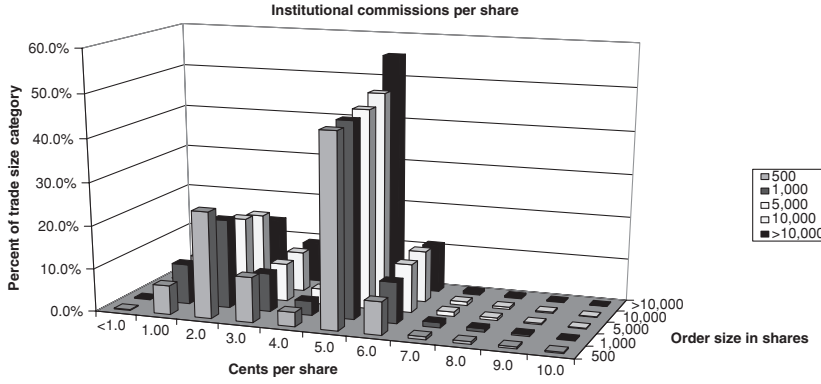

Figure 4
Per share institutional commissions for the NYSE-listed stocks in 1999–2003 by trade size
All commissions per share are rounded to the nearest cent. Zero cents per share commissions are not included, and the distribution is truncated above ten cents per share, losing only a few observations. Overall frequency of trades at each commission price is presented for five trade-size categories.

6 cents per share. Consistent with our point that commissions are not negotiated trade-by-trade, Sofianos (2001) contends that this variation in commission rates across trade size is likely due to the choice of trading venue by the client and not by client-broker negotiations over the commission rate on a particular trade.

An extensive literature treats (explicitly or implicitly) commissions as a *marginal* execution cost, including Copeland (1979); Loeb (1983); Roll (1984); Berkowitz, Logue, and Noser (1988); Brennan and Hughes (1991); Dermody and Prisman (1993); Chan and Lakonishok (1993, 1995); Livingston and

O'Neal (1996); Keim and Madhavan (1997); Bertsimas and Lo (1998); and Conrad, Johnson, and Wahal (2001). These studies find that commission costs, while smaller than price impact costs, are still significant, and thus should have a material impact on various decisions by investors.[4] Since the continuous distribution of percentage commissions is an artifact of price variability and has little to do with the determination of actual commissions, the interpretations of these findings may require revision. A significant part of what these studies consider a marginal cost of execution is not, in fact, a marginal cost at all. For example, Brennan and Hughes (1991) argue that firms can affect the level of analyst coverage they receive by splitting their stock. In their model, splits increase the potential commission revenue generated by trading the stock. However, if the total institutional commissions paid to a particular broker are predetermined, then the broker receives little or no marginal revenue benefit from the split, so the results of Brennan and Hughes (1991) may require an alternative explanation. We expand on this idea below.

2. Commission in the Context of a Long-Term Contract

Why should institutional commissions on the NYSE-listed stocks be charged using a few discrete cents-per-share prices? Given the downward trend in average commissions (Figure 1), the market for institutional execution appears competitive (Blake and Schack 2002). Why then is the distribution of per-share commissions in Figure 3 largely bimodal, a trend that accelerates as average per-share commissions fall? One possibility is that the discrete distribution of commissions in Figure 3 is consistent with the extant claim that commissions depend on the cost of executing a trade, as long as trades come in two discrete categories of difficulty. We test the prediction of this hypothesis using Abel/Noser data.

At the same time, we propose an alternative hypothesis that per-share commissions are determined by the broker as part of a long-term contract and are not subject to change or negotiations on a trade-by-trade basis. We conjecture that brokers use the commissions not only to charge for basic execution, but also for the premium services they provide. Each broker can provide several levels of service, each for its own total price per period. Given these prices, institutions decide which level of service fits their needs. If they choose to remain with the standard level of services, they are under no obligation to the broker. If, however, they choose to acquire premium status, then they pay for it by routing enough trading volume to this broker and paying the prespecified per-share commission.[5] Brokers and their client institutions then monitor the

[4] Commissions costs also have a significant impact on the cost of owning mutual funds. Hechinger (2004) reports that Lipper Inc. studied two thousand funds for *The Wall Street Journal* and found that brokerage commissions can more than double the cost of owning fund shares.

[5] For example, 3 million shares sent during a quarter at 5 cents per share generate a payment of $150,000 from the institution to its broker, as opposed to only $60,000 at 2 cents per share.

level of services received and the volume traded, keeping detailed accounts of each.[6] Under this hypothesis, per-share full-service commissions do not represent the marginal cost of trading for an institution, but rather serve as a metering device representing the average cost of services.

Marginal transaction costs reduce trading volume, which makes the use of high commissions by brokers puzzling. However, if the total commission payment per period is largely predetermined and the basic execution is available at competitive prices, then the effect of commissions on volume and trade size should be minimal. As long as an institution can trade with a discount broker or an ECN, its desired trading volume is set using the ECN's low transaction costs. Higher commissions, which include payment for other services, are inframarginal for the institution and thus should not affect the trading decision. This implies that bundling services with execution is not detrimental to trading because per-share commissions in excess of 1–3 cents are payments for broker services; therefore, they should have a minimal effect on volume.

The practice of paying for investment-related services out of commissions (rather than the management fee) was explicitly permitted under the safe harbor provision of Section 28(e) of the 1975 Amendments to the Securities Act. This significant advantage allows institutions to keep their management fees low. Paying a commission arranged in advance is also attractive to institutions relative to negotiating commissions on a trade-by-trade basis, which takes time and impacts immediacy of execution in a volatile market. Kavajecz and Keim (2005) argue that negotiations are costly, because they reveal details about each particular trade. Once the details of a trade are revealed to a broker, the institution cannot withdraw the information if the commission is unacceptable. A prearranged commission charge avoids these costs.

Institutional brokerage is not the only competitive industry that charges for services through something similar to the commissions relationship we have described. An analogous market mechanism is found in the airlines' frequent-flier programs. Airlines possess valuable assets that they cannot (or prefer not to) sell outright, such as empty first-class seats. These seats are often allocated to valuable customers based on the number of miles the customer has flown with the airline. The level of services is a step function of the accumulated miles. Travelers tend to concentrate their trips on their frequent-flier airline to ensure continued access to the airline's premium services. Both miles flown and total commissions represent easy-to-compute (for both parties) metrics that efficiently measure the importance of a client to each business.[7]

---

[6] Our conversations with market participants suggest that this is the way commissions are set and monitored. As per-share commissions are relatively constant, each broker must only measure the total number of shares received from an institution over the contract period to ensure that it receives enough revenue to continue providing the agreed-upon level of service. Where these institutions execute the rest of their trades is immaterial, as institutions have no incentive to reduce their level of trade with a broker unless they are dissatisfied with the services they receive.

[7] Another puzzling example of a linear contract based on a measure unrelated to performance is found in advertising. Advertising agencies receive revenues proportional to total media billings for their campaign. As in brokerage

The cent-per-share denomination of commissions, common in the United States, is not necessary for the long-term contract to exist and is not central to our argument. Commissions in Europe and Japan are traditionally quoted in basis points of trade value. The standard full-service institutional commission in Europe today is 15 bps, which yields similar revenue for the broker on a stock priced around $25–$30 as a U.S. commission of 4–5 cents per share.[8] Similarly, electronic execution in Europe is priced at 5 bps, which is comparable to 2 cents per share in the United States. Thus, the key element of the commission contract described here is not the basis on which the metering is done, be it cents or basis points, but the metering itself, which explains the continued existence of a premium commission in the presence of cheaper alternatives.[9]

Our conjecture also encompasses soft dollars, which represent an explicit charge for services purchased from outside vendors. They have been studied by regulators (SEC 1998), practitioners (Bennett 2002), and academics (Blume 1993; Conrad, Johnson, and Wahal 2001). While applicable to soft dollars, our conjecture also extends to all full-service commissions, whether or not they are recorded in a separate soft-dollar account. The difference in our emphasis is not merely semantic. First, according to Bennett (2002), in his report for Greenwich Associates, explicit soft-dollar commissions constitute only 27% of all full-service commissions, while the SEC (1998) reports that the seventy institutions it surveyed direct only about 8% of their total commissions to soft-dollar accounts. However, our data indicate that in 1999 over 70% of all commissions were above discount commission levels (in 2003 this number falls to 58%), implying that the market for premium commissions is much larger than conventional definitions of soft-dollar payments. Second, our argument applies in regulatory regimes, such as the U.K., where explicit soft-dollar arrangements are ruled out but where informal contracts for premium services still exist. Finally, explicit soft-dollar payments are predominantly used to buy third-party services: according to the SEC (1998), the most common use of soft dollars is as a payment to data vendors such as Standard and Poor's, First Call, and Bloomberg. Thus, soft dollars do not necessarily yield the same predictions regarding the allocation of order flow as the premium service hypothesis.

Viewing commissions as dependent on the cost of executing a trade implies that commissions should be mostly determined by individual trade

---

services, the quality of a single campaign is hard to quantify and contract upon, and thus the parties cannot base a payment on an objective performance measure. Instead, payments are based on an easily measurable variable that is under the full control of the client, who, therefore, determines the total payment. It is well known that firms frequently change their advertising agencies in search of better creativity. What is less known is that it is not uncommon for an agency to dismiss the firm if its billings are too low for the required effort.

[8] When commission deregulation finally arrived in Japan in October 1999, the Japanese commission contract changed from a function of price and volume (similar to that in the pre-deregulation NYSE) to European percentage commissions.

[9] Our hypothesis implies that the institutional commissions in Europe, as represented in cents per share, should be distributed continuously, whereas the distribution of commissions in basis points should be discrete. While we do not have data available to test this hypothesis directly, from the limited data that we have seen and conversations with industry practitioners, this seems to be the case.

characteristics, such as difficulty, price, and size. On the contrary, if brokers and their institutional clients predetermine commissions in a long-term agreement, there is no reason to negotiate commissions on a trade-by-trade basis, and the same commission can be charged repeatedly. Therefore, the main testable prediction of our conjecture is the persistence of commissions on trades between the same institutional client-broker pair. In an environment with little or no trade-by-trade negotiation over commissions, variables normally used to proxy for the execution cost of a trade should be relatively unimportant in determining per-share commissions. Our conjecture is consistent with previous empirical studies that find no significant correlation between the commission costs with execution costs (Berkowitz, Logue, and Noser 1988; Chan and Lakonishok 1993, 1995; Conrad, Johnson, and Wahal 2001). We proceed to test these alternatives directly using a large dataset of institutional trades.

2.1 Data

Our primary data source consists of 25,643,364 trades for NYSE-listed stocks by 683 institutional investors executed between January 1, 1999, and December 31, 2003. The proprietary data are obtained from Abel/Noser Corporation, an NYSE member firm and a leading provider of transaction cost analysis to institutional investors. Abel/Noser most often receives direct feeds from institutional investors' compliance departments; therefore, the database represents a complete record of an institution's trading. The database includes several unique items: the executing broker; an institutional client identification number, which permits us to track trades associated with each of the 683 institutions; and a buy/sell trade indicator. In addition, the database contains the commission cost of each trade, its size, date, and the average execution price.[10]

We next identify the broker used for each trade. There are 1064 brokers in the database; however, many brokers appear infrequently.[11] To concentrate on the most important participants, we restrict the sample to brokers who execute at least fifty trades in a calendar quarter. After imposing this restriction, only two hundred seventy active brokers remain in the Abel/Noser data, yet they account for over 98% of the original observations. We have further truncated the sample by deleting observations with commissions above 10 cents per share (2.01% of the sample), as well as those with zero commissions (2.77%). The resulting sample consists of 24,093,939 trades.

The size of the institutional client appears in several hypotheses, therefore we sort the clients into five quintiles, ranked by trading volume. We present the aggregate trading statistics by quintile in Table 1, which indicates that trading activity is highly skewed toward the largest clients. The highest-volume quintile

[10] Client orders can be executed in a single trade or broken into multiple trades. All results in the article are at the trade level. Robustness tests consolidating trades into orders produce similar results.

[11] We account for broker mergers. We track them using Ljungqvist, Marston, and Wilhelm (2006), Corwin and Schultz (2005), and several news and information services.

Table 1
Description of institutional client trading activity in the sample

	Client quintile by trading volume				
	1 = low	2	3	4	5 = high
Aggregate trading					
Total share volume (000s)	185,638	879,143	2,441,521	7,959,280	215,753,230
Total commission ($ 000s)	9,076	41,486	115,954	366,122	9,414,552
Trades	109,430	377,670	872,102	2,039,570	20,695,167
Average per client trading					
Volume per client ($ 000s)	46,370	207,776	588,440	1,854,271	53,809,910
Commissions per client ($ 000s)	67	303	846	2,672	69,225
Trades per client	805	2,756	6,365	14,887	152,170
Average commission/share	4.83	4.70	4.77	4.49	3.90
Average commission $/trade	82.95	109.85	132.96	179.51	454.92
Average trade size	1,696	2,327	2,800	3,902	10,425
Average price $/share	40.19	40.03	40.50	39.59	42.69

This table presents summary information on the trading activity of 683 institutional clients in the Abel/Noser dataset for 1999-2003. Institutional clients are sorted into five quintiles by total trading volume (shares executed). *Total share volume*, *Total commission,* and the number of *Trades* are sum totals for each client quintile. *Volume per client*, *Commissions per client,* and *Trades per client* represent the average across all clients in a quintile. Average commissions per share, per trade, trade size, and price per share are averages of all trades in each quintile.

dominates the other quintiles in terms of total trading volume, number of trades, and total commissions paid to brokers. As a robustness check, we verify that the average stock price per trade is roughly equal across quintiles, which indicates that differently sized institutions are not trading vastly differently priced stocks.

2.2 Results

We initially demonstrate that institutional commissions behave as if they were generated in a long-term contract using two empirical tests. First, for every calendar quarter in our sample, we identify the trades of client-broker pairs (keeping only pairs with at least five trades in that quarter), and calculate the mode of the commission distribution for each client-broker pair.[12] In our framework, where institutional per-share commissions are part of a long-term contract and are relatively constant over time, we expect the modal commission to dominate traditional measures of execution costs in predicting commissions in the subsequent quarter.

Table 2 presents the transition matrix between the mode of the commission distribution for each client-broker pair in the prior quarter and the mode for the same pair in the posterior quarter. The number of client-broker pairs in the prior (post) quarter is presented on the extreme right (bottom); altogether there are 4,776 client-broker pairs. The post-period row at the bottom of Table 2 is the number of post-period pairs that execute at a particular commission and therefore represents the unconditional distribution of the modal commission

[12] Our results are robust to alternative definitions of client-broker pairs. Using a minimum of five trades increases the noise in the mode estimate relative to larger cutoff points and, therefore, our presented results are conservative.

Table 2
Transition matrix of commission mode

Mode of cents per share in the prior period	Mode of cents per share in the post period										Mean no. of client broker pairs at that commission in prior period
	1.0	**2.0**	**3.0**	**4.0**	**5.0**	**6.0**	**7.0**	**8.0**	**9.0**	**10.0**	
1.0	**71.7**	11.7	3.8	2.0	8.9	2.0	0.0	0.0	0.0	0.0	59
2.0	2.6	**82.9**	3.6	1.1	7.8	1.8	0.1	0.0	0.0	0.0	302
3.0	1.0	5.9	**65.5**	2.3	20.2	5.8	0.1	0.0	0.0	0.0	281
4.0	0.9	3.3	5.3	**61.7**	23.3	5.3	0.2	0.0	0.1	0.0	144
5.0	0.2	0.9	1.9	1.6	**89.4**	5.8	0.1	0.0	0.0	0.0	2,850
6.0	0.1	0.7	1.8	1.1	23.2	**72.6**	0.3	0.1	0.0	0.0	1,097
7.0	0.0	1.6	1.4	1.8	17.6	19.2	**56.6**	1.9	1.4	0.0	30
8.0	0.0	0.1	0.1	0.0	9.7	14.9	7.1	**66.9**	0.0	0.0	8
9.0	0.0	0.0	2.5	2.5	12.5	25.0	0.0	0.0	**57.5**	0.0	2
10.0	0.0	0.0	1.6	0.0	0.0	12.7	0.0	7.9	6.4	**65.0**	3
Mean no. of client broker pairs at that commission in post period	63 (1.3%)	309 (6.5%)	279 (5.8%)	158 (3.3%)	2,928 (61.3%)	1,002 (21.0%)	24 (0.5%)	8 (0.2%)	2 (0.04%)	3 (0.1%)	4,776

This table presents the mode of the commission distribution between a specific institutional client and a specific broker. Commission modes are calculated quarter-to-quarter beginning in the first quarter of 1999 and ending in the third quarter of 2003. *Mode of cents per share in the prior period* is the mode of the client-broker commission distribution in the initial quarter. *Mode of cents per share in the post-period* is the mode of the commission distribution between the same broker-client pairs for trades executed in the following quarter. The mean number of client-broker pairs for all initial (following) quarters at each commission price is at right (bottom). Pairs with less than five trades executed in a quarter are omitted, as are pairs with fractional modes (7.3% of the sample) for clarity.

in the post-period. If commissions are negotiated on a trade-by-trade basis, then the distribution of modal post-period commissions should be independent of the prior-period commission. Instead, the data show that the actual transition probabilities depend heavily on the mode of the prior-period commission and are dramatically different from the unconditional distribution probabilities, as demonstrated by comparing the conditional probabilities along the main diagonal and the unconditional probabilities along the bottom row. To verify the importance of the prior-period commission on the frequency of post-period commissions, we perform a likelihood ratio test (Greene 1997). In each case, the hypothesis that the conditional probabilities are equal to the unconditional probabilities is strongly rejected. Hence, the observed frequencies of post-period commissions are significantly affected by prior-period commissions.

The fact that the prior-period commission between a client-broker pair is a strong predictor of the future modal commissions between that client-broker pair is consistent with the conjecture that per-share commissions represent average costs in long-term client-broker agreements. Next, we extend our tests of this hypothesis by contrasting the ability of standard measures of execution costs to predict trade-by-trade commissions against the ability of the prior-period mode.

Similar to many other authors, Roll (1984) assumes that brokerage commissions are negotiated on the basis of execution difficulty. Table 3 presents regressions, which estimate Equation (2) with and without the prior-period modal commission:

$$\begin{aligned} \textit{Commission per share} = {} & \alpha + \beta_1 \textit{Price} + \beta_2 \textit{Shares} + \beta_3 \textit{Mkt\%} \\ & + \beta_4 \textit{Mode} + \beta_5 \textit{Cvol} + \beta_6 \textit{Bvol} + \eta. \end{aligned} \tag{2}$$

In Equation (2), commission per share on a trade in the post-period is the function of the following: *Price*, the execution price; *Shares*, the trade size in log shares; *Mkt%*, the trade size as a percentage of that day's trading volume in the stock; *Mode*, the mode of the prior-period commission distribution for each client-broker pair; *Cvol*, the volume-based quintile size rank (smallest (1)–largest (5)) of the institutional client; and *Bvol*, the volume-based quintile size rank of the executing broker.

The explanatory power of the prior-period *Mode* relative to the explanatory power of the execution cost variables—*Price*, *Shares*, and *Mkt%*—is the key to interpreting Equation (2). Our hypothesis suggests that the *Mode* will have a strong explanatory power and a positive coefficient. Alternatively, if the execution costs of a particular trade really do affect commissions, then we expect the execution cost variables to influence the post-period commission. The effect of *Price* on commissions per share should be positive as higher priced stocks may require higher capital commitments from facilitating brokers. Larger trades may be more difficult to execute, so *Shares* should be positively related to commissions per share. *Mkt%* is a measure of trade difficulty: the

Table 3
Determinants of institutional commissions

Sample	$N =$	Intercept	Price	Shares	Mkt %	Prior Mode	Cvol	Bvol	OLS Adjusted R^2 (%)	Log likelihood from a logit regression
All	16,495,326	3.32	0.0000	0.089	−0.104				0.98	−1, 330, 695
		(0.001)	(0.001)	(0.001)	(0.001)					
All	16,495,326	0.406	0.0000	0.037	−0.026	0.83			65.80	−930, 933
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)				
Low cost	6,120,536	1.96	−0.0000	0.0015	0.047				0.01	−400, 575
		(0.001)	(0.001)	(0.001)	(0.001)					
Low cost	6,120,536	1.31	−0.0000	0.0004	0.038	0.28			26.03	−370, 531
		(0.001)	(0.003)	(0.001)	(0.001)	(0.001)				
Low cost	6,120,536	2.04	−0.0000	−0.0006	0.037	0.28	−0.0083	−0.173	26.62	−367, 941
		(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		
High cost	10,374,790	5.20	0.0000	−0.005	−0.007				0.03	−487, 434
		(0.001)	(0.262)	(0.001)	(0.001)					
High cost	10,374,790	3.86	0.0000	0.0005	−0.003	0.26			12.84	−436, 740
		(0.001)	(0.045)	(0.011)	(0.001)	(0.001)				
High cost	10,374,790	4.05	0.0000	0.0022	−0.027	0.26	−0.049	−0.002	13.03	−432, 881
		(0.001)	(0.008)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)		

This table presents the results of regressions using commissions per share in each quarter from the second quarter of 1999 to the last quarter of 2003 as the dependent variable. Commissions per share are truncated at ten cents a share and rounded to the nearest 1/10 of a cent. Zero cent commissions are not analyzed. *Shares* is the trade size, *Price* is the trade price. *Mkt %* is the size of the trade divided by the daily volume in the traded stock. *Prior Mode* is the mode of each client-broker pair commission per share cost in the preceding quarter. *Cvol* is the institution's quintile rank among all institutions in the sample. *Bvol* measures the brokers' quintile rank among all brokers in the sample. *Low cost* commissions are those trades with executed commissions per share less than or equal to 3 cents per share (Low cost). *High cost* commissions are those trades executed with executed commissions per share between 4 and 10 cents per share (High cost). *All* combines both low-cost and high-cost commissions. Log likelihood presents the goodness of fit statistic from an ordered logit regression specification of each regression with coefficient p-values reported in parentheses below the coefficient estimates.

larger the trade relative to daily volume, the greater total liquidity the trade demands. Hence, *Mkt%* should be positively related to commissions per share. *Cvol* and *Bvol* are included as control variables that measure potential effects in commission rates related to the size of the client or the size of the broker.

The first two regressions in Table 3 present two specifications of Equation (2) for all 16.5 million trades in the client-broker pairs sample (All). Under the null hypothesis that commissions can be represented as a continuous distribution of marginal transaction costs, OLS estimation is appropriate. However, as Figure 3 demonstrates, the distribution of commissions per share is discrete, not continuous. Thus, we also present the loglikelihoods from ordered Logit regressions to confirm that the OLS inferences about the economic significance of each regression are robust.[13]

Execution cost variables do not explain much: although trade size (*Shares*) has the predicted sign, trade difficulty ($Mkt\%$) does not, and the regression only manages an R^2 of 0.01. However, adding the prior *Mode* as an additional explanatory variable increases the R^2 dramatically to 0.66. This striking result shows that past commissions dominate the trade-specific characteristics in explaining the trade-by-trade commissions.

Given the relatively bimodal distribution of commissions per share presented in Figure 3, it is possible that the prior mode simply proxies for differences between the commission levels at full-service brokers as opposed to ECNs and discount brokers. To check the robustness of our results, we examine three regression specifications, estimating commissions per share separately for low cost (per-share commissions $\leq$3 cents) and high cost ($>$3 cents) markets separately. Again, we are primarily interested in the relative explanatory power of the execution cost proxies against our long-term agreement proxy (*Mode*). It turns out that in both subsamples, the execution cost variables do very poorly, obtaining R^2 of 0.01%–0.03%. In both markets, adding the *Mode* to the specification significantly increases the explanatory power of the regression to 26% and 13%, respectively. Client size (*Cvol*) and Broker size (*Bvol*) are both negative and significant but do not significantly increase the explanatory power of the regressions. This evidence indicates that individual trade commission costs are not driven by the characteristics of particular trades and indicates how far the market has evolved from the regulated commission market of Equation (1). Being invariant to the costs of trade execution, commissions are unlikely to represent marginal execution costs.[14]

---

[13] To save space we do not report the coefficient results from the logit specification. The results are similar and are available upon request. We also ran an OLS regression using log commissions as the dependent variable. The results were similar to the specifications presented.

[14] Nor is the trade-by-trade commission rate sensitive to the actual measures of execution costs that we can determine using these data. We calculate execution prices relative to the value-weighted average price (VWAP) and include this cost in unreported regression specifications. On a trade-by-trade basis, there is no significant correlation between execution cost and commissions per share ($\rho = -0.008$), nor do costs have a significant effect on the results in Table 3.

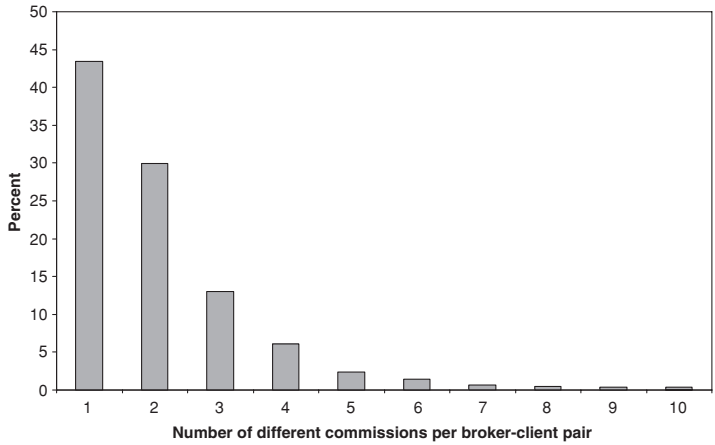

Figure 5
Frequency of number of different commissions per share by institutional client-broker pairs in 1999–2003
Percent represents the average over 20 quarters of the distribution of the number of commissions executed by a client-broker pair. Each institutional-client executed at least five trades with each broker during a quarter. Truncation at ten commissions removes 2.2% of the data from this graph.

Thus, the commission charge in both markets is best explained by the prior modal commission between a client-broker pair. This result is not driven by the safe-harbor provision: under Section 28(e), brokers could charge the marginal cost per trade plus a fixed markup. In this case, the fixed markup would be captured in the intercept of these regressions and the varying marginal cost of the trade should be captured by the coefficients of the independent variables and the prior mode would not much matter. However, we do not observe this result in Table 3.

The prior-period mode explains post-period commissions well because commissions are rarely negotiated on a trade-by-trade basis. Figure 5 presents the frequency distribution of commissions between our 4,776 client-broker pairs. Overall, 43.5% of all client-broker pairs in the sample only use a single per-share commission on all the trades they transact. An additional 30% of client-broker pairs pay only two commission prices and over 92.6% of all client-broker pairs use four or fewer commission prices. Clearly, trade-by-trade negotiation of commissions must play a relatively minor role in the institutional market. Yet more than one per-share commission can, in some instances, be used to fulfill the terms of the long-term contract.

The relationships between four prominent full-service brokers and their larger clients (with at least fifty trades per client per quarter) in the first quarters of 1999 and 2003 serve as an illustration of their response to low-cost competition. For example, Morgan Stanley offers low-cost (not exceeding 3 cent) executions to only thirty-one of these clients in 1999, but extends the arrangement to fifty-eight clients by 2003. Bear Stearns' low-cost commission relationships rise from thirty-one to forty-one during the same period. The

change at Goldman Sachs is more dramatic: low-cost commission alternatives are charged to twenty-two of their largest clients in 1999, and to fifty-five in 2003. Finally, Merrill Lynch's low-cost relationships almost triple from twenty-one in 1999 to sixty-one in 2003. These market changes are reflected in the decline of average commissions presented in Figure 1 and the shift in the distribution of commissions toward ECN prices presented in Figure 3. Overall, it seems that the data are consistent with our conjecture and provide little support to the commission being determined by the cost of execution.

3. Value of Premium Services

3.1 Equilibrium in the premium service market

Full-service brokers provide many services, the most prominent are timely information provision, the reduction of market impact on trade execution, and IPO access. We argue that even if these services were sold separately, the equilibrium would not take the form of a spot market, where these services are paid for on a *quid pro quo* basis, but would rather evolve into a long-term contract between brokers and institutions.[15] Clients observe only very crude proxies for the quality of these services on a daily basis. Removing the noise by averaging over a large number of events makes performance measurement far more accurate and easy to evaluate.

We conjecture that the equilibrium in the market for premium services takes the following form. The contract between a broker and a client is set for a specific period. Each broker provides a level of service corresponding to each client's choice and ensures that it gets appropriate payment. Premium clients receive top priority from the research department in providing timely information, the trading desk gives the foremost attention to their trades, and investment banking provides them with large IPO allocations upon demand. The absolute price for this level of service is high, as evidenced by the size of the commission market. At the end of the period, the institution evaluates the average quality (value) of services it received from the broker and decides at what level to continue the relationship.[16] At the opposite end of the scale, there are clients that demand only basic execution without any additional services, and so no long-term contract is required.

Our conjectured equilibrium is a variation on the Klein and Leffler (1981) equilibrium of product quality assurance. In their model, a high-quality producer prices the product above its marginal cost. The customers are willing

[15] Fulghieri and Spiegel (1993) present a model of IPO underpricing wherein broker services are complements. In their model, large clients received the most profitable IPO allocations.

[16] Conversations with institutional traders and research directors indicate that the quality of a broker's services can be considered fixed over a quarter or six months. Over longer periods, a broker's relative quality can deteriorate, in which case the institution pays a high price for inferior service, or it can improve, in which case the broker would wish to be compensated.

to pay more relative to cheaper substitutes as long as the quality is maintained above some predetermined level. The producer could cheat and produce a cheaper product, but this behavior would stop the stream of future positive profits associated with producing the high-quality good and receiving the premium price. Thus, the equilibrium yields a high-quality, high-price market even in the presence of low-cost substitutes. Applied to the brokerage services market, the model suggests that institutional clients can use repeated interactions to ensure high-quality service provision from their full-service brokers even without a formal contract and in the presence of discount brokers. Below, we specify the nature of the most important premium services that fall in that category.

3.1.1 Timely information provision. Timmons (2000) quotes an anonymous sell-side analyst as saying, "I kept my Buy rating, but I told my favorite investors to sell." Clearly, some clients are getting better information from their analysts than others. From any single client's perspective, the *value* of information the client receives crucially depends on the *timing* of its transmission from the broker. As prices adjust to reflect information imbedded in trades (Glosten and Milgrom 1985; Kyle 1985; Easley and O'Hara 1987), information loses its value upon revelation to additional market participants. Thus, the scarce resource in this context is the client's place in the information queue: those called first by the broker get the most valuable information.[17] This feature of information provision implies that clients have strong incentives to purchase a place near the head of the broker's queue. However, information quality that reflects one's place in the queue is hard to verify in any specific instance, as it is based on realized returns in a volatile market. Idiosyncratic effects tend to cancel over many independent observations, which suggests that the quality of research services provided by brokers can be best evaluated over a long period.

3.1.2 Trade execution. Institutional clients frequently demand that their brokers minimize the price impact of their trades. The time, skill, effort, and capital allocated by the broker to provide a counterparty for a trade determine the degree of its price impact. The sheer number of variables that could potentially affect execution on a particular trade suggests that ascertaining execution quality on a trade-by-trade basis is practically impossible. However, the idiosyncratic variables affecting execution quality on a particular trade tend to cancel out over time, so the precision of estimates of broker's performance improves over longer horizons. Indeed, the extensive use by institutional investors of such firms as Abel/Noser, which specialize in providing comparative analysis

[17] Historically, information was delivered by telephone and the broker determined the ordering of the queue, hence the name First Call for a well-known research distribution network. More recently, electronic dissemination of analysts' research notes ensures that most clients receive some information at approximately the same time. Today's queue revolves around a race to receive elaboration from the analyst on the brief First Call note to ascertain the value of the analyst's information.

of brokers' execution costs over time, suggests that the agreements based on execution cost measures are likely long term as well.[18,19]

3.1.3 IPOs. A broker's best institutional clients get larger allocations of "hot" IPOs, and the larger profits associated with them (Fulghieri and Spiegel 1993; Nimalendran, Ritter, and Zhang 2007).[20] It is obvious that brokers cannot explicitly charge for this service, and so they allocate shares to those who pay for them implicitly. Reuter (2006) finds confirming evidence through the correlation between mutual-fund commissions paid to underwriting brokers and post-IPO fund holdings. The fact that Reuter (2006) finds a significant relation despite the relatively infrequent reports from his data sources is strong evidence that the IPO allocation decision is at least partly based on long-term relationships between brokers and their clients.

The difficulty in measuring the quality of these premium services on a *quid pro quo* basis suggests that long-term agreements, which fix the level of service and the required payment over a long period, are appropriate in the institutional market.

3.2 Value of information provision: An illustration

Our hypothesis assumes tangible benefits from brokers' premium services, in particular, the timeliness and precision of sell-side analysts' information. We provide an illustration of the value of such service by investigating institutional trading around changes in analysts' recommendations.

We use a sample of 7010 analysts' recommendations changes from First Call during the 1999–2003 period for the NYSE stocks for which we also have Abel/Noser data. Panel A of Table 4 presents average event-day abnormal returns for the analysts' recommendation changes and finds them in line with prior research (Elton, Gruber, and Grossman 1986; Womack 1996). Upgrades produce an average abnormal return of 1.93% (t-statistic = 17.8) and downgrades produce an average abnormal return of -3.73% (t-statistic $= -24.0$). These recommendation changes seem to be informative, and hence timely trading in these stocks on these days may provide profit opportunities.

---

[18] Aitken, Garvey, and Swan (1995) and Foucault and Desgranges (2002) also discuss long-term relationships for trading services.

[19] Our analysis of execution costs relative to VWAP provides suggestive evidence supporting this idea. In preliminary tests, not presented here, we find that the actual execution costs had no significant effect in determining commissions. At the same time, when we examine aggregate commissions, we do find a suggestive pattern. Specifically, when we examine only low-cost discount commissions (not exceeding 3 cents), we find institutions execute trades at almost exactly the VWAP, while high-cost trades earn price improvement of about 1 cent per share on average. The average difference in commissions is 3.1 cents, indicating that, in aggregate, high-cost commissions are receiving about one-third of the benefit from their higher commissions through improved execution.

[20] There is a consensus in the IPO literature that underwriters compensate institutions that consistently provide them with information about the fundamental values of the issuing firms (Jenkinson and Ljungquist 2000). Production of this information requires institutions to invest in research capabilities, which is not economical if institutions are awarded small positions in IPOs. Consequently, there are imbedded economies of scale in the IPO pre-issue market.

Table 4
Analysts' recommendation changes and the subsequent trading

Panel A

	N	Mean	T-statistic
All upgrades	3,125	1.93%	17.81
upgrades to strong buy	1,805	1.95%	13.96
upgrades to buy	1,180	1.95%	10.49
upgrades to hold	140	1.67%	10.80
All downgrades	3,885	−3.73%	−24.04
downgrades to buy	1,321	−2.83%	−13.87
downgrades to hold	2,313	−4.06%	−19.02
downgrades to sell	251	−4.54%	−6.14

Panel B

	Trade through Other Brokers	Trade through the Changing Broker	Difference
Number of trades	219,320	3,965	
Improvement over VWAP – cents	2.34	6.76	**4.42**
	(0.0013)	(0.0082)	(5.27)
Improvement over Close – cents	9.07	19.31	**10.23**
	(0.0020)	(0.015)	(6.62)
Commissions per share – cents	4.35	5.05	**0.70**
	(0.00003)	(0.0002)	(37.79)
Commissions paid per trade ($)	745.8	965.6	**219.8**
	(8.61)	(70.80)	(3.08)
Share volume	15,967	18,709	**2,742**
	(173.19)	(1,324.74)	(2.05)
Mkt %	0.49	0.70	**0.21**
	(0.0002)	(0.0005)	(3.87)

Abnormal returns for 7010 NYSE-listed analyst recommendation changes in stocks that appeared in both First Call and Abel/Noser between 1999 and 2003. The abnormal returns reported are market-adjusted returns. The sample of brokerage recommendation changes consists of 7010 NYSE-listed analysts' recommendation changes in stocks that appeared in both First Call and Abel/Noser between 1999 and 2003. *Improvement over VWAP* is I (a buy-sell indicator variable) times the difference between the value-weighted average price for the day and the execution price of the trade. *Improvement over Close* is I times the difference between the closing price for that day and the execution price of the trade. *Share volume* is the number of shares in a trade. *Mkt %* is the size of the trade divided by the daily volume in the traded stock. Standard errors are reported in parentheses. T-statistics for the difference in means test are presented in parentheses below the mean difference in the Difference column. Significant differences (5%) are in boldface.

Panel B of Table 4 presents an analysis of institutional trades in the recommended stock on the day analysts change their recommendations. We test whether the recommending broker's clients receive superior information by comparing the profitability of client trades in the recommended stock on the day of the release of the report against the profitability of trades by nonclients. This is a powerful and direct test of the informational value of being a client of a full-service broker. Institutions trading through the recommending broker are by definition clients of that broker. Although not required to trade with the recommending broker, many clients apparently do, perhaps to reward the analyst whose bonus is often tied to the commission revenue generated by their recommendations.[21]

[21] Irvine (2004) reports that brokerage-firm trading volume in the recommended stock rises significantly after Buy recommendations. Similarly, Green (2006) notes that recommendation changes for NASDAQ stocks result in

Institutions that trade through the recommending broker obtained prices that average 19.31 cents per share better than the closing price, while trades through nonrecommending brokers received price gains of 9.07 cents relative to the close. These profits are comparable to existing evidence of trading gains for clients receiving notification of reports before they are broadly disseminated.[22]

Examining commission costs, we find that trades through the recommending broker on the day of the recommendation change paid higher average commissions—5.05 cents per share—than trades in the same stock on the same day executed through any other broker—4.35 cents per share. This difference reflects the fact that research providers are primarily full-service brokers who usually charge commissions of 5 or 6 cents per share. In this case, the extra commission payment is profitable as institutions that trade through the recommending broker gain significant price improvement in return for the higher commission. Thus, clients of the recommendation-changing broker made more profitable trades, despite the fact that these trades were, on average, significantly more difficult to execute, as measured by the size of the trade relative to that day's trading volume.

The profitability results support our assertion that brokers' services are valuable. Since a large portion of the gain from trading on analysts' recommendations is likely to dissipate quickly, access to early and precise information from the brokers' research department is a valuable asset.

4. Institutional Trading Patterns

How do institutions allocate their volume across various brokers? Several decisions are involved: how many brokers to use; which volumes to allocate; and how to allocate them among the chosen set. We present two ways to address this question. One is based solely on the cost of execution; the other focuses only on the effects implied by the payment for premium services hypothesis.

4.1 Hypotheses based on cost minimization

We identify three types of costs that institutions must take into account when allocating their order flow to brokers (we assume that prices of brokers are competitive).

1. **Fixed cost**: Cost of adding an additional broker to a client's list of brokers to cover electronic connections, billing, clearing, other back-office services, as well as regulatory compliance costs. This cost provides incentives for

aggressive quoting behavior from the affiliated market makers, as if accommodating customer order flow. Neither a cost-minimization framework nor the existence of soft dollars explains this result as directly as our conjecture that brokers reward profitable clients with premium services.

[22] For example, Kim, Lin, and Slovin (1997) find that traders who execute before Dow Jones widely disseminates an initial recommendation earn 32.3 cents-per-share intraday profit. Green (2006) finds that early traders on the day's First Call reports analysts' recommendation changes earn 45 cents-per-share profit when buying on upgrades and 52 cents-per-share profit by shorting on downgrades.

institutions to limit their number of brokers. If no other costs are present, this would imply a single broker for each market for each institution.

2. **Cost of frontrunning**: An institution may not want to send too much trading volume to a broker to prevent the broker from frontrunning.[23] This problem is much more significant for large institutions that occasionally submit very large orders. This cost induces the client to distribute volume evenly across brokers (yet randomize on every transaction), to split orders among brokers, and to increase the number of brokers used.
3. **Trading strategy recognition**: Institutions do not want the market to recognize that their trades represent a significant proportion of a particular broker's volume. They are afraid of market participants recognizing their trading patterns and increasing their price-impact costs.[24] As it is easier to hide a small volume than a large one, small institutions should not be worried about this cost. Large institutions can minimize this cost by increasing the number of brokers they use and allocating volume proportionate to the broker's size. Thus, an institution minimizing this cost would send comparable trading volumes to two of its equally sized brokers.[25]

We postulate several hypotheses based on institutions minimizing these costs (the letter "c" indicates that these are derived from cost considerations):[26]

Hypothesis 1c: *Institutions allocate higher percentage of their volume to larger brokers among the brokers they use, so to hide their order flow.*

Hypothesis 2c: *Smaller institutions allocate their volume more evenly than larger institutions, as they are less worried about frontrunning and strategy recognition.*

Hypothesis 3c: *Smaller institutions employ fewer brokers than larger institutions, since they face the same fixed costs, but lower frontrunning and strategy recognition costs.*

Hypothesis 4c: *Minimizing trading costs should not increase the turnover of smaller institutions relative to that of larger institutions. If small institutions pay higher per-share commissions to cover the fixed setup costs, then this trading impediment should reduce the turnover of smaller institutions*

---

[23] A broker who knows that a client has a large buy (sell) order can start buying from (selling to) the book, driving prices up, and then selling to (buying from) the institution at higher prices. Shwartz and Steil (2002) survey twenty-seven major investment management firms and conclude that frontrunning costs are important to buy-side institutions; such costs are a primary driver of the buy-side's demand for trading immediacy.

[24] Chan and Lakonishok (1993, 1995) conclude that the most important determinant of the price impact of an institutional trade is the identity of the institution behind the trade.

[25] An interesting case is Fidelity, which could easily dominate any broker's volume, but then the market will know that this broker's trades have a high probability of being Fidelity's trades. Market participants actively try to determine Fidelity's trading patterns, which it actively tries to hide (Pethokoukis 1997).

[26] Predictions below can be derived formally in a context of a model with the above assumptions. Details are available upon request.

relative to that of the larger ones. If the commissions are the same across institutions, there should be no difference between the turnover of larger and smaller institutions.

Hypothesis 5c: *Similar-sized brokers should receive similar proportions of volume from the same institution, while similar-sized institutions should send similar proportions of volume to a particular broker they work with.*

In addition to these costs, we have outlined the benefits available through client-broker relationships. These benefits produce an alternative set of hypotheses.

4.2 Hypotheses based on payment for services

We have argued that an institution willing to pay the price of a premium service package receives early access to analysts' research, priority in difficult trade executions, more capital committed to its trades, and a disproportionately larger share of IPOs. To obtain these services, institutions pay a fixed fee charged through a relatively constant per-share commission. The total payment is a product of the number of shares and the commission per share. In this framework, institutions allocate volume strategically so as to obtain premier status at as many service-providing brokers as possible. To do so they must concentrate their order flow with a subset of brokers to generate sufficient revenue within this subset. Under our conjecture, institutional trading patterns must, therefore, reflect a pattern of concentration ("bunching") of trades with particular brokers.

Note that larger institutions can automatically become premier clients with more brokers due to their size. Smaller ones should use fewer brokers to generate more volume per broker, as well as bunch more extensively. Small institutions could also increase their payments to brokers by increasing their turnover or by agreeing to pay higher commissions per share.[27] This is particularly relevant for smaller institutions that may want to increase their service above the level they would receive based on their size.

We postulate a second set of testable hypotheses based on our conjecture (indicated by "s"):

Hypothesis 1s: *Institutions disproportionately "bunch" their order flow with particular brokers to receive a premier level of service. Bunching is not necessarily related to a broker's size.*

Hypothesis 2s: *Smaller institutions bunch more than larger institutions due to their desire to obtain premier status with at least some brokers.*

Hypothesis 3s: *Smaller institutions employ fewer brokers than larger institutions since the sheer size of large institutions allows them to attain premier*

[27] This agency cost argument has been made relative to soft dollars by Berkowitz and Logue (1987) and Logue (1991). On the other hand, Johnsen (1994) and Horan and Johnsen (2004) argue that soft dollars may ameliorate agency cost issues.

status with more brokers, while small institutions need to concentrate their trades on fewer brokers.

Hypothesis 4s: *Smaller institutions may be willing to increase turnover and pay a higher per-share commission to generate higher commission revenues and receive additional services.*

Hypothesis 5s: *Similar-sized brokers may receive vastly different allocations from the same institution, depending on whether the institution wishes to become an important client and receive a particular broker's premier level of service. At the same time, two similar-sized institutions may send vastly different allocations to a particular broker for the same reason.*

While the intuition behind the two sets of hypotheses is completely different, some of them generate the same predictions. Under the cost-minimization alternative, institutions allocate a higher percentage of volume to the larger brokers they use to hide their order flow more effectively. It is not possible to distinguish between Hypotheses 1c and 1s with an analysis of institutional allocation of volume alone. Similarly, small institutions could employ fewer brokers (Hypothesis 3c and 3s) due to their desire to achieve premier services from some brokers, the fixed costs involved, or both reasons.

However, several predictions of the cost-minimization alternative are contrary to our hypotheses. Under the cost minimization alternative, small institutions, because of their small size, are less concerned with frontrunning and trading strategy recognition and would allocate their volume more evenly than large institutions. This prediction is contrary to Hypothesis 2s. Increasing the turnover would obviously increase costs for small institutions contrary to Hypothesis 4s. Finally, because the institution's total trading volume relative to the broker's total trading volume determines the ability to hide institutional trades, similar-sized brokers should receive similar proportions of volume from the same institution, while similar-sized institutions should send similar proportions of volume to a particular broker. This prediction is contrary to Hypothesis 5s.

As our hypotheses and the cost-minimization alternative are not mutually exclusive, we cannot claim that by testing these hypotheses we can reject one of these ideas. Instead, we interpret our findings as indicating which conjecture provides stronger empirical effects.

4.3 Results

Table 1 shows that small institutions spend significantly less in terms of total commission dollars. How do the four smallest quintiles compete for broker services? Hypothesis 4s suggests that smaller institutions may pay higher per-share commissions.[28] Table 1 shows that institutions in these quintiles indeed

[28] Recall that we expect large institutions to be high-revenue, high-cost customers. This contention implies that smaller institutions need not compete on total commission revenue, but rather on net profitability to the broker.

pay 15%–20% higher commissions per share than the institutions in the largest quintile. However, the difference in per-share commissions across size quintiles is dwarfed by the large differences in average trading volume. Thus, the total commission payment to a broker is potentially driven much more by the allocation of the trading volume (bunching) than by the size of the per-share commission.

4.3.1 Concentration of institutional trading. That order flow is the primary determinant of broker revenue has consequences for an institution's trading volume allocation decisions. Panel A of Table 5 presents institutional concentration of order flow as a function of institution size (quintile).[29] We examine both versions of Hypotheses 1 and 2 by calculating broker concentration as the average market share (percentage of each client's total commission dollars) that clients in each quintile send to their highest-revenue brokers.

Both versions of Hypothesis 1 predict order-flow bunching: a skewed allocation of client trades toward their most important brokers, which is precisely what we observe in the data. The largest institutions send 20.8% of their commission dollars to their top broker, whereas an evenly distributed allocation, which would best disguise their trading strategies, would allocate only about 1% of their order flow to each broker. The largest institutions concentrate their order flow with a few top brokers: 37.8% of their commission dollars goes to their top three brokers, 49.4% to their top five brokers, and 68.2% to their top ten brokers.[30]

Hypothesis 2s predicts that small institutions concentrate their trading more than large institutions, while 2c predicts the opposite. We show that the bunching of order flow with an institution's most important brokers increases as the size of the institution decreases. Panel A reveals that the percentage of commission dollars executed with their top broker increases monotonically with client size to a maximum of 40.7% for the smallest quintile. The null hypothesis that order flow executed with a top broker is independent of institution size is rejected at the 1% level with an F-statistic of 12.4. The top three, top five, and top ten broker categories show the same pattern of institutional bunching and similar rejections of equality across quintiles. This pattern is consistent with large institutions having the flexibility to become premier clients to many brokers, while small institutions are forced to concentrate their trading with only a few. This finding is inconsistent with the predictions of the pure cost-minimization

[29] Using information from the Securities Industry Association, and company websites, we classified our two hundred seventy active brokers into five types: full service, discount, ECN, wholesaler, and other brokers. Full-service brokers (144) are the most frequent broker type. Discount brokers, ECNs, and wholesalers generally do not provide premium services, while other brokers usually provide a single premium service. Tests of institutional trading patterns using only full-service brokers produce similar results to those presented.

[30] Table 5 reports institutional averages by commission dollars because commission dollars represent the important economic variable: brokers' revenue. Similar conclusions are obtained from share volume, but the reader should note that using commission dollars represents the low-cost market as a relatively less important execution method.

Table 5
Institutional concentration: bunching of order flow

Panel A: Institutional concentration of order flow

	Client quintile by trading volume					
	1 = low	2	3	4	5 = high	F-test
All brokers						
Average number of brokers per client	30.7	52.1	61.2	71.6	79.3	
Broker Concentration (% of client commissions)						
Top broker	40.7	27.9	24.7	22.2	20.8	12.4**
Top 1-3	60.2	47.2	43.4	40.6	37.8	15.4**
Top 1-5	69.9	57.8	54.3	51.6	49.4	15.6**
Top 10	83.3	74.1	71.1	69.1	68.2	15.3**
Low-cost trading						
Average number of low-cost brokers per client	8.8	7.5	8.9	9.9	13.3	
Percentage of client commissions paid to all low-cost brokers	11.4	9.7	11.4	12.7	17.3	2.4*

Panel B: Institutional concentration of order flow – Robustness test

Herfindahl Index	23.6	21.0	19.7	17.9	17.8	4.46**
Zeta Regression						
Intercept - α	−1.12	−1.21	−1.25	−1.32	−1.46	
	(0.050)	(0.031)	(0.029)	(0.023)	(0.025)	
Coefficient - θ	1.09	0.96	0.90	0.82	0.76	
	(0.062)	(0.039)	(0.033)	(0.027)	(0.037)	
R^2	91.7	91.8	91.8	91.5	90.9	
% of institutions with $\theta > 1$	41.9	32.0	25.7	21.3	14.0	
% of institutions with $\theta < 1$	46.8	48.4	64.7	70.6	75.0	

Panel C: Institutional average commissions by institutions

Commission Cost (cents per share)						
All brokers						
Top broker	5.01	4.94	4.76	4.53	4.46	2.95*
Top 1-3	5.04	5.02	4.88	4.67	4.45	10.81**
Top 1-5	5.07	5.04	4.93	4.72	4.48	19.45**
Top 1-10	5.11	5.08	5.01	4.80	4.55	40.09**
All brokers	5.04	5.07	5.01	4.90	4.80	71.63**

Panel D: Average broker rank for institutional clients' top brokers

Average broker rank - [median] (out of 261 active brokers)						
Top broker	35.4	32.5	33.1	31.7	27.4	
	[14]	[15]	[16]	[15]	[11]	
	(5.90)	(5.08)	(5.29)	(5.44)	(4.08)	
Second broker	29.8	30.7	29.7	29.4	26.6	
	[13]	[14]	[11]	[12]	[10]	
	(4.51)	(4.99)	(4.86)	(4.88)	(4.43)	
Third broker	29.8	29.2	30.1	25.2	28.5	
	[12]	[12]	[10]	[9]	[9]	
	(5.21)	(4.78)	(5.00)	(4.45)	(4.48)	

(*continued overleaf*)

Table 5
(Continued)

Panel D: Average broker rank for institutional clients' top brokers

	Client quintile by trading volume				
	1 = low	2	3	4	5 = high
Fourth broker	27.0	27.6	26.0	27.2	29.3
	[9]	[10]	[10]	[10]	[9]
	(5.39)	(5.32)	(4.31)	(4.73)	(4.68)
Fifth broker	30.1	27.6	29.6	24.4	23.3
	[13]	[11]	[12]	[10]	[8]
	(5.43)	(4.86)	(4.99)	(4.31)	(4.12)

Panel A presents institutional client market share statistics by client quintile. The *Average number of brokers per client* represents the average across institutions in a quintile. *Broker concentration* is the average of the percentage of their total commission dollars each client sends to their highest-volume broker(s). *Broker concentration* statistics are presented separately for both all brokers and ECN trading only. F-tests examine the null hypothesis of equality along each row. *Low-cost trading* tracks commissions less than or equal to 3 cents per share. The *Herfindahl Index* represents the (normalized) sum of the squared market shares of a client's ten largest brokers. *Zeta estimation* presents the average intercept, coefficient, and R-squared values in each quintile for client-by-client estimation of: log (broker market share) = $\alpha + \theta$ log (broker rank) + ε. Larger θ indicates more concentration. Percentage significantly greater or less than 1 is the percentage of each quintile's θ coefficients that are *significantly* greater or less than one at the 5 % level. Standard errors are in parentheses below the panel B coefficient estimates. *Commission cost* in panel C is the average per-share commissions on trades sent by clients to their highest volume broker(s). Panel D presents quintile-averaged statistics on *Average broker rank*, out of 270, for institutional clients' five most important brokers. Below each category's average, medians are presented in brackets and standard errors are presented in parentheses. The symbols * and ** indicate that the variation across quintiles is significantly different than zero at the 0.05 or 0.01 level, respectively.

model. Nor can the bunching result be explained by the soft-dollar arrangements, which primarily provide for data services (SEC 1998). Institutions do not care which broker provides soft-dollar credits to the data vendor, and, therefore, have no incentive to bunch. Thus, while the competition for valuable services from the broker may encourage bunching, soft-dollar arrangements are not likely to do so.

Figure 6 tracks top broker's order flow annually in 1999–2003. The allocation of order flow is consistent across client quintiles throughout our sample period, and similar patterns also prevail in our other top broker classifications. This figure indicates that the institutional trading patterns we document (and thus our conjectured commission contract) is consistent throughout our sample period. For specific agents, these important client-broker relationships are stable as well. In untabulated results, we find that a top broker in a quarter has an 89.7% chance of remaining in that client's list of top ten brokers in the following quarter. For relationships that we can track over the entire sample period, a client's initial top broker remains their top broker throughout the sample period 25.2% of the time. Initial top brokers remain in a client's top ten brokers 72.0% of the time. These numbers show a level of competition for a client's revenue stream, as top brokers are occasionally displaced. Yet displaced top brokers often remain important to the client and remain a competitive threat to reassert their dominant position.

This competition to be a client's most important broker may contribute to the overall declining trend in commissions. In general, when top brokers are

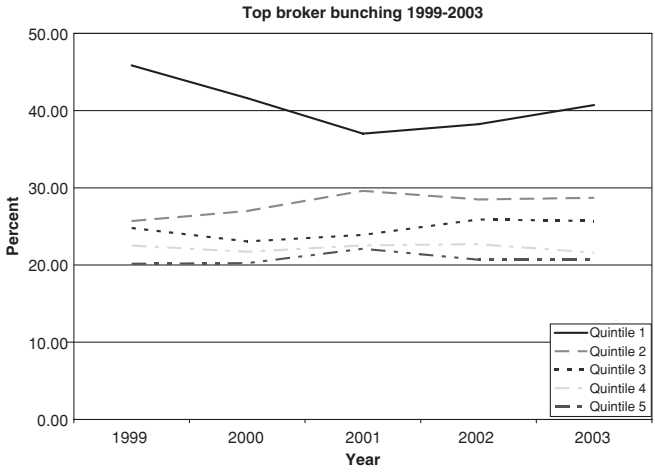

Figure 6
Institutional bunching over time
Top broker bunching is the average allocation to a client's most important broker. Cross-sectional averages for each client quintile and year are exhibited below.

replaced, we find that the replacement broker charges an average of 0.10–0.15 cent lower per-share commissions than the former top broker. When the former top broker is retained by the client, we find that they lower their commissions to conform to the replacement broker's price. Thus, despite the high retention ratios between specific clients and brokers, the threat of being replaced keeps commissions competitive.[31] Institutional bunching of order flow enhances the competitive threat; if institutions had dispersed their trades, a broker's rank with a client would be much less important.

Panel A of Table 5 also separates out low-cost trading commissions (not exceeding 3 cents per share). Low-cost execution does not vary much in the lower size quintiles, but the largest institutions use low-cost commissions for a greater proportion of their execution volume than do smaller institutions (F-statistic = 2.44). This is consistent with our first two hypotheses, as large institutions can easily pay for a broker's premium services with only a fraction of their total share volume, and so are free to execute a greater percentage of their trading at low prices.

Hypotheses 3c and 3s predict that large institutions will use more brokers than small institutions. As predicted, the average number of brokers used by institutions in each client quintile is increasing in the size of the institution. The smallest institutions use only 30.7 brokers on average, while the largest

[31] In unreported results, we also use Rule 606 (Dash-6) data to examine broker competition in trade execution. We find that newly promoted top brokers use more market centers and executed greater volume in alternative venues than existing top brokers. Based on the results in Boehmer, Jennings, and Wei (2007), who use Rule 605 (Dash-5) data to examine how execution quality affects order routing, we interpret this activity as greater effort on behalf of newly promoted brokers at seeking out low-cost execution for their clients. We thank the referee for this suggestion.

average 79.3 brokers. This pattern is also present in the low-cost market, where the smallest institutions use an average of 8.8 brokers, while the largest quintile uses an average of 13.3.

4.3.2 Additional tests of institutional bunching. We perform two additional tests on the degree of bunching. It could appear that large institutions do less bunching simply because they use more brokers, rather than choosing to concentrate their trading to earn premium services. To account for this fact, we conduct additional tests that restrict our attention to an institution's top ten brokers, which constitute the bulk of any institution's order flow and commission dollars.[32] First, we normalize to one the proportions of each institution's order flow to calculate each institution's Herfindahl-Hirschmann Index (HHI), which is the sum of squared proportions of volume sent to every one of the institution's top ten brokers times 100. The results are presented by client quintile in panel B of Table 5. It is clear that large institutions have a significantly more even distribution than the small ones, even after equalizing the number of brokers used. By comparison, a uniform distribution with ten brokers would yield an HHI of 10, compared to our findings of 17.8 for the largest institutions, and 23.6 for the smallest ones.

Next, we perform a parametric estimation of an institution's order flow allocation using the Zeta distribution, which is a discrete probability distribution commonly used in the natural sciences to measure concentration of types within a population. Let us denote by Z_k the proportion of volume that an institution sends to a broker ranked k ($k = 1$ being the largest) out of its K brokers. Zeta distribution implies that the proportion of volume allocated by an institution to its kth largest broker is

$$Z_k = \frac{C(K, \theta)}{k^{\theta}} \quad \forall k \leq K, \tag{3}$$

where $C(K, \theta)$ is a normalizing constant that increases in θ. Higher θ implies a less even distribution of volume allocation and hence a greater degree of order flow concentration. For example, $\theta = 0$ corresponds to a uniform distribution (HHI = 10), $\theta = 0.75$ corresponds to an HHI of about 14, whereas $\theta = 1.1$ corresponds to an HHI of about 20. Taking logs on both sides of Equation (3), we obtain an equation that allows us to estimate θ:

$$\log(Z_k) = \alpha - \theta \log(k) + \varepsilon, \tag{4}$$

where k is the rank of the executing broker for this institution. We perform the estimation separately for every institution in the sample, and then average the results by size quintile. Panel B of Table 5 clearly indicates that the Zeta

[32] We chose ten brokers to ensure that our tests include almost all clients. However, these results are also robust to alternative cutoffs.

distribution provides a good fit for client order flow. The distribution is significantly more concentrated for smaller institutions than for large ones, which is evident from the estimates of θ and the intercept. Order flow concentration declines with institution size and the effect is most pronounced for the smallest institutions. We also show the percentage of each quintile's θ coefficients that are significantly lower or higher than 1, which is the value of θ that coincides with the HHI of the median-sized institutions. The majority of the institutions in the largest quintile have θ significantly lower than 1 (75%), and only a small minority are significantly higher than 1 (14%), while the corresponding values for the smaller institutions are 42% and 47%, respectively. These results indicate that bunching is far more pronounced for smaller institutions, consistent with Hypothesis 2s, but inconsistent with 2c.

4.3.3 Commission size and broker rank. Hypothesis 4s suggests that, in addition to more extensive bunching, smaller institutions may also pay higher commissions to generate more profits for the broker and gain premier status. Panel C of Table 5 presents average per-share commissions for the institutions in each quintile.[33] This result is supportive of the hypothesis that smaller institutions are willing to pay more to get premium services from at least some of their brokers. While this difference is modest relative to the large differences in share volume across quintiles reported in Table 1, it may still affect small clients' relative positions with their brokers. As we have stated earlier, the high-commission share volume sent to a broker essentially determines the importance of an institution to a particular broker; nevertheless, consistent with Hypothesis 4s, the smallest institutions are willing to pay a per-share commission premium, particularly to their top brokers.

The cost minimization alternative suggests that volume allocation depends on broker size. Small institutions, which have no difficulty hiding their trades, should be indifferent between the large and the small brokers, while large institutions should strictly prefer larger brokers. Moreover, similar-sized brokers should get similar allocations from similar-sized institutions. If we assume that large brokers also provide more premium services, then the services hypothesis also suggests that large institutions will tend to use the largest brokers, as their volume ensures they will be important clients to any sized broker. For the smaller institutions, there is a trade-off between being a less important client for a large broker, or a more important client to a smaller broker. We do not know *a priori* their optimal choice.

Panel D of Table 5 examines the average and median broker size ranks (out of 270 active brokers) for an institution's five most important brokers, averaged within institutional size quintiles. The data reveal that each of the

[33] In Table 5, the structure of our tests mandates that per-share commissions are averaged quarterly; thus, they differ somewhat from the trade-weighted averages reported in Table 1. Given the overall decline in commissions over time (Figure 1) and the general increase in trading volume over time, the Table 1 averages weight the lower cost commissions later in the sample relatively more.

quintile average and the median ranks is below 50, which indicates that all institutions regardless of size concentrate their order flow with the largest brokers, presumably because this group provides the most valuable services.[34] Nevertheless, the comparison of means and medians also indicates that smaller institutions do tend to use somewhat smaller brokers as their top broker, which allows them to compete more effectively for these brokers' services.

Further, we know from panel A of Table 5 that an institution's top broker receives a much larger allocation of order flow than their fifth largest broker, yet the average size rank of the latter is lower than the average size rank for the top broker. These results show that similar-sized brokers receive vastly different allocations of order flow, and provide direct support for Hypothesis 5s, while contradicting Hypothesis 5c. Given the conclusions in Chan and Lakonishok (1993, 1995) that an institution's identity is the paramount factor in determining execution costs, the cost-considerations here are strong; thus, our results indicate that the institutions must place a high value on broker services to deviate from a strategy of hiding in the order flow as effectively as possible. Overall, our evidence is consistent with all clients concentrating their trades to capture the benefits from moving up higher in the queue for a broker's premium services. This pattern is most pronounced for small clients, where the benefits from bunching outweigh the potential costs.

Yet another way for small institutions to pay for services is to increase their volume of trading beyond what is required by their investment strategies, as stated in Hypothesis 4s. We test this hypothesis using Thompson's mutual fund quarterly holding data from 1997 to 2002. To avoid outliers, we first remove all the fund-quarter observations where the NAV was smaller than $10 million at the beginning of the quarter, or grew by more than 50% during the quarter. For each fund, we calculate the change in the number of shares of every security held over the course of the quarter, and treat it as the fund's trading volume in this security. We then multiply volume by the average quarterly price and aggregate over all securities, which yields an estimate of the total trading volume in dollars. Dividing trading volume by the NAV at the beginning of the quarter generates a turnover estimate. Each fund is then assigned to an NAV quintile and we calculate average turnover statistics by quintile. The annual averages are presented in Figure 7, which clearly shows that funds in the smallest quintile exhibit much higher turnover than funds in the two largest quintiles (the differences are significant at the 10% level).[35]

Although small institutions may have higher turnover for other reasons, these results are consistent with our interpretation of the market for brokers'

[34] This result is not tautological. Table 1 indicates that a broker's size rank in the sample is primarily determined by the largest institutions. Yet, the four smallest quintile institutions, whose allocations do not significantly affect brokers' size ranks, choose to concentrate their order flow with the same set of brokers as the largest quintile.

[35] Using the same technique, we also tested this hypothesis on the CDA/Spectrum data from 1994 to 2000, which provides quarterly holdings data on all investment managers with over $100 million in assets. We found that in four of Spectrum's five institutional type classifications, turnover significantly declines as size increases, as predicted.

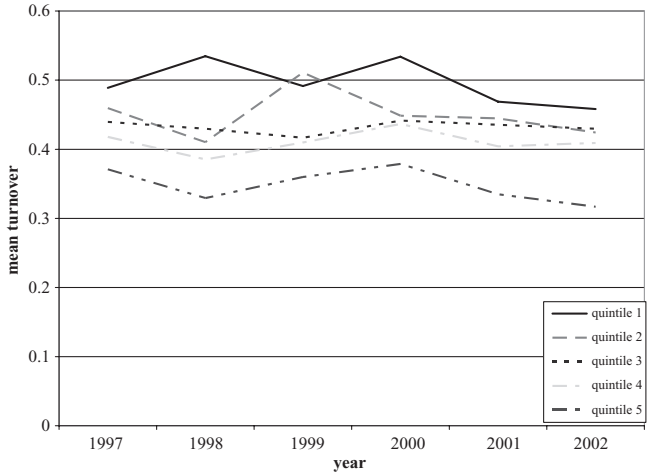

Figure 7
Mutual funds' turnover by year and size
This figure uses Thompson's mutual fund quarterly holding data from 1997 to 2002. We remove all observations where the fund was smaller than $10M at the beginning of the quarter, or grew by more than 50% during the quarter. For each fund we calculate the change in the number of shares of every security held by the fund over the course of the quarter, and treat it as the fund's trading volume. We then multiply trading volume by the average quarterly price and aggregate over all securities, yielding the estimate of total dollar trading volume. We then divide by the NAV at the beginning of the quarter to obtain an estimate of the turnover. Each fund is then assigned a NAV quintile and the average turnover statistics for each quintile are presented in the figure (Quintile 1 is the smallest institutions).

premium services. Small institutions that cannot generate sufficient brokerage revenues may attempt to increase their ability to procure premium services by concentrating their trading with only a few brokers, paying higher per-share commissions, and increasing their turnover to provide the required revenues to the chosen brokers.

5. Discussion and Conclusion

Brokers provide many valuable services that are difficult to sell explicitly. Moreover, the quality of these services is hard to evaluate over a short period, which calls for a long-term contract. Thus brokers need a simple mechanism that facilitates charging for these services over time. We conjecture that the total revenue a broker receives from a client during a period is a prearranged fixed fee for the level of services the client desires for that period. The client pays this fee through per-share commissions on trades sent to the broker. Thus, full-service commissions represent an average per-share cost of broker services. This framework sheds new light on trading volume allocation among brokers, and the possible future of the institutional brokerage industry.

First, we show that brokerage commissions are not set trade-by-trade, as assumed in the past, but rather determined in the context of a long-term contract. The distribution of institutional commissions indicates that proxies for the

execution costs of a trade are relatively unimportant determinants of per-share commission charges. Instead, past commissions are the strongest predictor of future commissions. This result is inconsistent with the view of commissions as a continuous execution cost negotiated on a trade-by-trade basis, yet supports our view that per-share commissions are a convenient way for institutions and brokers to track the revenues a client sends to a broker. Both parties need only concentrate on the volume of trade directed to a broker to calculate the payment rendered and gauge the importance of a client. Volume allocations accompanied by stable per-share commissions accumulate the fixed fee for the broker's services. A client that sends enough order flow to a particular broker expects to receive a premier level of service in return.

Viewing commissions as an average cost has important consequences for understanding the allocation of institutional order flow and the consequent payment of billions of dollars in commissions. We document that smaller institutions use fewer brokers than large institutions, at least partly due to the associated fixed costs, but also because it facilitates concentration (bunching) of their order flow with particular brokers. Institutions bunch their order flow with a small subset of brokers, from whom they receive premier services. We find that small institutions tend to concentrate their order flow significantly more than large institutions in order to become relatively important clients to a small set of brokers. These results are stable throughout our five-year sample period.

Bunching order flow is not an optimal strategy for hiding one's identity from the market. Therefore, if bunching partially reveals an institution's identity, it imposes significant price-impact costs on institutions. These costs must be offset by benefits to a bunching strategy. Understanding the costs and benefits of the commission contract is crucial for diagnosing the rapid changes in the full-service commission market. A soaring demand for liquidity has led to the emergence of alternative trading systems such as Liquidnet, UNX, and ITG, whose low-cost executions drive the significant decline in average commissions that we have documented. These alternative systems sometimes entirely bypass traditional brokers and many institutions implement order-routing programs that use brokers as one execution choice among many potential destinations for the order, diminishing the role of traditional full-service brokers.

At the same time, the value of the traditional premium services appears to be declining. The post-2000 IPO market offers fewer opportunities for brokers to allocate historically profitable IPOs than that of the late 1990s. Regulation Fair Disclosure, which restricts selective disclosure of management information, has reduced the precision of analyst information (Bailey et al. 2003). The 2003 adoption of the Global Settlement between the SEC and ten of the largest full-service brokers specifically restricts analysts' involvement in investment banking departments. This restriction, coupled with declining commissions, reduces the revenues supporting research departments. Kadan et al. (2006) document that from September 2002 to December 2004, the ten brokers covered

in the Global Settlement have among them discontinued coverage of 914 firms, or 12.2% of their covered firms. Smith and Linebagh (2006) contend that research cuts at Morgan Stanley are directly attributable to ECN competition for execution. Institutions have responded to the lower value of brokerage research and IPO services, as well as to the continuing pressure from their investors, by demanding lower commissions and more capital to facilitate their transactions.

Facing increased capital costs and lower commission revenues, brokers face a significant decline in the return on equity employed in the institutional trading business (Hintz and Tang 2003). As these changes continue, mid-size institutional brokers find it hard to maintain profitability in the current environment. This loss of profitability was the stated reason behind Wells Fargo's announcement in August 2005 of its complete withdrawal from the institutional equity business. Recently, Prudential Securities also announced their withdrawal from the institutional brokerage market.

At the same time, we show that even when faced with increasing competition and reduced ability to offer non-priced valuable services, the basic two-price commission market structure is still intact, which suggests that it provides economic benefits. Large brokers are able to maintain profitability through investment banking and by dramatically increasing the allocation of capital to proprietary trading. For a large broker, capturing order flow through internal execution (even at the discount commission levels) has the ancillary benefit of providing information to the broker's proprietary trading desk. Large brokers are thus involved in a race to provide low-cost, high-liquidity execution to their institutional clients, while at the same time extracting valuable information from the order flow they observe. Mid-sized brokers incur many of the costs of large brokers, but if they cannot invest in cutting-edge execution technology and proprietary trading, they may find the institutional equity business increasingly unprofitable and exit entirely. Consequently, the full-service segment of the institutional brokerage industry may become increasingly specialized, competitive, and concentrated in the near future.

References

Aitken, M., G. Garvey, and P. Swan. 1995. How Brokers Facilitate Trade for Long-Term Clients in Competitive Securities Markets. *Journal of Business* 68:1–33.

Bailey, W., L. Haitao, C. Mao, and R. Zhong. 2003. Regulation Fair Disclosure and Earnings Information: Market, Analyst and Corporate Responses. *Journal of Finance* 58:2487–514.

Bennett, J. 2002. A Closer Focus on Trading Costs. Report to Institutional Investors, Greenwich Associates, Stamford, CT.

Berkowitz, S., and D. Logue. 1987. The Portfolio Turnover Explosion Explored. *Journal of Portfolio Management* 13:38–46.

Berkowitz, S., D. Logue, and E. Noser. 1988. The Total Cost of Transactions on the NYSE. *Journal of Finance* 43:97–112.

Bertsimas, D., and A. Lo. 1998. Optimal Control of Execution Costs. *Journal of Financial Markets* 1:1–50.

Blake, R., and J. Schack. 2002. The Buy Side Wakes Up. *Institutional Investor,* April, 58–68.

Blume, M. 1993. Soft Dollars and the Brokerage Industry. *Financial Analysts Journal* 49:6–44.

Boehmer, E., R. Jennings, and L. Wei. 2007. Public Disclosure and Private Decisions: Equity Market Execution Quality and Order Routing. *Review of Financial Studies* 20:315–58.

Brennan, M., and T. Chordia. 1993. Brokerage Commission Schedules. *Journal of Finance* 48:1379–402.

Brennan, M., and P. Hughes. 1991. Stock Prices and the Supply of Information. *Journal of Finance* 46:1665–91.

Chan, L., and J. Lakonishok. 1993. Institutional Trades and Intraday Stock Price Behavior. *Journal of Financial Economics* 33:173–99.

Chan, L., and J. Lakonishok. 1995. The Behavior of Stock Prices Around Institutional Orders. *Journal of Finance* 50:713–35.

Conrad, J., K. Johnson, and S. Wahal. 2001. Institutional Trading and Soft Dollars. *Journal of Finance* 46:397–416.

Constantinides, G. 1986. Capital Market Equilibrium with Transaction Costs. *Journal of Political Economy* 94:842–62.

Corwin, S., and P. Schultz. 2005. The Role of Underwriting Syndicates: Pricing, Information Production and Underwriter Competition. *Journal of Finance* 60:443–86.

Copeland, T. 1979. Liquidity Changes Following Stock Splits. *Journal of Finance* 34:115–41.

Dermody J., and E. Prisman. 1993. No Arbitrage and Valuation in Markets with Realistic Transaction Costs. *Journal of Financial and Quantitative Analysis* 28:65–80.

Easley, D., and M. O'Hara. 1987. Price, Trade Size and Information in Securities Markets. *Journal of Financial Economics* 19:69–90.

Edmister, R. 1978. Commission Cost Structure: Shifts and Scale Economics. *Journal of Finance* 33:477–86.

Edmister, R., and N. Subramanian. 1982. Determinants of Brokerage Rates for Institutional Investors: A Note. *Journal of Finance* 37:1087–93.

Elton, E., M. Gruber, and S. Grossman. 1986. Discrete Expectational Data and Portfolio Performance. *Journal of Finance* 46:699–714.

Foucault, T., and G. Desgranges. 2002. Reputation-Based Pricing and Price Improvements in Dealership Markets. Working Paper, HEC and CEPR.

Fulghieri, P., and M. Spiegel. 1993. A Theory of the Distribution of Underpriced Initial Public Offers by Investment Banks. *Journal of Economics and Management Strategy* 2:509–30.

Glosten, L., and P. Milgrom. 1985. Bid, Ask and Transactions Prices in a Specialist Market with Heterogeneously Informed Orders. *Journal of Financial Economics* 14:71–100.

Green, C. 2006. The Value of Client Access to Analysts' Recommendations. *Journal of Financial and Quantitative Analysis* 41:1–24.

Greene, W. 1997. *Econometric Analysis,* 3rd ed. Upper Saddle River, NJ: Prentice Hall.

Hechinger, J. 2004. Deciphering Funds' Hidden Costs. *The Wall Street Journal,* March 17, D1.

Hintz, B., and K. L. Tang. 2003. U.S. Brokerage: Institutional Equities at a Crossroads, New York, NY: Sanford Bernstein.

Horan, S., and B. Johnsen. 2004. Does Soft Dollar Brokerage Benefit Portfolio Investors: Agency Problem or Solution? Law and Economics Working Paper 04-50, George Mason University.

Irvine, P. 2004. Analysts' Forecasts and Brokerage-firm Trading. *The Accounting Review* 79:125–49.

Jarrell, G. 1984. Change at the Exchange: The Causes and Effects of Deregulation. *Journal of Law and Economics* 27:273–312.

Jenkinson, T., and A. Ljungqvist. 2000. *Going Public: The Theory and Evidence on How Companies Raise Equity Finance.* Oxford, UK: Clarendon Press.

Johnsen, B. 1994. Property Rights to Investment Research: The Agency Costs of Soft Dollar Brokerage. *Yale Journal of Regulation* 11:75–113.

Kadan, O., L. Madureira, R. Wang, and T. Zach. 2006. Conflicts of Interest and Stock Recommendations—The Effects of the Global Settlement and Related Regulations. *Review of Financial Studies*, Advance Access published on January 12, 2009; doi:10.1093/rfs/hhn109.

Kavajecz, K., and D. Keim. 2005. Packaging Liquidity: Blind Auctions and Transaction Cost Efficiencies. *Journal of Financial and Quantitative Analysis* 40:465–92.

Keim, D., and A. Madhavan. 1997. Execution Costs and Investment Style: An Inter-exchange Analysis of Institutional Equity Orders. *Journal of Financial Economics* 46:265–92.

Kelly, K., and J. Hechinger. 2004. How Fidelity's Trading Chief Pinches Pennies on Wall Street. *The Wall Street Journal,* October 12, D1.

Kim, S. K., J. Lin, and M. Slovin. 1997. Market Structure, Informed Trading and Analysts' Recommendations. *Journal of Financial and Quantitative Analysis* 32:507–24.

Klein, B., and K. Leffler. 1981. The Role of Market Forces in Assuring Contractual Performance. *Journal of Political Economy* 89:615–41.

Kyle, A. 1985. Continuous Auctions and Insider Trading. *Econometrica* 53:1335–55.

Livingston, M., and E. O'Neal. 1996. Mutual Fund Brokerage Commissions. *Journal of Financial Research* 39:273–92.

Logue, D. 1991. *Managing Corporate Pension Plans.* New York: Harper Publishing.

Loeb, T. 1983. Trading Cost: The Critical Link between Investment Information and Results. *Financial Analysts Journal* 39:39–44.

Ljungqvist, A., F. Marston, and W. Wilhelm. 2006. Competing for Securities Underwriting Mandates: Banking Relationships and Analyst Recommendations. *Journal of Finance* 61:301–40.

Nimalendran, M., J. Ritter, and D. Zhang. 2007. Do Today's Trades Affect Tomorrow's IPO Allocation? *Journal of Financial Economics* 84:97–109.

Ofer, A., and A. Melnick. 1978. Price Deregulation in the Securities Industry: An Empirical Analysis. *The Bell Journal of Economics* 9:633–41.

Pethokoukis, J. 1997. The Thorn in Fidelity's Side. *U.S. News and World Report,* September 8.

Reuter, J. 2006. Are IPO Allocations for Sale? Evidence from Mutual Funds. *Journal of Finance* 61:2289–324.

Roll, R. 1984. A Simple Effective Measure of the Bid-Ask Spread in an Efficient Market. *Journal of Finance* 39:1127–39.

Schwartz, R., and B. Steil. 2002. Controlling Institutional Trading Costs. *Journal of Portfolio Management* 28:39–49.

Securities and Exchange Commission. 1998. Inspection Report on the Soft Dollar Practices of Broker-dealers, Investment Advisors and Mutual Funds. http://www.sec.gov/news/studies/softdolr.htm.

Smith, R., and K. Linebagh. 2006. Morgan Stanley Plans Reduction in Research Jobs. *The Wall Street Journal,* March 22, C1.

Sofianos, G. 2001. The Changing NASDAQ Market and Institutional Transactions Fees. Goldman Sachs Derivatives and Trading Research, May 31.

Timmons, H. 2000. I Kept on the Buy Rating, but I Told My Favorite Investors to Sell. *Business Week,* March 27.

Tinic, S., and R. West. 1980. The Securities Industry under Negotiated Commissions: Changes in the Structure and Performance of NYSE Member Firms. *Bell Journal of Economics* 11:29–41.

Vayanos, D. 1998. Transaction Costs and Asset Prices: A Dynamic Equilibrium Model. *Review of Financial Studies* 11:1–58.

Womack, K. 1996. Do Brokerage Analysts' Recommendations Have Investment Value? *Journal of Finance* 51:1137–67.