Trading Costs and Portfolio Trading Strategies

by

Richard O. Michaud, V.P., Ph.D.
State Street Asset Management
Boston, MA 02110

Draft: July 18, 1990

An equity portfolio, or "program," trading strategy is a procedure designed to trade a group of stocks in relatively close proximity in time. Because factors are more important for describing portfolio returns than for individual securities, there can be important differences between a program and an individual security trade.

Program trading strategies include using futures and options to hedge or control market risk and to allow for diversification to reduce the range of market impact costs. What is perhaps less obvious is that a trading cost factor model, which is not generally useful for traditional trading, can be useful for designing program trades. This is because a small but statistically significant factor model R-square when based on individual stock trades may be large when applied to a portfolio. The purpose of this report is to identify and quantify the factors that affect market impact trading costs and to develop portfolio trading models that may serve as a basis for defining program trading strategies.

This report is based on intra day equity portfolio trading cost data supplied by Kidder Peabody and Co. The availability of such data is very recent, providing a window into phenomena that has heretofore been closed to financial economists and investment professionals. Using regression analysis, a statistical model of the portfolio trading process is developed that can be applied to the analysis of some trading strategies, enhancement of trading efficiency and the design of investment strategies and portfolios that include a realistic estimate of market impact costs. Within the context of the limitations of the data, the results indicate that some common perceptions of liquidity and trading strategies may need to be revised.

Definitions and Assumptions

Total trading cost is defined as commissions plus price impact. Price impact is defined as the difference between the average price of the trade for the stock and the last sale price at the time of execution of the trade, for a buy, and inversely for a sell. The definition is convenient for assessing the effectiveness of a wide variety of trading decisions, such as the number of trade slices, by using a common basis -- last sale price at initial execution -- for measuring costs. Preliminary studies using the average of the bid and offer as the initial reference price did not reveal any substantial differences.

The report is entirely focussed on the efficiency of the trading decision. The point of view is that of the trader who is given a portfolio to trade: the decision is considered (essentially) irrevocable and the task at hand is to execute efficiently. Consequently, it does not include an evaluation of the success of the investment decision nor does it allow explicit consideration of the option not to trade.

Data

Kidder Peabody Program Trade Cost Database

The portfolio trading cost database consists of one hundred portfolio or program trades (15 or more stocks traded simultaneously) consisting of roughly 8000 individual stock trades of which 7248 have sufficiently complete data to be usable in the regression analysis. The data was accumulated from March 1988 to April 1989. All trades are market trades executed through dot. Discretionary trading decisions include dividing the trade into a number of blocks or "slices" of (approximately) equal size and the execution time for each slice.

Factor Description

Some of the factors are based on data available on the electronic order book (EOB) at the time of the trade. The data for OTC stocks is more limited than for NYSE stocks. The factors can be grouped as follows:

A. Execution time stock based information:

- SIZE: Size of the offer for a buy or the bid for a sell; not available for OTC stocks.
- BIDASK: Bid/ask spread.
- AVGVOL: Log of the average (10 day) daily trading volume.
- PRICE: Log of price.
- SHARES: Log of number of shares.
- TSIZEVLM: Trade size as a proportion of average daily trading volume.

B. Execution time trade information:

- FUTPREM: Futures premium relative to index value.
- SLICES: Number of slices of the trade.
- BUYSELL: Buy (1) or sell (0) indicator.

C. Subsequent trade information:

- MARKET: Percent change in index value from first to last execution.
- TRDTIME: Time to trade; difference between last and first execution time.
- TIMESQ: Square of trdtime.

Factor Discussion

The A Group can be further decomposed into information based on:

Stock -- price, average volume
Trade size -- trade size to volume, shares
EOB -- bid/ask spread, size

It is of interest to determine whether stock, trade size or EOB information dominates trading cost. Traders are often primarily concerned with trade size or futures information. Some stock data, such as price, can be considered a proxy for EOB bid/ask spread data. Until recently, when some remote access trading systems have been developed, EOB data has not been accessible to non exchange member traders.

The B group includes tools for defining a trading strategy: Futures premium is a potential timing indicator; trade slicing is presumed to reduce price impact.

The C group represents factors known after the conclusion of the trade that may have had a substantial impact on realized costs.

On an a priori basis, SIZE and AVGVOL should be negatively related, the remainder of the A group factors positively related, to market impact costs. For a buy program, MARKET and FUTPREM should have a positive relationship to cost; for a sell program the relationship should be negative. Since these two variables have a non constant sign relationship with costs, sign changes were made depending on the type of the trade so that the data could be pooled and/or the results made comparable.

In a preliminary study,[1] one of the more interesting results found was that the time factors -- TRDTIME and SLICES -- were positively associated with trading costs. While the databases are not strictly comparable, validation of these early results may have important implications on the value of trade time exposure intensive trading strategies such as crossing networks, sunshine trading and single price auctions and, consequently, was a major focus of the study.

Since SIZE was not defined for OTC stocks, the inclusion of OTC stocks reduces the statistical significance of the factor.

Table 1 provides basic descriptive statistics of the factors and computed trading costs in the database, with indicated units in parentheses. Among other things, the data show that 55.7% of the trades were buys, average market impact costs were 7.4 cents/share and average bid/ask spreads were 22.1 cents/share.

Multiple Regression Analysis

Tables 2-6 describe the multiple regressions of market impact trading costs with respect to the above factors. The signs of MARKET and FUTPREM are changed for sell programs to preserve comparability in the pooled data and across data sets. The regression tables include the regression coefficients, beta coefficients and t-statistics for each variable. Tables 2-4 display the results of regressions for pooled buy/sell portfolio trades; Tables 5-6 compare buy vs. sell portfolio trades. In Tables 2 and 5, all factors are included; in Tables 3, 4 and 6 the MARKET factor is excluded. Tables 4 and 6 constrain the regressions to time periods for which MARKET was positive for buys and negative (before sign change) for sells.

Trade Cost Analysis

Table 2 shows that the market is the dominant trading cost factor indicating that buy (sell) program trading costs increase when the market advances (declines). The investment meaning of the estimated regression coefficient (17.4) can be indicated by the following computation: If the market increases in value by one standard deviation (.005) during the trade, market impact costs due to the market will, on average, be 8.7 cents/share (.005*17.4). Note that the average time for execution of a trade is 2.5 hours.

While the signs of the coefficients in Table 2 are consistent with priors only EOB information -- SIZE and BIDASK -- and the trade size factor -- SHARES -- are also statistically significant. To illustrate the investment meaning of the SIZE factor, 100 shares above the mean share offer (350 shares) leads, on average, to a 3.9 cents/share reduction in market impact costs (-.039*1.0).

The beta coefficients can be used to indicate the relative importance of the independent variables in a multiple regression. They show that the market factor is more than three times as important as the next most important factor (SIZE), and that EOB information is roughly twice as important as the most important trade size information (SHARES).

Since market change is not known a priori, excluding the market factor provides a more practical setting for defining an optimal trading strategy. Table 3 displays the regression results of excluding the market as an explanatory factor and shows that a fourth A group factor -- AVGVOL -- and the time factor -- SLICES -- are also significantly related to trading costs. The beta coefficients indicate that SIZE, BIDASK AND SHARES are of roughly the same order of importance in explaining trading costs. The coefficient for SLICES indicates that each slice of a trade increases trading costs by 0.8 cents per share. It is also of interest to note that a popular measure of trade liquidity -- TSIZEVLM -- is not statistically significantly related to market impact.

The relationship of trading costs and time can be obscured due to the dominant effect of the market factor in Table 2 and to the fact that the sign of the impact of time in Table 3 may depend on market direction, making the relationship unobservable. Consequently, market direction needs to be controlled if the effect of time is to be quantified and understood. In Table 4, the regressions are constrained so that the market factor is always non negative (advances for buys and declines for sells).

The data in Table 4 show that the number of slices and trade exposure time are both positively associated with trading costs. The TIMESQ factor indicates that time increases trading costs at a decreasing rate. The results are consistent with the hypothesis that the impact of TRDTIME proxies for market direction.

The results concerning SLICES, based on Tables 3 and 4, indicate that dividing a trade into blocks may increase trading costs independent of market direction. However, there are two alternative hypotheses: 1) The trades were oversliced, reducing the possible benefits of more moderate trade slicing; 2) Trade slicing, commonly defined with respect to TSIZEVLM, may be based on a relatively ineffective or irrelevant measure of liquidity.

While FUTPREM is positively associated with trading costs in Table 4, its lack of significance in Tables 2 and 3 indicates that time of trade futures premium may not be very useful for implementing trade strategy. Alternatively, it may simply indicate that the variable is not sufficiently well defined and should be based on theoretical, rather than index, value. The issue remains an open one for future research since theoretical value at trade time is not available in the database.

Buy vs. Sell Programs

Do buy programs trade differently from sell programs? Tables 5 and 6 compare buy and sell program trading costs with regressions analogous to those in Tables 2 and 4. The results show that MARKET and SIZE are the most important factors in both sets of regressions. Note also that the buy/sell indicator in Table 2 is not statistically significant. While buy program trading costs are primarily affected by EOB information, sell programs may be more dominated by trade size factors -- TSIZEVLM and SHARES. On balance, there is little reliable evidence that buy and sell programs trade differently.

A Program Trading Cost Model

The Model
A model of market impact trading costs (C_p) for portfolio p consistent with the regression analysis is:

$$C_p = F_t + aM + \Sigma d_j D_{pj} + \epsilon_p + as_p \qquad (1)$$

where
- F_t Commission fees; depends on trade t
- M Market factor
- J Number of descriptors or indicators of liquidity
- D_{pj} Descriptor j for portfolio p
- ϵ_p Descriptor error term for portfolio p
- s_p Portfolio specific return
- a, d_j Estimated regression coefficients[2]

Define:
- α_k Market drift expected return per unit of time; k=B for a buy, k=S for a sell, $\alpha_B + \alpha_S = 0$
- σ_d Market drift volatility per unit of time
- σ_s Portfolio specific risk per unit of time

Therefore:

$$E(M) = \alpha_k T; \quad V(M) = \sigma_d^2 T; \quad V(s_p) = \sigma_s^2 T;$$

and expected trading costs and volatility are given by:

$$E(C_p) = F_t + a\alpha_k T + \Sigma d_j D_{pj} \qquad (2)$$

$$V(C_p) = a^2\sigma_d^2 T + V(\epsilon_p) + a^2\sigma_s^2 T \qquad (3)$$

Market Trades
For a market trade, T is small. Therefore, formulas (2) and (3) become:

$$E(C_p) = F_t + \Sigma d_j D_{pj} \qquad (4)$$

$$V(C_p) = V(\epsilon_p) \qquad (5)$$

Crossed Trades
The model can be applied to analyze the costs of other trading strategies. Assume a portfolio trade that uses a crossing network where the entire trade is crossed during the trade period. In this case, the trade characteristics of the portfolio and the fit of the trading cost model are unimportant ($D_{pj} = 0$ and $\epsilon_p = 0$ in (1)), since securities precisely matching the ones being bought (sold) are obtained. On the other hand, T may not be small.

$$E(C_p) = F_t + a\alpha_k T \qquad (6)$$

$$V(C_p) = a^2\sigma_d^2 T + a^2\sigma_s^2 T \qquad (7)$$

Discussion

Equations (4) and (5) show that, for a market trade, expected trading cost depends on the liquidity characteristics of the portfolio while volatility depends on the approximative power of the factor model for the portfolio. Market trade commissions will also depend on portfolio liquidity characteristics. For complete crossed trades, equations (6) and (7) show that expected costs are minimal but market impact cost risk depends on trade time exposure and the level of portfolio diversification. Figure 1 illustrates the basic characteristics of the results.

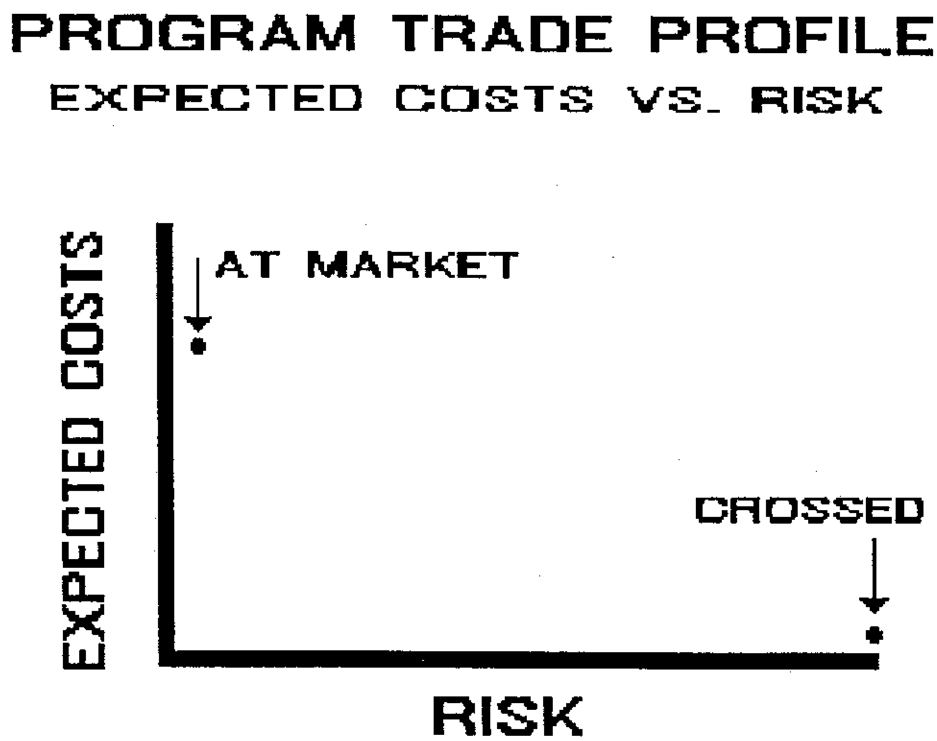

Figure 1

The fundamental notion that emerges from the simple model is that a strategy's trading costs must be understood in terms of two dimensions: expected costs and risk. A trader must weigh both dimensions to define an appropriate trading strategy. The optimality of a trading strategy depends on the utility of the trade-off between costs and risk.

The dimension of risk in Figure 1 requires further analysis. For crossed trades, trading cost risk is trade time dependent; i.e., trade time exposure affects the number of stocks that are likely to have been crossed. Figure 1 assumes that the portfolio is crossed completely during the trading period. In many cases this assumption is unrealistic. If only a portion of a portfolio is crossed, the residual portfolio is likely to have greater illiquidity than the original and may have to be traded in more traditional ways.

If we assume that the residual portfolio of a crossed trade is traded at-the-market at the end of the period, the formulas (6) and (7) will include the residual factors in (2) and (3) for the residual portfolio. As the time period lengthens, the amount of the portfolio that is crossed will likely increase, reducing expected costs but increasing risk. An illustration of the hypothesized relationship of expected costs to risk as a function of time is given in Figure 2.

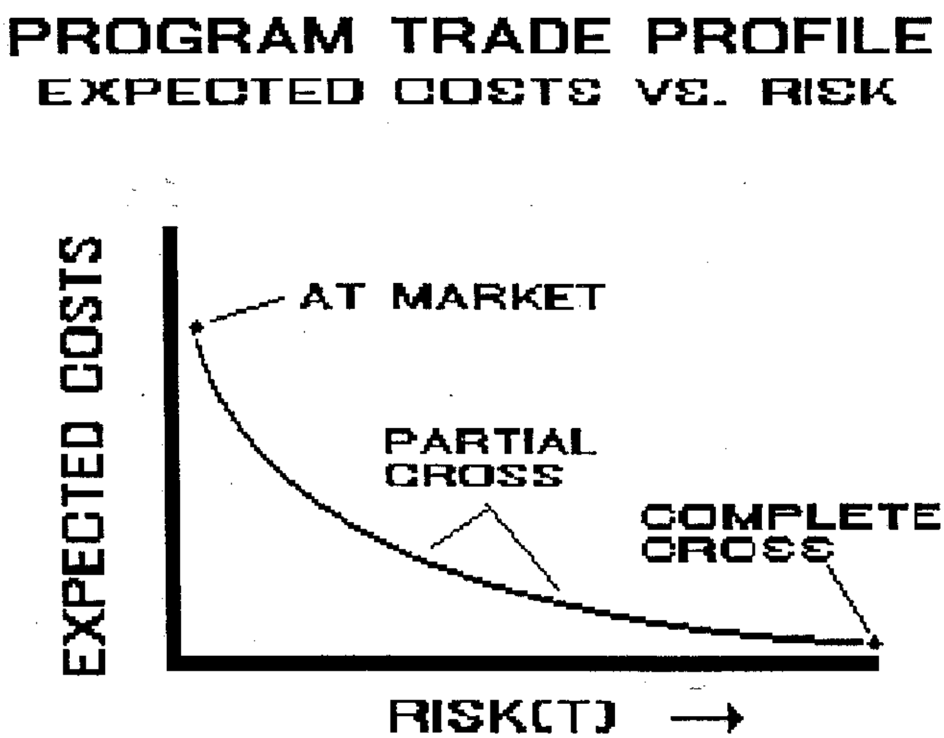

Figure 2

Applications

Market Factor Dominance

A reason for the dominance of the market factor is its volatility. Table 1 shows that the market has a 50 basis point standard deviation on a per trade basis, or approximately 30 basis points (bp) per hour (50/√2.5 bp/hour). Alternatively, using long term equity historical data, a 20% annual return standard deviation translates into a 50 bp/hour market volatility (2000/√(225*6.5) bp/hour). In contrast, a 15% annual market return corresponds to approximately a 1 basis point upward drift per hour.[3] The data demonstrate the substantial volatility that exists in the market on a micro basis and serves to rationalize the importance of the market factor in the regressions and the practice of using futures and other market hedging techniques as part of many program trading strategies.[4]

Electronic Order Book Information

Apart from the market, the most important information concerning trade costs in a well diversified portfolio is that in the electronic order book -- size and bid/ask spread. The results argue strongly that trading efficiency can be substantially improved with access to EOB data.

Traditional Definitions of Trade Liquidity

The results indicate that traditional notions of trade liquidity such as TSIZEVLM, the basis of many trading strategies, may have little real value. From a practical point of view, EOB factors appear to be the measures of liquidity of choice.

Time and Trading Costs

The results indicate that, in a well diversified portfolio, trade time exposure proxies for the market factor. Assuming, on average, an upward drift in the market, trade time exposure increases expected market impact costs for buy programs and decreases them for sell programs. The evidence indicates that trade slicing increases, not decreases, trading costs.

The results indicate that trade slicing should be kept to a minimum, or the basis for defining liquidity redefined. In the absence of a forecast of market direction or market risk hedging, trade time exposure should be kept to a minimum. This is because the possibility of large adverse costs due to portfolio volatility during trade time exposure, for both buyers and sellers, can be very substantial and dwarf expected cost considerations.

Crossing Networks

Crossing networks often imply significant trade time exposure. Assume a portfolio that is submitted to a crossing network at the open and crossed at the close.

Due to an expected market drift of 1 bp/hour, a buy program will have an expected 1 cent/share market impact cost for the trading day (17*.0001*6.5). Since the buyer and seller are in a zero sum game, the buyer's expected costs, before fees, is the seller's gain. However, crossing network commission fees are typically two to five times as large as the expected market drift return.

Assuming a MARKET standard deviation of 50 bp/hour, the trading day standard deviation is 125 bp. Consequently, a one standard deviation event is associated with 22 cents/share market impact (17*.0125). Implications for a well diversified portfolio trade include a one in three chance that market impact costs will be at least 10 cents/share and a one in five chance that costs will exceed 20 cents/share. Such volatility is a central fact of crossing network trading cost risk, which, of course, increases with the level of undiversification of the portfolio.

From a practical point of view, the major cost of crossing networks may be associated with trades that do not cross. A portfolio submitted to a crossing network is not guaranteed to fully cross during the course of the day. Resubmitting the residuals to the crossing network the following day exposes the trade to substantial additional trade cost risk. Implications of the results include holding trade time exposure to a minimum and pricing crossed stocks at, or as soon as possible after, the cross.

However, a combination of traditional and crossing network trades may often be attractive. Crossing networks could be used as a first step. This would take advantage of liquidity available at low cost while minimizing risk by keeping trade time exposure to a minimum. The residuals can then be traded using more traditional methods.

Crossing networks are representative of all trade time exposure intensive or patient trading strategies, including sunshine trading and single price auctions. While they may minimize expected market impact costs, they expose the trade to substantial market impact cost risk with a likelihood of only partial completion of the trade.

Passive vs. Active Investment Strategies

Trades involving passive investors often have little diversifiable risk. This is because the portfolio may closely resemble an index or much of the trade involves buying and selling stocks within the index in close proximity in time. Consequently, trade time exposure risk in a crossed trade (7) is less and may be worth assuming considering the likely reduction in expected trade costs. Also, a trade that is executed due to a restructuring of assets assumes less in terms of market change than one based on an asset allocation strategy which may forecast changes in the value of the market factor. Alternatively, active strategies may be far more sensitive to the need to control trade risk since they have more diversifiable risk and because they implicitly forecast a price change that is sensitive to time.

Market-on-Close Orders

The results imply that market-on-close orders should be avoided. This is because the trading strategy combines substantial trade time exposure risk as well as high at-the-market expected costs.

Summary

An optimal trading strategy depends on the risk tolerance of the investor. The program trader's task is to balance costs vs. risk suitable to the client. Relevant questions include the client's tolerance for large costs, fees and the liquidity and diversification characteristics of the portfolio.

Regression analysis using intra day portfolio trade data provides a new tool for the evaluation and design of trading strategies. Practical applications include: Forecast of market impact costs; Trade cost minimization; Benchmark for evaluation of trade efficiency.

The results redefine notions of trade liquidity, refocus attention on more relevant information, highlight the importance of trade analytics based on real time data and indicate reconsideration of widely used approaches such as crossing networks and other trade time intensive strategies that may not adequately consider the impact of time. By avoiding common trading errors, trading efficiency can be improved significantly, reducing trading costs and enhancing the capacity of active strategies.

Finally, it is important to note the limitations of the data. Further analysis is required for market factor hedging, worked orders, limit orders, and other trading strategies.

Footnotes

1 R. O. Michaud, Merrill Lynch, NY, March 1988.

2 The "a" coefficient is estimated for the market factor and assumed to apply to both components -- market and specific -- of portfolio risk.

3 The computation may more validly be based on the market premium (excess return over the risk free rate) which considers short term cost of capital. Assuming a 6% market premium, market drift is

roughly 1/2 bp/hour. This number can be compared to a 1 cent/share commission fee, typically charged by crossing networks, that implies a 2.5 bp/share cost (.01/$40).

[4] The effects of market factor hedging using futures or other instruments are outside the scope of this report.

Appendix

A more detailed analysis of the results based on the regression data supplied in the tables follows.

MARKET: The single most important explanatory factor; indicates the importance of dealing with market direction, including hedging, when defining a trading strategy.

SIZE: EOB factor; single most important explanatory factor apart from the market.

BIDASK: EOB factor; valuable indicator of trade liquidity.

SHARES: The most important trade size indicator.

PRICE: Some evidence of usefulness as an indicator.

TSIZEVLM: Some evidence of usefulness as an indicator.

AVGVOL: Little evidence of usefulness.

SLICE: Does not appear to reduce trade costs. Each slice estimated to increase costs by roughly 1 cent/share per slice.

TRDTIME: Appears to proxy for the market. Time is associated with increasing costs, but at a decreasing rate.

FUTPREM: Little evidence of usefulness as defined.

BUYSELL: No reliable evidence that buys trade differently from sells. Some evidence that buys more affected by market data while sells more affected by trade size data.

TABLE 1
DATABASE DESCRIPTIVE STATISTICS
Number of Observations -- N=7248

	AVERAGE	ST. DEV.
TRDCOST($/sh)	.074	.258
MARKET(retn)*	-.001	.005
SIZE(sh/100)	3.513	.653
BIDASK($/sh)	.221	.116
SHARES(logshs)	3.251	.534
PRICE(log$)	1.519	.258
TSIZEVLM(%)	3.243	10.440
AVGVOL(logvl)	3.203	.506
SLICES	3.345	1.877
TRDTIME(hrs)	2.518	2.439
TIMESQ	12.288	25.991
FUTPREM(retn)*	-.005	.079
BUYSELL	.557	.497

* Signs reversed for sell programs.

TABLE 2
MULTIPLE REGRESSION ANALYSIS
Pooled Buy/Sell Portfolio Trading Costs
N=7248, R-square=.14

	Reg. Coef	Beta Coef	T stat
MARKET	17.04	.327	29.0*
SIZE	-.039	-.098	-8.0*
BIDASK	.155	.069	5.6*
SHARES	.020	.042	3.1*
PRICE	-.014	-.014	-1.1
TSIZEVLM	.001	.028	2.1
AVGVOL	-.015	-.029	-1.9
SLICES	.004	.029	1.9
TRDTIME	-.004	-.042	-1.1
TIMESQ	.000	.008	.2
FUTPREM	.056	.017	1.6
BUYSELL	.004	.007	.6
CONSTANT	.181		5.5*

* Statistically significant at the .01 level.

TABLE 3
MULTIPLE REGRESSION ANALYSIS
Pooled Buy/Sell Portfolio Trading Costs
Market Factor Excluded, N=7248, R-square=.04

	Reg. Coef	Beta Coef	T stat
SIZE	-.043	-.110	-8.5*
BIDASK	.181	.081	6.2*
SHARES	.054	.111	8.0*
PRICE	-.006	-.006	-.5
TSIZEVLM	.001	.019	1.4
AVGVOL	-.030	-.059	-3.7*
SLICES	.008	.057	3.5*
TRDTIME	.001	.010	.2
TIMESQ	-.000	-.040	-.9
FUTPREM	.060	.018	1.6
BUYSELL	.029	.056	4.4*
CONSTANT	.076		2.2

* Statistically significant at the .01 level.

TABLE 4
MULTIPLE REGRESSION ANALYSIS
Pooled Buy/Sell Portfolio Trading Costs
Market Factor Excluded and Constrained, N=3354, R-square=.06

	Reg. Coef	Beta Coef	T stat
SIZE	-.054	-.119	-6.3*
BIDASK	.188	.073	3.8*
SHARES	.026	.043	2.2
PRICE	.102	.087	4.8*
TSIZEVLM	.001	.026	1.3
AVGVOL	-.002	-.003	-.1
SLICES	.016	.091	3.5*
TRDTIME	.062	.551	7.6*
TIMESQ	-.006	-.602	-7.9*
FUTPREM	.228	.086	5.0*
BUYSELL	.051	.083	4.1*
CONSTANT	-.104		-1.7

* Statistically significant at the .01 level.

TABLE 5
MULTIPLE REGRESSION ANALYSIS
Portfolio Trading Costs
Buy (N=4039, R-sq=.14) vs. Sell (N=3209, R-sq=.16)

	Reg Coef		Beta Coef		T stat	
	Buy	Sell	Buy	Sell	Buy	Sell
MARKET	16.95	19.13	.321	.370	20.6*	16.8*
SIZE	-.040	-.037	-.093	-.111	-5.5*	-6.2*
BIDASK	.194	.047	.084	.023	5.0*	1.3
SHARES	.015	.024	.026	.065	1.5	2.7*
PRICE	.004	-.028	.004	-.037	.2	-2.1
TSIZEVLM	.000	.003	.015	.089	.8	4.3*
AVGVOL	-.017	-.013	-.030	-.029	-1.5	-1.3
SLICES	.002	.009	.015	.076	.6	3.3*
TRDTIME	-.003	-.009	-.023	-.104	-.4	-1.6
TIMESQ	.000	.001	.000	.030	.0	.5
FUTPREM	.057	-3.56	.021	-.069	1.4	-2.4
CONSTANT	.181	.165			3.7*	4.0*

* Statistically significant at the .01 level.

TABLE 6
MULTIPLE REGRESSION ANALYSIS
Portfolio Trading Costs
Buy (N=1777, R-sq=.05) vs. Sell (N=1577, R-sq=.14)
Market Factor Excluded and Constrained

	Reg Coef		Beta Coef		T stat	
	Buy	Sell	Buy	Sell	Buy	Sell
SIZE	-.048	-.054	-.087	-.188	-3.3*	-7.1*
BIDASK	.221	.090	.076	.050	2.8*	1.8
SHARES	.002	-.011	.003	-.031	.1	-.8
PRICE	.137	.079	.092	.111	3.6*	4.3*
TSIZEVLM	.000	.007	.016	.144	.6	4.1*
AVGVOL	-.011	.046	-.016	.119	-.5	3.1+
SLICES	.008	.027	.047	.146	1.1	5.8*
TRDTIME	.082	.103	.704	.723	6.1*	7.5*
TIMESQ	-.007	-.011	-.737	-.580	-6.3*	-6.5*
FUTPREM	.242	25.38	.101	.255	4.1*	6.9*
CONSTANT	-.038	-.106			-.4	-1.9

* Statistically significant at the .01 level.
+ Wrong sign for statistical significance.