

# The Measurement and Control of Trading Costs

*Trading costs are probably not as inexpensive as some believe, nor as expensive as others have suggested. Trading costs are large enough to merit the focused attention of any investor. To some extent, they can be controlled. Such cost control (or lack of same) can profoundly affect long-term investment success.*

*Investment managers who perceive trading costs as small will formulate an investment process independent of the trading process. This encourages a convenient but false view that trading is an unimportant component of the investment process. Disappointing investment performance is often the result.*

*Insightful trading cost measurement requires a sequential tracing of investment ideas from manager through trader to implemented portfolio. The search is not for a meaningless proof of "best execution," but for areas that need improvement. The results will not necessarily show up in lower commissions or even low measures of trading costs. They will, however, translate into improved investment performance.*

**T**HE MEASUREMENT OF trading costs has created controversy for at least 20 years. Various studies have estimated total equity trading costs at 500, 100 and 10 basis points. Surely these vastly disparate costs would call for very different behavior on the part of portfolio managers. Investors need to examine the puzzle of trading costs from the vantage of both cost control and process control.

Early investigators of trading costs quickly determined that there is no unique "true" measure of trading costs. Ironically, the underlying truth is that accurate measurement is not possible. The true measure of trading costs is not the commission paid. Nor is it the difference between the trade price and the high, low, median or closing price that day. Nor can it be inferred from stock price movements in the hours before or after the trade is executed. The true trading cost is the difference between the execution price and the price that would have prevailed in the absence of the trade. Analogous to the Heisenberg principle in physics, the process of executing a trade masks what would exist without the trade taking place. The difference, forever unobservable, is the true trading cost.

A useful analogy can be drawn from the bond market. Most bond portfolios are priced from "matrix prices," which are thought to represent "fair value" better than actual transaction prices. Bond transaction prices, it is believed, are too dependent on the idiosyncratic meeting of one buyer and one seller. The same curious conclusion can be drawn for equities. Actual prices are set by lonely, marginal participants who represent not a consensus, but the strongest motivation to transact at a particular point in time.

The need to outbid competitive traders suggests that trading costs can never be negative. The price would always be the same or lower without the most motivated buyer's willingness to buy. Similarly, price would always be the same or higher without the most motivated seller's willingness to sell.

Even this definition of trading costs applies only to trades that actually execute. Trades that cannot be completed represent an even larger, and equally unmeasurable, trading cost. For uncompleted trades, we can observe unaffected prices—provided that our presence was not instrumental in shifting prices without our

## Glossary

**Trading Cost:** The overall cost of executing a transaction, including all commissions and the market impact of the trade.

**Matrix Prices:** On thinly traded markets, pricing is often based on a "matrix." With this approach, the value or price of a security is estimated on the basis of the price movements of similar securities that have recently traded.

**Immediacy Cost:** The market impact of any desire for immediate trading. Those who want to transact immediately may have to deal with a broker, who will take a risk position in the security, and may need to pay the broker's profit margin in order to effect a rapid trade.

**Opportunity Cost:** The cost of not participating in the success of an investment decision because one has not traded or not traded fast enough.

**Zero-Sum Game:** Any "game" in which each winner is balanced by a loser. A game of poker with friends (with no casino taking a cut) is a zero-sum game. Stock market investing is not, because the market has historically tended to provide favorable rewards of differing magnitudes to the players. In the long run, it is a "positive-sum game."

**Best Execution:** Transacting at the best possible net price inclusive of commissions.

transacting—but here we cannot observe actual completion prices.

Not only do our actions affect trading costs, the actions of like-motivated investors may also have an effect. Information moves stock prices as investors act in response. De Bondt and Thaler show that price movements tend to overshoot in response to significant news as an overabundance of buyers (or sellers) compete to be the ones who complete their trades.<sup>1</sup> As Treynor has pointed out, market-makers absorb temporary excesses in supply or demand with the expectation of trading out later at a profit.<sup>2</sup> The net effect is that trading costs get entwined in the mechanics of both the investment process and the stock market mechanism. Transaction costs may reflect not only the information that drives the transaction, but the competition between traders that causes prices to overshoot. No wonder bond investors prefer matrix prices!

Because unaffected prices are unobservable, true trading costs are inherently unmeasurable and can only be inferred. Importantly, this is

not a fatal flaw: Relative to the imprecision of other financial measurements (e.g., beta, value, future internal rates of return), trading costs can be estimated with a useful degree of accuracy.

It is useful for managers to measure their trading costs. Correctly constructed measurement provides valuable insights into the investment process. This is not to say that all transaction cost estimates are useful. Imperfect measures designed to validate a legal requirement for "best execution" must be viewed with a very large dose of skepticism. They fail to address the key question: "What are the fundamentals of the investment/transaction process we are trying to measure?"

Trading costs are probably not as inexpensive as some believe, nor as expensive as others have suggested. Trading costs are, we believe, large enough to merit the attention of any investor. To some extent, they can be controlled. Such control (or its lack) profoundly affects long-term investment success. Therein lies the challenge. The investor who views trading as comparatively unimportant exposes himself to a needless cost in the execution of his orders.

## Focusing on the True Costs of Trading

The tactical decision faced by every trader is whether to act immediately or wait and hope for more favorable terms. A trader incurs an *immediacy cost* by moving quickly, while exposing himself to a potential *opportunity cost* if he waits.

The immediacy cost is the cost a transactor incurs to trade now. A trader who demands immediate execution relies on the market structure to provide liquidity. But continuously available liquidity is not a free good; it must be manufactured. Where liquidity is frequently demanded, middlemen—dealers and specialists—provide liquidity whenever the offsetting natural trade cannot be found immediately. For this service, market-makers are well compensated. Markets without middlemen are often characterized by long delays and significant immediacy costs. Well conceived market structures enhance liquidity and increase the size of trades that can be completed without significant market impact.

The market-maker buys at a price lower than the price at which he expects to be able to sell the stock. If it is possible to [a] wait for that buyer to arrive and [b] find him when he arrives, a transactor can eliminate the market-maker's services and capture the market-

1. Footnotes appear at the end of article.

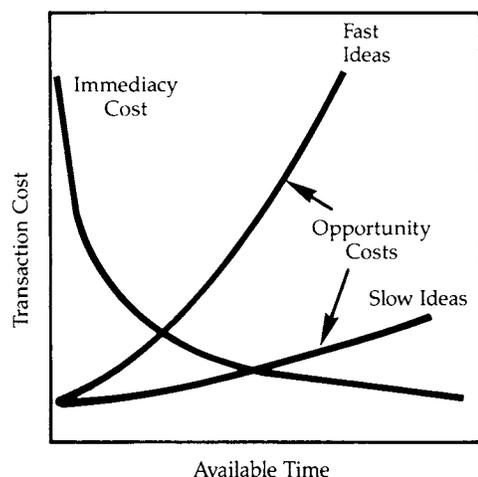
maker's compensation for himself (or at least split it with the other party). Thus the immediacy cost component of trading costs declines rapidly as the available time to complete the order increases. Figure A illustrates this.

Opportunity costs, represented by the two upward sloping curves in the figure, relate to the "half-life" of investment ideas. Investment ideas have an expected time to fruition. This time interval varies across investment strategies. If competing traders move the price prior to execution, the value of the investment idea will be lost. An opportunity cost is incurred.

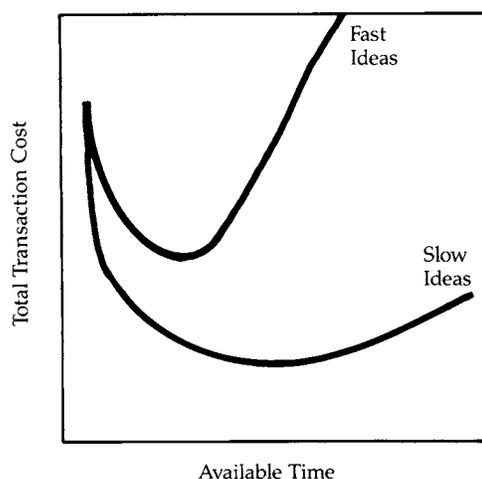
The immediacy cost decreases when there is time to work the trade. Unfortunately, as the immediacy cost goes down, the expected opportunity cost rises. Investment strategies based on "first call," fast information or momentum need rapid execution to avoid opportunity cost. These investors should be willing to pay the immediacy cost in order to avoid a steeply rising opportunity cost, as shown by the top-most curve in Figure A.

Investors with slower ideas (e.g., "hidden values," out-of-favor stocks, low P/E) can attempt to escape from immediacy costs because their opportunity cost curve slopes up less quickly. These investors should favor limit orders, crosses and other techniques designed to attract the natural other side. A trader who can wait for—and accommodate—the natural other side shifts from being a consumer of liquidity toward being a provider of liquidity, a role

**Figure A** Immediacy vs. Opportunity Cost



**Figure B** Total Cost



analogous to that played by the dealer. This trader avoids high immediacy costs.

To summarize, trading cost is a sum of two curves—a downward sloping immediacy cost and an upward sloping opportunity cost. Figure B illustrates this. Here total expected trading cost is represented by the U-shaped curves. The total cost curve has a "belly" of lowest expected cost, where immediacy cost has declined but opportunity cost has yet to have major impact. This is the area in which a trader should aim to trade.

This is not as straightforward as the diagram implies. This minimum is never directly observable, and changes from moment to moment and from trade to trade. Herein lie the subtleties of trading, subtleties that constitute the "art" of trading.

### Two Views of Transaction Costs

There seem to be two schools of thought on trading costs. Some observers believe that trading costs are insignificant and can be effectively ignored in designing and implementing an investment strategy. Others believe that costs are significant and that the shortfall between active manager performance and index performance can be attributed to hidden costs of trading.

### The High-Cost Case

Some practitioners and theoreticians have suggested that transaction costs are significant. They point to the inconsistency between low transaction cost estimates and the realized per-

**Table I** Median Equity Manager Performance vs. S&P 500 Index Returns

	Median Manager Performance	S&P 500		Percentage of Managers Outperforming
		Return	Difference	
1988	17.13%	16.83%	+0.30%	51%
1987	3.59	5.22	-1.63	39
1986	16.76	18.55	-1.79	36
1985	29.10	32.04	-2.94	28
1984	1.85	6.10	-4.25	28
1983	20.10	22.36	-2.26	36
1982	21.33	21.09	+0.24	51
1981	-1.47	-5.10	+3.63	71
1980	30.67	32.38	-1.71	43
Average			-1.16%	

Source: T. J. Loeb, "Trading Cost: The Critical Link Between Investment Information and Results," *Financial Analysts Journal*, May/June 1983; updated by the authors.

formance of actively managed funds. Table I compares annual median equity managed investment results for the decade with the S&P 500 index.

Where does the performance for such a large set of institutional equity managers go? The high-cost advocates suspect frictional costs incurred during the portfolio revision process. Three authors have pointed out differing aspects of high transaction costs.

Jack Treynor has focused on the process of market-making, identifying the costs of market-making and the costs of providing liquidity.<sup>3</sup> Treynor has concluded that the inside spread is not a true measure of transaction costs; the true measure is the spread between what the ultimate buyer pays and what the ultimate seller receives. Treynor has informally estimated the outside spread to be in the 10 per cent range!

In effect, Treynor argues that stock sold to a middleman is not the same as stock sold to an investor; it is inventoried, and is still for sale. It is in the hands of someone looking for a positive trading return and won't be "sold" until it is in the hands of an investor who seeks a long-term investment. Until it settles out of the market, it is competing with—and affecting the trade price of—the remainder of what the investor is trying to sell.

Thomas Loeb attempted to measure the cost of trading immediacy by requesting real dealer quotes on a carefully constructed sample of different trade sizes and different company capitalizations.<sup>4</sup> Loeb estimated the round-trip cost of trading immediacy to range from 1.1 to 43.8 per cent, with extravagantly higher figures associated with large blocks in small companies.

Loeb simulated a demand for instant liquidity from the brokers, sometimes for hazardous trading amounts. Loeb's data strongly support the idea that liquidity is not a free good, and that rational trading strategies will seek to avoid paying for unneeded immediacy.

Mark Edwards performed an analysis along lines suggested by André Perold.<sup>5</sup> Mark found that opportunity costs, the costs of uncompleted trades, are as high as or higher than simple execution costs. Edwards makes a distinction between *execution* costs and *implementation* costs. Edwards' conclusions are particularly interesting in that he investigates a transaction cost totally ignored by commercial evaluation services.

### The Low-Cost Case

Some traders believe that in every transaction there is a winner and a loser. If the stock trades on an uptick, then the buyer incurs the trading cost, while the seller benefits. If the stock trades on a downtick, then it is the seller who loses while the buyer wins. In this poker-game view, the market impact portion of a trade is a "zero sum game."

The zero-sum argument assumes that the buyers and sellers in the market are all "natural" transactors. (For the purposes of this discussion, we assume that a natural buyer or seller operates from an investment horizon in excess of one week.) In practice, however, the other party to a liquidity-demanding trade is frequently not another natural transactor. Most liquidity transactions involve a market intermediary—a dealer, a specialist, a middleman or similar party who trades *only* to accommodate

someone willing to pay up to complete a transaction quickly.

A recent quote from a trader suggests that there are six middlemen in the market for each natural transactor.<sup>6</sup> However populous they may be, all middlemen expect to make money accommodating natural buyers and sellers. Natural buyers demanding immediacy seldom encounter a natural seller in the open markets.

We might conclude that the middlemen, at least, enjoy negative transaction costs. They do not; they simply overcome their modest trading costs because of their short-term market expertise. Relative to the price they would like to trade at, even the middlemen must compromise to complete the transaction. They must accept a price that is simultaneously attractive enough to motivate the other side and better than any other transactor on the same side. Universal compromise occurs; competition for the trade implies that it goes to the transactor who is most anxious to trade, for whatever reason.

The truth is, *all transactors incur positive trading costs*, which reduce the ultimate success of their investment or trading strategy. In effect, all transactors must "give up" something in order to trade at all.

The low-trading-cost advocates disagree that transaction costs are enormous. They argue that every trade desk will have worked out strategies that fill orders with at least an intuitive minimization of cost. The argument, unfortunately, fails to explain adequately the average active manager shortfall vis-à-vis the indexes.

### **The Hazards of Underestimating Transaction Costs**

Investment ideas spring primarily from security analysis and portfolio concepts. At this level of conceptualization, transaction cost considerations are a nuisance. As we will show, however, ignoring transaction costs can lead to very ineffective portfolio implementation.

The principal problem with underestimating transaction costs is that it builds a false sense of security. If trading costs are perceived as small or even slightly negative, the investment process can be formulated independent of the trading process. This encourages the convenient but false view that trading is an unimportant component of the investment process. Disappointing investment performance is often the result.

Consider, for example, the gap between historical investment simulations and actual subse-

quent results. A major contributor to the gap can be found in trades that cannot be executed. *Carelessly conceived strategies often select issues that cannot be used in size in an institutional portfolio.* This is a true trading cost. It represents the cost of the unexecuted or inexecutable trade.<sup>7</sup>

The gap between simulated and actual results can be large because of factors that go beyond trading costs. The simulation may, for example, be based on prices established in low-volume trades. Just because a stock closes at a particular price does not guarantee the same price on the next trade, irrespective of trade size. This mismatch between simulation expectations and true investment management results can lead to leakage of an eighth or a quarter point *each way*, on *each trade*—a significant loss of expected performance, particularly for high-turnover strategies.

As André Perold has suggested, comparing a real-time simulation to a concurrent implementation is more telling than comparing historical simulations to subsequent actual performance.

### **Trading Cost Measurement: What Do We Do?**

The need for trading cost measurements is well recognized by practitioners and their regulators. Despite this, some sponsors believe that trading is part and parcel of the manager-hiring decision. It is no more appropriate, they argue, to inject their presence into the trading decisions than it is for them to opine on whose security research is to be used. Unfortunately, this arm's-length attitude is unworkable in the current investment management environment.

Managers traditionally rely on brokerage commissions (or direction of brokerage business) to cover many operating costs of running a portfolio. Because commissions (and other trading costs) result in payments out of the trust, managers have the ability to withdraw funds and spend them in ways that rightfully relate to the trust's long-term objectives.

The result has been a demand for accountability, one which, unfortunately, has generated some unpleasant side effects. As we have argued, transaction costs are forever unobservable. As a result, accounting for trading costs often focuses on the commission, the only observable part of the trading cost.

Managers are pressured to reduce the commission to bring it into line with the commission paid by other managers. They are also pres-

sured to supply "proof" that they secured "best execution." Managers, unable to disentangle themselves from soft-dollar dependency, and frequently only vaguely aware of how markets work, have acquiesced. Ever obliging, they have appeased their clients by altering trading procedures.

"Proof," if demanded, is easy to supply. A trader who does not wish to look bad relative to everyone else can simply trade like everyone else. Worse still, the trader who understands the vulnerabilities of the standard by which he or she is judged is tempted to game the measurement process.

Suppose, for example, that good trading is defined as selling on an uptick and buying on a downtick. If we measure a trader's effectiveness by this approach, we create an incentive to buy only on a downtick and to sell only on an uptick. Such a strategy gives the appearance of extremely well-controlled trading and negative trading costs. Unfortunately, such a strategy also leads to a high proportion of unexecuted trades and may actually *increase* true trading costs.

The result is tremendous growth in what has been described as "go-along" trading: Market orders, participate orders and other no-fault techniques have become prevalent methods of trading. Transaction measurement will never show these methods as bad trading, but the results do show up in investment performance.

The easiest way to comply with sponsor desires, then, is to reduce commissions. But reducing commissions will not reduce trading costs, particularly if it distorts trading practice. And it is likely to do so, for several reasons.

- The same methods that guarantee "best execution" are also the least expensive for a broker to provide. In essence, the order is not effectively represented, it is simply presented.
- Low commissions erode the agency market and work in favor of a dealer market.
- The brokers who concentrate on excellent handling of trades have become less profitable. Unless agency trading activity is covered by profits from proprietary trading or other unrelated sources, the overall business can become unprofitable. Some brokers respond by revamping their traditional cost structures, eliminating both execution capabilities and research facilities. Others

choose proprietary trading, which places them in a direct adversarial role with their customers.

- Most commercial trade measurement systems completely fail to measure opportunity costs. The trader evaluated under such a system faces no direct penalty for incurring opportunity costs and is thus likely to favor trading strategies that minimize immediacy costs.

Thus it is possible for a trader to look good—on paper—and his manager's stock picks to look good—on paper—at the same time the overall investment performance is a disappointment.

The problem with analyzing execution costs in isolation is that successful investment management does not derive from the efforts of a talented—but isolated—individual. It is the result of a coordinated effort to create investment ideas of intrinsic value and to assure that those ideas are effectively implemented in actual portfolios.

### How to Measure Trading Costs

Most services for measuring trading costs are as simplistic as the uptick/downtick approach mentioned above. The common approach for measuring market impact compares the trade price with (1) the prior close, (2) the subsequent close, (3) the mid-point between the high and the low for the day, (4) the trade-weighted average price during the day or (5) a quartile distribution of prices during the day.

Any trader held to this kind of comparison will have a powerful incentive to trade in a manner that gets good marks from the measurement process. A trader measured against these benchmarks can slant the trading strategy to improve his or her grades. A trader who "games" the measurement benchmark will likely couple inferior implementation of investment decisions to an *appearance* of superior trade execution.

Unfortunately, such gaming results in transaction decisions that are based on non-investment considerations. This interferes with successful investment management. The blame for subsequent inferior performance is then misleadingly attributed to the investment process, not the trading process.

Improperly constructed or applied measurement techniques cannot distinguish between investment strategy and trading strategy. For

example, an investor who evaluates recent news announcements for favorable or unfavorable surprises will usually buy on upticks and sell on downticks. This momentum-following behavior is inherent to the strategy. Yet most trade measurement systems will erroneously identify sloppy trading.

The trading game is commonly compared to a poker game, where the skills of one trader are pitted against the skills of the other players.<sup>8</sup> For traditional traders, the greatest sense of accomplishment comes from reducing the trade price by an eighth (or more). The problem with this "macho" view is that it focuses on immediacy cost, and thereby transforms the trading. Instead of being the last critical step in an investment implementation process, the trading becomes an end in itself. If the poker game analogy has any validity, it is only in the context of playing poker with the rent money: A lot more is at stake than is observed on the game table.

Insightful trading cost measurement, then, requires a sequential tracing of the *implementation* of investment ideas from manager through trader to portfolio. Potential leakages include improperly communicated trading instructions, postponed orders, unexecuted or substituted trades, and inappropriately fast or inappropriately slow trading. Useful insights require the active participation of the entire investment management team. In contrast to the strictures of the Department of Labor, the search is not for a meaningless proof of "best execution," but for areas that need improvement. The results will not necessarily show up in lower commissions or even low measures for trading costs. They will, however, translate into improvements in investment results.

The investor/sponsor's role is to encourage effective teamwork throughout the investment process. This includes asking the following questions.

- (1) What has been bought with my execution-cost dollars?
- (2) How is the optimal trading allocation among research, soft dollars and execution determined?
- (3) How is the appropriate level of execution cost determined?
- (4) How is implementation handled to minimize leakage between idea and implementation? How is the trader informed of

the nature of the trade? How is the portfolio manager informed of the execution limitations on his investment decisions?

## Conclusion

Trading costs have rightly become an area of concern and frustration for investors. Because of the inherent inability to measure trading costs accurately, attention has been focused on the level of the commission. The trader has been pressured to "look good" as measured by misleading and corruptible trading cost measurement reports. The real problems, (1) a lack of effective manager/trader communications and (2) a misunderstanding of markets and how traders can best use their facilities, have been swept out of sight.

The solution is to treat the trading process as the *implementation of investment ideas* and to recognize that excellence in trading is possible, subtle and rewarding. An investment idea, no matter how rich or insightful, is merely an idea until it is implemented. Effective implementation is not automatic. A poor implementation process can overwhelm the value of the idea. In too many organizations, the trading desk is viewed as a clerical function, little more important than the reception desk. Perhaps the trade desk will soon be known as "The Implementation Department" and receive the respect it merits. ■

## Footnotes

1. W. F. De Bondt and R. H. Thaler, "Does the Stock Market Overreact?" *Journal of Finance*, July 1985.
2. J. L. Treynor, "The Economics of the Dealer Function," *Financial Analysts Journal*, November/December 1988.
3. *Ibid.*
4. T. F. Loeb, "Trading Cost: the Critical Link Between Investment Information and Results," *Financial Analysts Journal*, May/June 1983.
5. M. Edwards, "Transaction Costs and the Search for Fool's Gold," *Trader Forum*, June 1989, and A. F. Perold, "The Implementation Shortfall: Paper vs. Reality," *Journal of Portfolio Management*, Spring 1988.
6. K. A. Logan, "The Institutional Equity Business: A House Divided" (Paine Webber, New York, November 30, 1988).
7. Factors unrelated to trading costs play an important role. Historical simulations are often constructed on an *ex post* basis on historically recorded data. An element of biased selection enters into this game. No practitioner will "go public" with a

historical simulation that yields dismal results; such models are rejected before they see the light of day. Furthermore, some over-zealous practitioners succumb to the temptation to use *ex post* simulations in which a model is tested on the same data used to develop the model, because stellar simulations are more likely to attract investor interest.

Furthermore, portfolio construction considerations might result in a gap between simulation and execution. Structured simulations often recommend very lopsided portfolios, resulting in excessive industry or factor exposure. Portfolio managers are frequently compelled to moderate

these structural biases, often muting strategy performance.

The authors do not suggest that simulations are useless. Properly conducted, they can lead—and have led—to profound insights and successful innovations. But biases in design (e.g., *ex post* simulations as described above) and theoretic strategies that cannot be traded introduce real crises of expectations. Some simulations must be taken with a grain of salt, others with a whole shaker-full.

8. G. L. Beebower and R. J. Surz, "Analysis of Equity Trading Execution Costs" (Center for Research in Security Prices Seminar, November 1980).

---

*Harlow and Brown footnotes concluded from page 62.*

- and *Clinical Psychology*, 1978; D. S. Robinson, J. M. Davis, A. Nies, C. L. Ravaris and D. Sylvester, "Relation of Sex and Aging to Monoamine Oxidase Activity of Human Brain, Plasma and Platelets," *Archives of General Psychiatry*, 1971; and Zuckerman, Buchsbaum and Murphy, "Sensation Seeking and Its Biological Correlates," *op. cit.*
16. Sharpe, "The Risk Factor," *op. cit.*
17. V. L. Smith, "Microeconomic Systems as an Experimental Science," *American Economic Review*, December 1982.
18. W. Vickery, "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, March 1961.
19. The theoretical development of the first price sealed bid auction can be found in J. C. Cox, B. Roberson and V. L. Smith, "Theory and Behavior of Single Object Auctions," in V. L. Smith, ed., *Research in Experimental Economics* (Greenwich, CT: JAI Press, 1982) and J. C. Cox, V. L. Smith and J. M. Walker, "Theory and Individual Behavior of First Price Auctions" (Working paper, University of Arizona, 1987).
20. Complete details of all experimental procedures are available from the authors upon request.
21. Each of the four subscales represents responses to 10 question items, scored on the basis of the number of relevant responses (0 to 10). The Gen scale is the sum of the four subscales, hence ranges in value from 0 to 40. See M. Zuckerman, *Sensation Seeking: Beyond the Optimal Level of Arousal* (Hillsdale, NJ: Lawrence Erlbaum Associates, 1979).
22. S. R. Hathaway and J. C. McKinley, *The Minnesota Multiphasic Personality Inventory*, rev. (New York: Psychological Corp., 1951).
23. See Zuckerman et al., "A Correlational Test," *op. cit.*
24. See J. C. Cox et al., "Theory and Individual Behavior of First Price Auctions," *op. cit.*
25. Financial support for this project was provided by The Research Foundation of the Institute of Chartered Financial Analysts, the University of Arizona Foundation (grant program of the Office of the Vice President for Research) and the Economic Science Laboratory at the University of Arizona. The opinions and analyses presented, however, are those of the authors, and not necessarily those of Salomon Brothers Inc, The Research Foundation or the University of Arizona Foundation.