The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

A Review of Trading Cost Models: *Reducing Transaction Costs*

ANDREW FREYRE-SANDERS, RENATE GUOBUZAITE, AND KEVIN BYRNE

ANDREW FREYRE-SANDERS
is a vice president in the Trading Analytics Group at JPMorgan Chase.
andrew.freyre-sanders@jpmorgan.com

RENATE GUOBUZAITE
is a researcher at London Business School.
renate_guobuzaite@ml.com

KEVIN BYRNE
is a vice president in Transition Management at JPMorgan Chase.
kevin.byrne@jpmorgan.com

Over the last few years, transaction cost analysis has been one of the biggest areas of investment for both the buy and sell side of the equity industry. This increased focus has led to intensification of research in this field. At the same time, there has been an enormous leap in the provision of tools to aid measuring and predicting costs.

As a precursor to modeling transaction costs, we felt it would be useful to the reader to compile a literature review that takes the reader from the very early attempts of modeling market microstructure through to some techniques used today. Although many of the early (classical) studies may seem inappropriate to today's electronic order books, many of the concepts developed are still relevant at present.

Until recently, the focus of the investment community has been on commissions, taxes, and spreads. In this article, we have not gone into any detail regarding these fixed costs, but we have focused on the so-called "hidden" costs of trading.

PRICE DISCOVERY AND SPREAD FORMATION

Price discovery and spread formation is a long established and well covered area of research in the market microstructure literature. Initial studies assumed that the effective bid-ask spread (i.e., difference between the mid-quote and the price actually paid by the investor) is a good measurement of transaction costs and that it therefore remained a prime focus of their analysis. Three main factors shaping the spread included: 1) order processing; 2) inventory control; and 3) adverse selection—each of them uniquely contributing to the development of the bid-ask spread.

This section briefly reviews empirical and theoretical studies as related to each category of the spread formation factor. Exhibit 1 provides a summary of major theories that correspond to these approaches.

Order Processing Costs

Definition. A liquidity provider incurs order processing costs, as well as additional risk, by providing intermediation services to the rest of the investment community and, therefore, should be compensated for them in terms of the bid-ask spread.

Liquidity providers (e.g., market makers, dealers, or specialists) assume a leading role in setting prices at which a particular security can be bought or sold, naturally resulting in the bid-ask spread. Due to the high visibility of the liquidity provider, most of the early literature (e.g., Demsetz [1968], Tinic [1972]) viewed the liquidity provider's position as critical in the price-setting process.

In essence, the liquidity provider is any trader who takes the other side of the trade for orders randomly arriving in the market. The most prominent example of such practice is the specialist on the New York Stock

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

Exchange (NYSE). The specialist is required to maintain a constant market presence by posting quotes, ready to buy and sell the given security at any time. If no one else is willing to trade, the specialist accommodates marginal orders through his inventory, often acting as "the liquidity supplier of last resort." Other examples might include security dealers, proprietary traders, or traders submitting orders in electronic limit-order books.

Demsetz [1968] was the first to introduce the idea that transaction costs arise as a natural outcome of the liquidity provider's business. He argued that the liquidity provider incurs an additional risk, as well as order processing costs, by supplying a "predictive immediacy" service to the market and should be compensated for it in the form of a bid-ask spread. This resulted in naming the first factor, i.e., order processing costs, which accounts for development of the bid-ask spread. Of particular interest, therefore, are the factors that determine the magnitude and variability of the spread. Demsetz's 1968 study provided first empirical evidence that a critical factor in determining it was *the level of trading activity*. The cross-sectional regressions, with two measures of trading activity (the daily trading frequency on two non-adjacent days and the number of shareholders as a proxy for the long-run trading activity), confirmed a significant negative relationship between the spread and both measures of trading activity.

The order processing cost analysis suggests that a liquidity provider's services should be the most valuable and, therefore, that resulting transaction costs should be highest in lower capitalization, less liquid securities (Madhavan and Sofianos [1998]). Exhibit 2 illustrates this point. The market capitalization of U.K. stocks traded on the London Stock Exchange (LSE) SETS system clearly coincides with the order book participation rate (therefore, it is negatively related to dealers' presence in the market) and indicates higher spread charges for less liquid names, where market makers predominate.

Practical points to remember:

1. Using a liquidity provider's services reduces the risk of executing the trade; however, it also imposes an extra cost. *Alternative trading systems* (e.g., crossing systems, matching networks) can significantly reduce transaction costs by eliminating the role of the "middle man."
2. A liquidity provider's services are particularly valuable and, as a result, spreads are the widest in *thinly traded, illiquid stocks*.

Inventory Control Costs

Definition. The liquidity provider has an optimal inventory position that he prefers to hold based on his risk/return preferences. Accommodating other market participants' trades forces him away from this position; therefore, liquidity providers change their bid-ask prices to attract trades and to restore an initial position.

A second factor contributing to the bid-ask spread that attracted researchers' attention was inventory control exercised by the liquidity provider. Despite apparent differences in various approaches, the common concept underlies basic inventory-based models that can be described as follows. Every time the market maker acquires a security from other market participants, he incurs a risk of carrying additional inventory. The risk comes from two sources: 1) the liquidity provider does not know how long the security will remain in his inventory; 2) he cannot anticipate security price changes during that period.

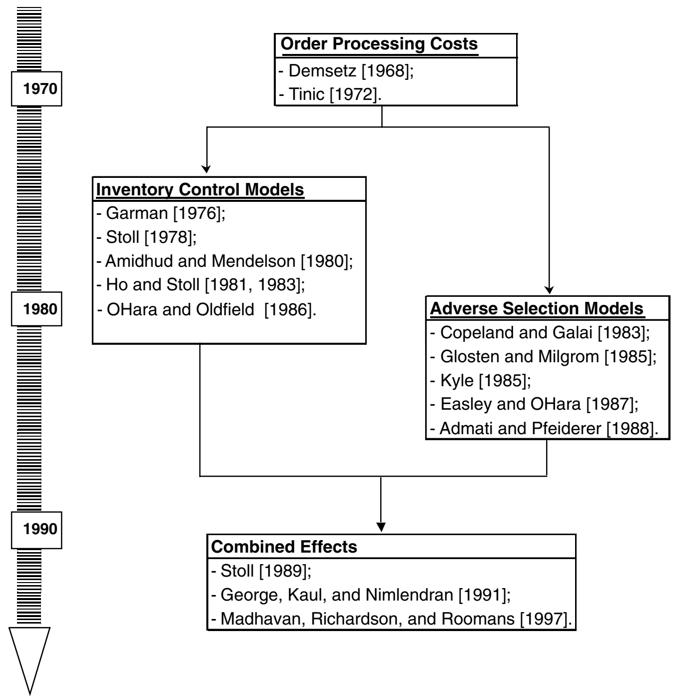
The market maker then sets an optimal/preferred inventory position in each security traded, and when his actual holdings deviate from it, he is likely to adjust the bid-ask spread, or even the price level, to restore that position. Theoretical studies that follow this approach include Garman [1976], Stoll [1978], Amidhud and Mendelson [1980], Ho and Stoll [1981], Cohen, Maier, Schwartz, and Whitcomb [1981], and O'Hara and Oldfield [1986], among others. These effects are predominant today in the principal bidding of portfolios or single stock blocks.

Example. Let's introduce the market maker in XYZ stock with an initial inventory of I_0 in XYZ and access to the finite capital K. Based on risk preferences, expectations, and/or other considerations, he sets a preferred inventory position at I* to be held in XYZ. At the beginning of trading, he quotes the bid at 48 and the ask price at 52. However, following a buy from the market maker, his inventory decreases ($I_t < I^*$) and the market maker raises the quotes up to 49 bid and 54 ask to attract incoming sell orders (*see Exhibit 3*). Note that the spread becomes wider when the market maker is further away from his preferred position.

Garman [1976] was the first to formally model the relationship between the bid-ask spread and inventory levels. In his analysis, inventory fluctuations follow a random walk with a zero drift, resulting in a sequence of trades that is likely to eventually deplete either his stock or cash position. In order to avoid a certain failure or bankruptcy, the market maker is required to set a lower price when he buys and a higher price when he sells. This justifies the emergence of the bid-ask spread in the

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

EXHIBIT 1
Development of Market Microstructure Models



Source: JPMorgan.

market. Amidhud and Mendelson [1980] followed this logic by expecting the market maker's inventory level to be mean-reverting, as the market maker prefers to sell when he is long and buy when he is short inventory. In addition, the spread is expected to be minimal when the market maker is at his preferred inventory position and to widen as the deviations become larger.

Stoll [1978] views the market maker as an investor who would otherwise prefer a diversified portfolio based on his risk-return characteristics. However, the requirement to accommodate the trades of other investors in the security in which he specializes forces the market maker to hold a portfolio that is sub-optimal. His next course of action is then to determine an appropriate compensation resulting in a bid-ask spread for carrying an additional risk. Ho and Stoll [1981] further extended this analysis to a multi-period framework. An important result is that the bid-ask spread here represents a compensation for the risk of bearing unwanted inventories.

Empirical studies prove that inventory levels do indeed exhibit some reversion towards their mean, albeit at a very slow rate. For example, in an empirical analysis of the New York Stock Exchange (NYSE) specialist inventories, Madhavan and Smidt [1993] find that it requires, on average, 7.3 trading days to reduce the imbalance in inventory by 50%. The corresponding figure for the London Stock Exchange (LSE) dealership market (Hansch, Naik, and Viswanathan [1998]) is only 2.5 trading days. In addition, Ho and Macris [1984] confirm inventory effects in the AMEX option market; Manaster and Mann [1996], however, find little evidence of inventory effects in futures markets.

Practical points to remember:

1. Transaction costs are expected to be highly correlated with *the risk* incurred by the liquidity providers by carrying excess inventories.
2. Mean reversion in inventory levels induces *serial dependence* on security price changes; therefore, statistical analysis of price changes might provide insights regarding the liquidity provider's inventory level and vice versa.
3. Role of *interdealer trading* increases when inventories diverge significantly, as it provides an efficient mechanism for market makers to unload their unwanted inventories.
4. Inventory effect depends on the extent to which the liquidity provider is *capital constrained*; therefore, larger inventory effects are likely to occur when dealing with marginal market makers with less capital.

Adverse Selection Costs

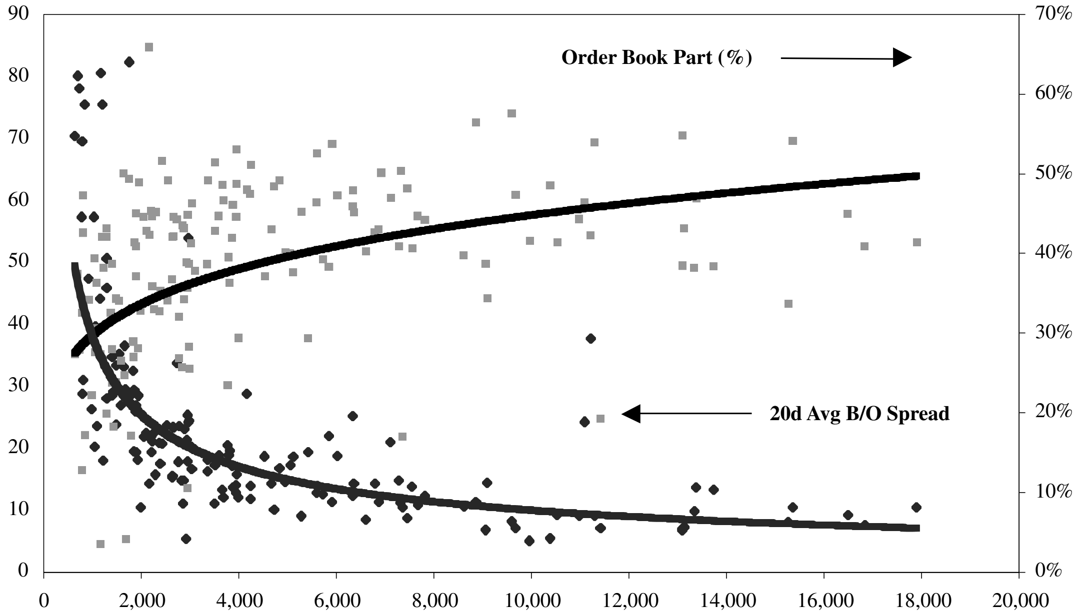
Definition. Some market participants (i.e., informed traders) possess special information that allows them to constantly profit from their trades. As a market maker will consistently lose to this type of trader, he charges a larger bid-ask spread to everyone in order to cover his potential losses.

The third set of models (Copeland and Galai [1983], Glosten and Milgrom [1985], Kyle [1985], Easley and O'Hara [1987], Admati and Pfeiderer [1988]) argue that a bid-ask spread can arise due purely to information-based trading, even if other costs incurred by the liquidity provider (including previously discussed order-processing and inventory-control costs) are non-existent.

In 1971, Jack Treynor published a short but influential study ("The Only Game in Town") introducing many basic principles that were later applied in adverse selection models. The study makes it clear that trading is a "zero-sum" game, where, for one player to win, another must be losing. It follows that some traders with superior information or analytical/trading ability are potentially to profit at the expense of the liquidity provider or other

EXHIBIT 2

Market Capitalization versus Average-Weighted Spread/U.K. Order Participation Rate for LSE SETS Traded Stocks



Source: JPMorgan Tic Database.

traders. Based on this logic, Bagehot [1971] and many other adverse selection models assume that the liquidity provider (or market maker) faces two types of traders in the market place:

- *informed traders*, with some private information that allows them to know or to better estimate the "true" value of the security (e.g., information on corporate events, superior analytical skills, etc.).
- *uninformed (or liquidity) traders*, who have no private information, but frequently transact for reasons determined outside the financial markets (e.g., trades resulting from mutual fund savings and redemption flows, or, in particular, asset liability changes for pension funds, or even simple income manufacture).

Informed traders who know the value of an underlying security with certainty will buy only if the price is lower than its actual value and sell only when it is higher. Considering that, in all other cases, the informed traders stop trading, the market maker will always lose to this type of trader. However, in order for a market to function, the certain market maker's losses to informed investors should be offset by the gains from the rest of the market. As the market maker can hardly distinguish informed from uninformed traders, he would charge the spread to everyone just to break even. The trading cost is then the price that liquidity traders are paying to compensate the market maker for losses incurred through trading with informed traders.

An important implication of this analysis is that, although the market maker cannot determine the true value of the security himself, he can "learn" it from informed traders by observing the sequence of trades and updating his beliefs accordingly (Glosten and Milgrom [1985], Easley and O'Hara [1987]). Following the buy from his inventory, the market maker revises upwards the probability that true value is actually higher, and therefore increases the price and vice versa—the phenomenon is known as *the Bayesian learning process*.

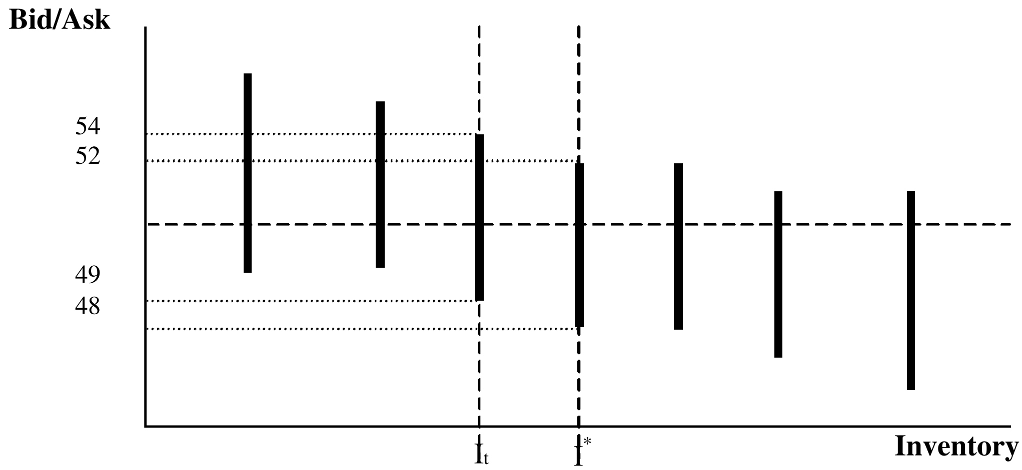
Given that all informed investors always trade on the same side of the market, they create a relative pressure on a security price towards its "true" value. Over time, the value estimates of the market maker and of the informed traders tend to converge, and private information is fully disseminated into market prices.

Interestingly, the presence or absence of a trade may provide information to the market as well (Easley and O'Hara [1992]). Since a no-trade outcome is more likely to occur if there has been no new private information in the market, the market maker believes it is less likely that he will be trading against an informed trader and, therefore, he reduces the spread. In particular, if trade signals the direction of new information, the lack of trade signals

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

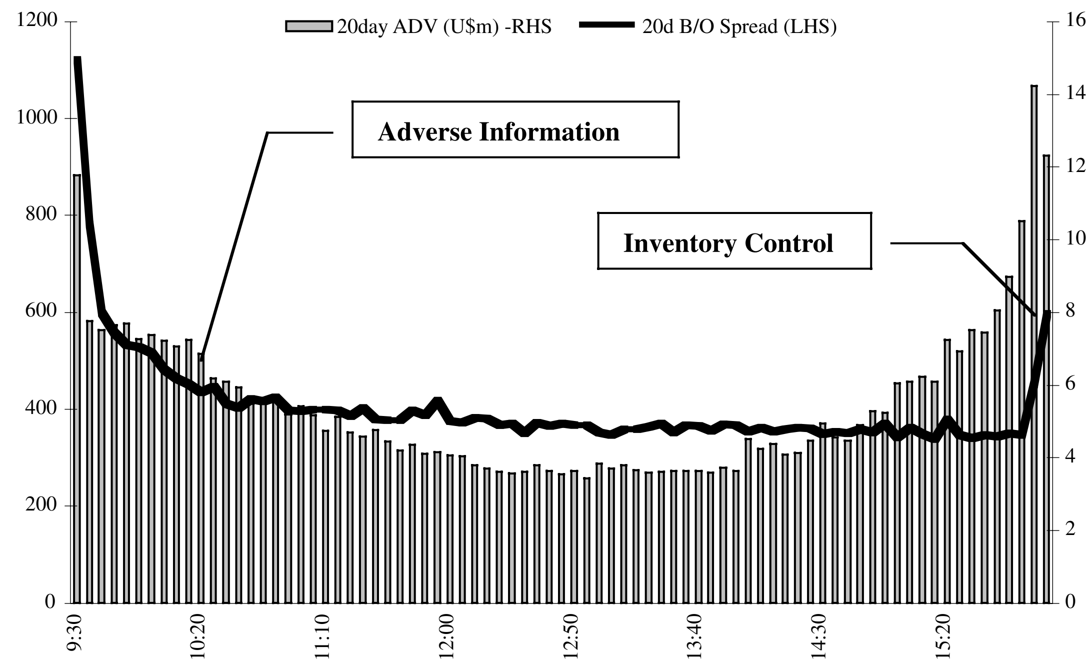
The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

EXHIBIT 3
Relationship Between Optimal Portfolio Position and Bid-Ask Spread



Source: Amidhud and Mendelson [1980].

EXHIBIT 4
Intraday Trading Patterns for the S&P 500 Index



Source: JPMorgan Tic Database.

the existence of such information. As a result, this model indicates that the spread narrows as the time interval between trades increases.

The strategic behavior of informed traders is the focus of analyses presented in Kyle [1985], Easley and O'Hara [1987], Admati and Pfeiderer [1988], and Foster and Viswanathan [1990]. The authors note that, considering that informed traders often receive low liquidity and poor execution from the market makers, it is within their best interests to strategically "hide" their trades behind the activity of uninformed or liquidity traders. It might be achieved by: 1) spreading the trades over time (as in Kyle [1985]), or 2) trading when trading volume is high (Admati and Pfeiderer [1988]).

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

Another strategic option of informed traders is to choose trading at certain times of the day. Admati and Pfeiderer [1988] show that, having no specific information, all uninformed investors prefer to trade when the trading volume is high and the market is most liquid. Clustering of their trades also attracts informed traders, as a higher liquidity trading volume provides an opportunity to "hide" informed trades more effectively. The tendency of market participants to pool their orders results in intraday volume patterns of trading, where higher trading volume at certain periods of the day is associated with higher adverse selection costs, as well as more informative prices (Wood, McInish, and Ord [1985], Jain and Joh [1988]).

A well-known example of this phenomenon is the so-called U-shape of trading volume, volatility, and average bid-ask spreads in the U.S. securities markets (*see Exhibit 4*). Theory suggests that, at the start of the day, information asymmetry is large and that both bid-ask spread and price volatility are high. Over the day, asymmetries are resolved through the price discovery, causing spreads to narrow. By the end of the day, the asymmetric information cost component is reduced, but the risk of carrying inventory overnight causes both the trading activity and spreads to widen again. In practice, however, other reasons might also contribute to this pattern, as well as, for example, increased institutional trading activity at closing prices (Cushing and Madhavan [2001]).

Subrahmanyam [1991] suggests that uninformed or liquidity traders can significantly reduce their adverse selection costs by trading in the basket of securities, rather than in individual stocks. Considering that most informed traders possess security-specific information (it is unlikely for market-wide information to exist), their orders submitted in the basket tend to offset each other when the number of securities comprising the basket increases. In this case, the impact of asymmetric information is minimized. It is consistent with the fact that proportional bid-ask spreads in index futures are approximately one-tenth of those in individual stocks.

Practical points to remember:

1. Markets with high numbers of *uninformed (liquidity) traders* are likely to have lower spreads; however, they are expected to be more volatile.
2. Larger trades can be broken up into a *sequence of smaller orders* to reduce their impact on market prices.
3. Clustering of trades might result in certain *calendar patterns* that exist consistently for prolonged periods of time.
4. Trading in a basket of securities (e.g., index, futures) reduces the adverse selection costs, as market-wide private information is unlikely to exist.

Combined Effects

Definition. Quoted spreads are likely to reflect all (i.e., order processing, inventory, control, and adverse selection) effects; some regression techniques using trading data allow separate spread components attributable to each factor.

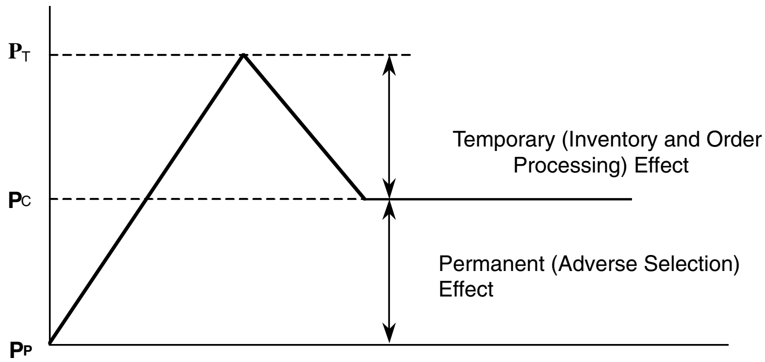
Until the late 1980s, most of the market microstructure models were focused on a single criterion or factor (i.e., order processing, inventory control, or adverse selection) that determined the quoted price of the security. Although this approach is useful in that it allows an in-depth, thorough analysis of each factor, in practice, the bid-ask spread is likely to incorporate an effect of all of these factors, at least to some degree.

In testing this hypothesis, Stoll [1989] first concluded that the quoted spread indeed reflects all three factors and that the components of the bid-ask spread, as contributed by each factor, appear to be constant as a percentage of the quoted spread. The results showed that order processing, inventory control, and asymmetric information cost components are approximately equal to 47%, 10%, and 43% of the quoted spread, respectively. In most studies, the order processing costs are found to be a major contributing component, and the effect of inventory control costs is relatively weak.

The analysis of spread components and their variation can be used for comparing liquidity and transaction costs across securities and exchanges, as well as for describing general trading conditions. For example, a larger adverse selection component would indicate a higher level of informed trading that, in turn, results in wider spreads; however, it is likely to exhibit less volatility due to more informative prices.

Combined adverse selection, inventory control, and order processing cost effects are also used to explain permanent and temporary price changes (Kraus and Stoll [1972], Holthausen, Leftwich, and Mayers [1987], Keim and Madhavan [1996]) following large trades. In general, the price increase (drop) following a large buy (sell) is expected to partially recover to its pre-trade level after the trade is completed. Here, the adverse selection component reflects new information brought to the market through the trade and, therefore, is associated with the permanent effect. On the other hand, the temporary component or price rebound after the trade represents a tran-

EXHIBIT 5
Permanent versus Temporary Price Impact



Source: JPMorgan.

sitory price movement (including the inventory control and order processing effects), which is necessary to provide sufficient liquidity to accommodate larger trades. Therefore, decomposition of the spread might be useful in forecasting security price behavior before and after trades.

Example. Temporary and permanent effects on a price in XYZ stock are illustrated in Exhibit 5. P_P is the market price prior to the transaction and represents the equilibrium price of the stock before any new information is revealed to the market prior to the trade. P_T is the actual transaction price at which the trade was executed, and P_C is the equilibrium price after the temporary price effect of the trade was dissipated. The difference between P_T and P_C is the temporary effect (i.e., price rebound) associated with the transaction. The permanent price effect is the difference between P_P and P_C. The total price impact cost incurred by the buyer in this case is the difference between P_P and P_T—the sum of the temporary and permanent price effects.

Practical points to remember:

1. In markets with significant *informed investor* presence, the adverse selection component is substantial, potentially resulting in higher market impact costs.
2. Following a large trade, the price is likely to *partially recover* to its pre-trade level.
3. The price impact is more persistent for buys than for sells, according to literature.

MANAGING TRANSACTION COSTS

In today's complex trading environment, many traders would agree that a successful execution of large trades has become more of an art form. Investment performance depends on both the investment strategy characteristics (i.e., risk/return) and the execution costs incurred in realizing the objectives. If ignored, transaction costs might significantly reduce the realized returns of an otherwise optimal portfolio strategy. Although the majority of transaction costs are incurred during the execution stage of trading, both pre-trade (e.g., cost estimation, implementation strategy selection) and post-trade analysis (e.g., cost measurement, execution quality evaluation) have proved to be essential in minimizing market impact and other related transaction costs.

COST COMPONENTS

Definition. Following the implementation shortfall approach, total transaction costs can be categorized into execution and opportunity costs (see Exhibit 6). Execution costs comprise explicit/fixed costs (including brokerage fees and various taxes applicable in some equity markets) and implicit/market impact components. The opportunity costs arise owing to adverse price movements in the market when part of the order is still unexecuted.

Example. A fund manager receives confirmation of the executed trade: "7,500 XYZ bought at \$20.25." However, he had actually submitted an order for 10,000 shares of XYZ based on an opening price of \$20 earlier that day. Before his order reached the desk, the price jumped up to \$20.15. As XYZ's price increased, the broker was able to cross 10,000 shares at \$20.25, and submitted a limit order at \$20.45 for the rest of the order. XYZ closed at \$20.50 and the limit order was canceled. The broker charges a commission of 0.05% and taxes are 0.50%.

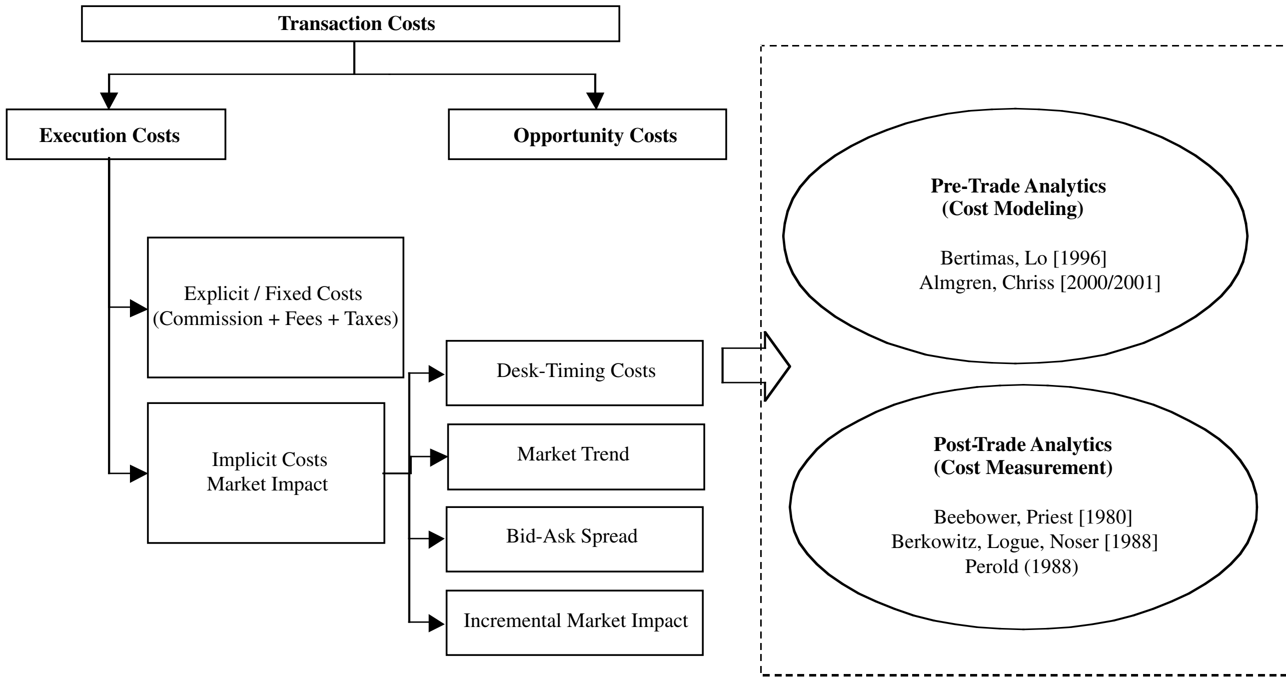
It is helpful to visualize the time-line of the trade (*see Exhibit 7*).

The cost components of the trade are estimated as follows:

$$\textit{Explicit costs} = 0.75 \times (0.05\% + 0.5\%) = 0.41\%$$
$$\textit{Delay costs} = 0.75 \times ((20.15 - 20)/20) = 0.56\%$$
$$\textit{Market impact} = 0.75 \times ((20.25 - 20.15)/20) = 0.37\%$$
$$\textit{Execution costs} = 0.56\% + 0.37\% + 0.41\% = 1.34\%$$
$$\textit{Opportunity costs} = 0.25 \times ((20.50 - 20)/20) = 0.62\%$$
$$\textit{Total costs} = 1.34\% + 0.62\% = 1.96\%$$

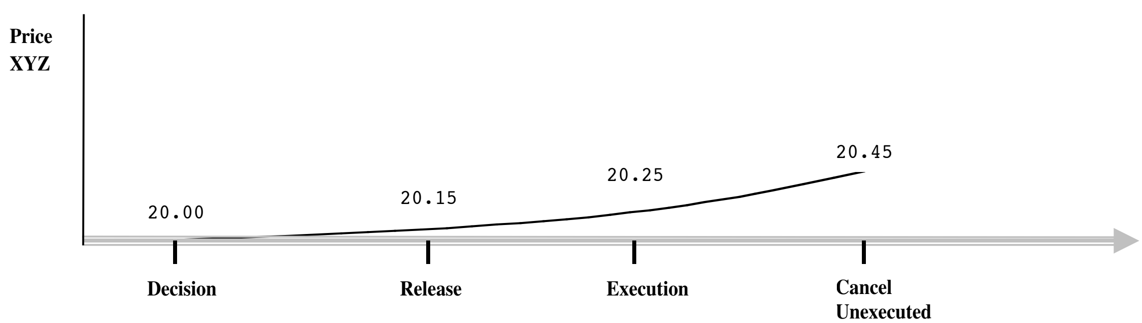
The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

EXHIBIT 6
Cost Components, Cost Analysis, and Modeling Studies



Source: JPMorgan.

EXHIBIT 7
Timeline of a Trade



Source: JPMorgan.

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

Explicit/Fixed Costs

Explicit (or fixed) costs include commissions charged by brokers, custody fees, stamp duties, taxes and other regulatory charges, depending on the particular country and exchange. Keim and Madhavan [1997] found that commission costs are relatively low, accounting for only about 0.20% of trade value (*Exhibit 8*). In addition, commissions paid by the investors declined over time, generally. This may be attributed to increased institutional presence in the market, a more competitive environment for trading services, and technological innovations, such as increased use of the electronic crossing systems.

Fixed cost measurement is complicated by the fact that commissions are often paid for a bundle of services, not just for order execution. The payments for exchanging other non-execution services are commonly referred to as "soft dollar" practices and are commonly used to: 1) obtain in-house research, such as fundamental research or macroeconomic forecasts; 2) purchase third-party research for investment managers' use; 3) simply compensate for directing order flow to a particular brokerage

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

EXHIBIT 8
Estimated Spread Components

			EFFECTS (%)			
			Order processing	Inventory control	Combined - order processing and inventory	Adverse selection
STUDY	YEAR	EXCHANGE	[TRANSITIONARY]			[PERMANENT]
Stoll	1989	NASDAQ/NMS	47%	10%	57%	43%
George, Kaul and Nimalendran	1991	NYSE/AMEX	-	-	88-90%	10-12%
		NASDAQ	-	-	87-92%	8-13%
Foster and Viswanathan	1990	NYSE/AMEX	-	-	88%	12%
Afleck-Graves, Hegde and Miller	1994	NYSE/AMEX	12%	29%	41%	59%
		NASDAQ/NMS	41%	24%	65%	35%
Madhavan, Richardson and Roomans	1997	NYSE	-	-	49-65%	35-51%
Huang and Stoll	1997	NYSE (20 most active stocks)	62%	29%	91%	9.60%
		Hong Kong Stock Exchange				
Brockman and Chung	1999		44%	21%	65%	32%
		Paris Bourse				
Declerk	2000	Swiss Stock Exchange	82%	8%	90%	10%
Ranaldo	2002		49%	19%	68%	32%
Average:			48%	23%	71%	29%

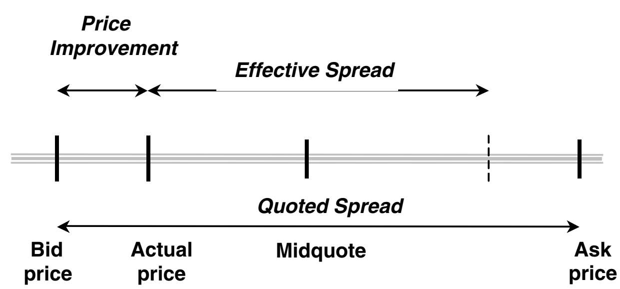
(Blume [1993], Easley, Kiefer, and O'Hara [1996], Blume and Goldstein [1997], Batalio and Holden [2001]). In effect, "soft dollars" can inflate costs and make it difficult to compare execution costs among brokers.

Implicit/Market Impact Costs

Desk-Timing Costs. Edwards and Wagner [1993] first pointed out the desk-timing costs that arise from delays in placing orders with the dealer. The delay between decision and execution may occur for a variety of reasons: for example, the portfolio might not be tradable immediately upon the decision, owing to unreconciled positions, required compliance checks, market closure, or investors hesitating in selecting the best execution venues, etc. In this case, desk-timing costs account for all price movements, from the point at which a decision is actually made to the earliest moment that the portfolio is available for trade.

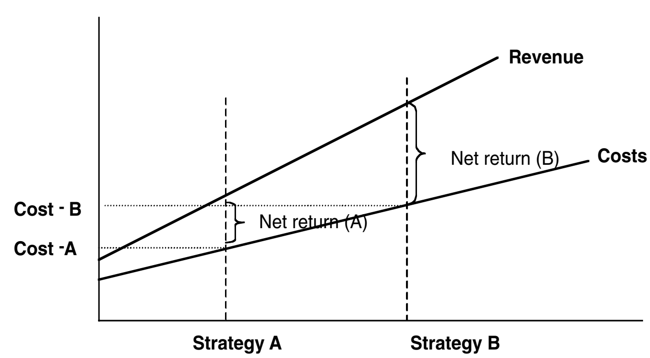
Bid-Ask Spread. A quoted bid-ask spread cost component refers to a difference between the mid-quote and the bid (for sells) or the ask (for buys), at which investors are offered a pre-specified number of shares (i.e., quoted depth) in dealership markets. In electronic limit-order book markets, the quoted spread represents half the difference between the best bid and the best offer. In fact, the definition captures only one-half of the quoted spread. Although easy to estimate, the quoted spread may overstate actual transaction costs, as, in dealership systems such

EXHIBIT 9
Price Improvement



Source: JPMorgan.

EXHIBIT 10
Net Returns versus Execution Cost Measures



Source: JPMorgan.

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

EXHIBIT 11
Factors Affecting Market Impact Costs

Investor-specific factors:	Market-specific factors:
▪ Order size	▪ Market capitalization
▪ Investment style	▪ Turnover (avg. daily trading volume)
▪ Traders reputation, skills	▪ Volatility
▪ Technology	▪ Market momentum
▪ Order submission strategy	▪ Trading venue
▪ Buy vs. sell orders	▪ Stock price
▪ Participation fraction	▪ Trading frequency
▪ % of daily volume traded	▪ Analyst coverage

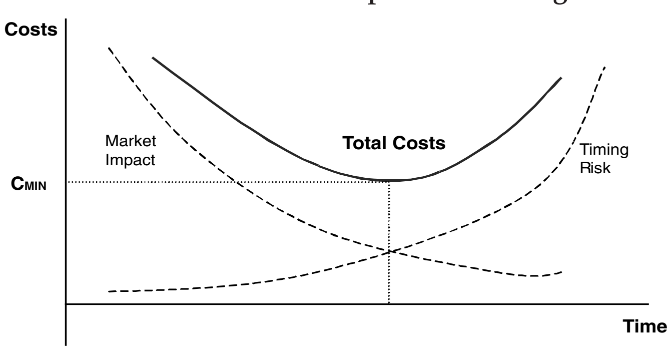
Source: JPMorgan.

EXHIBIT 12
Regression Analysis for Predicting Market Impact Costs

Variable	All orders	
	Estimate	Standard error
Intercept	0.687	0.269
D^{NASDAQ}	0.239	0.045
Trsize	0.165	0.005
Logmcap	-0.076	0.016
1/P	9.924	1.029
D^{TECH}	0.607	0.035
D^{INDEX}	0.451	0.040
Adjusted R^2	*0.060*	

Source: Keim and Madhavan [1997].

EXHIBIT 13
Trade-off Between Market Impact and Timing Risk



Source: JPMorgan.

as Nasdaq or LSE SEAQ, trades are often executed within the quoted spread, resulting in a price improvement for a particular trade.

The quoted bid-ask spread can often be misleading in markets where off-order book trading is available. Several authors have proposed a more practical measure, which accounts for price improvement practices, called *an effective bid-ask spread*. The effective spread differs from the quoted spread, as it represents the difference between the mid-quote and the *actual* price paid by the investor (or again—to be precise—this definition captures one-half of the effective spread). Exhibit 9 illustrates the difference between two measures of the spread.

One of the most appealing effective spread measurement models in market microstructure was introduced by Roll [1984]. It enables us to estimate an effective spread directly from time-series data of market prices, eliminating the need for recording and matching quoted and actual prices for a particular transaction. Based on the assumption that spreads consist of only order-processing costs, the model proves that the presence of a bid-ask bounce (execution of orders at bid and ask prices, instead of mid-quote) induces negative serial autocorrelation in security returns. Therefore, an effective spread can be calculated from the observed serial correlation of transaction prices, as follows:

$$S = 2 \text{ sqrt } (-\text{cov}(\Delta p_t; \Delta p_{t+1}))$$

Here, S is the spread, and $\text{cov}(\Delta p_t; \Delta p_{t+1})$ is the serial covariance of price changes. Choi, Salandro, and Shastri [1988] extended Roll's model, incorporating the probability of a transaction at a bid (ask) following the previous transaction.*

The serial covariance estimator of this and similar models, as a rule, resulted in the effective spread being smaller than the quoted spread. For example, based on daily returns of a sample of all NYSE and Amex stocks for the 1963-82 period, Roll finds that an average effective spread across all stocks is 0.298%, which is lower than the average quoted spread even for most liquid stocks.

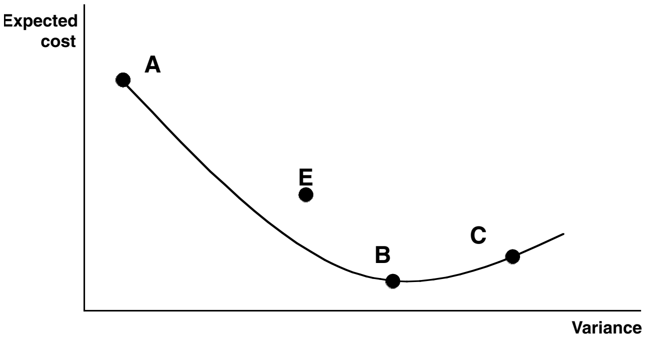
Market Trend. The market trend refers to the change in the stock price, which results from the overall market movement. It can be negative if the market trend is contrary to the trade direction and the order is supplying liquidity to the market, or positive if the investor is trying to buy (sell) in a rising (falling) market. The market trend does not depend on a particular decision to trade, and it should therefore be eliminated when comparing execution quality across brokers. Patel [2001] suggests using the market-adjusted implementation shortfall measure (MAIS), which removes the market-related cost component from both execution and opportunity costs:

$$\text{MAIS} = (\text{Execution Costs} - \beta_e * \Delta M_e) + (\text{Opportunity Costs} - \beta_o * \Delta M_o)$$

where:

EXHIBIT 14

Efficient Frontier of Trading Strategies



Source: Almgren and Chriss [2000/2001].

β_e = the historical beta of the traded stock or portfolio versus the market;

ΔM_e = the percentage market movement between the start and end of the execution period;

β_o = the historical beta of the unexecuted stock or portfolio versus the market; and

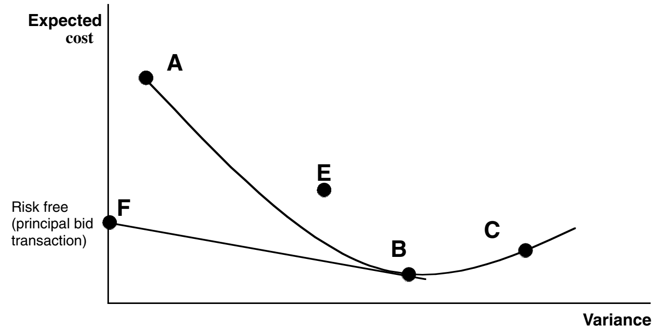
ΔM_o = the percentage market movement between the start and closing time of the trade.

Incremental Market Impact. If the trade size falls within the quoted depth, the order is likely to be executed at the quote and, therefore, the total market impact would equal half of the bid-ask spread. However, if the order exceeds the depth available at the quote, then the price for larger trades increases and the market impact, typically, would be greater than the half-spread. The difference between the total market impact and the half-spread adjusted for market trend effects results in the incremental market impact measure.

Large block trades are especially likely to move market prices in the trade direction. Edwards and Wagner's [1993] study shows that roughly two-thirds of institutional managers' orders average more than 50% of an average day's volume, and four out of 10 exceed a full day's volume. Executing orders this large as a single trade will surely constrain the liquidity of the market. However, postponing execution, searching for liquidity, or executing in smaller portions is also a risky activity, as the knowledge of unfilled large orders is a highly valuable commodity in a closely watched market. In any case, a price increase is likely to occur in order to create enough liquidity for larger orders in the market.

EXHIBIT 15

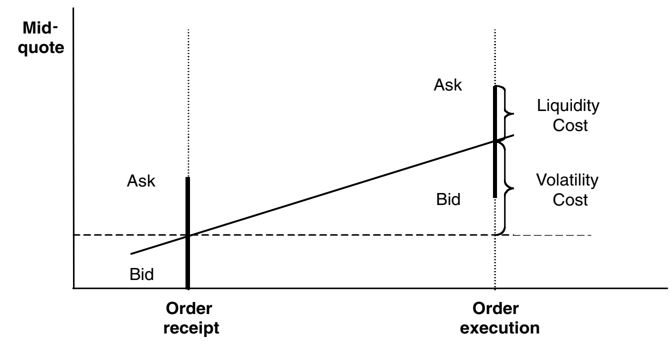
Efficient Frontier of Trading Strategies with Capital Trade Line



Source: Kissell and Glantz [2003].

EXHIBIT 16

Risk Inventory Model Components



Source: JPMorgan.

Opportunity Costs

Opportunity transaction cost accounts for the cost of not trading or only partially completing a desired order. It refers to the price change on an unexecuted portion of the order from the time a trading decision is made until cancellation or some other pre-specified time after the last completed order trade in the stock. This type of cost may occur for one of two reasons: 1) the trader was not able to locate shares required to complete the order, or 2) the price has moved out of the range that the investor is willing to transact. In either case, potential losses might be substantial.

The opportunity (also referred to as "hidden") costs are not directly recorded and therefore difficult to estimate. The measurement requires detailed data on 1) the time that a decision to trade is made; 2) total number of shares intended to be bought/sold; and 3) the expected trade horizon. These costs are, however, less significant in the case of institutional investors, owing to high completion rates of institutional trades. Empirical studies (Keim [1995],

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

Keim and Madhavan [1998]) indicate that institutional orders are fully completed more than 95% of the time.

Practical points to remember:

1. Both *pre-trade analysis* and *post-trade performance measurement* are critical in securing best execution.
2. Smaller-sized trades in dealership markets are often executed within the quoted spread, resulting in *price improvement*.
3. *Market trends* can hardly be influenced by an individual trader; therefore, they should be excluded when comparing execution quality among the brokers. (However, where an order is a large part of the market, then removing market trend will not necessarily add value.)
4. Executing large orders is likely to constrain the liquidity of the market, resulting in *incremental market impact costs.*
5. *Opportunity costs* represent an "invisible" element of costs; however, they still account for a significant part of transaction costs.

COST MEASUREMENT

Definition. In general terms, the execution costs of the trade are measured as the difference between the price prevailing when the decision to trade was made and the price at which the trade was actually executed.

In broad terms, measuring transaction costs is defined as estimating the difference between the security price prevailing at the time the trading decision was made and the average price at which the trade was actually executed (Kissell and Glantz [2003]). A related concept—an investment performance measurement—refers to the difference between the average execution price and the benchmark price set for the trade. The most commonly used benchmarks include open/close prices (Barnea and Logue [1978], Beebower and Priest [1980]), volume-weighted average prices (VWAP) (Berkowitz, Logue, and Noser [1988]), and open-high-low-close average (OHLC) and implementation shortfall (Perold [1988]). Several studies suggest using multiple benchmarks and incorporating some broader industry-wide or relative analysis (e.g., broker-by-broker or peer group comparisons) as execution performance measures.

It is important to consider that transaction cost estimates have little value if analyzed separately from corresponding returns earned by the trading strategy. In fact, higher transaction costs are not bad news per se; they could simply prove to be a necessary condition for generating superior returns. For example, Strategy A, presented in Exhibit 10, has a lower cost than Strategy B and so is supposed to be superior from a cost perspective. However, its generated returns are also lower, which results in lower returns on investment, i.e., representing lower net return in this case. Therefore, an optimal measure for evaluating trading (or, for that matter, any) strategy is the net return earned after completing the trade (*see Exhibit 10*).

Another factor to consider is a non-random order flow allocation to execution brokers. One of the main objectives for measuring transaction costs is to ensure a fair comparison of execution costs among execution venues. However, if the best-performing brokers are often used as a "last resort," they are likely to systematically receive order flow that is more difficult to execute. Therefore, their performance measures should reflect the fact that they are handling more difficult orders, on average. In summary, each individual case should be thoroughly analyzed before applying corresponding performance evaluation measures.

Open/Close Prices

The benchmark set as a fixed price at a certain time of the day, such as opening or close prices, is both easy to estimate and interpret. As a convenient performance measure, it is widely used by brokers and investment managers. The method was first introduced by Barnea and Logue [1976] and then extended and tested by Beebower and Priest [1980].

A combination of open and close prices can also be used to estimate temporary and permanent effects of the trade. Chan and Lakonishok [1993] examine percentage returns calculated from the opening price to the trade and from the trade to the day's closing price. The figures obtained correspond with total (from opening to trade), temporary (from trade to close), and permanent (from open to close) effects of the trade prices. This approximates the portion of price movement that was actually informative (i.e., revealed new information to the market) and that which simply arose owing to liquidity effects (*see Exhibit 10*) of the trade.

The main problem with these measures is, however, that both open and close prices are drawn from distributions that differ from the average security price distribution over the course of a day and, therefore, may not be representative of prices quoted when the order was actually submitted. An open price is an appropriate benchmark if the decision to trade is made before the market opens. A condition for using

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

the closing price is that liquidity effects caused by the trade must have dissipated by the end of the day.

Finally, as with any benchmark that is known to the trader prior to trading, both open and close prices are subject to "gaming" by the traders who are being evaluated against it. For example, in the case of opening prices, the trader can choose to buy stocks on the order list whose prices have fallen since the receipt of the order and sell stocks whose prices have increased, resulting in favorable performance measures. All other orders are dismissed as too expensive. The trading objective might become "beating the benchmark" and undermine brokers' efforts to search for more favourable execution.

Volume-Weighted Average Price

Berkowitz, Logue, and Noser [1988] proposed using a value-weighted average of transaction prices around the trade as a proxy for the objective price. The volume-weighted average price (VWAP) is calculated by multiplying a stock transaction price by its volume traded, adding all transactions and dividing by the total shares traded. By definition, it should represent an unbiased estimate of the price facing a "naïve" trader. A popular alternative includes comparing the price with the mean of open, high, low, and closing prices (OHLC) during the day. Willoughby's study [1998] shows that price impact computed using both methodologies are relatively close.

The VWAP measure is best used for smaller trades that have little or no impact on the existing market prices. In all other cases, such as large trades or illiquid stocks, the order might comprise a significant portion of the daily trading volume over the trading period and the trade itself might become the VWAP. Even for more liquid stocks, if the broker's share of overall trading volume is sufficiently large, it still might drive VWAP in the direction of the trade.

In addition, several more limitations should be considered when using the VWAP as the performance measure, as follows.

- VWAP judges the trade solely by a relative rank—despite the difference in size or difficulty of the trade, the standard remains the same.
- It is specific to a particular trading period, which might include prices irrelevant to the analyzed transaction.

One further important point is the quality of the VWAP number. This, in itself, may well be subject to different calculation methods. For example, large crosses are often included in market VWAPs, even though this volume may not be available to market participants. However, in the U.K., it is common for portfolio trading users to look only at order book VWAP rather than the total market. There is still not yet 100% consistency in determining VWAP prices.

Again, as a pre-specified benchmark, the VWAP might become the subject of gaming practices. For instance, the trader might postpone trading until late in the trading period and then execute the trade only if prices are favorable relative to the VWAP benchmark; or alternatively, he might spread the order into small portions over the day, so that the price received closely resembles the VWAP.

Implementation Shortfall

Perold [1988] introduced the "implementation shortfall" approach to measuring transaction costs. The method suggests that the investor runs two portfolios simultaneously—a "paper" portfolio, with no transaction costs, and a real portfolio, based on the trades actually executed. At the initial decision point to trade, both portfolios contain a target list of stocks, which should be bought (or sold) during the course of trading. The paper portfolio is then "executed" immediately, at the mid-quotes of prices prevailing when the original decision was made. In reality, the stocks are purchased at various prices, while trading evolves, and some of the orders might not be executed at all. If, at the beginning of the trading horizon, both portfolios are assumed to have the same value, at the end they will differ by the amount of the implementation shortfall.

The implementation shortfall consists of two components—execution costs (i.e., commissions, bid-ask spread, and market impact costs) and opportunity costs (due to non-executed trades). Using this methodology, the trader cannot influence trading results simply by delaying or by not executing difficult trades, as they would appear in the opportunity costs category. In order to reduce the implementation shortfall, he should consider an optimal trade-off between the execution and opportunity costs.

The initial concept of implementation shortfall attracted significant attention from industry professionals and, as result, a number of adjusted versions were introduced, including (but not limited to) the following.

- *Market-adjusted implementation shortfall*—is the implementation shortfall measure adjusted for a market-

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

related movement in the portfolio value, as it cannot normally be influenced by the trader and can distort performance measurement results.

- *Execution implementation shortfall*—is the difference between the execution price and the price when the broker received the order (instead of when the actual decision to trade was made); it eliminates the impact of desk-timing costs, which again are not within the responsibilities of the broker.
- *Post-execution implementation shortfall*—uses the post-trade price instead of the price at the beginning of the trade for estimating the shortfall; as such, it controls for the informational content of the trade.

Each modified version of the basic implementation shortfall approach emphasizes different aspects of execution quality, and therefore allows for a more in-depth comparison of trading costs among execution venues.

Practical points to remember:

1. Each benchmark is designed for a particular aspect of execution quality; therefore, multiple benchmarks are best used for evaluating trading performance.
2. In evaluating execution quality, analysis should focus on *net returns*, rather than on estimates alone, as higher costs are often correlated with higher returns.
3. *A combination of open/close prices* can be used to estimate temporary and permanent effects of the trade.
4. Order flow is often *selectively allocated* to brokers, with the best brokers likely to receive more difficult orders for execution. This tendency should be considered when comparing costs among brokers.
5. An option to *game benchmarks* might direct trading practices from "best execution" towards "beating the benchmark" objectives.

PRE-TRADE ANALYTICS AND COST MODELING

Definition. Forecasting transaction costs and designing optimal trading strategies allows minimization of market impact for the selected level of risk and requires consideration of a number of factors, such as trade characteristics, market conditions, and investors' risk preferences.

A successful execution of trades starts long before the order actually reaches the trading desk. The pre-trade cost analysis incorporates a number of tools that are designed to eliminate costly decisions, reduce timing risk and associated opportunity costs, as well as develop optimal trading strategies. The market impact has already proved to be the most significant, although the most difficult, cost component to forecast. Theory suggests that the market impact cost is dependent on a number of factors, which might include market capitalization of stock, volatility of returns, trading frequency, etc. (Keim and Madhavan [1997], Kissell and Glantz [2003]).

Considering that execution cost is subject to a large number of dynamic factors (which include an ever-changing market environment, uncertain behavior of other market participants, macroeconomic or firm-specific factors), it is, by itself, a random variable, and is best described as a distribution with a corresponding mean and variance of random variables. To capture the uncertainty of execution cost, the models proposed by Bertsimas and Lo [1996] and, more recently, by Almgren and Chriss [2000/2001] and Kissell and Glantz [2003] balance the trade-off between expected value and the uncertainty of costs.

In this framework, a more aggressive trader is likely to incur higher costs (and less uncertainty), and more patient trading is rewarded by lower transaction costs (and higher uncertainty), on average. Alternatively, the traders that are highly averse to the variance in costs may alter their trading towards strategies that let them predict and control costs. Examples include crossing systems or guaranteed bids, where the trading costs are known prior to trading.

Finally, Leinweber [1995] suggests some general "rules of thumb" for minimizing transaction costs:

- skillful execution reduces cost;
- patient trading reduces cost;
- crossing reduces cost;
- some trades produce transaction profits, which potentially offset costs.

Market Impact Modeling

The market impact cost is defined as the difference between the execution price and the price that would have prevailed had this order not been submitted to the market. Prior theoretical and empirical research suggests that market impact costs are positively related to the trade difficulty—quite a general concept, which is not that easy to evaluate, let alone to quantify. Nevertheless, relevant factors that make the trade more or less difficult might be classified into two broad categories:

1. *investor-specific factors* (such as investment style, trade size, trader reputation and skills);

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

2. *market-specific factors* (including market capitalization, volatility, market momentum, and market design).

The first category (investor-specific factors) includes trade characteristics that can be pre-selected or altered by the investor. For example, the investor can change the order size or the order submission strategy based on the prevailing situation in the market. The second category (market-specific factors) refers to market-wide factors that are "set" and can hardly be modified by a single investor. Exhibit 11 provides a more detailed list of commonly used factors in both categories. In many studies (Loeb [1983], Edwards and Wagner [1993], Keim and Madhavan [1997]), trade size, market capitalization, and volatility are referred to as major factors that affect market impact costs.

Kissell and Glantz [2003] introduced a general model, where the market impact (κ) depends on the following factors:

$$(\kappa) = f(\text{imbalance, volatility, trading style, market conditions})$$

The market impact is further modeled by allocating the total market impact into permanent and temporary components and relating it to market imbalance, price volatility, and trading style factors.

Keim and Madhavan [1997] proposed using a regression-based approach for estimating market impact costs. In the regression model, average bid-ask spread expressed as a fraction of price is regressed against several explanatory variables. In this particular case, the explanatory variables' set includes the market capitalization, the reciprocal of the price, and the trade size, and additional variables control for the influence of trade venue (NYSE vs. Nasdaq) and investment style (value, technical, or index style).

$$C = \beta_0 + \beta_1 D_i^{NASDAQ} + \beta_2 Trsize + \beta_3 Logmcap + \beta_4 1/Pi + \beta_5 Di^{TECH} + \beta_6 Di^{INDEX}$$

where:

Ci = is the total cost of transaction;
D_i^{NASDAQ} = the dummy variable for Nasdaq stocks;
$Logmcap$ = the log of the market capitalization of stock traded;
$Trsize$ = the size of the trade as measured by order size divided by shares outstanding;
P_i = the price per share of the stock traded;
D_i^{TECH}, D_i^{INDEX} = dummy variables for technical and index traders.

The results of regression analysis using trading costs for 21 institutional traders in the period January 1991 to March 1993 are presented in Exhibit 12.

As expected, the trade cost is found to be inversely related to market capitalization (as a proxy for liquidity), so the price impact is found to be smaller for more liquid stocks with larger market capitalizations. The reciprocal of price was used as a proxy for the bid-ask spread, and is therefore highly correlated with transaction costs. The last factor—trade size—is a primary determinant for the trade difficulty and is expected to be directly correlated with costs.

Transaction costs vary significantly according to investment style, market design, and trade direction (i.e., buy or sell side); therefore, separate regression analyses were used in each case. For example, the investment style (e.g., active or passive, value or growth) is a proxy for unobserved factors, such as a trader's time horizon and aggressiveness. Index and technical traders often use active strategies that require immediate execution and, as a result, may incur high execution costs. The value traders, whose trades are motivated by considerations of long-term fundamental value, may incur lower costs because of more patient trading strategies (e.g., limit orders, crossing networks). They might even become suppliers of liquidity and incur negative transaction costs. Keim and Madhavan [1997] show that technical and index traders incur execution costs that are, on average, 0.61% and 0.45%, respectively, higher than value traders' trading costs.

The cost measurement is complicated by the fact that cost is a non-linear function of the trade size, so the sum of costs of smaller trades will differ from the cost of their cumulative order. However, there is an extensive discussion in academic literature regarding the shape of the function of the trade size. Some studies (Hasbrouck [1991], Madhavan and Cheng [1997]) argue that it is convex, especially for smaller orders, while Loeb's [1983] study, using data from block trades, concludes that the result is a concave function of the trade size.

Barra's Market Impact Model proposes a solution for this problem. In this model, the market impact cost k is regressed against the trade size V, using the following function: $\kappa = V^e$ for different values of exponent e. The best fit appears for e = 0.5, so, in this case, the exponent functional relationship is best described by one-half power law. However, a similar experiment can be replicated with other data sets, where the values of e < 1 would indicate a concave function and e > 1 a convex function.

Kissell and Glantz's [2003] model suggests several

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

methods using market impact functions based on two contributing variables—order imbalance (z) and volatility (δ)—that can be used for estimating market impact cost (I):

Linear Model: $I_{BP} = a_1 z + a_2 \delta + a_3$
Non-Linear Model: $I_{BP} = a_1 z^{a2} + a_3 \delta + a_4$
Power Function: $I_{BP} = a_1 z^{a2} + d^{a3}$

In this model, the non-linear and power functions result in smaller regression errors, implying that the true relationship between cost and size is non-linear. Finally, in estimating the market impact cost function, it is important to consider that it is comprised of both permanent and temporary components (see also discussion on "Combined Effects"). The former represents the information revealed to the market through the trade, while the latter refers to the transitory liquidity conditions required to accommodate the trade. In modeling market impact costs, it is critical to estimate the impact of both effects.

Cost Optimization Strategies

A trading strategy is defined broadly as how orders are to be executed over the trading period. As shown above, a large order is naturally associated with higher market impact costs. From both a theoretical and practical point of view, the market impact cost can be substantially reduced by breaking up the order and executing it more discretely in a series of smaller trades over time. The optimal strategy, then, might lie somewhere between two extremes: trade everything now at a known but potentially high cost, or spread the trade over the entire trading horizon, executing it in smaller packages, but bearing the risk of adverse price movements.

Therefore, one of the most important decisions facing investment managers is how quickly they should implement their buy and sell decisions. With regard to choosing a trading strategy, the trader faces two requirements: 1) on the one hand, he can reduce the market impact by trading patiently, partially supplying liquidity to the market, but facing a greater uncertainty with respect to final results, and 2) on the other hand, he can choose to trade aggressively to reduce the risk, but at the expense of higher prices, owing to a substantial market impact. The goal of allocating the trades then becomes an optimization problem relating to the trade-off between the market impact and associated risk.

Exhibit 13 illustrates the trade-off between the market impact and timing risk. The additional costs that should be considered include commission costs; however, these are often fixed and therefore relatively easy to estimate. The exhibit shows how all cost components add up to the total cost of transacting—this results in a convex total costs function with a minimum point that corresponds to an optimal trading strategy (i.e., corresponding minimum transaction costs).

The transaction cost model, based on cost optimization strategy logic, as described above, was first introduced in Bertsimas and Lo [1998] and then extended in Almgren and Chriss [2000/2001]. Both models assume that a rational trader will always seek to minimize the expected transaction costs for a given level of variance (or risk) associated with each strategy. The analysis suggests that the trading cost estimate cannot be projected as a single number (as in the case of the post-trade cost measurement), but rather will yield a distribution of the expected (mean) cost and corresponding risk (volatility) parameters. Therefore, the optimization problem is set to find the trading strategy that minimizes the following function

$$\min\ [E(x) + \lambda^* V(x)]$$

where:

E(x) is the expected value of transaction costs;
V(x) is the variance of cost values; and
λ is an investor's risk aversion parameter.

Here, the risk aversion parameter λ is a unique number corresponding to a particular investor or trading strategy. It defines how many units of risk an investor is ready to assume for a one unit increase in returns. For example, if λ is set to be equal to 1, it indicates that the investor is equally concerned with risk and return. The case of $\lambda > 1$ refers to the more risk-averse investor and corresponding more aggressive strategies, and $\lambda < 1$ is likely to result in more patient trading.

An interesting result of the Almgren and Chriss [2000/2001] study is that, for all levels of risk aversion (apart from risk-neutral), optimal executions of trades have a "half-life," which represents an optimal horizon for completing the trade. The security's "half-life" is independent of the pre-specified trading horizon and is a function of the intrinsic factors, such as the security's liquidity, volatility, and the trader's risk aversion. As such, a "half-life" of a security can be regarded as an idealized or optimal time horizon for execution, which results from each specific stock characteristic.

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

Although the basic Almgren and Chriss [2000/2001] model results in a closed-form solution, this is, however, owing only to a number of simplifying assumptions (e.g., linear price impact function, discrete arithmetic price walk, no drift). In more complex cases, a closed-formed solution might not be available, but numerical approximations can be used in each case (for a more extensive analysis on incorporating drift, serial correlation, and parameter shift factors, see Almgren and Chriss [2000/2001]). The subsequent studies extend the initial model by introducing non-linearity into the market impact function (Almgren [2001]), and analyze the market impact of principal bid trading (Almgren and Chriss [2002]).

Although both the Almgren and Chriss [2000/2001] and Bertsimas and Lo [1998] models provide the optimal solutions for balancing transaction costs and risk, they do not specify or provide more detailed analysis for forecasting its main component—the market impact cost. In Almgren and Chriss's [2000/2001] model, the market impact is simply assumed to be a linear function of average trading speed, while Bertsimas and Lo's [1998] model assumes it to be a linear function of trading size. In practice, deriving the market impact function for the cost optimization strategy requires an in-depth analysis of a number of factors reflecting a particular trade and situation in the market.

To illustrate the results, Almgren and Chriss [2000/2001] use the efficient frontier approach, where each level of an investor's risk aversion is represented by a unique optimal trading strategy. Exhibit 14 depicts the efficient frontier as an AB portion of the curve. Here, Strategy A represents a trader who executes all trades immediately to reduce exposure to volatility risk, but incurs high transaction costs. The minimal point, or Strategy B (which Bertsimas and Lo [1998] referred to as the "naïve" strategy), corresponds to trading equally-sized packages, using all allocated trading horizons. Neither E nor C are optimal trading strategies, because the lower expected cost can be obtained for the same level of risk and/or a lower amount of risk will result in the same expected cost.

Kissell and Glantz [2003] extend the analysis by introducing a risk-free asset such as a principal bid transaction. Here, by allocating the execution between the principal bid and agency execution (*see Exhibit 15*), the investor can shift the cost profile from AB to FB, which provides superior execution in terms of risk or cost, or both, in some cases. It is found by drawing a line from the return value corresponding to principal bid transaction (F) on the y axis tangent to the minimum point (B) on the efficient frontier. This enhanced set of points is named the *capital trade line* (CTL). This closely resembles the capital market line commonly used in the capital asset pricing model (CAPM).

Portfolio Trading Costs

Portfolio trading cost analysis represents a more complex case—fortunately, it can extend from the basic model of single-stock trading. Although the analysis in both cases is conceptually close, it will differ in one important aspect—portfolio trading costs capture important cross-effects between prices and over time. In particular, unless the securities in a portfolio are independent in terms of transaction costs—i.e., trades in one security have no impact on other security prices—the market impact of a portfolio trade will differ from a weighted average of the individual stocks. Whether or not the portfolio execution costs are greater than the sum of the individual securities costs will largely depend on the signs and magnitude of the cross-effects.

A good starting point for analyzing portfolio transaction costs is the volatility of portfolio returns, as compared with those of individual securities. The portfolio diversification effect is a well-known concept and is based on the fact that correlation between stocks causes the portfolio volatility to differ from the simple-weighted sum of stock volatilities (it usually tends to be lower). Therefore, portfolio volatility is made up of two components—its variance (δ_i) and covariance effects (δ_{ij}), i.e., $\delta_{portfolio} = (\Sigma\omega_i^2\delta_i^2 + \Sigma\Sigma\omega_i\omega_j\delta_{ij})^{1/2}$. Similar effects might be expected in other variables, such as duration of the trade.

Another important consideration in the portfolio case is that the volatility of the portfolio returns changes over the trading horizon, as some of the portfolio positions are liquidated. In this case, the risk needs to be estimated for each period *t*, and the total variation of returns is computed as follows:

$$\begin{aligned}\delta(\Delta p(t)) &= ((\delta_1^2(\Delta p) + (\delta_2^2(\Delta p) + \ldots + (\delta_t^2(\Delta p))^{1/2} \\ &= (\Sigma\delta_2^2(\Delta p_j))^{1/2}\end{aligned}$$

Kissell and Glantz [2003] note that *risk contribution* (RC) is an important measure in analyzing the portfolio trading case. This study indicates how much risk a particular stock or position adds to or deducts from the overall portfolio risk. Such a measure is helpful in prioritizing a sequence of portfolio trades—the stocks with high RCs should be traded more aggressively, and stocks that have

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

EXHIBIT 17
Comparison of Industry Models

	PLEXUS	BARRA	ITG	Risk-Inventory
Factors	▪ Buy/sell balance (product of size and 2-day momentum) ▪ market capitalization ▪ % av. daily volume ▪ total order shares ▪ a small-trade "dummy" (equals 1 if the trade is under 10,000 shares)	▪ trade volume ▪ elasticity ▪ volatility ▪ intensity ▪ distribution of trade sizes ▪ market tone (price of risk) ▪ investors skill	▪ security (ticker, CUSIP, exchange) ▪ closing price ▪ volatility ▪ av. trading volume (21-day median value) ▪ bid/ask spread (5-day av, time weighted) ▪ distribution of trading volume	▪ volatility ▪ % av daily volume ▪ participation fraction ▪ crossing fraction ▪ bid-ask spread ▪ capitalization ▪ av. market depth ▪ etc.
Benchmark	implementation shortfall	pre-trade price	implementation shortfall	pre-trade price
Opportunity costs	Included	excluded	included	N/a
Buys vs. sells	Included	included	included	Varies
Regions covered	NYSE, Nasdaq, Japan, UK, Canada, Europe, Emerging Europe, Asia, Emerging Asia, Latin America	US	US, UK, Canada, Australia, UK, France, Germany, Italy, Japan, Hong Kong, Singapore	Varies
Cost optimization/trading strategy	-	-	+	-
Underlying system	TransPort™	Aegis™	ITG ACE™	Varies

Source: JPMorgan.

lower RCs, or even reduce the total risk of the portfolio, can be postponed until a later period in the trading horizon.

Practical points to remember:

1. Market impact is related directly to trade difficulty, which can be approximated by a number of factors (trade size, volatility, market momentum, and market capitalization are the most commonly used).
2. Owing to uncertainty surrounding final estimates, the execution cost is best described as *a distribution*, with a corresponding expected mean and standard deviation of the variable.
3. Estimated cost is expected to be a *non-linear function* of the trade size; however, exact functional dependency depends on prevailing market conditions.
4. The *trade-off* exists between trading patiently, to reduce market impact costs, but with greater uncertainty in the final results, and trading more aggressively to reduce timing risk, but at the expense of substantial market impact.
5. Estimated *portfolio trading costs* should capture important cross-effects among corresponding security prices.

INDUSTRY MODELS

At this point in the study, our analysis switches from mostly theoretical models to their practical applications. If previous sections provided an in-depth overview of theoretical studies as related to transaction costs and their components, the remainder of the study deals with more practical issues, such as market impact models developed by industry practitioners, the impact of the market structure on transaction costs, and an analysis of other significant market factors and changes in regulatory compliance.

The increasing importance of transaction costs resulted in several market impact models being developed by the brokers and third-party services providers. The risk inventory model (which, as the name implies, directly relates transaction costs to the liquidity provider's risk of carrying excess inventories) is commonly used in the industry. Other third-party applications may include ITG, Barra, and Plexus market impact models, which are designed to assist their clients in searching for optimal trading solutions (*see Exhibit 17*).

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

Risk Inventory Models

Owing to the simplified structure of the risk inventory model, it is quite commonly used in the investment industry. The risk inventory model is a total-impact model, combining both opportunity cost (volatility risk) with market impact. It incorporates the basic principle of theoretical inventory control models—the liquidity provider should be compensated for the risk of bearing extra inventories. In its simple form, the inventory cost model consists of just two components, representing liquidity and volatility costs. The first component usually deals with an effective spread or, in fact, any concession paid to the liquidity provider for offering the liquidity, while the second component refers to the movement of the mid-quote itself, representing the volatility of the stock (*see Exhibit 16*).

So, the total market impact cost can be expressed as follows:

Market Impact Cost = Liquidity Cost + Volatility Cost

where:

The *liquidity component* is the normalized cost of trading stocks in the market, as defined by the difference between the mid-price and the bid (for sell) or the ask (for buy) prices. By definition, it approximates one-half of the effective spread. Typically, liquidity cost is set at one-half of the quoted spread, but it might be higher for illiquid stocks or larger orders. The liquidity cost can be modeled based on a number of parameters that describe liquidity conditions in the market, such as: 1) historical average bid-ask spread; 2) stock capitalization; 3) % of average daily trading volume; 4) market depth, etc.

The *volatility cost* arises because, during the course of execution, the investor is exposed to the movements in stock prices. Since the total risk of unfavorable price fluctuations increases with time, the expected mid-quote of a transaction can be estimated as a function of duration to trade (e.g., the size of the trade expressed as a fraction of the typical daily volume). One of the shortcomings of the model is that the volatility is commonly assumed to be constant over time.

Plexus PAEG/Ls

PAEG/L (Plexus Average Execution Gain/Loss) estimates represent a good example of how aggregated market data of prior trade executions can be used in forecasting future market impact costs. PAEG/L equations are derived from the past execution results accumulated over the extensive client database, using comparable trades. In essence, they provide a peer-based comparative standard for a typical experience of executing similar trades in similar situations.

Each quarter, Plexus aggregates client data (over a million orders go into the computation of US PAEG/Ls and over 300,000 orders are used for derivation of non-U.S. equations) for the previous six months, including both completed trades and corresponding desk orders. The data is then screened for possible distortions (e.g., outlier observations, very small trades). For example, the unusually large orders, or orders with excessive gains/losses, are excluded from the sample, as they might skew the results and lessen PAEG/L's predictive capacity.

Normalized data are fitted with a linear regression, using the following five independent variables:

- buy/sell balance (as a product of size and two-day momentum);
- market capitalization;
- percentage of average daily volume;
- total number of order shares; and
- a small-trade "dummy" that equals 1 if the trade is under 10,000 shares.

Note that all variables are closely related to the liquidity conditions in the market. The buy/sell balance or two-day momentum clearly indicates whether the order will be liquidity-demanding or liquidity-providing in the market. The market capitalization is an "intrinsic" liquidity measure of a particular stock, and both the percentage of average daily volumes and the total order shares refer to the size of the order, where larger orders are naturally less liquid and more difficult to execute.

The equations are updated quarterly to reflect constantly changing liquidity conditions in the market. Also, to reflect a trading environment specific to each market, Plexus allocates the data and calculates separate benchmarks for the NYSE, Nasdaq, Japan, U.K., Canada, Europe, Emerging Europe, Asia, Emerging Asia, and Latin America regions. In addition, distinctive data sets are used for buys and sells, resulting in specific PAEG/L equations corresponding to the direction of the trade.

Both the major strength and weakness of PAEG/L equations is that the equations are closely tied to the historical data of trade executions. Using actual data assures that they incorporate a "real-world" experience. Plexus reports that, as the number of observations grows, the fit

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

improves and, over time, PAEG/Ls closely approximate the actual data. Therefore, PAEG/Ls are best applied for analyzing the big picture and are better used as benchmarks in cost measurement than in cost estimates in pre-trade analysis (as they can fail to capture an uncertainty, related to a particular transaction). In this case, estimating the market impact for unusually large or difficult trades might prove to be problematic, simply because of a lack of comparable trades.

Barra Market Impact Model

Unlike PAEG/Ls estimates, purely based on market data, Barra's Market Impact Model (MIM) introduces a modeling strategy that combines both analytical and empirical approaches. Barra suggests that this combination is more suitable for market impact modeling, where the analytical approach can overcome limitations of data and where empirical tests will ensure that the results are closely related to the realities of the market.

A main theoretical concept underlying Barra's Market Impact Model is that the market impact cost of a trade should be proportional to the risk borne by the liquidity provider. The model uses four factors (volatility, elasticity, intensity, and shape) that characterize a specific asset, one factor (market tone) that approximates overall market conditions, and one factor (investor's skill) that relates the forecast to a specific investor's trading ability for estimating market impact costs. In general form, the final framework of Barra's MIM can be shown as follows:

$$\kappa = F(V, \varepsilon, \delta, \phi, \zeta, \tau, \chi) = (1 - X)\, \delta(V, \varepsilon, \delta, \phi, \zeta, \tau)V$$

where:

κ = the forecast market impact of the trade;
V = the trade size measured by value;
$F(\bullet)$ = the function that integrates the factors into the final forecast.

asset-specific factors:
ε = the elasticity, or an order flow response to price changes;
δ = the volatility of asset price;
ϕ = the intensity, describing how often the asset is traded;
ζ = the shape, or distribution of trade sizes.

market-specific factor:
τ = the market tone, or price of liquidity.

investor-specific factor:
χ = the investor's skill.

Barra's model uses the implementation shortfall approach (i.e., measuring cost from a pre-trade reference price) for estimating market impact costs. However, owing to the lack of data on the preliminary intentions of market participants to trade, it does not calculate opportunity costs explicitly. It is important to note that Barra uses forward-looking estimates of market characteristics and therefore employs four submodels of volatility, elasticity, intensity, and shape, providing the input data required for the final model.

A market-specific factor τ, named as *the market tone*, has a significant role in Barra's model. Theoretically, it represents a price for each unit of risk borne by the liquidity provider and, owing to the competitive pressures in the market, it is assumed to be uniform for all market participants. Given that τ is assumed to be a market-wide number, it should be possible to estimate in practice from trade results using cross-sectional regressions. Indeed, Barra finds that the market tone τ is a relatively constant number and tends to change by only a few percent on a day-to-day basis.

The market impact cost function $F(\bullet)$ is an econometric function that incorporates all estimates into the final forecast. Given that the full function $F(\bullet)$ is quite complex and time-consuming to estimate, in Barra's Market Impact Model it is actually replaced by a piecewise linear approximation. The piecewise linear curve is a continuous curve formed by joining straight-line segments together. In principle, it can approximate any smooth curve with a relatively high degree of accuracy, and provides a time-efficient version for the final $F(\bullet)$ function.

The last factor—investor's skill—incorporates investor-specific characteristics, such as style, technology, experience, etc., which reflect how efficient a particular investor is in searching for/purchasing the liquidity.

The MIM results are presented in the form of short-range and long-range forecasts of trading costs. The short-range forecast refers to current market conditions and is valid for trading on that particular day and/or over the next few days. The long-term estimate includes an average of short-term forecasts over the past two months and is provided for more general strategic planning purposes. It reflects the forecast costs under typical market conditions. All estimates of Barra's Market Impact Model are updated daily.

The natural question is then how well an econometric model can approximate the real trading conditions in the market. Barra's statistical tests show that the model

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

correctly identifies the more expensive trade of two assets 92% of the time, and, similarly, that forecast impact costs for a trade are accurate within a 10% value range.

ITG Agency Cost Estimator (ACE)

The ACE (Agency Cost Estimator) is an econometric model that forecasts transaction costs based on the optimal balance between market impact and opportunity costs. Considering theoretical research in this area, it most closely resembles the model structure originally introduced by Almgren and Chriss [2000/2001]. The ACE model considers that a purely cost-minimizing strategy might not necessarily be an ideal solution. For example, a trader can minimize costs by breaking up the trade over a very long time horizon; however, in this case, he will face the risk of significant market movements. Therefore, the optimal strategy can be derived by balancing these two considerations—opportunity and market impact costs.

ACE measures execution costs using the implementation shortfall approach, where the execution cost is defined as the difference between the average execution price for the trade and the prevailing price at the start of the order execution. This measure of execution costs includes the bid-ask spread, the market trend, and the incremental market impact of the trade. ACE also considers opportunity costs, which arise owing to price volatility and uncertainty in the realized cost of trading.

The ACE model is based on stock-specific econometric models of volatility, price impact, and price improvement, which are required inputs to solve the optimization problem. ACE uses market parameters, such as security master information (ticker, CUSIP, exchange), closing prices, volatility, trading volume, bid-ask spread, and distribution of trading volume volatility for estimating these parameters. The model also uses the forward-looking intraday volatility based on the standard deviation of return for the most recent 60 trading days.

SUMMARY

One of the most important decisions facing investment managers is how quickly they should implement their buy and sell decisions. On the one hand, trading slowly can mitigate impact, while trading aggressively will reduce uncertainty and opportunity costs of unexecuted orders.

Early literature on the subject developed a conceptual framework which is still applicable to market microstructure forces today. For example, effects on price discovery due to broker inventory levels and market participant trading patterns can still be rationalized by using these ideas. More recent publications have set out an excellent framework for optimising execution, while developing further practical market impact models.

> *"I have heard of people who amuse themselves conducting imaginary operations in the stock market to prove with imaginary dollars how right they are. Sometimes these ghost gamblers make millions. It is very easy to be a plunger that way. It is like the old story of the man who was going to fight the duel next day."*
>
> *His second asked him, "Are you a good shot?"*
>
> *"Well," said the duelist, "I can snap the stem of a wineglass at twenty paces," and he looked modest.*
>
> *"That's all very well," said the unimpressed second.*
>
> *"But can you snap the stem of the wineglass while the wineglass is pointing a loaded pistol straight at your heart?"*
> —Edwin Lefèvre, "Reminiscences of a Stock Operator"

Lefèvre makes a very clear point that the market is alive, and that, in trying to implement any strategy, one is brought face to face with the desires of other participants. Because trading is driven by individuals and their psychology, a quantitative approach will never truly replace the skill of the individual trader. However, developing trading strategies and knowing where costs come from at the outset will certainly contribute to minimizing overall costs and maximizing investor alpha.

REFERENCES

Admati, A., and P. Pfeiderer. "A Theory of Intraday Patterns: Volume and Price Variability." *Review of Financial Studies,* 1 (Spring 1988), pp. 3-40.

Almgren, R. "Optimal Execution with Nonlinear Impact Functions and Trading-Enhanced Risk." *Applied Mathematical Finance,* 10 (2003), pp. 1-18.

Almgren, R., and N. Chriss. "Value Under Liquidation." *Risk Magazine*, December 1999.

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

——. "Optimal Execution of Portfolio Transactions." *Journal of Risk*, 3 (Winter 2000/2001), pp. 5-39.

——. "Competitive Bids for Principal Program Trades." Working paper, 2002.

Amidhud, Y., and H. Mendelson. "Dealership Market: Market Making with Inventory." *Journal of Financial Economics,* 8 (1980), pp. 31-53.

Bagehot, W. [J. Treynor.] "The Only Game in Town." *Financial Analysts Journal*, vol. 27, no. 2 (March/April 1971), pp. 12-14, 22.

Barnea A., and D.E. Logue. "Stock Trading and Portfolio Performance." *Journal of Bank Research,* 7 (Summer 1976), pp. 150-157.

Batalio, R., and C. Holden. "A Simple Model for Payment for Order Flow, Internalization, and Total Trading Cost." *Journal of Financial Markets*, 4 (2001), pp. 33-71.

Beebower, G., and W. Priest. "The Tricks of the Trade." *The Journal of Portfolio Management*, vol. 6, no. 2 (Winter 1980), pp. 36-42.

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

Bertsimas, D., and A. Lo. "Optimal Control of Execution Costs." *Journal of Financial Markets*, 1 (1996), pp. 1-50.

——. "Optimal Control of Execution Costs." *Journal of Financial Markets*, 1 (1998), pp. 1-50.

Biais, B., L. Glosten, and C. Spatt. "The Microstructure of Stock Markets." Working paper, L'Institut d'Economie Industrielle, 2002.

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

Blume, M., and M.A. Goldstein. "Quotes, Order Flow, and Price Discovery." *The Journal of Finance*, vol. 52, no. 1 (March 1997), pp. 221-244.

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

Choi, J., D. Salandro, and K. Shastri. "On the Estimation of Bid-Ask Spreads: Theory and Evidence." *Journal of Financial and Quantitative Analysis,* 23 (1988), pp. 219-230.

Cohen, K., S. Maier, R. Schwartz, and D. Whitcomb. "Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread." *Journal of Political Economy*, 89 (1981), pp. 287-305.

Copeland, T., and D. Galai. "Information Effects and the Bid-Ask Spread." *Journal of Finance,* 38 (1983), pp. 1457-1469.

Coughenour, J., and K. Shastri. "Symposium on Market Microstructure: A Review of the Empirical Evidence." *Financial Review*, 1999.

Cushing, D., and A. Madhavan. "The Hidden Cost of Trading at the Close." In Brian Bruce, ed., *Transaction Costs: A Cutting Edge Guide to Best Execution*, New York: Institutional Investor, 2001.

Demsetz, H. "The Cost of Trading." *Quarterly Journal of Economics*, 82 (1968), pp. 33-53.

Easley, D., N. Kiefer, M. and O'Hara. "Liquidity, Cream-skimming or Profit Sharing? The Curious Role of Purchased Order Flow." *Journal of Finance,* 51 (1996), pp. 811-833.

Easley, D., and M. O'Hara. "Price, Trade Size, and Information in Securities Markets." *Journal of Financial Economics*, vol. 19, no.1 (September 1987), pp. 69-90.

——. "Time and the Process of Security Price Adjustment." *Journal of Finance,* 47 (1992), pp. 577-606.

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

Foster, F.D., and S. Viswanathan. "A Theory of the Intraday Variations in Volume, Variance and Trading Costs in Securities Markets." *Review of Financial Studies,* 3 (1990), pp. 593-624.

Garman, M. "Market Microstructure." *Journal of Financial Economics,* 3 (1976), pp. 257-275.

George, T., G. Kaul, and M. Nimalendran. "Estimation of the Bid-Ask Spread and Its Components: A New Approach." *Review of Financial Studies,* 4 (1991), pp. 623-656.

Glosten, L., and P. Milgrom. "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics*, vol. 14, no. 1 (March 1985), pp. 71-100.

Hansch, O., N. Naik, and S. Viswanathan. "Do Inventories Matter in Dealership Markets? Evidence from the London Stock Exchange." *Journal of Finance*, 53 (1998), pp. 1623-1656.

Harris, L. *Trading and Exchanges: Market Microstructure for Practitioners.* Oxford University Press, 2003.

Hasbrouck, J. "The Summary Informativeness of Stock Trades." *Review of Financial Studies,* 4 (1991), pp. 571-594.

Ho, T., and R. Macris. "Dealer Bid-Ask Quotes and Transaction Prices: An Empirical Study of Some AMEX Options." *Journal of Finance*, 39 (1984), pp. 23-45.

Ho, T., and H. Stoll. "Optimal Dealer Pricing Under Transactions and Return Uncertainty." *Journal of Financial Economics,* 9 (1981), pp. 47-73.

——. "The Dynamics of Dealer Markets Under Competition." *Journal of Finance,* 38 (1983), pp. 1053-1074.

The Journal of Investing 2004.13.3:93-115. Downloaded from www.iijournals.com by NEW YORK UNIVERSITY on 02/19/15.
It is illegal to make unauthorized copies of this article, forward to an unauthorized user or to post electronically without Publisher permission.

Holthausen, R., R. Leftwich, and D. Mayers. "The Effect of Large Block Transactions on Security Prices." *Journal of Financial Economics,* 19 (1987), pp. 237-267.

Jain, P.C., and G.-H. Joh. "The Dependence Between Hourly Prices and Trading Volume." *Journal of Financial and Quantitative Analysis,* 23 (1988), pp. 269-284.

Keim, D.B., and A. Madhavan. "Anatomy of the Trading Process: Empirical Evidence of the Behaviour of Institutional Trades." *Journal of Financial Economics*, vol. 37, no. 3 (March 1995), pp. 371-398.

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

——. "Transaction Costs and Investment Style: An Inter-Exchange Analysis of Institutional Equity Trades." *Journal of Financial Economics*, 3 (1997), pp. 265-292.

——. "The Costs of Institutional Equity Trades: An Overview." *Financial Analysts Journal,* vol. 54, no. 4 (July/August 1998), pp. 50-69.

Kissell, R., and M. Glantz. "Optimal Trading Strategies." AMACOM American Management Association, 2003.

Kraus, A., and H. Stoll. "Price Impacts of Block Trading on the New York Stock Exchange." *Journal of Finance,* 27 (1972), pp. 569-588.

Kyle, A.S. "Continuous Auctions and Insider Trading." *Econometrica*, 53 (1985), pp. 1315-1336.

Leinweber, D. "Nerds on Wallstreet." In *Blending Quantitative and Traditional Equity Analysis*. AIMR ICFA Continuing Education Series, 1995.

Loeb, T.F. "Trading Cost: The Critical Link between Investment Information and Results." *Financial Analysts Journal*, May/June 1983, pp. 39-43.

Madhavan, A. "A Bayesian Model of Intraday Specialist Pricing." *Journal of Financial Economics,* 30 (1991), pp. 99-134.

——. "An Analysis of Changes in Specialists Quotes and Inventories." *Journal of Finance*, vol. 48, no. 5 (December 1993), pp. 595-1628.

——. "Market Microstructure: A Survey." *Journal of Financial Markets,* vol. 3, no. 3 (August 2000), pp. 205-258.

——. "Market Microstructure: A Practitioner's Guide." *Financial Analysts Journal*, September/October 2002.

Madhavan, A., and M. Cheng. "In Search of Liquidity: An Analysis of Upstairs and Downstairs Trades." *Review of Financial Studies,* vol. 10, no. 1 (Spring 1997), pp. 175-204.

Madhavan A., M. Richardson, and M. Roomans. "Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks." *Review of Financial Studies*, 10 (1997) pp. 1035-1064.

Madhavan, A., and G. Sofianos. "An Empirical Analysis of NYSE Specialist Trading." *Journal of Financial Economics*, 48 (1998), pp. 189-210.

Manaster, S., and S.C. Mann. "Life in the Pits: Competitive Market Making and Inventory Control." *Review of Financial Studies*, 9 (1996), pp. 953-975.

O'Hara, M. *Market Microstructure Theory*. Cambridge, MA: Blackwell, 1995.

O'Hara, M., and G. Oldfield. "The Microeconomics of Market Making." *Journal of Financial and Quantitative Analysis,* 21 (December 1986), pp. 361-376.

Patel, B. "Measurement of Transaction Costs in Equity Trading." In Brian Bruce, ed.,*Transaction Costs: A Cutting Edge Guide to Best Execution*. New York: Institutional Investor, 2001.

Perold, A. "The Implementation Shortfall: Paper versus Reality." *The Journal of Portfolio Management*, vol. 14, no. 3 (Spring 1988), p. 49.

Roll, R. "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market." *Journal of Finance,* 39 (1984), pp. 1127-1139.

Stoll, H. "Dealer Inventory Behaviour: An Empirical Investigation of NASDAQ Stocks." *Journal of Financial and Quantitative Analysis,* 11 (1976), pp. 359-380.

——. "The Supply of Dealer Services in Securities Markets." *Journal of Finance,* 33 (1978), pp. 1133-1151.

——. "Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests." *Journal of Finance*, vol. 44, no. 1 (March 1989), pp. 115-134.

Subrahmanyam, A. "A Theory of Trading in Stock Index Futures." *Review of Financial Studies,* 4 (1991), pp. 17-51.

Tinic, S. "The Economics of Liquidity Services." *Quarterly Journal of Economics,* 66 (1972), pp. 79-97.

Willoughby, J. "Execution's Song." *Institutional Investor*, vol. 32, no. 11 (1998), pp. 51-56.

Wood, R.A., T.H. McInish, and J.K. Ord."An Investigation of Transaction Data for NYSE Stocks." *Journal of Finance,* 25 (1985), pp. 723-739.

To order reprints of this article, please contact Ajani Malik at amalik@iijournals.com or 212-224-3205.