The Cost of Institutional Equity Trades

Donald B. Keim and Ananth Madhavan

Presented are an overview of the findings from the recent literature on the cost of U.S. equity trades for institutional investors and new evidence on trading costs from a large sample of institutional trades. The findings discussed have important implications for policymakers and investors: Implicit trading costs are economically significant; equity trading costs vary considerably and vary systematically with trade difficulty and order-placement strategy; and whether a trade price represents "best execution" depends on detailed data for the trade's entire order-submission process, especially information on pretrade decision variables, such as the trading horizon

his article provides an overview of the empirical evidence on the magnitude and determinants of equity trading costs. The focus is primarily on the trades of institutional investors. This topic has immediate practical value for investors, portfolio managers, exchange officials, and regulators. In addition, these groups have considerable interest in the relationship between the structure of security markets and trading costs. Indeed, the growth of alternative trading systems may be linked to efforts by large traders to reduce their trading costs. In this respect, institutional traders are of special interest because they account for a significant portion of equity ownership and trade larger volumes than retail traders.2

The increased interest in these issues has stimulated rapid growth in the literature on trading costs, much of which deals with methodological issues in cost measurement. Unfortunately, the data necessary to analyze many questions of interest are difficult to obtain. In particular, publicly available databases do not indicate whether a trade was a buy or a sell or whether a trade represented all or part of the desired order quantity. Furthermore, identifying the trades of institutional investors is difficult to impossible with publicly available data.

Recently, however detailed trading data from institutional traders have become available, which

Donald B. Keim is professor of finance at The Wharton School, University of Pennsylvania. Ananth Madhavan is professor of finance at the Marshall School of Business, University of Southern California.

greatly expands researchers' understanding of the trading process and costs. The objective of this article is to summarize the findings of the recent literature on equity trading costs. Specifically, we aim to (1) summarize the main methodological issues in measuring transaction costs, (2) review the current state of knowledge regarding the trading costs of institutional traders, (3) augment those findings with new evidence on trading costs from a large sample of institutional equity trades, and (4) outline the practical implications of the recent findings on trading costs for portfolio managers and for public policy.

Measuring Trading Costs

Analysts commonly decompose trading costs into two major components: explicit costs and implicit costs. Explicit costs are the direct costs of trading, such as broker commissions and taxes. Implicit costs represent such indirect costs as the price impact of the trade and the opportunity cost of failing to execute in a timely manner. Whereas explicit costs are associated with visible accounting charges, no such reporting of implicit costs occurs. As a result, considerable disagreement surrounds how best to measure implicit trading costs.³

The recent availability of high-quality transaction-level data on institutional trades has permitted more accurate measurement of trading costs than previously, which has increased considerably researchers' understanding of institutional trading costs. Exhibit 1 contains a brief description of the main findings of some recent studies of equity trading costs. The articles vary in both the data used and the empirical methods used, but the list indicates how thinking about equity trading

Exhibit 1. Recent Studies of Equity Trading Costs

Study	Data Source and Period	Sample Type	Major Findings For institutional trades, small sizes and low price impacts of 0.34% for buys and -0.04% for sells.		
Chan and Lakonishok (1993)	SEI Corporation, 1986–88	Individual trades			
Chan and Lakonishok (1995)	SEI Corporation, 1986–88	115,000 constructed "packages"	Packages are common. Investment style is important. Price impacts are 1% for buys and -0.3% for sells.		
Edwards and Wagner (1993)			Implicit timing and opportunity costs are significant and vary with market conditions. Total costs in neutral markets are \$0.85.		
Keim and Madhavan (1997)	Plexus Group, 21 U.S. institutional equity traders in 1991–93	25,732 orders	Costs vary with investment style, trader skill, trade size, and market capitalization. Total costs range from 0.20% to 2.87%.		
Keim and Madhavan (1996)			Pretrade price movements are important. Price impacts in small stocks are 3%–5%.		
Madhavan and Cheng Audit-trail data on large (1997) block trades		21,000 trades in DJIA stocks	Price impacts are very low, 0.16%–0.19%. Reputation affects upstairs market costs.		
Perold and Sirri (1993)	State Street Investments, international, 1987–91	International equity orders	Significant international variation. No corrections for trade difficulty.		
Stoll (1995)	Brokerage firm revenue, 1982–92	Indirect estimates of costs	Total costs range from 0.35% to 0.63% of the market value of equity.		

costs has evolved in the past few years. This section reviews the major components of trading costs in the context of the results of the studies in Exhibit 1.

Explicit Trading Costs. The main explicit cost is the commission paid to the broker for execution.4 Commission fees averaged 4-5 cents a share for shares trading below \$5 and increased with share price to as much as 15 cents a share in the 1991-93 period. Keim and Madhavan (1997) found that commission costs overall are low, about 0.20 percent of trade value. Stoll (1995) reported that commissions in 1992 averaged 7.9 cents, 0.24 percent of the market value of the trade. Edwards and Wagner (1993), in their comprehensive examination of trading costs, reported slightly lower commission costs, an average 5.6 cents a share. Commission costs also vary systematically by broker type and by market mechanism. For example, crossing systems charge commissions in the range of 1-2 cents a share, whereas the commissions charged by upstairs brokers may be as high as 10-15 cents. The variation in commission costs among studies is probably a result of differences in the sample institutions in terms of the types of trades, their degree of difficulty, and where the trades are executed.

Commissions paid by institutional investors have declined over time. Stoll (1995) reported that commission costs in 1982 were 17.8 cents a share, or 0.58 percent of market value, more than double the commission costs in 1992. The decline may be explained by the increasing institutional presence in the market, which may have produced a more competitive environment for trading services, one

in which institutions commonly negotiate lower commission rates. The decline in commission costs is also related to technological innovations in trading—for example, the increased use of low-cost electronic crossing networks by institutional traders. In addition, simple averages of stated commission costs may overstate the explicit trading costs for institutions. For example, brokers often return a portion of the stated commission in the form of soft-dollar payments to institutional investors, so the net costs are lower than stated (see Blume 1993 and Easley, Kiefer, and O'Hara 1996). This practice has grown since the late 1980s, which suggests that the true decline in commission costs is even larger than that documented by Stoll.

Implicit Trading Costs. Of primary interest to researchers and practitioners, and much more difficult to measure than explicit costs, are implicit trading costs—bid—ask spreads, price impacts, and opportunity costs.

Quoted bid-ask spreads. Early studies of implicit trading costs focused on the bid-ask spread as the relevant cost. The quoted bid-ask spread was considered the market maker's compensation for providing liquidity; thus, it was analogous to the commission cost charged by brokers (see Demsetz 1968). That the percentage bid-ask spread is related to the stock's liquidity, typically proxied by the stock's price per share or market capitalization, is well established in the literature. Estimates of the quoted spread as a percentage of the stock price vary widely, from less than 0.5 percent for the most liquid (largest market cap)

stocks to 4-6 percent for the most illiquid (smallest market cap) stocks (see Loeb 1983; Keim 1989; Sinquefield and Severoglu 1989; and Huang and Stoll 1996).

For several reasons, however, quoted spreads may be imprecise estimates of the true cost of transacting with a market maker. First, the quoted bidask spread may overstate the true spread because trades are often executed inside the quoted spread. This aspect is especially important for exchangelisted stocks. Second, both the bid and ask prices have a systematic tendency to rise (fall) following a buy (sell) order, so true round-trip trading costs are less than the quoted spread suggests. Third, large block transactions need not occur at the quoted bid or ask prices. For example, on exchanges, upstairs intermediation may lead to crosses between the quoted bid and ask prices, and other block trades may occur outside the quoted spread.⁶ On Nasdaq, large traders may negotiate prices directly with dealers, which again would lead to trades that occur outside or inside the quoted spread. These biases in quoted spreads are likely to be especially important for institutional traders, whose trades are, on average, much larger than those of retail traders.

Effective bid—ask spreads. To avoid the problems with quoted spreads, several authors have proposed measures of the "true" spread, often referred to as the "effective bid—ask spread." The effective spread is based on transaction prices, which may be more representative of market reality than quoted prices. Roll (1984) proposed one such measure, and George, Kaul, and Nimalendran (1991), and others, later extended it.

Implicit spread estimates exploit the fact that transaction prices tend to alternate between bid and ask prices (Niederhoffer and Osborne 1966). This bid-ask bounce induces a negative serial covariance between successive price changes. Thus, the serial covariance of successive price changes or returns can be used to derive an estimate of the underlying bid-ask spread and its components.

These serial covariance estimators of the effective spread tend to be smaller than the quoted spread. For example, based on daily returns for a sample of all NYSE and Amex stocks for the 1963–82 period, Roll found an average effective spread (across all stocks) of 0.298 percent, lower than the average quoted spread for the most liquid stocks. Madhavan, Richardson, and Roomans (1997) extended the approach to take into account mid-quote transactions and autocorrelation in the order flow and also found that effective spreads are significantly smaller than quoted spreads. Their estimates of the effective spread for a sample of 274 NYSE stocks in 1990 ranged from 7.3 to 8.6 cents over the day. In contrast, the average quoted spread for these stocks was almost three times larger, ranging from 21 to 22.8 cents over the day.

With quotation data, an alternative way to measure the effective spread is to use the average absolute price deviation from the prevailing midquote. Using this approach, Lee (1993) also found the effective spread to be significantly smaller (by as much as 50 percent) than the quoted spread. Lee estimated an effective cost of 9.6 cents a share on the NYSE but higher costs in most other markets, including Nasdaq. These studies reported that the same cross-sectional patterns documented for quoted spreads are also evident for effective spreads. Example 1 illustrates calculation of implicit costs using quoted spreads versus using the average absolute price deviation from the prevailing mid-quote.

Price-impact costs. Bid—ask spread estimates, although informative, fail to capture the fact that large trades, those that exceed the number of shares the market maker is willing to trade at the quoted bid and ask prices, may move prices in the direction of the trade. The resulting market impact or price impact of the transaction can be thought of as the deviation of the transaction price from the "unperturbed price" that would have prevailed had the trade not occurred. This definition also captures one-half of the bid—ask spread. Note that

Example 1. Calculation of the bid–ask spread component of implicit costs using quoted spreads versus using the average absolute price deviation from the prevailing mid-quote

Consider a buy order of 20,000 shares when the prevailing ask price is \$30.125 and the bid price is \$29.875. The trade is executed at the ask price, and the broker charges a commission of \$0.05 a share. The explicit cost is thus \$0.05 a share, or 0.2 percent of trade value. If quoted spreads are used to measure implicit costs, the one-way price-impact cost is measured by half the quoted bid-ask spread—that is, \$0.125, or 0.415 percent of trade value. If the average absolute price deviation from the prevailing mid-quote as in the Lee method is used, the effective spread measured relative to the mid-quote is still \$0.125, or 0.417 percent of trade value. Total trade costs are approximately 0.6 percent of trade value in either method.

the price impact of a trade can be negative—for example, if a trader buys at a price below the unperturbed price. Presumably, liquidity providers will enjoy negative costs whereas liquidity demanders will face positive costs.

Although conceptually simple, the price impact is difficult to measure because the unperturbed price is not observable. The unperturbed price is usually defined as a weighted average of the prices and quotes surrounding the trade. Differences in the weights placed on the pre- and posttrade prices yield different estimates of the unperturbed price and, therefore, different measures of the price impact

In the simplest weighting scheme, the unperturbed price is defined as the previous transaction price or the previous closing price. This measure is especially common in the literature on large-block trading, which documents significant price impacts associated with trades of 10,000 shares or more. The resulting price impact is then typically decomposed into permanent and transitory components, which provide estimates of the information and liquidity costs of the trade (see Example 2).

Block trades are likely to originate from institutional traders, which trade in larger volumes than individual traders. Block trades now account for almost 54 percent of NYSE trading volume, compared with 3 percent in 1965 (Schwartz and Shapiro 1992). The price impacts of block trades have been shown to be related to trade size and market capitalization (see Loeb; Holthausen, Leftwich, and Mavers; and Keim and Madhavan 1996). For example, Loeb, using quotations of block brokers, found that one-way trading costs can be significant for large trades in low-market-cap stocks. Loeb reported that the market impact of a large block transaction for stocks with market caps less than S25 million in 1983 often exceeded 15 percent. For large trades in liquid, large-market-cap stocks, Loeb found significantly smalle: market impacts, as low as 1 percent. Similarly for a more recent period, Madhavan and Cheng (1997) examined 21,000 block trades in the (very liquid) DJIA 30 stocks and found relatively small price impacts, 15-18 basis points. Keim and Madhavan (1996) developed and tested a model of large-block trading. They showed that block price impacts are a concave function of order size and a decreasing function of market capitalization (or liquidity). These findings are consistent with Loeb's results.

Keim and Madhavan (1996) also showed that the choice of pretrade benchmark price makes a large difference in the estimated price impact for large blocks. Using a sample of block trades from an institutional trader, they found that the average price impact for a seller-initiated transaction is -4.3 percent when the benchmark (unperturbed) price is the closing price on the day before the trade. When the benchmark is the price three weeks before the trade, however, the measured price impact is -10.2 percent, after adjustment for market movements. Although part of this difference in price impacts might be explainable by the initiating institutions placing the sell orders after large price declines, Keim and Madhavan (1995) found little evidence to suggest that institutional traders act in such a manner. Indeed, Nelling (1995), using the same sample of institutional transactions as Keim and Madhavan (1996) but with additional evidence on the length of time that the block was being "shopped," found no evidence that the trades were conditioned on prior price movements. Rather, Keim and Madhavan (1996) attributed the difference to information leakage arising from the process by which large blocks are shopped in the upstairs market. If leaks are the cause, their results suggest that previous estimates of permanent price impacts for block trades (i.e., the information components of the impacts) are downwardly biased. Furthermore, their findings suggest that a pretrade benchmark based on the date on which the decision to trade is made, if available, should be used when measuring block-trading costs. If the actual decision date is not available, the pretrade benchmark should attempt to capture any leakage related to the block trade.

Weighting schemes that place weight on past prices or quotes are subject to two important criticisms. First, if the proxy for the unperturbed price is known to the trader, this knowledge may affect the order-placement strategy. For example, a trader who knows his or her trades are measured against

Example 2. Decomposition of price impact into permanent and transitory components

Suppose a trader sells a block at \$97 and the pretrade price (e.g., the previous closing price) is \$100. The price impact of the trade is -3 percent. If the posttrade price (e.g., the next day's closing price) is \$99, the total price impact can be decomposed into a permanent component (which reflects the information content of the trade) of -1 percent and a temporary, or transitory, component (associated with the discount demanded by the block broker to accommodate the trade) of -2 percent.

a benchmark given by the previous closing price can achieve negative trading costs by placing buy orders only if prices have fallen from the previous day's close. Measured costs may be negative in this case, but that fact does not necessarily imply that the trader's prices do not move trades; hence, it does not imply superior performance. Of course, if a trader has little latitude over the trading decision or the timing of trades, gaming is not an issue, but this information is not readily available to an outside researcher. Second, in a dynamic context, weighting past prices may produce problematic findings. For example, a large trader who breaks up his or her orders into a sequence of subtrades may move prices over a long period of time. A naive pretrade benchmark may understate the true trading costs associated with the entire order because the benchmarks for the subtrades are moving in the direction of trading.

Some studies place no weight on past prices; instead, the researchers compare the trade prices to posttrade prices. Beebower and Priest (1980) used a weighting scheme that places all the weight on the closing price on the day following the trade. This approach overcomes the gaming problems, but it assumes that any liquidity effects arising from the trade are dissipated in a day. And again, this method may be appropriate for a single trade but perhaps not for a sequence of subtrades.

Berkowitz, Logue, and Noser (1988) suggested using a weighted average of transaction prices on both sides of the trade as a proxy for the unperturbed price. For example, the Abel/Noser Corporation uses a volume-weighted average of all transaction prices on the trade day to estimate this notional price. The rationale is that a weighted average of pre- and posttrade prices is an unbiased estimate of the prices facing a nonstrategic trader (i.e., a trader who places orders without paying attention to either intraday price dynamics or time of day) during the day of the trade. The Abel/Noser approach is compared with the posttrade approach in Example 3.

The previous criticisms regarding gaming, however, also apply to volume-weighted average price measures. The use of a volume-weighting scheme is also questionable for large block trades, especially in illiquid securities, for which the estimated benchmark essentially reflects the trade itself.

Opportunity costs. The final component of implicit costs is the opportunity cost associated with missed trading opportunities. The notion of an opportunity cost assumes the trade is motivated by information that has value that decays over time, so timely execution is necessary to capture the value. Such opportunity costs can arise for two reasons. First, some orders incur an opportunity cost because they are only partially filled or are not executed at all. Second, some orders are executed with a delay, during which the price moves against the trader. For index or passive investment managers, such opportunity costs are zero.

Development of a trade cost metric incorporating opportunity cost has proved to be difficult because its measurement requires knowledge of the date of the decision to trade. In one example, Treynor (1981) proposed measuring trading costs as the difference in performance between a portfolio based on the trades actually made and a hypothetical, or paper, portfolio whose returns are computed under the assumption that the transactions were executed at prices observed at the time of the trading decision and ignoring commissions, taxes, spreads, and so on. Perold (1988) termed this measure the "implementation shortfall." This measure accounts for the total trading costs associated with a package of trades, including the opportunity costs of failing to execute in a timely manner (see Example 4).

Unfortunately, researchers rarely have sufficiently detailed data to accurately compute opportunity costs. For example, to accurately measure the true costs of trading for a crossing system, researchers would need data on trades that never took place. Such information is unavailable. The measurement of opportunity costs is difficult also

Example 3. Comparison of price impact measured by trade price versus volumeweighted average price with impact measured by trade price versus posttrade price

Suppose in Example 1 that the value-weighted average price during the day is \$30.125, the previous day's closing price was \$29.50, and the next day's closing price is \$30.25. Relative to the volume-weighted average price, the price impact is zero, so total trading costs measured by the Abel/Noser approach are approximately 0.2 percent. Relative to the previous day's closing price, the price impact of the trade is (\$30.125 – \$29.50)/\$29.50, or 2.118 percent of trade value, so total costs are about 2.3 percent. Relative to the next day's closing price, the price impact is -0.413 percent. The posttrade price rise in the stock implies negative implicit costs. Total costs measured by the Beebower and Priest method, which involves comparing the trade prices to posttrade prices, are roughly -0.2 percent

Example 4. Trading costs calculated using implementation shortfall

In Example 3, suppose the price at the time the order was placed was \$27. Using Perold's method to compute the total implementation shortfall as the difference between the notional and actual returns produces total trading costs of 11.59 percent of trade value, which also happens to be the cost as measured by Keim and Madhavan (1997). Note that the opportunity costs from failing to execute in a timely manner are a substantial portion of this total cost.

because of "dynamic inconsistency." As illustrated in Example 5, the investor may have initially directed the trader to try a cross but then insisted on a partial execution at market prices. The change in stock price over the course of the trading period may have led to the investor scaling back his or her desired demand by 50 percent, so the true opportunity costs are lower than those measured by implementation shortfall.

These considerations illustrate the need for information on the underlying motivations for the trade (such as the investment objectives, target price, and trade horizon) to accurately measure opportunity costs. Several recent studies (e.g., Bodurtha and Quinn 1990; Perold and Sirri 1993, Keim and Madhavan 1996; and Leinweber 1995) have used the methods described with proprietary data to compute total trading costs that include opportunity costs.

Edwards and Wagner distinguished between desk timing costs, which arise from delays in placing an order, and opportunity costs, which arise from nonexecution. They estimated the timing costs as a function of whether the trade is liquidity demanding, liquidity supplying, or liquidity neutral, as proxied by market (momentum) conditions at the time of the trade. In markets where momentum is little (i.e., liquidity-neutral markets), their estimate of timing cost was \$0.07 (or about 0.20 percent of value). In markets where the order is liquidity demanding (i.e., a buy order in a rising market), they estimated the timing cost to be \$0.99 (or 3.56 percent).

The opportunity cost associated with failing to execute the entire order could be especially important for institutions using passive trading strategies, such as crossing systems or limit orders, where the risk of nonexecution is potentially significant. Because nonexecution is more likely to occur for traders trying to buy (sell) in up (down) mar-

kets, this cost could be quite high. Keim and Madhavan (1995) and Perold and Sirri found high rates of completion in institutional trades (typically about 90 percent), however, which suggests that the opportunity costs from failing to execute are low. Edwards and Wagner, in contrast, estimated the average opportunity cost to be 1.8 percent and found that it is significantly higher for large trades or trades in low-market-cap stocks. Edwards and Wagner estimated the total cost of trading to be \$0.85 per trade in neutral markets and \$1.87 in liquidity-demanding markets. For a stock trading at \$35, these figures translate to, respectively, 2.43 percent and 5.34 percent. They also found that the opportunity and desk timing costs represented a large percentage of total trade costs in their sample—85 percent in neutral markets and 90 percent in liquidity-demanding markets.

Practical Issues in Measuring Trade Cost.

Before turning to a detailed review of the empirical evidence from these many ways of measuring trading costs, this section summarizes the important considerations that should be kept in mind when assessing the various cost estimates reported in the literature.

The importance of measuring total costs. The previous discussion suggests that the individual components of transactions costs are economically significant. Obtaining a measure of total cost by simply adding up researchers' separate cost estimates is misleading, however, when those costs have been obtained from different studies, different institutions, or even different trades made by the same institution. The aggregation should take place at the transaction or order level. Keim and Madhavan (1997) estimated a positive correlation between implicit and explicit costs and noted that

Example 5. The effect of dynamic inconsistency on measurement of opportunity cost

In Example 4, suppose the original order is for 40,000 shares and the trader first directs the order to a crossing system but does not obtain execution. The trader then fills halt the order by buying at the market price. By failing to execute 20,000 shares of the order, the investor incurs an opportunity cost of approximately 56 percent in addition to the direct costs. Perold's implementation shortfall measure correctly captures the opportunity cost from failing to execute in this case.

explicit and implicit costs are jointly determined.⁸ Thus, making inferences about a trade's likely cost by simply adding unconditional estimates of implicit and explicit costs is misleading. Rather, the focus should be on total costs, especially when making intermarket cost comparisons.

The unit of observation. Another important issue in measurement is the unit of observation. Numerous studies have focused on trade costs at the level of the individual trade (see, e.g., Berkowitz, Logue, and Noser and Chan and Lakonishok 1993). Individual trades, however, are often part of a larger package of trades. The size of the package expresses the trader's desired order quantity more accurately than any individual trade; therefore, the package should be used to assess the associated price impact and opportunity costs. To see the rationale for this approach, suppose a trader wishes to buy 20,000 shares in an illiquid stock and, to fill the desired order quartity, makes five 4,000-share trades over a five-day period. If each individual trade occurred at the prevailing ask price, the implicit costs measured at the individual-trade level would be simply the spread costs. But if prices rose 15 percent over the five-day trading period, the transaction costs for the entire order would reflect the opportunity cost associated with that adverse price movement, so they would be much larger than the sum of the transaction costs of each individual component trade.9

Magnitude and Determinants of Costs

Given the importance of measuring total costs at the order level rather than the isolated cost components of individual trades, Perold's implementation shortfall approach is the most natural method to measure transaction costs. The data required to compute such costs were not available until recently, but studies by Edwards and Wagner, Perold and Sirri, Leinweber, Chan and Lakonishok (1995, 1997), and Keim and Madhavan (1997) provide estimates of total transaction costs associated with the entire order. Furthermore, these articles generally have two additional advantages over previous studies. First, the trades they examined are known to have been placed by institutional traders and, therefore, are not contaminated by possible differences between individual and institutional trade costs. Second, the data they used contained detailed information on the process by which the order was presented to the market, including the critical information of whether the order was to buy or sell. Although seemingly the most fundamental information, data on trade initiation have often been lacking, and researchers typically inferred the information indirectly from a comparison of prices with prevailing quotes. ¹⁰ Although this method may ensure a high degree of accuracy overall, the possible misclassifications may induce biases in cost estimates.

In this section, we present evidence on institutional trade costs based on the data on institutional trades used in Keim and Madhavan (1997). The data were obtained from the Plexus Group and identify the decision date, the actual order quantity, and component trades. That is, the 25,732 orders in the data file used here are *ex ante* expressions of desired trade quantities rather than the *ex post* approximations in other studies. Heim and Madhavan found that orders in this sample typically have a duration of one to two days.

We computed total trading costs using an approach similar to the implementation shortfall approach of Perold. For a buyer-initiated order, the implicit cost is the ratio of the volume-weighted average price of the component trades in the order to the closing price on the day before the decision to trade was made, minus 1. The implicit trade cost for a seller-initiated order is the negative of this price change. The implicit costs reported are not adjusted for market movements. The explicit cost for an order is in percentage form measured as the ratio of the dollar value of the commissions paid to execute the entire order to the total value of the order at the time of the decision to trade. Because 95 percent of the orders in our sample were filled entirely, we (unlike Perold) did not assign a cost to any portion of the desired order that was not executed.

Trade Difficulty. Previous research has shown that trading costs depend on trade difficulty. Simple estimates of average realized trading costs are relatively uninformative without some idea of the level of difficulty involved, as shown by Example 6.

The problem facing empirical researchers is to find adequate proxies for trade difficulty. Trade difficulty can be thought of as a function of two factors: (1) decision variables—that is, factors that are determined by the choices of the investor and trader—and (2) exogenous factors—that is, stock-specific factors outside the control of the individual trader. Deviously, the two types are interdependent, but this distinction is often useful in practice.

The literature on trading costs is a guide to the factors that determine trade difficulty. A list of some of the most frequently cited factors is given in Exhibit 2. Note that several studies identified market cap and trade size as the most important factors. Trades in large-cap stocks, which are more

Example 6. Relationship of trade cost to trade difficulty

Consider two trades in different stocks. In the first stock, which is illiquid, the trade is given to a full-service broker, who slowly "works" the order and incurs explicit costs of 0.9 percent and implicit costs of 2 percent. In the second stock, which is liquid, the trade is given to a discount broker whose explicit costs are 0.2 percent and implicit costs are 0.4 percent. Without controlling for the level of trade difficulty, a researcher cannot conclude that the broker with the higher total costs is a worse broker than the broker with the lower total costs. What if the assignments had been reversed? Having the higher-cost full-service broker execute the trade in the liquid stock would clearly not be optimal. And if the low-cost discount broker had been asked to effect the sale of the illiquid stock, that broker could have incurred implicit costs as high as, say, 3 percent.

This example also shows why simply adding estimates of implicit and explicit costs to predict realized costs is misleading. In this particular case, the two cost components are positively correlated.

liquid than small-cap stocks, have lower implicit and explicit costs than trades in small-cap stocks. Similarly, large orders, by demanding more liquidity, result in higher costs than small orders.

Trade size. Our analysis shows that trading costs are consequential and that a clear relationship exists between total trade cost and trade size for common stocks traded in the United States. Table 1 reports average trading costs by trade-size quartiles for trades of common stock by 21 institutions for the period January 1991 to March 1993. The results are reported separately for exchange-listed and Nasdaq stocks and for buy and sell transactions. There are several important findings.

First, trade costs are economically significant. For even the smallest (least difficult) trades of exchange-listed stocks, the round-trip costs are 0.64 percent. A 10 percent reduction in such costs would represent an economically significant boost to portfolio performance.

Second, a distinct relationship exists between trade size and total trade costs. The percentage costs for the largest trades are much larger than for the smaller trades.¹³ For example, the average round-trip cost for the largest trades ir. Nasdaq stocks is

4.43 percent. Furthermore, after controlling for trade size, we found that the average costs of trading tend to be larger for Nasdaq stocks than for exchange-listed stocks. The findings reported in Table 1 do not reveal whether this result represents differences in other measures of trade difficulty (e.g., market liquidity), differences in market design (for example, the difficulty of using passive trade strategies, such as limit orders, in a dealer market structure like Nasdaq [Greene 1996]), differences in dealer behavior (Christie and Schultz 1994 and Christie, Harris, and Schultz 1994), or differences in the way the order is presented to the market. To draw such inferences, one has to control for other determinants of trade costs. We discuss results when such controls are exercised later in the article.

In the only other set of comparable results for institutional traders, Chan and Lakonishok (1997) examined SEI Corporation data for a sample of 33 institutions for the 1989–91 period and found round-trip trade costs to be similar in magnitude to those reported in Table 1. They also found a similar relationship between cost and trade size. Chan and Lakonishok did not find a significant difference in costs among exchanges.

Decision Variables	Exogenous (Stock-Specific) Variables Relative illiquidity (Loeb 1983; Keim and Madhavan 1996, 1997; Chan and Lakonishok 1995)			
Order size relative to average daily volume or shares outstanding (Loeb 1983; Chan and Lakonishok 1988; Keim and Madhavan 1996, 1997)				
<i>Trade duration</i> or number of subtrades (Chan and Lakonishok 1995; Keim and Madhavan 1997)	Market momentum or stock volatility (Edwards a Wagner 1993)			
Order type—e.g., market, limit, working, crossing; Bodurtha and Quinn (1990)	Market design (Chan and Lakonishok 1997; Keim and Madhavan 1997)			
Investment style (Chan and Lakonishok 1995; Keim and Madhavan 1997)	The trader's skill (Chan and Lakonishok 1995; Kein and Madhavan 1997)			
Upstairs intermediation, which may reduce the price impact of the trade (Keim and Madhavan 1996; Madhavan and Cheng 1997)	The trader's reputation or ability to signal that the trade is not information motivated (Madhavan and Cheng 1997)			

Table 1. Average Trading Costs by Trade-Size Quartile for Common Stock
Trades by 21 Institutions, January 1991–March 1993
(standard errors in parentheses)

	Exchange-Listed Stocks ^a				Nasdaq Stocks	
Trade-Size Quartile	Total Cost	Implicit Cost	Explicit Cost	Number of Trades	Total Cost	Number of Trades
Buyer-initiated trades						
1: Smallest	0.31%	0.18%	0.13%	7,392	0.76%	1,755
	(0.02)	(0.02)	(0.00)		(0.06)	
2	0.36	0.19	0.17	6,577	1.01	2,571
	(0.03)	(0.03)	(0.00)		(0.07)	
3	0.53	0.32	0.21	6,503	1.08	2,645
	(0.04)	(0.04)	(0.00)		(0.09)	
4: Largest	0.90	0.65	0.25	5,570	1.80	3,577
	(0.05)	(0.05)	(0.00)		(0.10)	
Seller-initiated trades						
1: Smallest	0.33	0.15	0.18	5,736	0.29	696
	(0.03)	(0.03)	(0.00)		(0.12)	
2	0.31	0.11	0.20	5,291	0.50	1,142
	(0.04)	(0.03)	(0.00)		(0.11)	
3	0.38	0.17	0.21	4,766	0.71	1,666
	(0.04)	(0.04)	(0.00)		(0.11)	
4: Largest	1.42	1.13	0.29	3,830	2.63	2,602
	(0.08)	(0.08)	(0.00)		(0.14)	

Notes: Implicit trading costs were defined as (P_a/P_d) – 1, where P_a is the average price of all the executed trades in the order and P_d is the closing price for the stock on the day before the decision to trade the stock. Explicit trading cost was defined as (Commissions per share/ P_d). Trade-size quartile was defined as number of shares traded divided by total outstanding shares; quartile cutoffs were determined separately for buy and sell transactions.

Market liquidity. Market capitalization of the company is also directly related to the costs of trading that company's stock. Table 2 reports average trading costs by market-cap quintile for trades of common stock by 21 institutions for the period January 1991 to March 1993. Again, the cost estimates are economically significant and the unconditional costs of trading in Nasdaq, except for the largest quintile of Nasdaq stocks, appear higher than costs for trading exchange-listed stocks.

Most pertinent to this discussion is that for both exchange-listed and Nasdaq stocks, the total trading cost decreases monotonically with market capitalization. For example, the average total round-trip cost for the smallest-market-cap quintile for exchange-listed stocks is 3.81 percent. The same cost for the largest-market-cap quintile is 0.57 percent. Chan and Lakonishok (1997) reported similar magnitudes for their sample of 33 institutions.

Finally, Table 2 indicates that sell costs generally exceed buy costs. This result may be driven by differences in trading size; sell orders are often larger than buys (most block trades are sells) for reasons that are not well understood.

Investment Style, Trading Skill, and Repu-

tation. Recent evidence shows that trade costs are also affected by other factors in addition to trade difficulty. As shown by Chan and Lakonishok (1995), Keim and Madhavan (1997), and Leinweber, an important variable is the trader's investment style or strategy. Investment style (e.g., index, momentum) affects trading costs because it proxies for unobservable factors, such as the trader's time horizon or aggressiveness in order submission. Thus, aggressive traders following short-run technical trading strategies will, other things being equal, have higher expected costs than less-aggressive traders whose strategies are based on fundamental analysis. The reason is that the technical traders demand (and are willing to pay for) immediacy.

Keim and Madhavan (1997) found significant differences in average costs among traders with different styles. They distinguished between "value," "technical," and "index" investment managers. Value managers, defined as those who use fundamental analysis, tend to have low costs because their investment horizons are relatively long, which provides them the luxury to trade patiently. Indeed, if value traders are willing to supply liquidity, they

aNYSE and Amex.

Table 2. Average Trading Costs by Market-Cap Quintile for Common Stock
Trades by 21 Institutions, January 1991– March 1993
(standard errors in parentheses)

		Exchange-L	Nasdaq Stocks			
Market-Cap Quartile	Total Cost	Implicit Cost	Explicit Cost	Number of Trades	Total Cost	Number of Trades
Buyer-initiated t	trades					
1: Largest	0.31%	0.17%	0.13%	10,960	0.24%	1,155
	(0.02)	(0.02)	(0.00)		(0.11)	
2	0.43	0.28	0.17	7,989	0.51	1,934
	(0.03)	(0.03)	(0.00)		(0.09)	
3	0.64	0.41	0.24	4,137	0.92	2,929
	(0.06)	(0.06)	(0.00)		(0.08)	
	1.00	0.70	0.30	2,115	1.52	2,720
	(0.07)	(0.08)	(0.00)		(0.09)	
5: Smallest	1.78	1.35	0.42	834	2.85	1,801
(0.	(0.12)	(0.12)	(0.01)		(0.13)	
Seller-initiated t	rades					
1: Largest	0.26	0.11	0.15	10,901	0.16	960
	(0.02)	(0.02)	(0.00)		(0.12)	
2	0.63	0.41	0.23	4,738	0.85	853
	(0.04)	(0.05)	(0.00)		(0.18)	
	1.02	0.72	0.30	2,296	1.18	1,517
	(0.00)	(0.09)	(0.00)		(0.12)	
4	1.33	0.92	0.41	1,112	1.73	1,613
	(0.16)	(0.15)	(0.01)		(0.15)	
5: Smallest	2.03	1.36	0.67	568	2.91	1,106
	(0.23)	(0.23)	(0.02)		(0.23)	

Notes: The sample is partitioned by market capitalization with cutoffs determined by NYSE quintile break points as of December 1991. See also the notes for Table 1.

may even have negative trading costs as measured in thisarticle. Technical traders and indexers, whose investment styles rely on rapid execution, have higher costs. Value managers should have the lowest costs, and the average costs of the technical and index managers should be substantially higher. Keim and Madhavan's findings are consistent with these predictions: They found est mated round-trip costs for value traders to be 0.45 percent, for index traders, 1.09 percent, and for momentum (or technical) traders, 2.04 percent.

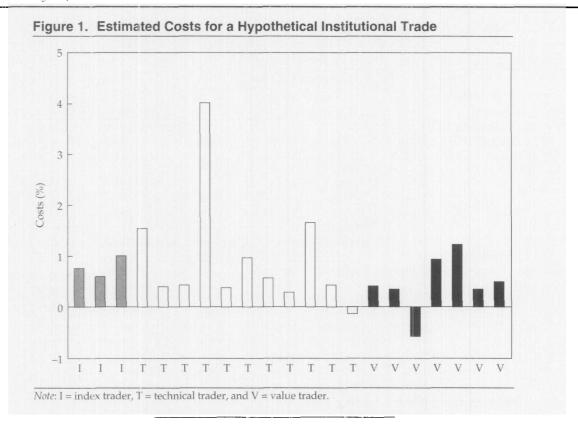
Chan and Lakonishok (1995) also found that the style of the manager affects measured costs. High-turnover growth-stock managers have high costs, and low-turnover value-stock managers have negative costs. This finding is also consistent with the preceding predictions.

Even within a particular investment style, differences in order-submission strategy may have significant effects on costs. For example, two traders using fundamental- (or value-) based trading strategies may have significant differences in the number of component trades necessary to fill an

order, which may translate into cost differences. Using the regression techniques described in the next section, Keim and Madhavan (1997) found significant differences in trading costs among traders within the same investment style even after correcting for trade difficulty. Based on this regression approach, Figure 1 shows the estimated trading costs separately for each of the 21 institutions in our sample for a typical trade in the sample. The computations assume an order size of 35,000 shares (on a base of 30 million outstanding shares) in a Nasdaq stock with a price of \$34. The figure illustrates the wide variation in costs among institutions and among styles. The differences in costs even within investment styles shown in Figure 1 may reflect such unobservable factors as traders' different skills and abilities.

Reputation also affects trading costs. Traders who have a reputation for liquidity trading may be able to obtain better prices because the adverse-selection costs associated with their trades are likely to be minimal. This advantage is especially likely for those trades that are negotiated in the

aNYSE and Amex.



upstairs market, because the upstairs market is less anonymous than the exchange floors or Nasdaq (Keim and Madhavan 1996). Madhavan and Cheng (1997) partitioned a sample of 22,000 block trades in DJIA stocks into upstairs and "downstairs" trades. Using an econometric model that corrects for selectivity biases, they found strong evidence that trading costs reflect the effect of an unobserved reputational variable.

Studies by Chan and Lakonishok (1995, 1997); Loeb; Edwards and Wagner; Leinweber; Perold and Sirri; Stoll; and others, complement the findings reported in Tables 1 and 2, and those authors came to similar conclusions. Readers should note several key points, however, about the cost estimates in the studies. The first point is the wide variation in cost estimates. Some of this variation is undoubtedly a result of differences in cost measurement and/or data, but even relatively homogenous samples produce a high degree of unpredictability in costs (see Leinweber, for example). Also, mean cost estimates are difficult to interpret without reference to the complexity of the underlying order—the subject to which we now turn.

Style. The discussion so far clearly mandates that cost estimates be interpreted in the overall context of trade difficulty and investment style. If a trader believes that a stock will appreciate 10 percent in

the next few days, the trader may be willing to bear substantial costs to ensure the execution of large trades of that stock within a short period. In such circumstances, trade difficulty and costs will both be high, but even with total costs of 4 percent, the investor will probably be happy with the trade. If a trader has no private information but simply seeks a position for liquidity reasons, the trader will be patient. In this case, the trading costs will be smaller, but the cost of the trade will probably represent a large portion of the expected return from the planned trade. Thus, one cannot conclude that higher-cost trades simply provide poorer execution; "best execution" for an informed trader is not identified solely by lower cost. In short, the quality of a trade can be assessed only when its cost is measured relative to a benchmark cost that incorporates the difficulty of the trade, the market environment, and the investment style motivating the trade. This notion is analogous to using riskadjusted performance measures in performance evaluation (the Jensen measure, for example).

Keim and Madhavan (1997) proposed using a regression-based approach to measure such relative costs. The idea was to assess a trade's cost relative to a benchmark that corrects for trade difficulty, stock-specific factors (such as exchange listing), and investment style. To understand this approach, consider a trade whose realized execu-

tion cost is 3 percent. Although this figure may appear large, suppose the predicted, or benchmark, cost for this trade (based on a regression model that takes into account trade difficulty, market liquidity, and order-placement strategy) is 3.5 percent. In this case, the *trader-specific cost* is -0.5 percent, which indicates that the trader outperformed the benchmark.

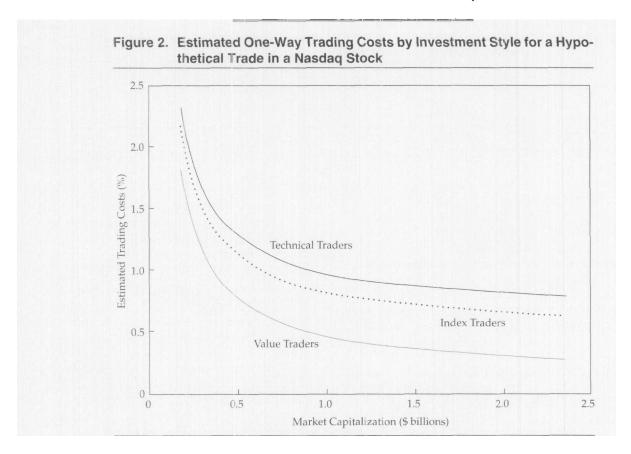
An illustration will clarify the practical application of this approach. Figure 2 shows the predicted percentage trading costs for a hypothetical buy order in a Nasdaq stock by investment style as a function of the stock's market capitalization. It is based on the regression analysis by Keim and Madhavan (1997) of the trades of 21 institutional traders from January 1991 through March 1993. The regression specifically controls for the influence of trade venue (Exchange versus Masdaq), trade size (computed relative to total outstanding shares), log of market capitalization, share price, and investment style. Figure 2 assumes a trade horizon of one day and trade size equal to 35,000 shares on a base of 30 million outstanding shares. Little difference shows up between the predicted trading costs of the index traders and the technical traders; the costs of the value traders are considerably lower. In all cases, the predicted costs decrease as market capitalization increases.

Figure 3 presents a similar plot for execution on the NYSE. The costs are lower than shown for trading on Nasdaq for all three trader types, but the effect is dramatic for the value traders, who are headed for negative trading costs for stocks of very large market capitalization. This cost fall presumably reflects the increased availability, and use, of low-activity order-submission strategies on the NYSE.

The figures can be used to assess the performance of traders after correcting for trade difficulty and style. Suppose a trader incurred total trading costs of 1 percent for the hypothetical Nasdaq trade described for Figure 1 for the stock of a company with market capitalization of \$1.7 billion. Based on Figure 1, if the trader was following a value strategy, the 1 percent cost was abnormally high, by about 80 basis points. But if the trader's objective was to mimic an index, the cost was abnormally low, by approximately 20 basis points.

Implications of the Findings for Public Policy

The findings on trading costs and the recognition that trading costs must be measured relative to a benchmark that takes into account trade difficulty and investment style as well as execution costs have



July/August 1998

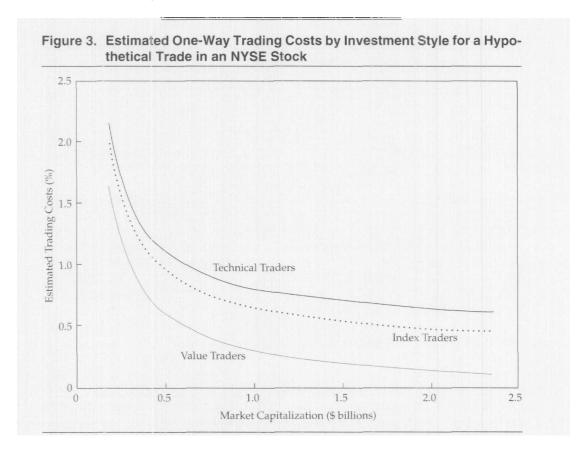
implications for defining and interpreting best execution, for analyzing the efficiency of current trading systems, and for the design of new market trading mechanisms.

Best Execution. Brokers assume an agency responsibility to assure best execution for the customer trades for which they have fiduciary responsibility. ¹⁴ Although the term "best execution" is not well defined, it is typically interpreted to mean trading at the most favorable price available in the market. The findings reported here, however, raise questions about the practicality of this objective for individual and institutional investors. Specifically, to enforce compliance with the best execution requirement, policymakers and plan sponsors must be able to measure "best price," which is not straightforward.

For small trades, the total costs of trading consist of the commission costs and the bid-ask spread costs. In general, such trades rarely incur significant price-impact costs, and if they are executed in a single market order, the opportunity costs (from timing and failure to execute) are also negligible. Thus, for retail trades, the best price available is often the best bid or offer quote prevailing in the entire stock market. As noted by Harris (1996) and Macey and O'Hara (1996), however, even this definition is too simplistic in today's equity markets.

Such factors as payment for order flow, "preferencing" arrangements (arrangements to direct portions of the order flow to certain market makers), and insufficient exposure of limit orders greatly complicate the measurement of best price even for retail orders, as Example 7 spells out.

Best execution is even harder to define for institutional traders, whose orders tend to be much more complicated than retail traders'. First, because of the size of their orders, institutions typically follow dynamic order-placement strategies and break their orders up into several component trades (see Chan and Lakonishok 1995 and Keim and Madhavan 1995). As a result, and as pointed out by Bertsimas and Lo (1996), best price at the time of a trade must be defined in the context of the overall orderplacement strategy---especially because market movements make price-impact costs difficult to measure in a dynamic context. Furthermore, as shown by Keim and Madhavan (1996), information leakage about the order throughout the trading horizon may significantly affect measured price impacts. Second, as shown by Chan and Lakonishok (1995), large differences may exist between costs measured at the individual trade level and those measured at the order level. These differences arise from timing and opportunity costs (see Edwards and Wagner), which are typically ignored in computations of best execution. Third, as shown



Example 7. Order preferencing and best execution

Suppose the best intermarket quotes are \$50.00–\$50.25 and a retail investor directs a broker to buy 100 shares. The broker sells the order to a third party for \$0.01 a share, and the trade is executed at \$50.25. Although the broker may have met the usual standards for best execution, the investor might have done better if the broker had directed the order to an exchange, where it might have been executed (with, say, a 30 percent probability) at the mid-quote.

by Keim and Madhavan (1997), and others, institutions vary greatly in their willingness to bear costs. For example, institutions following active momentum-based investment strategies that demand immediacy are willing to trade off higher execution costs against the expected performance from their investment ideas. For institutions supplying immediacy, execution costs should be small or even negative. Consequently, a workable definition of best execution for institutional traders must consider the investment objectives of the investor. Indeed, Edwards and Wagner argued that the definition of best execution for institutional traders must be expanded "to include prudent control of the entire implementation of the investment idea" (p. 65). Unfortunately, because trading objectives are difficult to quantify and need not be constant over time, actual implementation of this sophisticated concept of best execution appears unlikely.

In summary, a simple and meaningful definition of best execution is unworkable, especially for the largest traders. Therefore, oversight of best execution may best be left to the competitive pressures of the market.

Costs under Alternative Market Systems.

To use the estimates of trading costs to make inferences about the relative efficiency of alternative trading venues is natural. In particular, researchers have shown considerable interest in comparing execution costs on the NYSE and Nasdaq. The NYSE operates as a specialist auction market, where immediacy is provided by public limit orders and an exchange-designated specialist. Nasdaq is a dealer market, where multiple market makers post quotes prior to trading. The extent to which these differences in market structure affect execution costs is an important issue.

Some evidence (see Lee; Blume and Goldstein 1992; and Huang and Stoll, among others) suggests that quoted and effective spreads on Nasdaq stocks are generally wider than on comparable exchange-listed stocks. Huang and Stoll compared execution costs along several dimensions (including effective and quoted spreads) and found that costs for Nasdaq stocks are almost twice as high as for a comparable sample of exchange stocks. They attributed this difference to institutional features of the Nasdaq market, especially arrangements to

direct (or preference) portions of the order flow to certain market makers (see also Christie and Schultz; Christie, Harris, and Schultz; and Dutta and Madhavan 1997). Lee found that spreads for non-NYSE trades are 0.7–1.0 cent a share greater than for NYSE trades, with Nasdaq offering worse execution in all size categories. For retail traders, whose trades are typically quite small, the price impact and opportunity costs in both markets are likely to be negligible. Thus, the findings for spreads suggest that retail investors face higher execution costs on Nasdaq than on the NYSE.

These results do not apply to large traders (typically institutions), which may negotiate prices that differ from the posted bid and ask prices through, in essence, bilateral negotiations. To the extent such price "discounting" occurs, a naive comparison of quoted bid—ask spreads among market structures may be misleading, especially for institutional traders. Complicating matters is the fact that for many institutional trades on Nasdaq, the commissions are built into the price, so researchers need to examine total costs.

Keim and Madhavan (1996) and LaPlante and Muscarella (1997) compared the price impacts of comparable block trades on listed exchanges and Nasdaq. Their results suggest that large traders may obtain better liquidity on the exchanges. Chan and Lakonishok (1997) and Keim and Madhavan (1997) compared the total execution costs to large institutional traders on the NYSE and Nasdaq. Both studies controlled for trade size and market capitalization and found mixed evidence. Chan and Lakonishok found that the costs of trading are lower on the NYSE for large firms but that the opposite is true for small firms. Controlling for industry, the stock price level, volatility, and trade duration, they did not find costs to be uniformly higher on one exchange than on the other. The evidence in Keim and Madhavan (1997), who also controlled for investment style and trader skill, is somewhat more conclusive. Their unconditional average cost estimates (see also Table 2) indicate that trading costs are higher on Nasdaq than on the NYSE and Amex for all but the largest stocks, but their regression analyses indicate that, whereas Nasdaq buy trades are more expensive than comparable exchange buy trades, sell trades have no

significant differences in costs between Nasdaq and the exchanges. Keep in mind that these results apply only to institutional trades, not to small retail trades that typically execute at the quoted bid or ask price.

New Trading Systems. The estimates of costs have implications for the design of market trading mechanisms. The estimates may also help explain some developments in the U.S. equity markets in recent years.¹⁷ In particular, the rapid growth of electronic crossing systems, such as Posit, Instinct Corporation's crossing system, and the NYSE's after-hours crossing system, is often linked to efforts to reduce transaction costs. Electronic crossing networks function by trying to match the natural buyers and sellers of a security at predetermined prices. Unlike traditional exchanges, the transactors themselves provide the liquidity; middlemen or dealers are not required. Electronic crossing networks typically do not provide independent price discovery; rather, the predetermined prices at which buvers and sellers trade are usually determined in other markets.

The evidence suggests that crossing systems offer substantially lower execution costs than traditional exchanges. Crossing commissions are usually below 2 cents a share, much less than the charge of full-service brokers on exchanges. Moreover, participants also obtain substantially lower implicit costs because of the lack of any bid-ask spread (because the traders provide the liquidity) and the lack of any impact cost (because the trade price is independent of order size).

Readers should note some important aspects of crossing systems, however, in interpreting the findings about costs. First, the crossing systems provide no guarantee that an investor's order will be executed. The cost estimates researchers attribute to an observed transaction of this type often understate the true cost of trading because the opportunity costs of failing to execute are ignored. The situation is even more complex when multiple trading opportunities are available. For example, suppose a trade did not fill at all on the crossing system and the trader later submitted it as a set of market orders for execution at the posted prices. If the price of the stock rose in the interim, the costs of failing to execute through a cross are part of the opportunity costs or timing costs of the order trades. Researchers lack sufficient data to measure the opportunity costs on crossing systems from partial (or failure of) executions, so the economic significance of these costs is unclear.

Second, executions on crossing networks may be associated with adverse-selection costs that affect

subsequent investment performance. For example, consider an investor who places a buy order for 50,000 shares on a crossing system for execution at the prevailing mid-quote. If the trader is liquidity motivated, an execution when the stock price rises is less likely than when it falls because informed traders will place competing buy orders if the stock is likely to appreciate in value. (Incidentally, this phenomenon suggests that researchers can correlate the estimated opportunity costs of failing to execute in a timely fashion with future performance as a way to assess the trader's information.)

Finally, even if crossing networks do provide lower execution costs than traditional exchanges, the lack of an intrinsic price-discovery mechanism suggests natural limitations for this type of trading system. Externalities compound the problem; if the traders using crossing systems are liquidity traders, the primary market may experience higher adverse selection costs and, hence, lower liquidity as these traders migrate to cheaper trading mechanisms. Conversely, the existence of low-cost crossing systems allows institutional traders to trade with one another without adversely affecting primary market liquidity. For example, large block trades placed by a pension fund rolling over its portfolio may generate significant temporary price impacts if sent directly to the downstairs market, because floor traders and market makers cannot distinguish between information and liquidity motives for the trades. Such transitory price volatility may discourage trading by small retail investors and may take away liquidity from the market. In contrast, the existence of a crossing system allows the fund to trade without large price movements and may contribute to a more liquid and efficient market in the long run. The net impact of these factors is an important issue for regulators, exchanges, and policymakers.

Implications for Investors and Portfolio Managers

Leinweber noted that the Value Line Group I stocks had an annualized return of 26.3 percent for the 1979–91 period but the Value Line mutual fund that contains the same stocks returned only 16.1 percent for the period. The difference between the paper return and the actual portfolio return is the cost of trading. Obviously, transaction costs are economically significant issues in portfolio management.

Active and Passive Management Styles.

Investment management styles are often grouped into two broad genres, active and passive. Passive management may involve the design of an investment vehicle and control of its risk but does not involve trading on speculation, information, or momentum. Underlying a passive strategy is the idea that the value added by trading stocks in response to possible market inefficiencies does not outweigh the transaction costs associated with such an active style of management. Passive portfolios are often simply proxies for an underlying index (e.g., the Russell 2000 Index or S&P 500 Index), but a strategy may be further tailored to suit a specific investment style (e.g., value or growth).

Active portfolio managers seek to identify mispriced securities or economic sectors; their underlying belief is that the value from exploiting such market inefficiencies can be large.

Active strategies generally involve substantially more trading than passive strategies, and the trade immediacy often demanded by active managers may be inferred by market makers as information motivated. Passive managers, on the other hand, may be able to signal that their trades are liquidity motivated by using limit orders or upstairs intermediation (see Keim and Madhavan 1996 and Madhavan and Cheng 1997). In combination, these two attributes of active management informationally motivated trades and a demand for immediacy—result in trading costs that are substantially higher than for passive managers, as illustrated in Example 8. Therefore, when transactions costs are considered, passive indexing strategies may dominate active management strategies even if active managers can add value by identifying mispriced securities. Given the voluminous evidence on the unprofitability of active portfolio management, we suggest that resources would be put to better use in attempting to understand and reduce trade costs than in trying to exploit scarce market inefficiencies.

Construction of Passive (Indexed) Portfolios. Another area where transactions costs are important is in the construction of passive or

indexed portfolios. The traditional method of constructing an index fund is by duplication; that is, seeking exact replication of the target universe, the fund holds all the stocks in the underlying index, weighted by their market capitalizations. Fund inflows or outflows give rise to trades as the portfolio quickly adjusts to mimic the benchmark. Duplication helps minimize tracking error, but the transaction costs incurred can be significant. Keim and Madhavan (1997) found one-way trade costs for index managers in their sample to be 0.37 percent for buys and 0.38 percent for sells.

Trade costs may be of secondary importance for an indexed or passive portfolio containing liquid securities (such as a fund indexed to the S&P 500), but passive portfolios mimicking an index of illiquid securities (small-cap stocks or some value indexes), given the high costs of trading illiquid stocks, can incur trading costs that are large enough to create significant performance shortfalls. An alternative to such a "pure" indexing strategy is one that sacrifices tracking accuracy by allowing actual portfolio weights to deviate from the underlying index, thereby reducing the volume of trading and corresponding trading costs. Sinquefield (1991) examined the performance of four indexed small-cap funds to illustrate the impact indexing techniques can have on the investment performance of a passive portfolio of illiquid securities. He found that for his sample period, the pure indexed portfolios had the lowest returns whereas the portfolios for which trading costs were minimized at the expense of tracking accuracy had the highest returns. In an analysis of a well-known passive small-cap fund that pursues a more flexible strategy, Keim (1998) documented that the fund's trading strategy added 17 basis points a month to performance. This impressive performance is in contrast to the average drag on performance of 191 basis points associated with the one-way trade costs of exchange-listed stocks in the smallest quintile of market capitalization reported in Table 2.

Example 8. Cost effect for an active portfolio versus cost effect for a passive portfolio

Consider an active portfolio with an expected return of 10.5 percent, portfolio turnover of 60 percent, management fee of 0.25 percent, and trading costs of 0.75 percent of value. For the active manager, the total (two-way) turnover represented by purchases and sales is 2×60 percent = 120 percent, so total costs are 0.75 percent \times 120 percent = 0.90 percent of portfolio value. The net expected return of this portfolio is 10.5 percent – 0.90 percent – 0.25 percent = 9.35 percent.

Now consider a passive portfolio with a lower expected return, 10 percent, but turnover of only 4 percent, management fees of 0.10 percent, and trading costs of 0.25 percent. The total turnover for this portfolio is 8 percent, which implies transaction costs of only 0.25 percent \times 8 percent = 0.02 percent of portfolio value. The net expected return is 10 percent = 0.02 percent = 0.10 percent = 9.88 percent, which is higher than the net expected return from the active portfolio.

Predicting Trading Costs. The divergence in cost estimates reported in previous studies is troubling. It suggests that investment professionals' ability to predict trading costs is poor. And, indeed, even those studies using the most detailed trading data available do poorly in terms of predicting costs. For example, when Keim and Madhavan (1997) regressed estimated costs on proxies for trade difficulty, market-specific factors, and dummy variables for trader identity and style, they found that the regression R^2 s ranged from 0.10 to 0.15. In other words, more than 85 percent of the variation in trading costs is idiosyncratic and cannot be explained by trade venue, market liquidity, trade difficulty, or investment style. Similar results were reported by Chan and Lakonishok (1995), Leinweber, and others.

The unpredictability of costs is a particularly distressing fact to inst tutional traders and portfolio managers, who would like to predict costs in real time. If they knew the costs, traders who are averse to the high variance in costs might choose trading strategies that would let them predict and control costs. Examples of such strategies are crossing systems (where the crossing price is predetermined), automated limit-order book systems, and guaranteed principal bids (where the trading costs are known prior to trading).

Improving investment professionals' ability to predict execution costs requires understanding of why the previous estimates are so noisy. Two factors complicate the task of estimating and predicting trading costs. First, although some elements of trading costs (e.g., commissions and taxes) are highly predictable, others (e.g., opportunity and timing costs) are highly variable and depend heavily on prevailing market conditions. Opportunity costs are also a function of the trader's investment style. For example, a pure index trader may incur low opportunity costs but high price-impact and commission costs; a value trader may face large opportunity costs but small commission and price-impact costs.

Second, many unobservable factors that are not easily measured may explain the large variation in execution costs, including trader reputation, skill, investment objectives, and subtleties of the trading process (e.g., upstairs intermediation). The construction of empir cal proxies for these factors might considerably facilitate the prediction of trading costs. Edwards and Wagner suggested that the inclusion of variables for market momentum helps explain opportunity costs, which are especially difficult to estimate. A variable for market momentum might be a good proxy for certain unobserved elements of the investment style of the trader, which

is an important determinant of costs.

Proxies can go only so far toward improving cost predictions, however, especially when traders follow dynamic policies. Indeed, one lesson that emerges from the recent literature on equity trading costs is that data on the order-submission process alone are not enough to accurately predict trading costs. Rather, researchers need detailed information on the investor's motivations and goals prior to the order-submission process. They need data on the investor's trading style, size of the trade, investment objectives, trading horizon, and estimate of the stock's fundamental value. They also need the precise details of the instructions at each point in time to brokers and traders. This information would allow researchers to examine execution costs in relation to traders' underlying investment strategies—a promising avenue for future research.

Conclusions

An understanding of the magnitude and determinants of execution costs is crucial to resolving many practical and academic issues. A list of the issues includes predicting the trading costs of alternative trading strategies, determining the effect of execution costs on realized ("live") portfolio performance, understanding the behavior of institutional equity traders, making intermarket cost comparisons, and assessing arguments about the nature and causes of market fragmentation. The increased availability of detailed data on institutional equity trades in recent years has allowed increased understanding of equity trading costs, and our aim in this article was to summarize the main findings of the research and discuss their practical implications.

Although transaction costs can be measured in many ways, two important considerations must be kept in mind. First, any reasonable method to measure costs must capture both the implicit and explicit trading costs because the two components are jointly determined. Second, to capture the overall price movement associated with the individual transactions in an order, costs must be measured at the level of the entire order. Perold's implementation shortfall approach is the most natural method to meet these two necessities in measuring transaction costs. Only recently, however, have the data required to compute the total costs at the level of the order become available. Studies by Edwards and Wagner; Perold and Sirri; Leinweber; Chan and Lakonishok (1995); and Keim and Madhavan (1997) provide estimates of total costs that use information on the entire order-submission process.

The recent studies allow several conclusions about transaction costs:

- Although considerable debate still surrounds how to measure trading costs, the consensus is that implicit trading costs (such as the price impact of a trade and the opportunity costs of failing to execute) are economically significant relative to explicit costs (and relative to realized portfolio returns).
- Equity trading costs vary systematically with trade difficulty and order-placement strategy.
- Differences in market design, investment style, trading ability, and reputation are important determinants of trading costs.
- Even after researchers control for trade complexity and trade venue, trading costs are found to vary considerably among managers.
- Accurate prediction of trading costs requires more detailed data on the entire order-

- submission process than are available, especially information on pretrade decision variables.
- Finally, the recent literature on equity trading costs offers important lessons for policymakers and investors. For example, it suggests that the concept of "best execution" for institutional traders is difficult to measure and hence to enforce.

We thank Mark Edwards, Larry Harris, and George Sofianos for their helpful suggestions and comments. An earlier version of this article was presented at the NYSE conference "Search for the Best Price." The comments and opinions contained in this article are those of the authors alone.

Notes

- 1. This issue attracted considerable attention following the allegations by Christie and Schultz (1994) and Christie, Harris, and Schultz (1994) that Nasdaq dealers engage in "implicit collusion" to keep spreads above competitive levels.
- Schwartz and Shapiro (1992) reported that U.S. institutional equity holdings in 1990 were approximately 50 percent of total NYSE capitalization and institutional trades were 72 percent of NYSE share volume.
- 3. See, for example, Beebower and Priest (1980); Treynor (1981); Berkowitz, Logue, and Noser (1988); Arnott and Wagner (1990); Collins and Fabozzi (1991); and Wagner and Banks (1992)
- 4. Other explicit fees include, for example, the New York State tax, the U.S. SEC fee, and a clearing charge from The Depository. Trust Company. These charges are relatively small (e.g., the New York State tax is 3.75 cents per share for stocks trading between \$10.00 and \$19.99) and are typically levied against the seller of securities.
- 5. Kothare and Laux (1995) argue, however, that institutional trading has widened bid-ask spreads.
- Burdett and O'Hara (1987) and Keim and Madhavan (1996) provide theoretical models of upstairs intermediation; Hasbrouck. Sofianos, and Sosebee (1993) discuss the institutional features of upstairs trading; and Keim and Madhavan (1996) and Madhavan and Cheng (1997) provide empirical evidence on upstairs trades.
- See Kraus and Stoll (1972); Holthausen, Leftwich, and Mayers (1987); and LaPlante and Muscarella (1997), among others, for unconditional estimates of the price impacts of large block trades.
- This correlation may be explained by the fact that the more difficult trades, which tend to have higher price impacts, are given to full-service brokers, who charge higher commissions.
- 9. Chan and Lakonishok (1995) found that the costs associated with packages are higher than cost for individual trades. They found that price impacts (measured by the principal-weighted price change from the open on the first day of the package to the close on the last day) are 1 percent for buys and -0.3 percent for sells. In contrast, an earlier study by

- Chan and Lakonishok (1993) found that the principal-weighted price change from open to close on the day of the trade itself is only 0.34 percent for buys and –0.04 percent for sells. These results clearly document the need to consider the total order as the relevant unit for cost analysis. Note also that the cost differential between buys and sells found in their study is not anomalous. Kraus and Stoll; Holthausen, Leftwich, and Mayers; Madhavan and Smidt (1991); and Keim and Madhavan (1996) noted similar asymmetry.
- Much of the previous literature on block trading and the literature on implicit and effective spreads followed this approach.
- 11. Notable in the approximation category are Chan and Lakonishok (1995, 1997), who constructed trade "packages" by combining trades in a particular stock that occurred on adjacent days and thus appeared to be part of the same trading decision.
- 12. Within the class of decision variables, some decisions (e.g., the stock, order size, and horizon) are made by the investor and others (e.g., number of subtrades and type of order) are made at the trade desk. This distinction may be important in the prediction of trading costs, but it is not of primary importance.
- 13. Using a sample of trades for the DJIA 30 stocks, Madhavan and Cheng (1997) found that the expected relationship between price impact and order size may be confounded by the upstairs-facilitated large block trades. In particular, if a significant fraction of upstairs block trades are crossed within the prevailing quotes (because upstairs brokers have already arranged counterparties to the trades) or if a substantial portion of the price impact has already been impounded in previous prices because of leakage in the upstairs market, the measured impact of these trades can be very low. Thus, large trades may appear to have less of a price impact than some small trades. Indeed, Leinweber found that small trades are responsible for a disproportionate share of costs whereas large trades have lower-than expected costs
- 14. Harris (1996) provides an analysis of the economics of best execution; Macey and O'Hara (1996) discuss the legal and

- economic issues related to best execution; and Bertsimas and Lo (1996) propose a method to measure best execution for institutional traders.
- Such inferences have been made by Reinganum (1990); Neal (1992); Blume and Goldstein (1992); Kothare and Laux; Bessembinder and Kaufman (1996), and Huang and Stoll, among others.
- Seppi (1997) provides a theoretical model of specialist behavior. Madhavan and Smidt (1991, 1993); Hasbrouck and Sofianos (1993); and Madhavan and Sofianos (1997) provide empirical analyses of NYSE specialist trading.
- 17. See Shapiro (1993) for a comprehensive summary of recent changes in the structure of U.S. equity markets.

References

Arnott, R., and W. Wagner. 1990. "The Measurement and Control of Trading Costs." "inaucial Analysis Journal, vol. 46, no. 6 (November/December):73–80.

Beebower, G., and W. Priest. 1980. "The Tricks of the Trade." *Journal of Portfolio Management*, vol. 6, nc. 2 (Winter):36–42.

Berkowitz, S., D. Logue, and E. Noser. 1988. "The Total Cost of Transactions on the NYSL." *Journal of Finance*, vol. 43, no. 1 (March) 97–112.

Bertsimas, D., and A. Lo. 1996. "Optimal Control of Execution Costs." Working paper, Massachusetts Institute of Technology.

Bessembinder, H., and H. Kaufmar. 1996. "Quotations and Trading Costs on the Domestic Equity Exchanges." Working paper, Arizona State University.

Blume, M. 1993. "Soft Dollars and the Brokerage Industry." *Financial Analysts Journal*, vol. 49, no. 2 (March/April):36–44.

Blume, M., and M. Goldstein. 1992. "Displayed and Effective Spreads by Market." Working paper, University of Pennsylvania.

Bodurtha, S., and T. Quinn. 1990. "Does Patient Program Trading Really Pay?" *Financial Analysts Journal*, vol. 46, no. 3 (May/June):35–42.

Burdett, K., and M. O'Hara. 1987. 'Building Blocks: An Introduction to Block Trading." *Journal of Banking and Finance*, vol. 11, no. 2 (June):193–212.

Chan, L., and J. Lakonish k. 1993. "Institutional Trades and Intra-Day Stock Price Behavior." *Journal of Financial Economics*, vol. 33, no. 2 (April):173–200.

———. 1995. "The Behavio of Stock Prices around Institutional Trades." Journal of Finance, vol. 50, no. 4 (September):1147–74.

———. 1997. "Institutional Equity Trading Costs: NYSE versus Nasdaq." *Journal of Finance* vol. 52, no. 2 (June):713–35.

Christie, W., and P. Schultz. 1994. "Why Do NASDAQ Market Makers Avoid Odd-Eighth Quotes?" *Journal of Finance*, vol. 49, no. 5 (December):1813–40.

Christie, W., J. Harris, and I. Schultz. 1994. "Why Did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes?" *Journal of Finance*, vol. 49, no. 5 (December):1841–60.

Collins, B., and F. Fabozzi. 1991. "A Methodology for Measuring Transaction Costs." *Financial Analysts Journal*, vol. 47, no. 2 (March, April):27–36.

Cuneo, L., and W. Wagner. 1975. "Reducing the Cost of Stock Trading." *Financial Analysts Journal*, vol. 31, no. 6 (November/December):35–44.

Demsetz, H. 1968. "The Cest of Transacting." Quarterly Journal of Economics, vol. 82, no. 103–53.

Dutta, P., and A. Madhavan. 1997. "Competition and Collusion in Dealer Markets." *Journal of Finance*, vol. 52, no. 1 (March):245–76

Easley, D., N. Kiefer, and M. O'Hara 1996. "Cream Skimming

or Profit-Sharing? The Curious Case of Purchased Order Flow." *Journal of Finance*, vol. 51, no. 3 (July):811–833.

Edwards, M., and W. Wagner. 1993. "Best Execution." *Financial Analysts Journal*, vol. 49, no. 1 (January/February):65–71.

George, T., G. Kaul, and M. Nimalendran. 1991. "Estimation of the Bid–Ask Spread and Its Components: A New Approach." *Review of Financial Studies*, vol. 4, no. 4 (Winter):623–56.

Greene, J. 1996. "The Impact of Limit Order Executions on Trading Costs in NYSE Stocks: An Empirical Examination." Working paper, Indiana University.

Harris, L. 1996. "The Economics of Best Execution." Working paper, University of Southern California.

Hasbrouck, J., and G. Sofianos. 1993. "The Trades of Market-Makers: An Analysis of NYSE Specialists." *Journal of Finance*, vol. 48, no. 5 (December):1565–95.

Hasbrouck, J., G. Sofianos, and D. Sosebee. 1993. "New York Stock Exchange Systems and Trading Procedures." Working paper, New York Stock Exchange.

Holthausen, R., R. Leftwich, and D. Mayers. 1987. "The Effect of Large Block Transactions on Security Prices: A Cross-Sectional Analysis." *Journal of Financial Economics*, vol. 19, no. 2 (December):237–67.

Huang, R., and H. Stoll. 1996. "Dealer versus Auction Markets: A Paired Comparison of Execution Costs on NASDAQ and the NYSE." *Journal of Financial Economics*, vol. 41, no. 3 (July):313–57.

Keim, D.B. 1989. "Trading Patterns, Bid–Ask Spreads, and Estimated Security Returns: The Case of Common Stocks at Calendar Turning Points." *Journal of Financial Economics*, vol. 25, no. 1 (November):75–98.

———. Forthcoming 1998. "An Analysis of Mutual Fund Design: The Case of Investing in Small-Cap Stocks." *Journal of Financial Economics*.

Keim, D.B., and A. Madhavan. 1995. "Anatomy of the Trading Process: Empirical Evidence on the Behavior of Institutional Trades." *Journal of Financial Economics*, vol. 37, no. 3 (March):371–98.

——.. 1996. "The Upstairs Market for Large-Block Transactions: Analysis and Measurement of Price Effects." *Review of Financial Studies*, vol. 9, no. 1 (Spring):1–36.

——. 1997. "Transaction Costs and Investment Style: An Inter-Exchange Analysis of Institutional Equity Trades." *Journal of Financial Economics*, vol. 46, no. 3 (December):265–92.

Kothare, M., and P. Laux. 1995. "Trading Costs and the Trading Systems for Nasdaq Stocks." *Financial Analysts Journal*, vol. 51, no. 2 (March/April):42–53.

Kraus, A., and H. Stoll. 1972. "Price Impacts of Block Trading on the New York Stock Exchange." *Journal of Finance*, vol. 27, no. 3 (July):569–88.

LaPlante, M., and C. Muscarella. 1997. "Do Institutions Receive

Comparable Execution in the NYSE and Nasdaq Markets? A Transaction Study of Block Trades." *Journal of Financial Economics*, vol. 45, no. 7 (July):97–134.

Lee, C.M.C. 1993. "Market Integration and Price Execution for NYSE-Listed Securities." *Journal of Finance*, vol. 48, no. 3 (July):1009–38.

Leinweber, D. 1995. "Using Information from Trading in Trading and Portfolio Management." *Journal of Investing*, vol. 4, no. 1:40–50

Loeb, T.F. 1983. "Trading Cost: The Critical Link between Investment Information and Results." Financial Analysis Journal, vol. 39, no. 3 (May/June):39–43.

Macey, J., and M. O'Hara. 1996. "The Law and Economics of Best Execution." Working paper, Cornell University.

Madhavan, A., and M. Cheng. 1997. "In Search of Liquidity: Block Trades in the Upstairs and Downstairs Markets." *Review of Financial Studies*, vol. 10, no. 1 (Spring) 175–204.

Madhavan, A., M. Richardson, and M. Roomans. 1997. "Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks." *Review of Financial Studies*, vol. 10, no. 4 (Winter):1035–64.

Madhavan, A., and S. Smidt. 1991. "A Bayesian Model of Intraday Specialist Pricing." *Journal of Financial Economics*, vol. 30, no. 1 (November):99–134.

——. 1993. "An Analysis of Daily Changes in Specialist Inventories and Quotations." *Journal of Finance*, vol. 48, no. 5 (December):1595–628.

Madhavan, A., and G. Sofianos. 1997. "An Empirical Analysis of NYSE Specialist Trading." *Journal of Financial Economics*, vol. 48, no. 2 (May):189–210.

Neal, R. 1992. "A Comparison of Transaction Cost between Competitive Market Maker and Specialist Market Structures." *Journal of Business*, vol. 65, no. 3 (July):317–34.

Nelling, E. 1995. "Large Block Trades and Positive Feedback Investment Strategies." Working paper, Georgia Institute of Technology.

Niederhofter, V., and M. Osborne. 1966. "Market Making and Reversal on the Stock Exchange." *Journal of the American Statistical Association*, vol. 61, no. 316:897–916.

Perold, A. 1988. "The Implementation Shortfall: Paper versus Reality." *Journal of Portfolio Management*, vol. 14, no. 3 (Spring):4–9.

Perold, A., and E. Sirri. 1993. "The Cost of International Equity Trading." Working paper, Harvard University.

Reinganum, M.R. 1990. "Market Microstructure and Asset Pricing: An Empirical Investigation of NYSE and NASDAQ Securities." *Journal of Financial Economics*, vol. 28, no. 1–2 (November–December):127–48.

Roll, R. 1984. "A Simple Implicit Measure of the Effective Bid—Ask Spread in an Efficient Market." *Journal of Finance*, vol. 39, no. 4 (September):1127–39.

Schwartz, R., and J. Shapiro. 1992. "The Challenge of Institutionalization for the Equity Markets." In *Recent Developments in Finance*, edited by A. Saunders. New York: New York University Salomon Center.

Seppi, D. 1997. "Liquidity Provision with Limit Orders and a Strategic Specialist." *Review of Financial Studies*, vol. 10, no. 1 (Spring):103–50.

Shapiro, J.E. 1993. "Recent Competitive Developments in U.S. Equity Markets." Working paper, New York Stock Exchange.

Sinquefield, J.C., and C. Severoglu. 1989. "Trading Illiquid Stocks." In *The Complete Guide to Securities Transactions*, edited by W. Wagner New York: John Wiley & Sons.

Sinquefield, Rex A. 1991. "Are Small-Stock Returns Achievable?" *Financial Analysts Journal*, vol. 47, no. 1 (January/February):45–50.

Stell, H. 1995. "The Importance of Equity Trading Costs: Evidence from Securities Firms' Revenues." In Global Equity Markets: Technological, Competitive, and Regulatory Challenges, edited by R. Schwartz. New York: New York University Salomon Center.

Treynor, J. 1981. "What Does It Take to Win the Trading Game?" *Financial Analysts Journal*, vol. 37, no. 1 (January/February):55–60

Wagner, W., and M. Banks. 1992. "Increasing Portfolio Effectiveness via Transaction Cost Management." *Journal of Portfolio Management*, vol. 19. no. 1 (Fall):6–11.