The Pennsylvania State University
The Graduate School
The Mary Jane and Frank P. Smeal College of Business Administration

THREE ESSAYS IN FINANCE

A Thesis in
Business Administration
by
Edward Glidewell

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2004

The thesis of Edward Glidewell was reviewed and approved* by the following:

Charles Cao
Associate Professor of Finance
Thesis Co-Advisor, Co-Chair of Committee

Herman Bierens
Professor of Economics
Thesis Co-Advisor, Co-Chair of Committee

Oliver Hansch
Assistant Professor of Finance

Timothy Simin
Assistant Professor of Finance

William Kracaw
Professor of Finance
Department Chair

*Signatures are on file in the Graduate School.

Abstract

I discuss three topics in this thesis. The first two concern topics centering on addition to the S&P 500 Index. In chapters two and three, I examine the price and liquidity effects (respectively) of additions to the S&P500 Index from January 1993 to December 2000. The results from chapter two indicate that previous inferences about long run buy and hold returns are fallaciously obtained from a test statistic with low power. Consequently, while the short run demand curve for newly added S&P stocks appears to slope down, the data suggest that no S&P effect exists in the form of a permanently downward sloping demand curve.

In chapter three, I present the lasting liquidity improvements that arose from index addition. Permanent reductions in several static and dynamic measures of liquidity are statistically significant but are confined to NYSE firms. The spread decomposition and static trade analyses agree with previous research and imply a perennial increase in liquidity through a 15% reduction in the quoted and effective spreads. The dynamic liquidity measure indicates a permanent increase in market resilience. Post addition, smaller quote revisions obtain. Cumulative mid-quote revisions to buy side shocks fell by 25% and revisions to sell side shocks fell by 27%, implying that private information in the order flow fell after addition.

The final chapter is a spread decomposition model that allows for nonlinearities in the data. Most spread decomposition models utilize volume and/or trade signs in a linear specification to explain observed price changes. This paper illustrates the importance of another variable, unexpected trade intensity, in explaining and forecasting effective spreads. I use Engle's (1996) WACD(p,q) model and other proxies for unexpected trade intensity in a generalized version of Glosten and Harris' (1988) spread decomposition model. I adapt Hansen's (1999) technique for testing for the presence of nonlinearities in the relationship between price changes and trade volume and between price changes and unexpected trade intensity. The data suggest that volume and intensity are priced independently and on average place a 22% premium on large volume and a 8% premium on immediacy conditional on a given volume. Selecting a suitable proxy for trade intensity and incorporating the threshold values into the model not only allows for enhanced estimation of the effective spread, but also for more precise forecasts.

Table of Contents

List of Figures

Chapter 1 Introduction

Chapter 2 BHAR Analysis

Chapter 3 S&P Dynamic Liquidity Analysis

List of Tables

Acknowledgments

I most gratefully thank Herman Bierens, Charles Cao, Ian Domowitz, Laura Field, Oliver Hansch, J. Harold Mulherin, Dennis Sheehan and Tim Simin and seminar participants at Auburn University and Southwest Texas State University for helpful remarks and suggestions. All remaining errors are mine.

Dedication

To the Horwitz brothers and Mr. Fienberg, GL and Larry Gelbart.

Chapter 1

Introduction

This thesis is divided in to three parts. The first discusses abnormal returns around the time of adoption into the Standard and Poors 500 Index (S&P 500). The second examines liquidity changes measured three months before and after the addition and then discusses the cause of the perceived changes in liquidity. The last chapter breaks from the index adoption theme and examines the relationship between volume and trade duration and expected trading costs of orders on the New York Stock exchange.

The question central to the first chapter is: what would one expect the buy and returns to be, ceteris paribus, three months after an arbitrary date if there were no addition? This is a slight refinement over the conventional question asked in the literature which is: do permanent BHARs obtain after inclusion into the S&P 500? One could have positive returns 60 trading days after inclusion for may reasons, most notably if one had a cohort of positive momentum stocks. If there were no addition, then the stocks would still register economically and statistically positive returns due the momentum factor driving returns. If one fails to take this into consideration, then one would attribute the permanent price increase to the inclusion, which would be interpreted as saying the stock's long run demand curve slopes downward. On the other hand, if the long run BHARs were positive (as in the case of a positive momentum stock) then only the short run abnormal returns would from inclusion. Hence, the short run demand curve would slope down, but in the long run the stocks would still be perfect substitutes for each other.

In order to address adequately the first question, one must have a methodological tool to ascertain what null distribution obtains from the given sample of firms (1993-2000, inclusive). I set up several bootstrap procedures to establish a null. The crucial element in selecting the hinges on conditioning expectations contingent on the properties of the sample and not on a cohort created to match on size or other common quantifiable characteristic. The reason for this goes back to the positive momentum quality of the sample. A firm's stable financial condition is a key criterion that must be satisfied before the S&P committee will elect to add the firm to their 500 index. Thus, a firm must not only have a scope of activities that is representative of its industry, but also must have a sound historical performance. Bootstrapping the null BHAR to take the expected BHAR to account (account for the large, sustained return stream) should control for that that we should otherwise expect from this (or any) cohort.

The liquidity examination in Chapter 3 fits into the S&P literature by bolstering the notion that liquidity improves after addition even though float is reduced. Past findings indicate that spreads lower after addition because the order processing component of the spread wanes. The adverse selection lower only slightly (this is a very slight part of the spread to begin with). However, other veins of research in finance indicate that the models used to decompose the spread often give spurious results. Chapter 3 contributes to this literature by assessing the liquidity of the sample through measuring liquidity in a dynamical fashion. If liquidity increases, then one effect of this increase would be that surprise order flow shocks would have a smaller impact on quotes than they would have before the liquidity enhancement. The cumulative generalized impulse response function measures this effect.

The GIRF is used in place of the traditional vector moving average expression for the impulse response function. Nonlinearities in the data prohibit the linear model from producing the needed shock responses. Identifying the level and significance of the thresholds (the nonlinearities are approximated by piece wise linear expressions) is performed by borrowing techniques from the threshold and changepoint literature in economics. Inference is nonstandard due to the presence of a nuisance parameter that is identified only under the null. Hence simulations were performed to approximate the test statistic's true distribution. The results show that a significant reduction in the cumulative price impact obtains post addition. This interpreted as evidence that there is less private information in the order flow post addition, a sign of improved liquidity.

Chapter 4 borrows much of the methodology from Chapter 3 and applies it in the context of a spread decomposition model. I generalize Glosten and Harris' (1988) model to allow both time and signed order volume to proxy for adverse selection in the order flow and also allow for data driven thresholds. Theory suggests that such thresholds will exist, but does not place any bound on what their level should be or what relation they should have other variables in the trading environment such as quoted depth. Using the same general methodology as in Chapter 3, I estimate the level and significance of volume and time thresholds. I significant source of concern centers on whether the two threshold variables are correlated. If the two variables are independent, then the estimation of each is relatively straightforward. If not, then one would need to perform further simulations to measure accurately one threshold variable in the presence of the other. Analysis through a correlation integral (an asymptotic test for dependence of an unspecified type) indicates that no dependence is present between the order flow and trade intensity variables.

Three different proxies were used to measure the normalized time between trades. Each was a function of the expected time between trades conditional on the time of day. All of the thresholds were very close and the forecasts that were created were all close as well. It turns out that the simplest variable for time between trades (trade duration) was the simple moving average specification with a lag length of four. All of the proxies had reasonable lag lengths associated with them, implying that a significant signal must be sent to the market before prices will change.

Chapter 2

S&P BHAR Analysis

Introduction

S&P 500 index additions present a unique opportunity for testing the semi strong form of the Efficient Market Hypothesis (EMH). The addition announcement is a rare economic phenomenon in that is an informationless event yet it causes large, positive excess returns that have grown over time.[1] Whether this stock boost, the so called "S&P Effect", persists is the crux of the efficient markets discussion. Hence, the ability to draw accurate inferences about abnormal returns surrounding the event is imperative.

One corollary of the EMH is that for a given risk tranche, all stocks are substitutes in so much as they promise cash flows discounted at the

[1] Informationless in the sense that no new data were released about the firm's financial state.

same rate. If the firm's stock prices do permanently change as a result of inclusion, then there must be a change in risk associated with index addition in order for the EMH to hold.[2] If the inflated prices do not persist, then the EMH predicts that the price reversal should be immediate modulo liquidity effects. Although some dispute has arisen over whether a meaningful portion of the announcement day abnormal returns dissipates, the consensus in the literature is that significant long run daily returns obtain after addition.

The three main hypotheses predict that price behavior around the time of an index addition are the Price Pressure, Imperfect Substitutes, and Liquidity hypotheses. Each ascribes the price increase at announcement to a different source and each has a (not necessarily mutually exclusive) prediction regarding how long the abnormal returns will last. Harris and Gurel[3] first proposed the Price Pressure hypothesis which predicts that the price impact of the addition is transient. Consequently, the returns that traders require in order to supply liquidity to the market explain the abnormal returns that arise around the time of an addition. In the absence of any trading frictions and holding all else equal, a gradual price reversion would constitute a violation of the EMH.

A second explanation for the positive price reaction at the announcement is the Liquidity hypothesis (LQ). The Liquidity hypothesis is derived from a

[2] Despite the lack of agreement on the cause of the price change, past studies have concluded that the positive, economically meaningful long run daily abnormal returns are statistically significant.

more general paper by Amihud and Mendelson[1] wherein they examine risk reduction in a general setting and conclude that a permanent price increase would result from a permanent decrease in the spread.[3] This circumstance is consistent with the EMH. However, recent work by Dmitri Vayanos[9] notes that the expected change in the required rate of return arising from changes in the relative spread translates to a price change of second order importance. It is therefore doubtful that the (LQ) could cause the large price reactions observed and is not considered a likely source of the index effect.

The third hypothesis, the Imperfect Substitutes hypothesis (IS), is built on the construct that two otherwise identical stocks become delineated as one is selected for the S&P 500 Index. From that point in time on, the two stocks are no longer identical, i.e. they are not perfect substitutes. Ceteris paribus, the market considers this permanent, therefore, the stock will realize a long term positive cumulative abnormal return. If no change in risk arises from the addition, then a permanent price change would indicate a violation of the EMH.

Two of the more recent papers on value gains around the time of S&P adoption are by Beneish and Whaley[2] and Lynch and Mendenhall[7]. Each paper differs slightly in sample composition but substantively in their respective inferences on price behavior after announcement. Beneish and Whaley

[3]The price change is a capitalization of the decrease in expected trading costs, which in turn lowers the required rate of return.

(BW) find that prices permanently increase with a 5.89% abnormal return 60 days after the addition date in their 1986-1994 sample. Lynch and Mendenhall (LM) look at additions and deletions in the 1990-1995 period and report a 2% permanent increase and conclude that the majority of the price effects is transitory. Each set of authors adheres to a convention in this literature and uses a buy and hold or net of market type methodology to measure abnormal returns over intervals that range from 10 to 60+ days.

This paper's contributions are twofold. First, I show how a generalized test statistic applied to long run daily buy and hold abnormal returns improves inferences. I present a general specification to the traditional abnormal returns test statistic in this literature, the t-test on the buy and hold abnormal return (BHAR). The BHAR is usually evaluated economically and tested statistically against a null of zero. In the case of a small one or two day window, this assumption is plausible as expected firm and market returns are each close to zero, the expected difference of the two would also lie close to zero.

An unvarying method of benchmarking such as this is clearly neither suitable for all subsets of stocks nor over all time intervals. The convention in the S&P literature is to examine returns from the announcement day to 60 trading days post addition, which in this sample produces event windows from 61 to 90. Given the nonstochastic composition of the sample and the multiplicative feature of the BHAR, a positive expected BHAR obtains from bootstrapped own sample history. Thus, if one retains this method out of convention (or because of its ability to represent investor experience) an unfortunate semantic consequence of this assertion ensues as a non zero buy and hold abnormal return may be *expected* and hence is not abnormal in any statistical or economic sense. Consequently, the sample's persistent and exceptional past performance magnifies greatly the problem of benchmarking. Matching the sample to the market or to a market value cohort neither explains past performance nor performance after addition. Only by comparing the sample's post addition performance to its past performance does an explanation arise.

Ceteris paribus, a firm's superior past performance will influence Standard and Poor's decision to add it to the 500 index over its industry rivals as sound financial performance is reported to be a key criterion for inclusion. Thus all else equal, Standard and Poor's would prefer to draw from the right tail of the performance distribution, thus introducing a strong degree of sample bias. This will create a positive bias in firm performance that should be incorporated into the firm's expected returns. Consequently, in order to assess adequately *only* the effect of addition on a firm's returns, one should control for all other elements related to a firm's risk. As the characteristics of this sample are highly idiosyncratic to S&P 500 addition, a customized (own sample historical) benchmark was added to the normal array of market value and market value/book to market control portfolios.

To address the benchmarking issue for this sample, I use three simple bootstrap procedures. I use own sample past returns and contemporaneous returns from two cohorts to calculate the null for five different intervals: the announcement day, the announcement day to the addition or ex day, the announcement day to the ex day + 20 days, announcement to ex + 40 days and the announcement to ex + 60 days. The size and size and book to market benchmarks have an expected null of zero while the own sample benchmark is positive and economically meaningful after only 20 trading days. In this sense, the previous literature may suffer from what Fama[11] terms the 'bad model problem'. That is, the abnormal returns observed may be a result of a misspecification of the expected returns, rather than a result of aberrant returns originating from a surprise event. The effect of incorporating this general method (general in that a firm's return history dictates the expected return for each firm and for each interval) is that the price effects of inclusion that earlier studies considered permanent now become transient. The appendix contains evidence of how a parametric estimation routine, like market model CARs, appear problematic to these data.

The outline of the chapter is as follows. In Section 2, I describe the data and discuss the process by which Standard and Poor's uses to make its decision to add/drop firms from the 500 Index. Section 3 features the various methodologies used. In Section 4, I discuss the results and conclude in Section 5.

Data

The announcement and execution date data were supplied by CRSP for the 1993-1996 interval and by Standard and Poor's for the years 1997-2000. Standard and Poor's has a decision making committee which meets periodically to make suggestions about the composition of the firms in the 500 Index. Choice of selection or deletion is not a function of expected future performance but rather is a function of the firm's financial stability, liquidity, industry, and the representativeness of the industry in the U.S. economy.

The raw sample consists of 249 firms. Seventy seven firms were deleted because of confounding events around the announcement or addition dates. Firm and index returns were drawn from the CRSP daily stock and index files. The index returns are value weighted and both returns files include distributions.[4] All of the intraday data are from the Trade and Quote (TAQ) database. All data were filtered for recording errors and were constrained to normal trading hours. Following previous studies, I selected the major

[4]The results were similar for estimations carried out with the equally weighted index.

exchange for each stock by summing the volume of each of the markets on which it was traded and chose the primary exchange as the one with the largest volume. This designation coincides with the CRSP index indicator. Lastly, all of the 'Day 0' event days mentioned in the paper are actually the trading day after the event as the announcement always takes place after trading hours.

Since a goal of this paper is to identify and examine the effects of a shift in liquidity for the firms added to the index, I identified firms that exhibited other phenomena that could be associated with changing the adverse selection component in the firm's spread in the three months that followed the addition and analyzed them separately.[5] Interestingly, the results reported are quantitatively very similar for both sets of firms and qualitatively the same.

[5]Such events include merger and acquisition activity over 10% of firm value, a change in the number of shares outstanding, or changes in corporate leadership. Changes in corporate leadership would include events such as the CEO and/or two or more positions such as that of the COO, CFO, Chairman, or key vice president personnel leaving their office.

Methodology

Measuring Abnormal Returns

Traditionally in the S&P literature, researchers draw inferences on the presence of a permanent price impact through analysis of the significance of the buy and hold abnormal returns (BHAR) (2.1) compounded from the announcement date to some post announcement date, T.

$$BHAR_i = \prod_{t=1}^{T}(1 + R_{i,t}) - \prod_{t=1}^{T}(1 + R_{m,t}) \tag{2.1}$$

where $R_{i,t}$ is the firm's raw return at time t and $R_{m,t}$ is the return on the market proxy at time t. An advantage to using the BHAR centers on its use as a proxy for investor experience. It is not without its methodological problems, as a strong return anywhere along the Day 0 to T interval can create the illusion of a permanent price impact arising from the event in question via the multiplicative nature of the compounding.

The traditional BHAR significance test works well with event studies with small (one day to one week) intervals because expected firm and market returns are small and there are few compoundings over such intervals. In a long term study that uses daily data, abnormal returns may arise from a single large return anywhere in the 60 to 90 day intervals examined in this literature. In past studies, the perdurable abnormal returns range from 2 to almost 6%. I seek to show that such BHAR levels are expected over such a long event window and it is important to modify the test statistic to account for this finding.

A natural alternative to this methodology might include one that accounts for risk. The fact that the sample consists of a set of nonrandom (over 95% of the firms in this sample were picked from the S&P MidCap 400 index), medium to large market value firms whose return windows last up to 90 days. This suggests some measure of risk adjustment is necessary. In a parametric abnormal return model such as the market model, this is accomplished by adding an intercept and a coefficient on the market proxy to calculate the abnormal returns. Ordinarily, this would suffice but the data show evidence of parameter shift around the event, so residual analysis of a parametric model becomes problematic until a year after announcement when post addition parametric results may be used.

The test for abnormal returns should include the multiplicative compounding that creates returns that parallel 'investor experience' while deflating the returns by the amount expected for a given firm over a particular interval. The qualitative results from this method should coincide with a parametric method using *post* event data (given there exists a structural break in the data), but without the delay for data acquisition and parameter estimation. Hence this method provides an accurate and timely inference on the magnitude of abnormal returns.

In order to draw more precise inferences about possible violations of the EMH, I borrow from Efron and Tibshrani[12] and bootstrap the null return distribution. I propose three general metrics of assessing abnormal returns through the calculation of the null in the test statistic (for each firm and for each interval) via resampling returns from three control groups. The first group controls for market value, the second for market value and the book to market ratio and the last for own sample idiosyncratic risk. The purpose of each is to create an average BHAR that proxies for the expectation of what one would otherwise observe in the absence of the addition.

To construct the size cohort, one creates a contemporaneous portfolio of like-sized firms. To do this, I calculate the market value for a given sample firm one week before addition and pull all NYSE firms in the CRSP database that have market values between 80% and 120% of the sample firm's. Then, for all firms meeting this criterion, I pull all observations from the sample firm's announcement date to its addition date plus 60 trading days as well as the returns on the market proxy for these days. So two N×T matrices result. The first is of firm returns and the other of corresponding market returns where N is the number of firms in the cohort and T is the event window. If a firm drops out of CRSP during this time, I exclude it from the cohort.

For each event day, I select a column at random from the firm and the market return matrices and resample from these cross sectional vectors 1000 times with replacement. So for a two day interval, I draw two 1x1000 row vectors and stack them on each other and then do this again for the market returns that coincide with the firm returns. Next, from the 2x1000 matrix of firm returns (two days by 1000 resampled observations), compound the 1000 pairs of firm returns and the 1000 pairs of the market returns. So now one has a 1x1000 vector of compounded returns for the firm and another one for the market. Take the average of each to get an estimation of the expected firm return over two days and then subtract the market average to get an average BHAR for the cohort. Do this for every interval of interest Day 0, Day 0 to ex date + 20, etc. The size and book to market cohort is formed in a similar manner, although only the subset of firms that also have a book to market ratio from 70% to 130% of the sample firms' during the quarter it was added to the index will be included in this cohort. Similar data pulls and calculations ensue from there.

The BHAR for the contemporaneous size cohort for firm i would then be

$$BHAR_{i,Size} = \frac{1}{N_j}\sum_{j=1}^{N_j}\prod_{t=1}^{T}(1+R_{j,t}) - \prod_{t=1}^{T}(1+R_{m,t}) \tag{2.2}$$

where $R_{j,t}$ is the return for firm j in the size cohort on day t and $R_{m,t}$ is the corresponding market return. So, the BHAR for firm i, over and above what one would expect based on a size matched cohort would be

$$BHAR_i = \prod_{t=1}^{T}(1 + R_{i,t}) - \prod_{t=1}^{T}(1 + R_{m,t}) - E_{Size}[BHAR_i] \tag{2.3}$$

or

$$\begin{aligned} BHAR_i &= \prod_{t=1}^{T}(1 + R_{i,t}) - \prod_{t=1}^{T}(1 + R_{m,t}) - \\ &\quad \frac{1}{N_j}\sum_{j=1}^{N_j}\prod_{t=1}^{T}(1 + R_{j,t}) + \prod_{t=1}^{T}(1 + R_{m,t}) \end{aligned} \tag{2.4}$$

which of course simplifies to

$$BHAR_i = \prod_{t=1}^{T}(1 + R_{i,t}) - \frac{1}{N_j}\sum_{j=1}^{N_j}\prod_{t=1}^{T}(1 + R_{j,t}) \tag{2.5}$$

where $R_{i,t}$ is the return for sample firm i on day t. Similarly, for the size and book to market controlled group,

$$BHAR_i = \prod_{t=1}^{T}(1 + R_{i,t}) - \prod_{t=1}^{T}(1 + R_{m,t}) - E_{Size,Size/BM}[BHAR_i] \tag{2.6}$$

where $E_{Size,Size/BM}[BHAR_i]$ is the expected BHAR for sample's size and book to market control group. (2.6) expanded

$$\begin{aligned} BHAR_i &= \prod_{t=1}^{T}(1+R_{i,t}) - \prod_{t=1}^{T}(1+R_{m,t}) - \\ &\quad \frac{1}{N_k}\sum_{k=1}^{N_k}\prod_{t=1}^{T}(1+R_{k,t}) - \prod_{t=1}^{T}(1+R_{m,t}) \end{aligned} \tag{2.7}$$

where $R_{k,t}$ is the return for firm k in the size and book to market cohort on day t. The analogous measure of abnormal returns to that of (2.7) is

$$BHAR_i = \prod_{t=1}^{T}(1+R_{i,t}) - \frac{1}{N_k}\sum_{k=1}^{N_k}\prod_{t=1}^{T}(1+R_{k,t}) \tag{2.8}$$

The own sample bootstrap is constructed by taking sample firm and market returns from the 241 day interval [-250,-10] before addition and applying the same method as above. So the bootstrap will be performed on two 110×241 blocks (one of firm returns the other the contemporaneous market returns). This will form a distribution of historical averages of the null in the same fashion as the other benchmarks.

The BHAR for the own sample is

$$\begin{aligned} BHAR_i &= \prod_{t=1}^{T}(1+R_i) - \prod_{t=1}^{T}(1+R_m) - \\ &\quad \frac{1}{N_l}\sum_{l=1}^{N_l}\left(\prod_{\tau=1}^{T}(1+R_{l,\tau}) - \prod_{\tau=1}^{T}(1+R_{m,l,\tau})\right) \end{aligned} \tag{2.9}$$

where $R_{l,\tau}$ is the own sample historical return for firm l on day τ. The change of time index from t to τ is done to emphasize that this is not a contemporaneous bootstrap like the previous two. $R_{m,l,\tau}$ is the corresponding market return for firm l on day τ. So to get an average one day expected BHAR, pick a column vector from the firm and market matrices. Resample from that vector 1000 times to get two 1000×1 vectors. Repeat the last step 1000 times and then take an average so that one is left with two 1000×1 firm and market average return vectors. If there were more than one day in the interval one would compound the number of days in the interval at this point, then take the difference of the average of the firm and market compoundings (in this case, with only one day, we would just take the difference of the firm and market averages) to get the expected BHAR.

The consequence of allowing the data to dictate expectations is that constructs a test of means between expected BHAR calculated from a control group and the event period BHAR to form a test statistic and p-value. Thus, given each simulated null distribution, I take the means and standard errors from the sample and the benchmark to create t-tests with unequal variances. It turns out that the expected BHARs formed from the size and size and book to market cohorts are statistically zero over every interval. The expected BHARs formed from the sample are large and significant after 20 trading days. Table I contains the BHARs and their t-statistics with and without the bootstrapped null in the test statistic. After accommodating a generalized null in the test statistic, the long term abnormal returns using the two size cohorts are not significant at the 5% level. However this only arises because the standard errors are inflated, not because the control group BHAR is economically meaningful (it is never statistically greater than zero). Using information from sample's history shows insignificant abnormal returns (over and above what is expected) after 20 trading days. Thus, after adjusting for the sample's idiosyncratic risk, no significant abnormal returns persist from addition to the S&P 500.

Results

The announcement and execution date abnormal returns align themselves with past studies. On average, inclusion into the S&P 500 has gained importance as measured by the first day price response.[6] As noted in Panel A of Table I and in Figure 1, the sizeable Day 0 response propagates through time using the traditional BHAR methodology[7]. Once one uses a general methodology that accounts for what is otherwise expected, the Day 0 effect dies out after 20 trading days as Panel B indicates. The size of the abnormal returns were similar to those of past studies but its significance disappears

[6] I also confirmed Beneish and Whaley's conclusion that investors cannot capture the inclusion gains as most of the price response occurs overnight.

[7] Similar to Beneish and Whaley and Lynch and Mendenhall.

when the test statistic is modified to account for the large expected component of the buy and hold abnormal returns. The false positive that was registered in many previous studies appears to result from inappropriately constraining the null hypothesis to zero, a common feature of event studies with shorter time frames. This constitutes an example of the bad model problem described by Fama. The BHARs from a representative event window ((-160.-100) days before announcement) shown in Figure 2 illustrates the nontrivial magnitude and statistical significance of the sample's returns from a 'typical' 3 month window. Given that a large, permanent BHAR obtains without S&P inclusion, the permanence of the returns is suspect, and one must conclude that the effect from inclusion is a temporary result. Transient price changes would indicate that the Price Pressure hypothesis best describes the data and the EMH is violated.

A series of power curves is provided (Figures 6-15) to illustrate the power of the BHAR test statistics over a range of null returns (-10% - 10%). There are two sets, one covering the sample the other of sample's the size cohort (the results were similar for the size and book to market cohort and were omitted). Each set is divided in five intervals: the announcement day, the announcement day to addition or ex day, the announcement day to ex plus 20 trading days, announcement to ex + 40, and the announcement to ex +60 for a total of 10 graphs.

The graphs for the sample were created in the following manner: for each firm in the sample, firm and corresponding market returns were drawn from the CRSP daily stock file and from the daily index file, respectively. Each firm had its returns and those of the market pulled from the interval -250 to -10 days before announcement. So two matrices of 110×241 (one for firms, the other for the market) were formed. For each interval described above, that number of days (columns) were pulled at random from the firm and market matrices. Let's start with an example for the one day interval. The firm and the market vectors were resampled 1000 times with replacement to form two 1000x1 vectors. The mean firm and market returns and their standard errors were calculated. The last step was repeated 1000 times and 1000 means and standard errors were collected. An interval range for the null hypothesis and an acceptance region for the test statistic were chosen to be [-10% to 10%] and 5%, respectively. Taking a particular null from that range, say -10%, we then subtract -0.10 from the first mean and then divide that amount by the first standard error to form the first test statistic. Repeat this over all of the 1000 means and standard errors. Count the number of times out of 1000 that the test statistic was less then -1.96 or greater than 1.96 and that's the probability of Type II error, given the null=-0.10. 1-Prob(Type II | null=-0.10) is the power of the test statistic. This procedure was conducted over the interval of nulls with an increment of 0.001 to form a range of 201 nulls.

The results for the sample's power test are contained in Figures 6-10. The power tests indicate that after about 20 trading days, tests that have the null expected BHAR of zero have very low power. Figures 11-15 show the results of the power test on the size cohort. These tests suggest that a test statistic with a null of zero has high power compared to a range of alternatives.

An interesting question centers on why the abnormal returns last as long as they do. If the market were to respond to the price changes over the course of two or three days, then one might dismiss this as a clear example of liquidity providers desiring a premium for supplying liquidity. The abnormal returns last for about a month. An interview with a sales manager (in a Wall Street Journal article about the S&P Effect(see Appendix C)) offers one possible explanation. The manager believes that a high degree of collusion exists among managers and brokers and this explains why the "S&P Effect" has grown through time. His theory is that profit sharing arrangements are made that encourage price inflation through the addition date. In this sense, the build up after announcement and the drop off after addition could be the result of something close to the 'laddering' practice between investment banks and their top customers. Here, fund managers buy up shares (before index fund managers are allowed) early on and bid up the price. On the addition date, the fund managers split the cost with the index fund managers and then in the ensuing weeks push their remaining shares off on retail customers, capitalizing on recent run up in price and media coverage. Whether exploring these arrangements can lead to full explanation of why prices are so slow to return to an equilibrium level is debatable, but the contention that this event is far more complex than a liquidity based run-up and sell-off is not.

Figures 1 through 3 feature the sample BHARs and its standard error bands (s.e. bands) for the 60 days post announcement, the sample BHARs and s.e. bands over a representative three trading month interval ([-160,-100] days pre announcement), and the bootstrapped mean BHAR and s.e. bands over a three trading month period. Large BHARs emerge for each of the sample's tests. The pre announcement BHAR was larger than the event BHAR, but the difference was not statistically significant. Figures 4 and 5 illustrate the BHARs and the bootstrapped BHARs for the contemporaneous size cohort. Little evidence of abnormal returns for the cohort exists.

Figures 6-10 examine the power curves for the sample pre announcement and Figures 11-15 examine those for the size cohort during th4e event window. Tests of the sample's BHARs during the pre announcement period show very low power at the null of zero, implying that a traditional t-test on the BHARs will over reject. The results from the power tests on the size cohort showed quite the opposite results. Their power curves indicate that a test statistic with a null of zero has high power against a range of alternatives.

This suggests that the source of risk for the S&P firms is not one that is captured by a conventional proxy like market value or even a further sort conditioned on book to market. This sample appears to contain a highly idiosyncratic type of risk that has driven returns for at least a year before addition and continues the three trading months after inclusion to the 500 index. The lack of identifiability implies that historical averaging may be the best metric by which to ascertain whether a return is unexpected or not.

Conclusion

The permanence of the positive BHAR emanating from the announcement day appears to be illusionary. Using a generalized approach to the measurement of the significance of the BHAR that allows for a non zero expected BHAR, I observe that the abnormal returns were significantly different from zero for only the first month after the addition date. This is to say that the abnormal returns measured three months after addition are no different than those otherwise expected in the absence of the event. Hence, my conclusion supports the essence of the Price Pressure hypothesis. The protracted price pressure may have been driven by money managers as they work with indexers in profit sharing agreements. As the number of indexers has grown significantly over time, the potential profit from such a strategy has grown as well.

Table I

BHARs with and without Bootstrapped Null Adjustment

The daily firm and market returns used in Table I were pulled from the CRSP daily stock and index files. All returns include distributions and the market proxy is the CRSP value weighted market return (vwret). Panel A contains the compounded net of market or buy and hold returns for the sample of 110 firms for the intervals shown to the left. The t-statistics are calculated with a null of zero. Panel B shows the original buy and hold returns that appear in Panel A next to the bootstrapped own sample returns. To calculate the own sample expected BHAR for one day, I collect a vector of 150 historical returns [-200,-51] days before announcement for each firm and stack them on top of each other to form a 110x150 firm returns matrix. I do the same thing for the corresponding market returns. I then pick an event day (column) at random and resample (with replacement) 1000 pairs of firm and market returns. I average each type of return. I repeat this (pick a different column at random) 1000 times and take an overall average of the firms' and market returns. Because there is no compounding over one day, I take the difference between the firm and market average. This is the expected own sample BHAR for one day. For a 20 day interval, I would do this 20 times and then compound the 20 firm returns and market returns and then take the difference to form the expected BHAR over 20 days. The bootstrapped expected returns are positive and statistically significant, but not economically meaningful until a month after the announcement date. Panel C and D show the results of a cross sectional bootstrap controlling for size and size and book to market, respectively. For each firm, contemporaneous size and size-(B/M) cohorts were formed over the 60 days post addition. 1000 draws (with replacement) from the cohort's firm and market returns were averaged to form expectations of each sample firm's BHAR for each calendar day and then compounded over the various intervals on the left. An average of those expectations were taken over all of the sample firms to get the expected sample BHAR over any particular interval. Qualitatively, the same result arises as from the own sample time series bootstrap i.e. the expected BHAR are economically insignificant until 1 trading month post announcement. After 60 trading days post addition, the sample BHARs are no different than those expected from the two benchmark cohorts or from the own sample benchmark. The size cohort had 16,624 firms; the size-B/M cohort had 5,010 firms.

Panel A: Buy and Hold Abnormal Returns with Expected BHAR Equal to Zero

	AR	t*
Day 0	0.0489	13.81
(0, +20)	0.0336	4.46
(0, +40)	0.0518	3.22
(0, +60)	0.0462	2.06

*All t-statistics have p values < 0.05

Panel B: Buy and Hold Abnormal Returns with Own Sample Bootstrapped Null

	AR	BSAR	t	p
Day 0	0.0489	0.0484	13.66	0.00
(0, +20)	0.0336	0.0154	2.04	0.04
(0, +40)	0.0518	0.0031	0.19	0.85
(0, +60)	0.0462	-0.0238	-1.06	0.29

Panel C: BHAR with Bootstrapped Null from Size Cohort

	AR	BSAR	t	p
Day 0	0.0489	0.0485	13.69	0.00
(0, +20)	0.0336	0.0208	2.76	0.01
(0, +40)	0.0518	0.0269	1.67	0.10
(0, +60)	0.0462	0.0072	0.32	0.75

Panel D: BHAR with Bootstrapped Null from Size and B/M Cohort

	AR	BSAR	t	p
Day 0	0.0489	0.0484	13.67	0.00
(0, +20)	0.0336	0.0197	2.61	0.01
(0, +40)	0.0518	0.0251	1.56	0.12
(0, +60)	0.0462	0.0050	0.22	0.82

Table II

Market Model CAR Analysis

The daily firm and market returns used in Table II were pulled from the CRSP daily stock and index files. All returns include distributions and the market proxy is the CRSP value weighted market return (vwret). Market model coefficients and cumulative abnormal returns (CARs) were estimated on all 110 firms before addition and on the remaining 94 firms after addition (see Panel C). The pre-addition estimation was run from -200 to -50 days before the announcement date. The post-addition estimation was run +100, +250 days after the announcement. Panel A contains the results of a parameter stability analysis of the market model slope and intercept coefficients. Both the intercept and slope terms show evidence of structural break around the time of addition. The cumulative abnormal returns (CARs) in panels B and C are the summed arithmetic errors arising from each estimation. Panel B, using pre announcement coefficients with post announcement data, shows a strong Day 0 response to the announcement but then trends negative after two trading months. The results from Panel A (an increase in slope and a decrease intercept) predict this nonsensical result. Panel C, using post announcement coefficients with post announcement data, reports results that are qualitatively the same as those report in Table I. While the Day 0 reaction is strong, the CARs' significance dies out after one trading month. Panel A and B have 110 observations. Panel C has 94 observations.

Panel A: Average Slope and Intercept Coefficients Before and After Addition

	Intercept*	p-value	Slope*	p-value
Before	0.00140	0.005	1.26	0.043
After	0.00013		1.42	

* p-values are all less than 0.05

Panel B: Market Model CARs with Pre-Addition Parameters

	CAR	t-statistic
Day 0	0.0557	15.399**
(0, +20)	0.0238	2.267
(0, +40)	-0.0171	-1.166
(0, +60)	-0.0486	-2.711*
(0, +80)	-0.0815	-3.948**
(0,+100)	-0.1007	-4.636**

Panel C: Market Model CARs with Post-Addition Parameters

	CAR	t-statistic
Day 0	0.0579	15.421**
(0, +20)	0.0540	4.518**
(0, +40)	0.0369	1.589
(0, +60)	0.0307	1.497
(0, +80)	0.0299	1.264
(0,+100)	0.0266	1.007

**p value < 0.01
* p value < 0.05

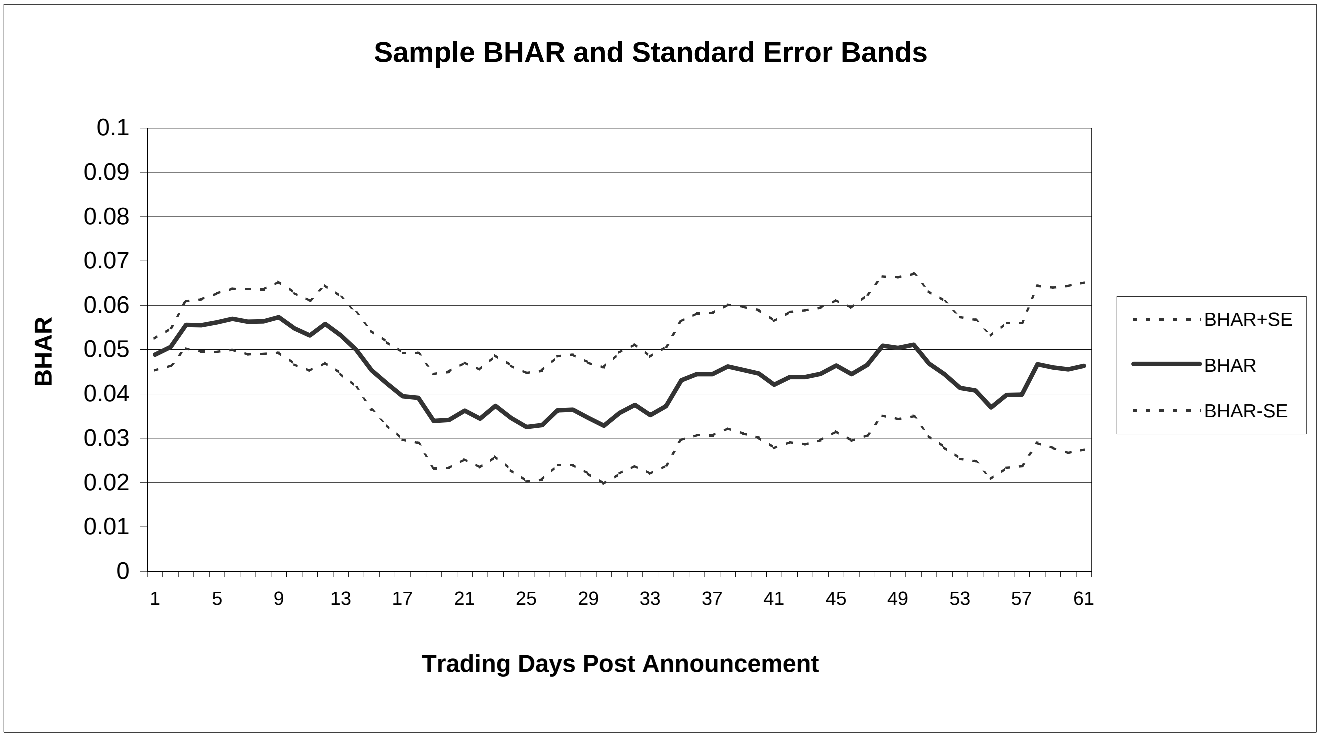


Figure 1. The BHAR of the Sample (BHAR) and its Standard Error Bands (BHAR+/-SE)

Figure 1. The buy and hold abnormal returns (BHAR) for the NYSE S&P sample firms (BHAR, N=110), the standard error bands around the sample BHAR. The daily firm and market returns were collected from the CRSP daily stock and index files. Returns over the interval [0,60] days after announcement were collected for each of the 110 sample firms. The matrix of all such intervals formed 110x61 matrices of stock and index returns from which the cross sectional means and standard errors of the buy and hold returns (BHARs) were calculated. The above graph shows that the sample BHAR is statistically and economically significant 60 days post announcement.

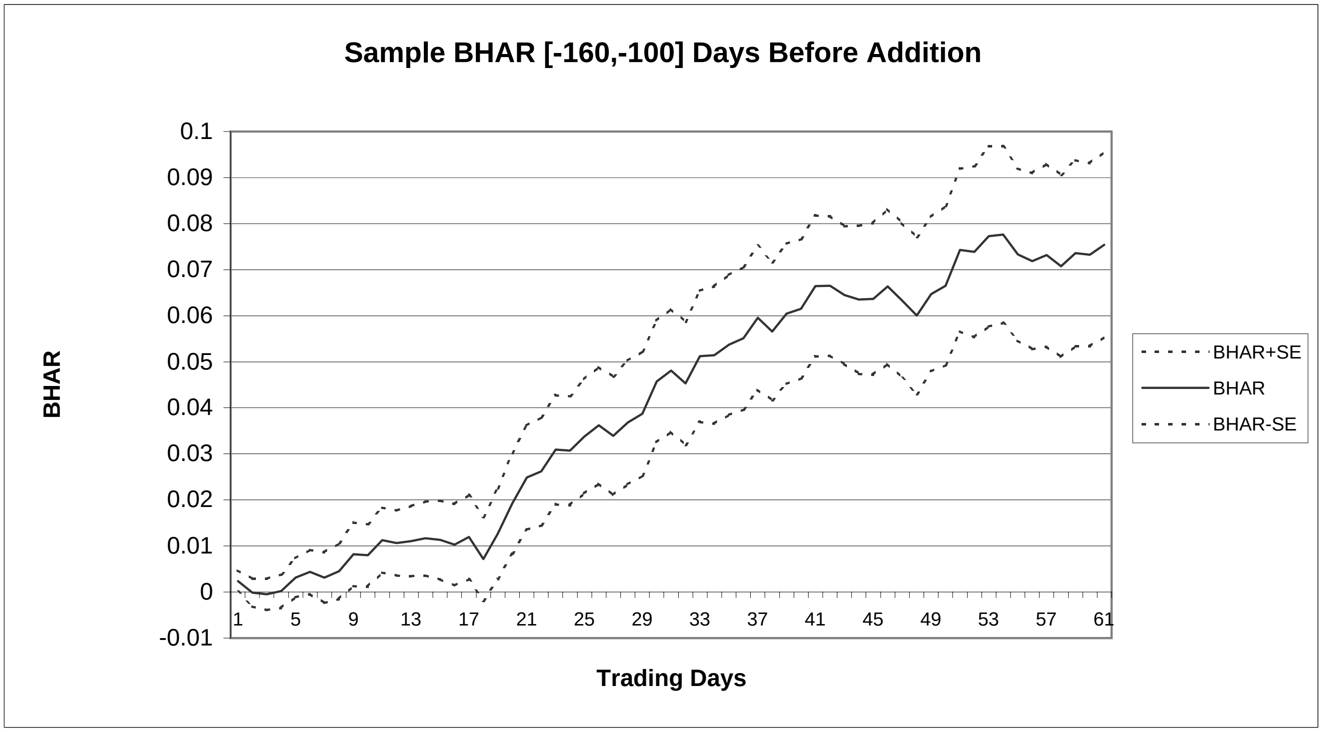


Figure 2. The Historical BHAR of the Sample Firms (BHAR) and its Standard Error Bands (BHAR+/-SE) for the period [-160,-100] Days before Announcement

Figure 2. The buy and hold abnormal returns (BHAR) for the NYSE S&P sample firms (BHAR, N=110), the standard error bands around the sample BHAR over the interval [-160,-100] days before announcement. The daily firm and market returns were collected from the CRSP daily stock and index files. Returns over the interval [-160,-100] days before announcement were collected for each of the 110 sample firms. The matrix of all such intervals formed 110x61 matrices of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The above graph shows that the historical BHAR over the [-160,-100] period before announcement is statistically and economically significant after twenty trading days. After 60 trading days, the sample BHAR is over 7% with a p-value significant at the 1% level.

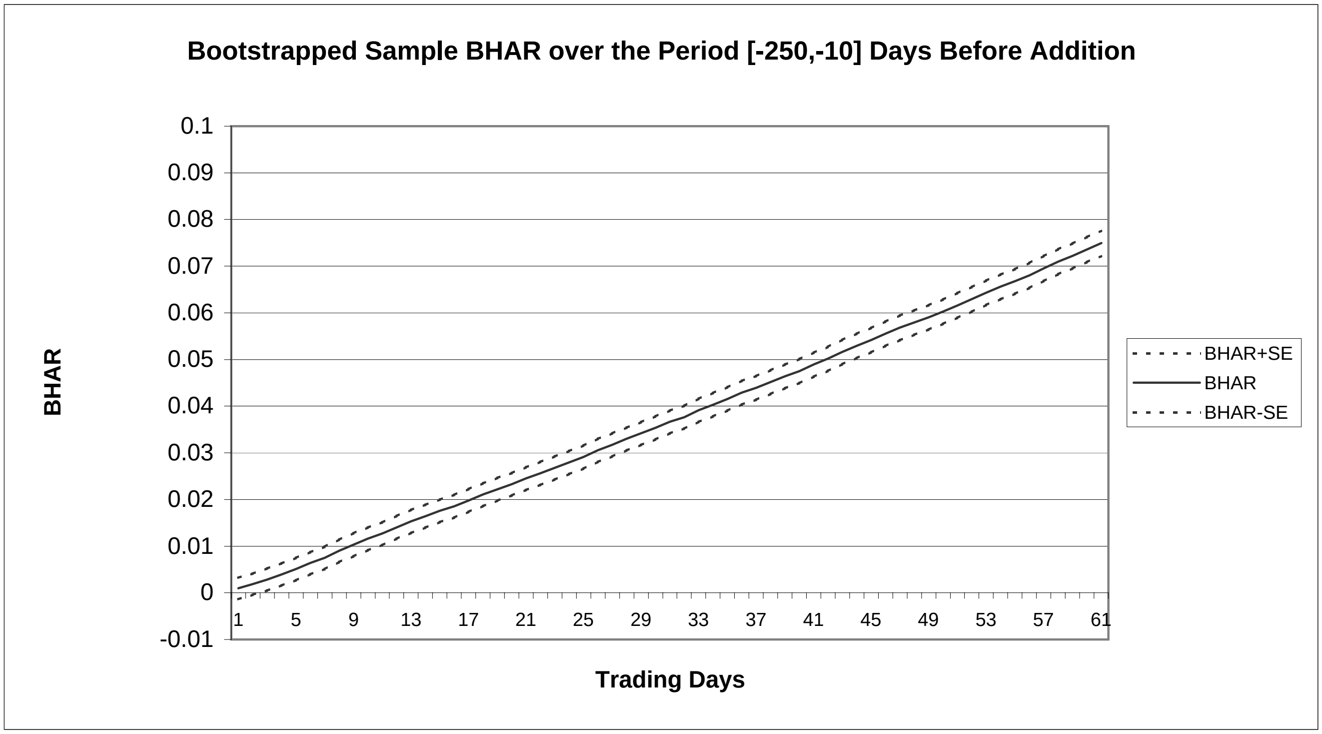


Figure 3. The Bootstrapped Historical BHAR of the Sample Firms (BHAR) and its Standard Error Bands (BHAR+/-SE)
Figure 3. The bootstrapped historical buy and hold abnormal returns (BHAR) for the NYSE S&P sample firms (BHAR, N=110) and the standard error bands around the sample BHAR over the interval [-250,-10] before announcement. The daily firm and market returns were collected from the CRSP daily stock and index files. Returns over the interval [-250,-10] before announcement were collected for each of the 110 sample firms. The matrix of all such intervals formed 110x241 matrices of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The bootstrap was calculated in the following manner: Over the 110x241 matrix, pick one day (column) at random. Resample 1000 times with replacement from the firm and market return vectors and calculate the mean BHAR. Repeat this last step 1000 times to get 1000 average daily firms and market returns. Compound over the 61 day interval to get all needed compounded firm and market returns. Take the difference of the average of the firm and market compoundings. These mean differences and their standard errors are plotted in the above graph. The graph shows that the bootstrapped average BHAR over a two trading week period is approximately 1%; after a trading month, the average or expected BHAR is slightly more than 2%. At the end of 3 trading months, the average BHAR is 7.5%.

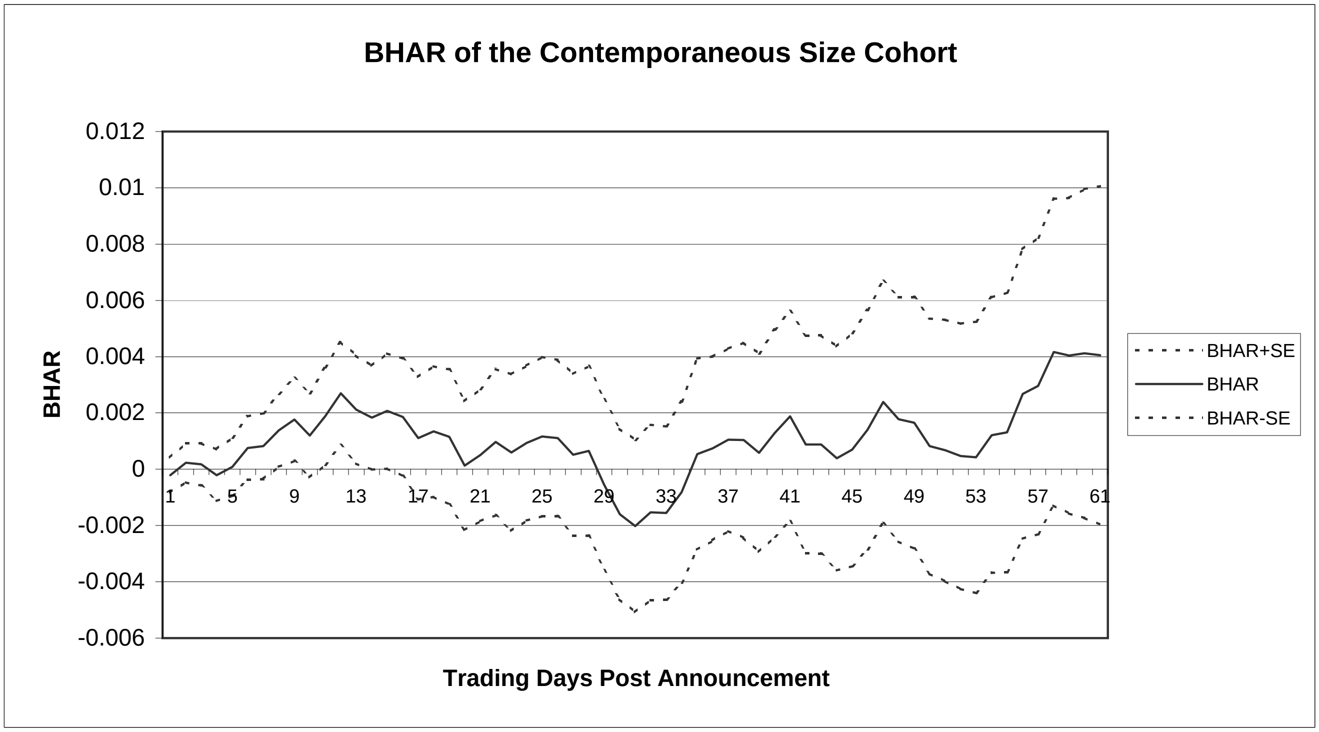


Figure 4. The Contemporaneous Size Cohort BHAR (BHAR) and its Standard Error Bands (BHAR+/-SE)
Figure 4 displays the mean BHAR and its standard error for a 61 day interval generated by the size cohort of the S&P sample using contemporaneous data. The daily firm and market returns were collected from the CRSP daily stock and index files. The size cohort was formed using the contemporaneous cohort of all firms traded on the NYSE which had between 80% and 120% of the market value for a given sample firm one week before that sample firm was added to the S&P 500. Returns over the interval [0,100] after announcement were collected for each of the size cohort's 15,452 firms. For a given sample firm, the matrix of all such firms formed an Nx101 matrix of stock and index returns from which the buy and hold return means (BHARs) were calculated. The standard errors were calculated across the sample's 110 firms. Only the first 61 days of the above-mentioned matrix were used for the above graph. Hence, for each sample firm, the cohort analysis covered the same 61 days post announcement as shown in Figure 1. The averages and standard errors over all 110 sample firms are shown above. The graph shows no significant BHAR 61 days post announcement.

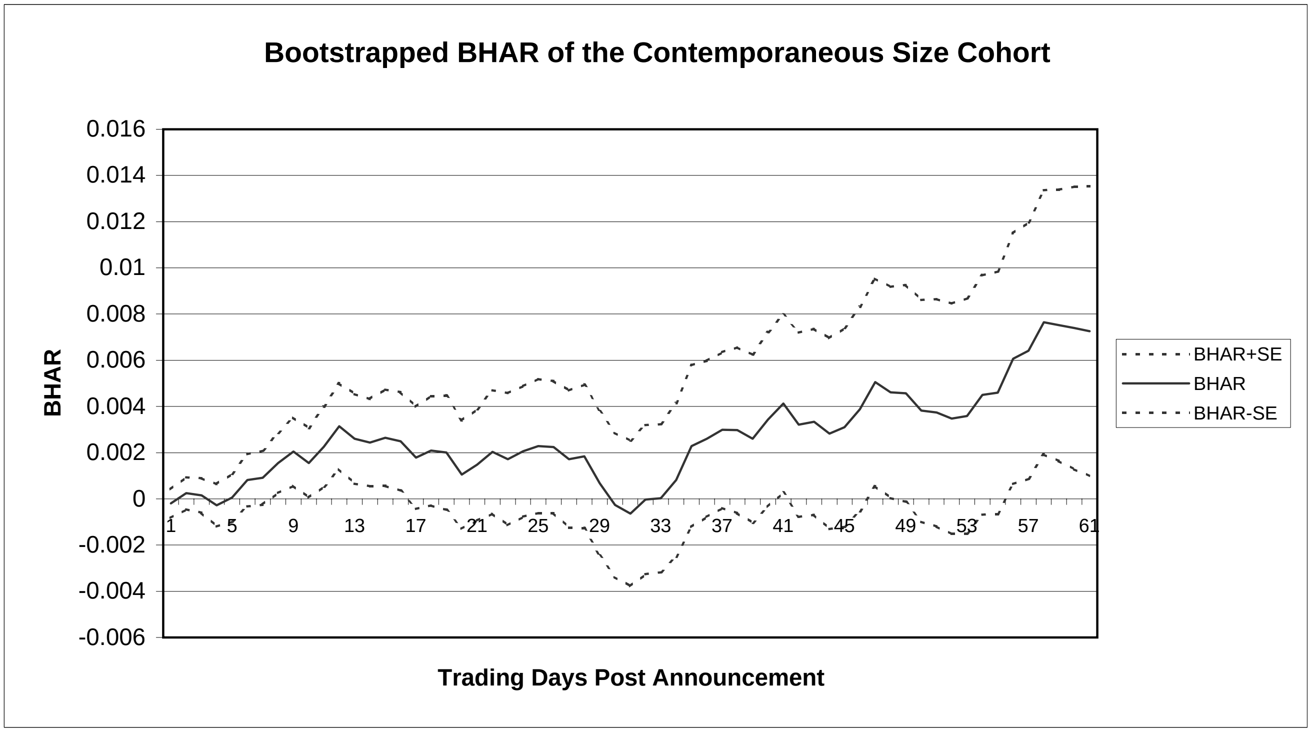


Figure 5. The Bootstrapped BHAR of the Size Cohort (BHAR) and its Standard Errors (BHAR+/-SE)
Figure 5 displays the bootstrapped BHAR for a 61 day interval generated by the size cohort of the S&P sample using contemporaneous data. The daily firm and market returns were collected from the CRSP daily stock and index files. The size cohort was formed using the contemporaneous cohort of all firms traded on the NYSE which had between 80% and 120% of the market value for a given sample firm one week before that sample firm was added to the S&P 500. Returns over the interval [0,100] after announcement were collected for each of the size cohort's 15,452 firms. For a given sample firm, the matrix of all such firms formed an Nx101 matrix of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. Only the first 61 days of the above-mentioned matrix were used for the above graph. Hence, for each sample firm, the cohort analysis covered the same 61 days post announcement as shown in Figure 19. The bootstrap was calculated in the following manner: Over the Nx61 matrix, pick one day (column) at random. Resample 1000 times with replacement over the firm and market return vectors and take the average of each to form the BHAR for one day. Repeat 1000 times to get average daily firm and market returns. Compound over the 61 day interval to get all needed compounded firm and market returns. Take the difference of the average of the firm and market compoundings. These differences and their standard errors are in the above graph. The graph shows no significant BHAR 61 days after announcement.

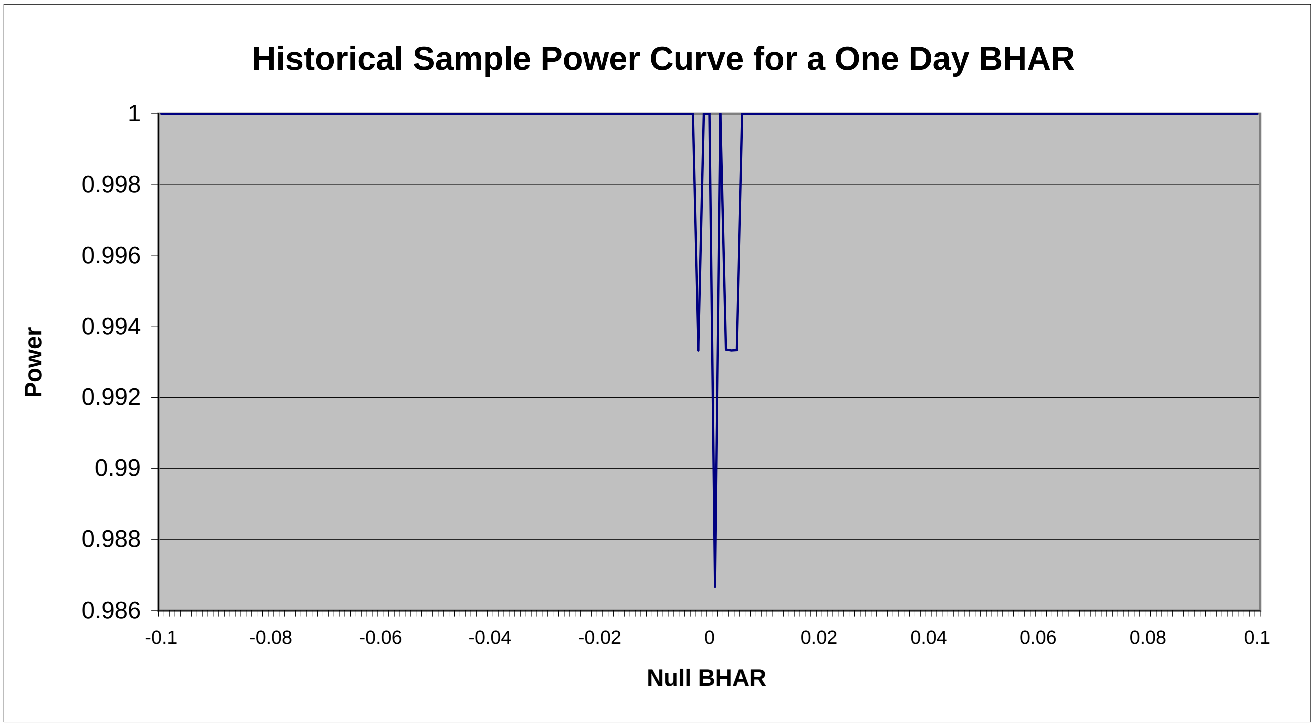


Figure 6. Bootstrapped Historical Sample Power Curve for a One Day Interval
Figure 6 displays the power curve for a one-day interval generated by the S&P sample using bootstrapped historical data. The daily firm and market returns were collected from the CRSP daily stock and index files. Returns over the interval [-250,-10] before announcement were collected for each of the 110 sample firms. The matrix of all such intervals formed 110x241 matrices of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: One day in the [-250,-10] sample was chosen randomly. For this day, the firm and market returns were resampled 1000 times with replacement and a BHAR mean and standard error were calculated. The last step was repeated 1000 times to get 1000 means and standard errors. An interval of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were then used to calculate the probability of Type II error. For example, one would subtract a particular null value from the range, say -0.10, from the BHAR mean and then divide by the standard error and see how many observations out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was repeated and the power curves averaged over the 110 sample firms to form the graph above. The graph shows a relatively large number of mean sample BHARs are larger than zero even with a one day window.

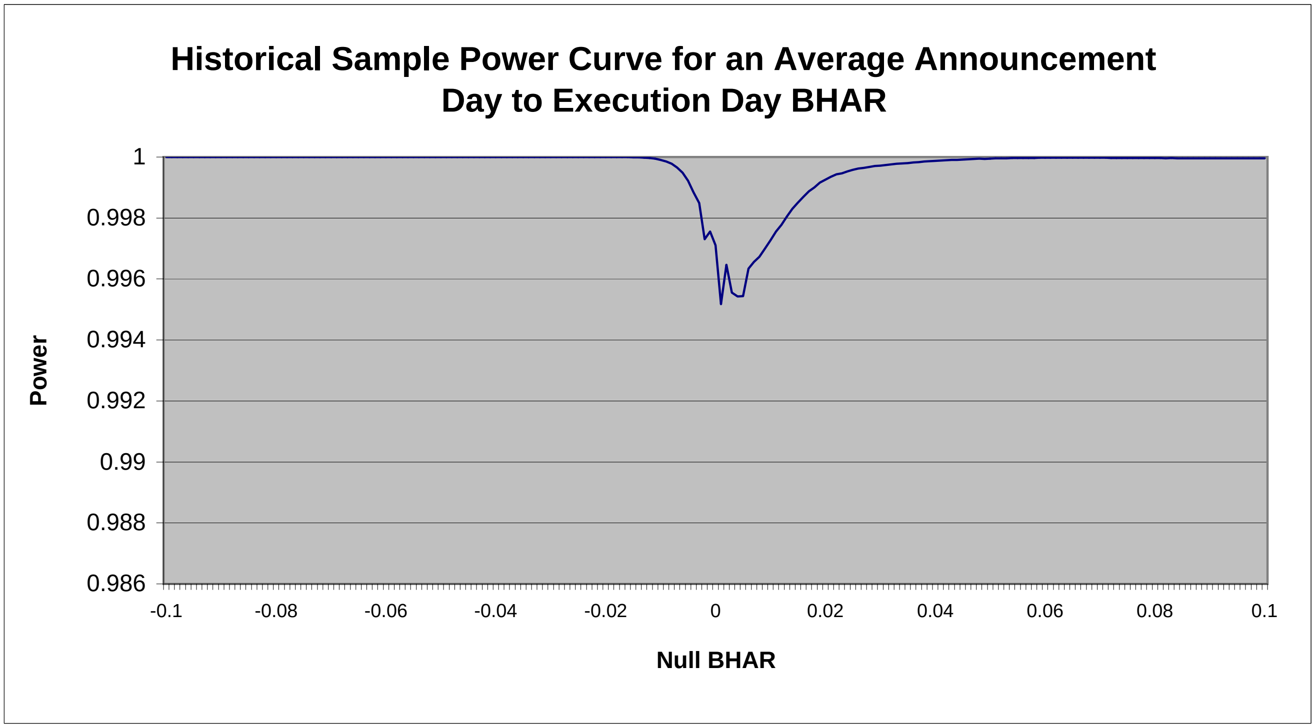


Figure 7. Bootstrapped Historical Sample Power Curve for the Announcement Day to Execution Day Interval
Figure 7 displays the power curve for an announcement day to execution day interval generated from the S&P sample using bootstrapped historical data. The daily firm and market returns were collected from the CRSP daily stock and index files. Returns over the interval [-250,-10] before announcement were collected for each of the 110 sample firms. The matrix of all such intervals formed 110x241 matrices of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: The number of days corresponding to a sample firm's announcement day to execution day interval are drawn at random from the [-250,-10] sample range. For each day in this interval, the firm and corresponding market returns were resampled 1000 times with replacement. After compounding for the appropriate number of days, the mean BHARs and their standard error were calculated. The above steps were repeated 1000 times to get 1000 means and standard errors per sample firm. A span of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were used to calculate the probability of Type II error. For example, one would subtract null value, say -0.10, from the range of 1000 means calculated above and then divide by the respective standard error and see how many values out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was repeated over the 110 firms in the S&P sample and averaged to form the graph above. The graph shows a very large number of mean sample BHARs are larger than zero even with a one day window thus indicating reduced power at the null of zero BHAR.

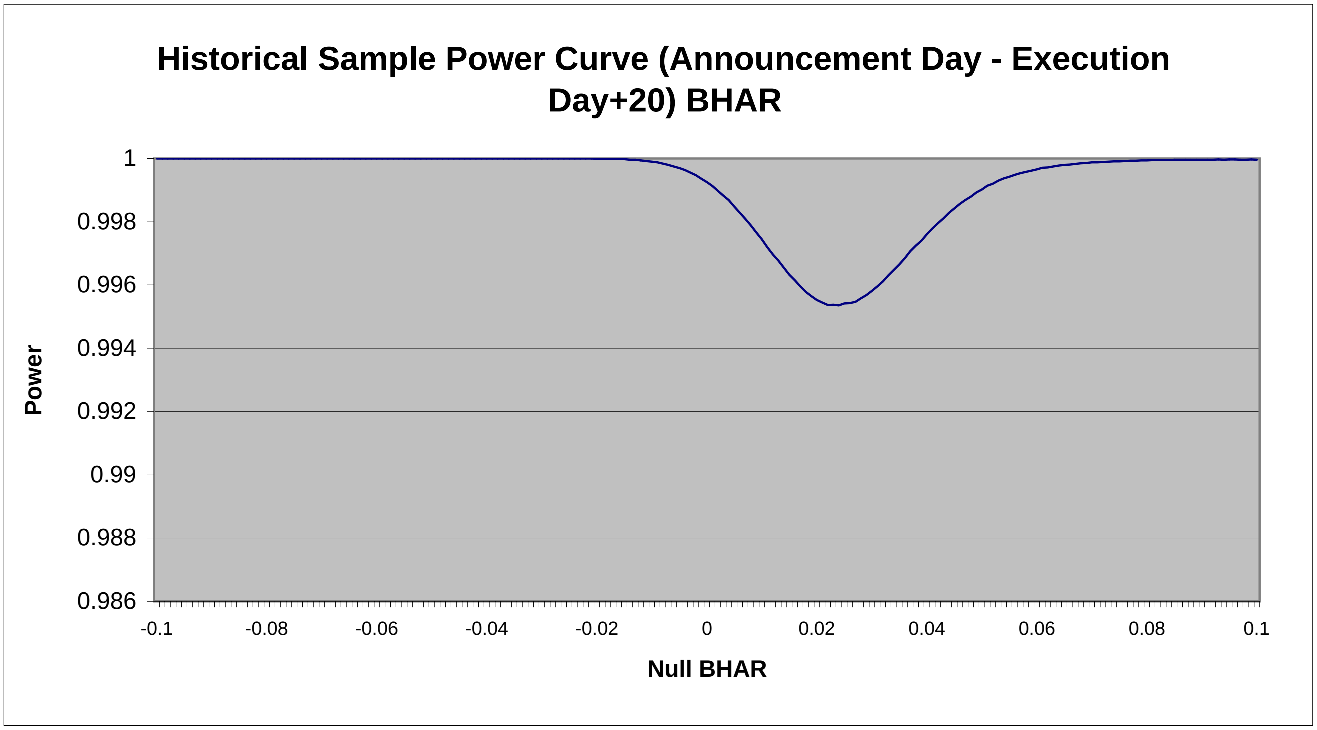


Figure 8. Bootstrapped Historical Sample Power Curve for Announcement Day to Execution Day+20 Interval
Figure 8 displays the power curve for an announcement day to execution day + 20 day interval generated from the S&P sample using bootstrapped historical data. The daily firm and market returns were collected from the CRSP daily stock and index files. Returns over the interval [-250,-10] before announcement were collected for each of the 110 sample firms. The matrix of all such intervals formed 110x241 matrices of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: The number of days corresponding to a sample firm's announcement day to execution day +20 day interval are drawn at random from the [-250,-10] sample range. For each day in this interval, the firm and corresponding market returns were resampled 1000 times with replacement. After compounding for the appropriate number of days, the mean BHARs and their standard error were calculated. The above steps were repeated 1000 times to get 1000 means and standard errors per sample firm. A span of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for for a two tailed test) were used to calculate the probability of Type II error. For example, one would subtract null value, say -0.10, from the range of 1000 means calculated above and then divide by the respective standard error and see how many values out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was repeated over the 110 firms in the S&P sample and averaged to form the graph above. The graph shows a very large number of mean sample BHARs are larger than zero with a mean of about 2.5% thus indicating very low power at the null of zero BHAR.

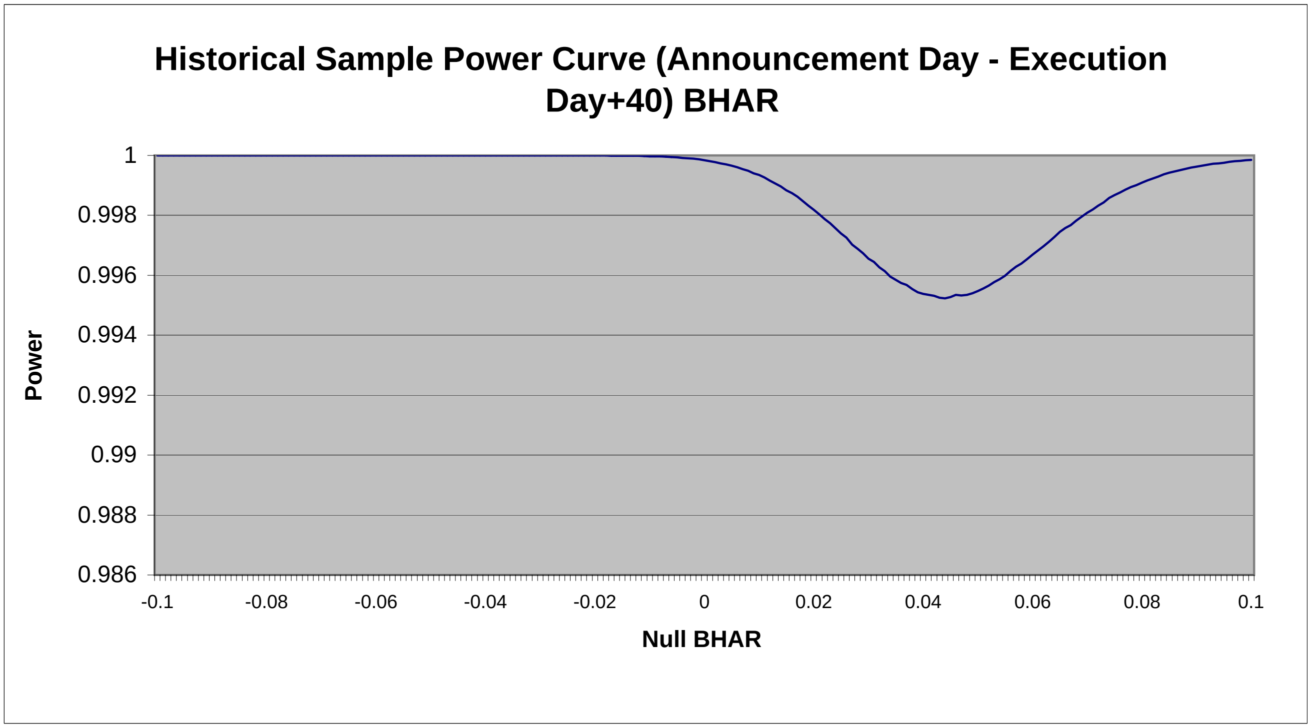


Figure 9. Bootstrapped Historical Sample Power Curve for Announcement Day to Execution Day+40 Interval
Figure 9 displays the power curve for an announcement day to execution day + 40 day interval generated from the S&P sample using bootstrapped historical data. The daily firm and market returns were collected from the CRSP daily stock and index files. Returns over the interval [-250,-10] before announcement were collected for each of the 110 sample firms. The matrix of all such intervals formed 110x241 matrices of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: The number of days corresponding to a sample firm's announcement day to execution day +40 day interval are drawn at random from the [-250,-10] sample range. For each day in this interval, the firm and corresponding market returns were resampled 1000 times with replacement. After compounding for the appropriate number of days, the mean BHARs and their standard error were calculated. The above steps were repeated 1000 times to get 1000 means and standard errors per sample firm. A span of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were used to calculate the probability of Type II error. For example, one would subtract null value, say -0.10, from the range of 1000 means calculated above and then divide by the respective standard error and see how many values out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was repeated over the 110 firms in the S&P sample and averaged to form the graph above. The graph shows a very large number of mean sample BHARs are larger than zero with a mean of about 4.5% thus indicating very low power at the null of zero BHAR.

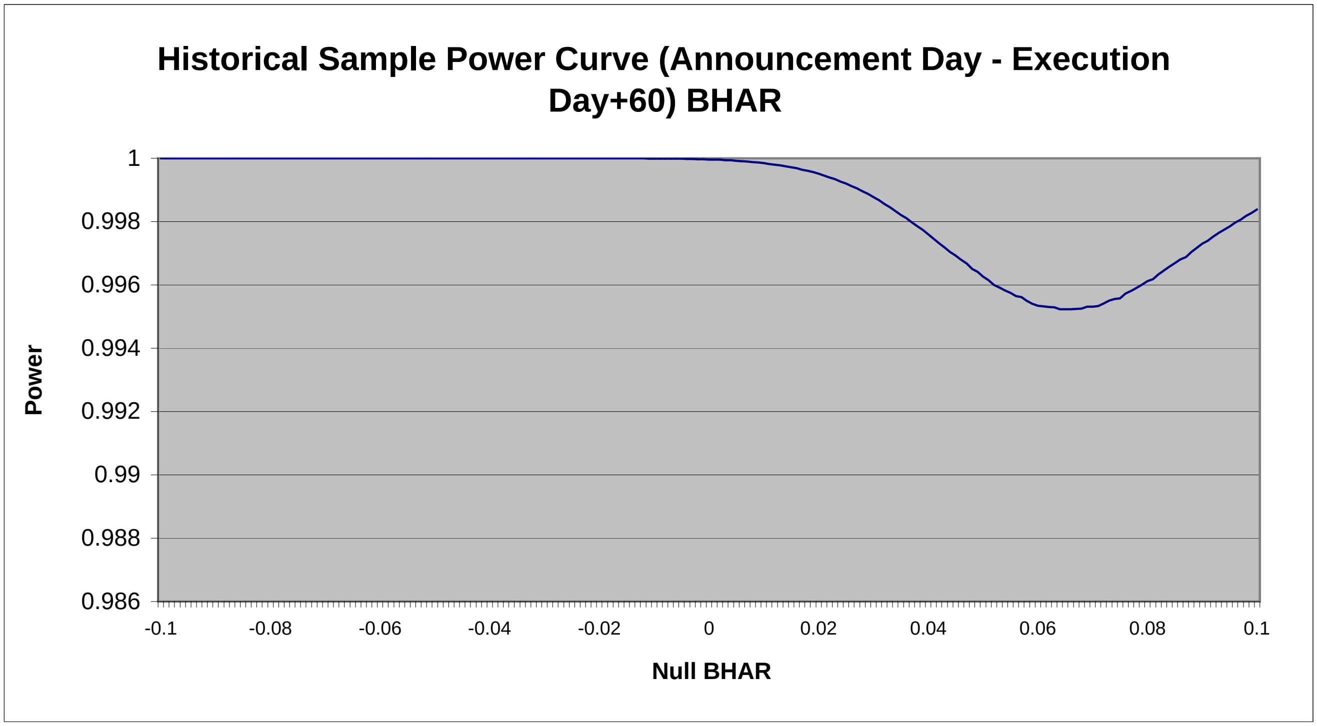


Figure 10. Bootstrapped Historical Sample Power Curve for Announcement Day to Execution Day+60 Interval
Figure 10 displays the power curve for an announcement day to execution day + 60 day interval generated from the S&P sample using bootstrapped historical data. The daily firm and market returns were collected from the CRSP daily stock and index files. Returns over the interval [-250,-10] before announcement were collected for each of the 110 sample firms. The matrix of all such intervals formed 110x241 matrices of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: The number of days corresponding to a sample firm's announcement day to execution day +60 day interval are drawn at random from the [-250,-10] sample range. For each day in this interval, the firm and corresponding market returns were resampled 1000 times with replacement. After compounding for the appropriate number of days, the mean BHARs and their standard error were calculated. The above steps were repeated 1000 times to get 1000 means and standard errors per sample firm. A span of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were used to calculate the probability of Type II error. For example, one would subtract null value, say -0.10, from the range of 1000 means calculated above and then divide by the respective standard error and see how many values out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was repeated over the 110 firms in the S&P sample and averaged to form the graph above. The graph shows a vary large number of mean sample BHARs are large than zero with a mean of about 7% thus indicating very low power at the null of zero BHAR.

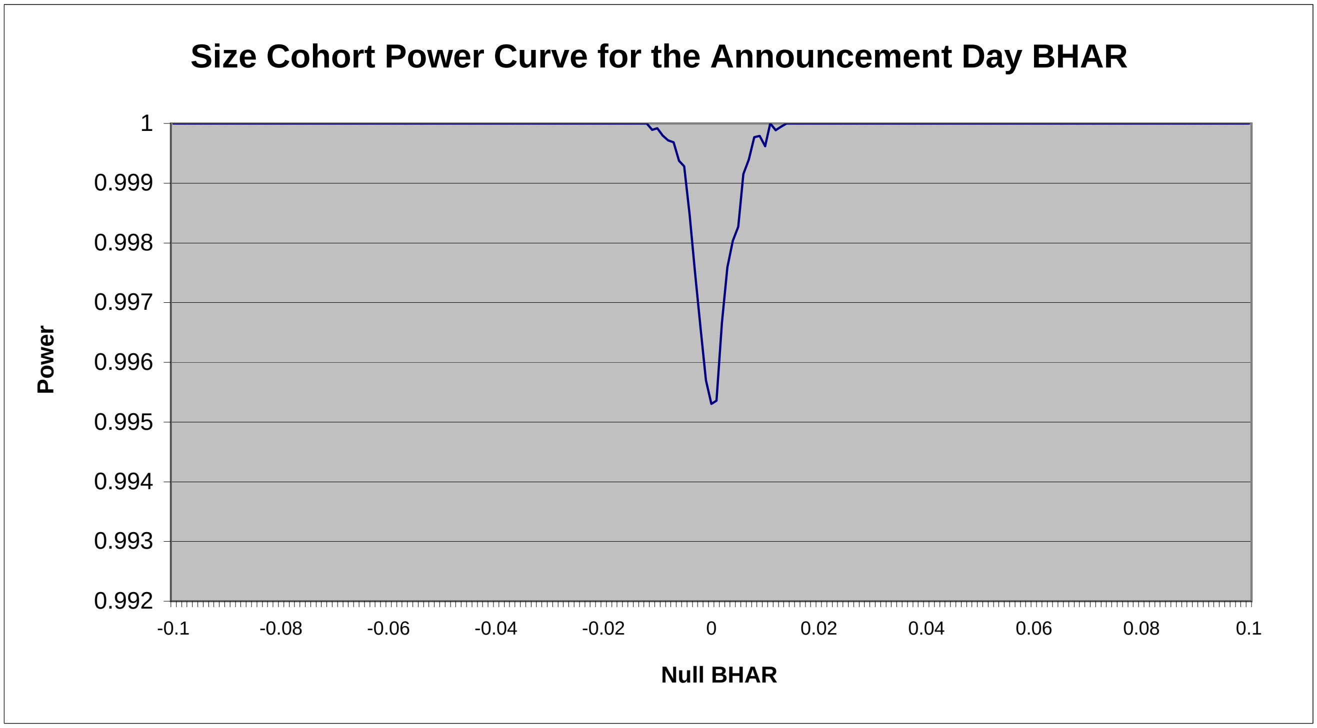


Figure 11. Bootstrapped Size Cohort Power Curve for Announcement Day

Figure 11 displays the power curve for the announcement day (one day interval) generated by the size cohort of the S&P sample using bootstrapped contemporaneous data. The daily firm and market returns were collected from the CRSP daily stock and index files. The size cohort was formed using the contemporaneous cohort of all firms traded on the NYSE which had between 80% and 120% of the market value for a given sample firm one week before that sample firm was added to the S&P 500. Returns over the interval [0,100] after announcement were collected for each of the size cohort's 15,452 firms. For a given sample firm, the matrix of all such firms formed an Nx101 matrix of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: For the addition day (Day 0), the firm and market returns for the cohort were selected. For this day, the firm and market returns were resampled 1000 times with replacement and a BHAR mean and standard error were calculated. The last step was repeated 1000 times to get a collection of 1000 means and standard errors. A span of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were then used to calculate the probability of Type II error. For example, one would subtract null value, say -0.10, from the range of the BHAR means and then divide by the respective standard error and see how any observations out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null= -0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was averaged over all of the sample firms to form the graph above. The above graph show that the BHAR null of zero appears to have good power for the size cohort.

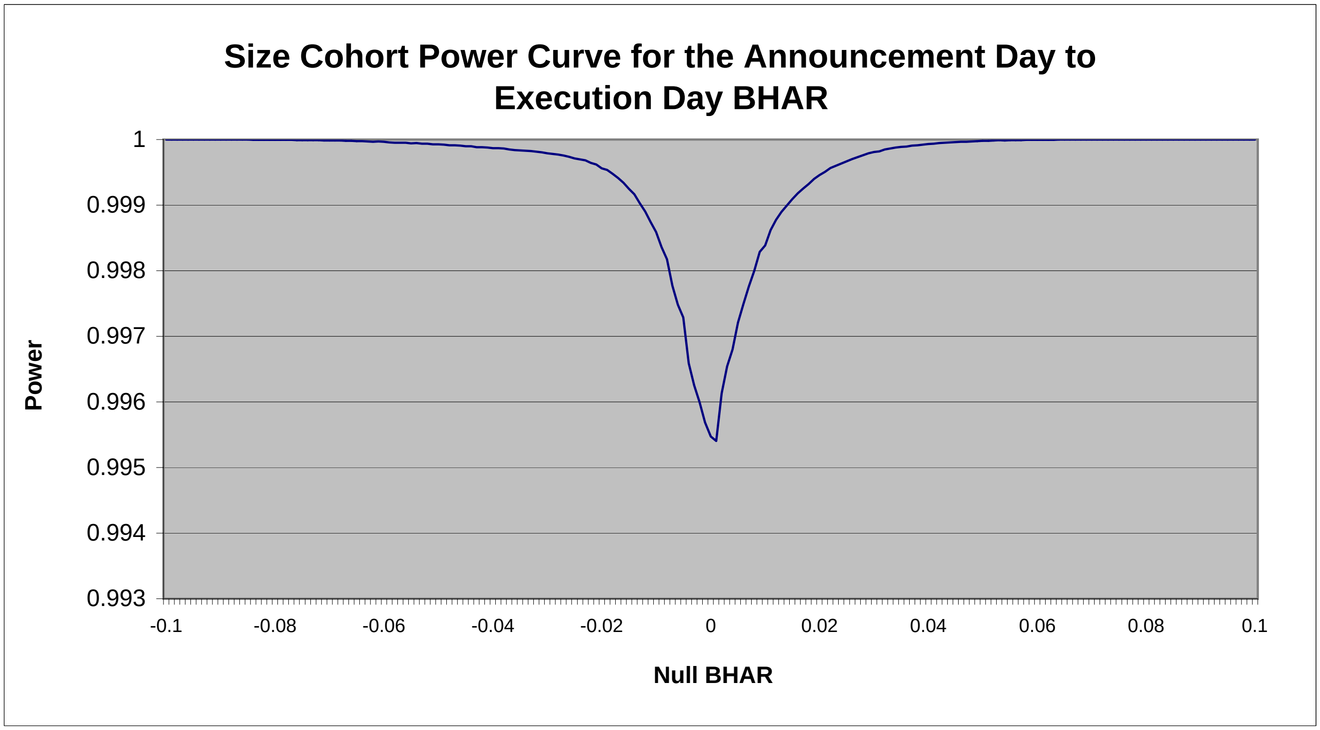


Figure 12. Bootstrapped Size Cohort Power Curve for Announcement Day to Execution Day Interval
Figure 12 displays the power curve for the announcement day to execution day interval + 20 generated by the size cohort of the S&P sample using bootstrapped contemporaneous data. The daily firm and market returns were collected from the CRSP daily stock and index files. The size cohort was formed using the contemporaneous cohort of all firms traded on the NYSE which had between 80% and 120% of the market value for a given sample firm one week before that sample firm was added to the S&P 500. Returns over the interval [0,100] after announcement were collected for each of the size cohort's 15, 452 firms. For a given sample firm, the matrix of all such firms formed an Nx101 matrix of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: From the addition day (Day 0) to the execution day + 20, the firm and market returns for the cohort were selected. For these days, the firm and market returns were resampled 1000 times with replacement and a BHAR mean and standard error were calculated. The last step was repeated 1000 times to get a collection of 1000 means and standard errors. An interval of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were then used to calculate the probability of Type II error. For example, one would subtract a null value, say -0.10, from the range of the BHAR means and then divide by the respective standard error and see how any observations out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was averaged over all of the sample firms to form the graph above. The symmetrical nature of the above graph suggests that the size cohort has good power against the null BHAR of zero .

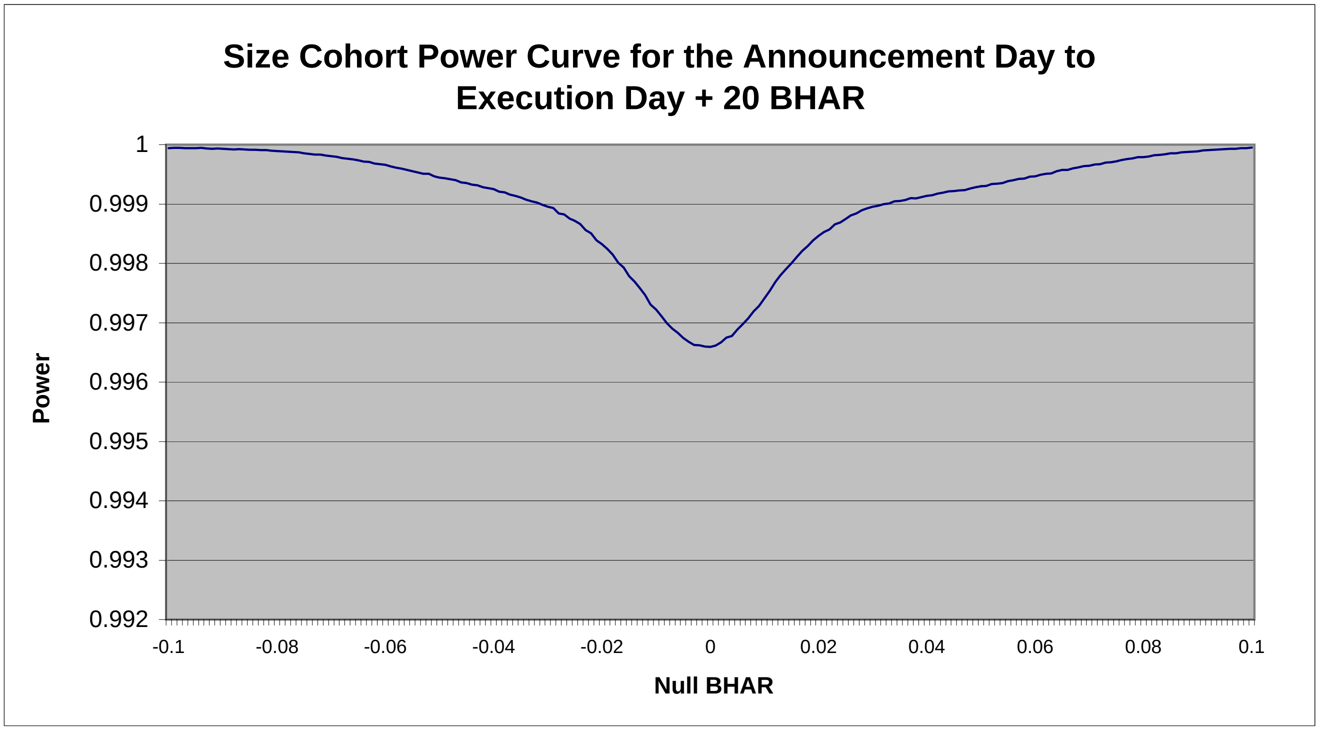


Figure 13. Bootstrapped Size Cohort Power Curve for Announcement Day to Execution Day + 20 Interval

Figure 13 displays the power curve for the announcement day to execution day interval generated by the size cohort of the S&P sample using bootstrapped contemporaneous data. The daily firm and market returns were collected from the CRSP daily stock and index files. The size cohort was formed using the contemporaneous cohort of all firms traded on the NYSE which had between 80% and 120% of the market value for a given sample firm one week before that sample firm was added to the S&P 500. Returns over the interval [0,100] after announcement were collected for each of the size cohort's 15, 452 firms. For a given sample firm, the matrix of all such firms formed an Nx101 matrix of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: From the addition day (Day 0) to the execution day, the firm and market returns for the cohort were selected. For these days, the firm and market returns were resampled 1000 times with replacement and a BHAR mean and standard error were calculated. The last step was repeated 1000 times to get a collection of 1000 means and standard errors. An interval of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were then used to calculate the probability of Type II error. For example, one would subtract a null value, say -0.10, from the range of the BHAR means and then divide by the respective standard error and see how any observations out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was repeated 1000 times and averaged and then repeated over all of the sample firms to form the graph above. The symmetrical nature of the above graph suggests that the size cohort has good power against the null BHAR of zero .

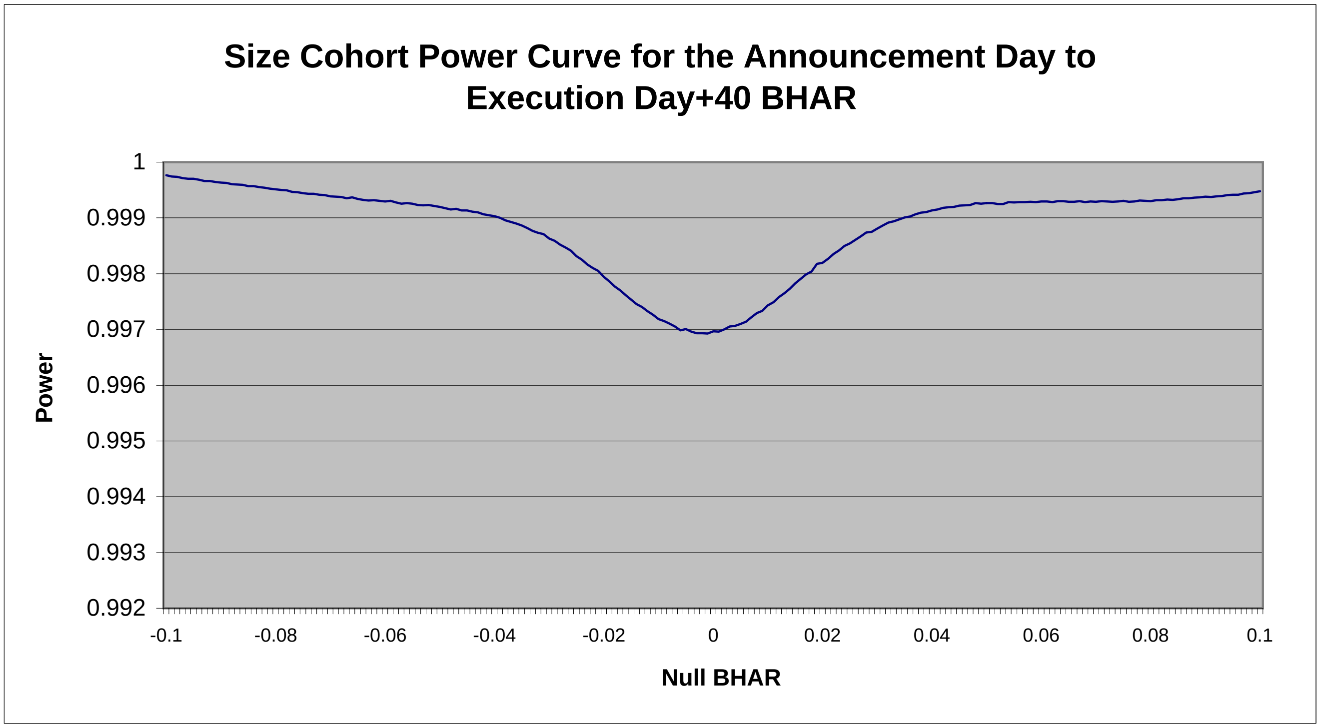


Figure 14. Bootstrapped Size Cohort Power Curve for Announcement Day to Execution Day+40 Interval
Figure 14 displays the power curve for the announcement day to execution day interval + 40 generated by the size cohort of the S&P sample using bootstrapped contemporaneous data. The daily firm and market returns were collected from the CRSP daily stock and index files. The size cohort was formed using the contemporaneous cohort of all firms traded on the NYSE which had between 80% and 120% of the market value for a given sample firm one week before that sample firm was added to the S&P 500. Returns over the interval [0,100] after announcement were collected for each of the size cohort's 15, 452 firms. For a given sample firm, the matrix of all such firms formed an Nx101 matrix of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: From the addition day (Day 0) to the execution day + 40, the firm and market returns for the cohort were selected. For these days, the firm and market returns were resampled 1000 times with replacement and a BHAR mean and standard error were calculated. The last step was repeated 1000 times to get a collection of 1000 means and standard errors. An interval of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were then used to calculate the probability of Type II error. For example, one would subtract null value, say -0.10, from the range of the BHAR means and then divide by the respective standard error and see how any observations out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was averaged over all of the sample firms to form the graph above. The symmetrical nature of the above graph suggests that the size cohort has good power against the null BHAR of zero.

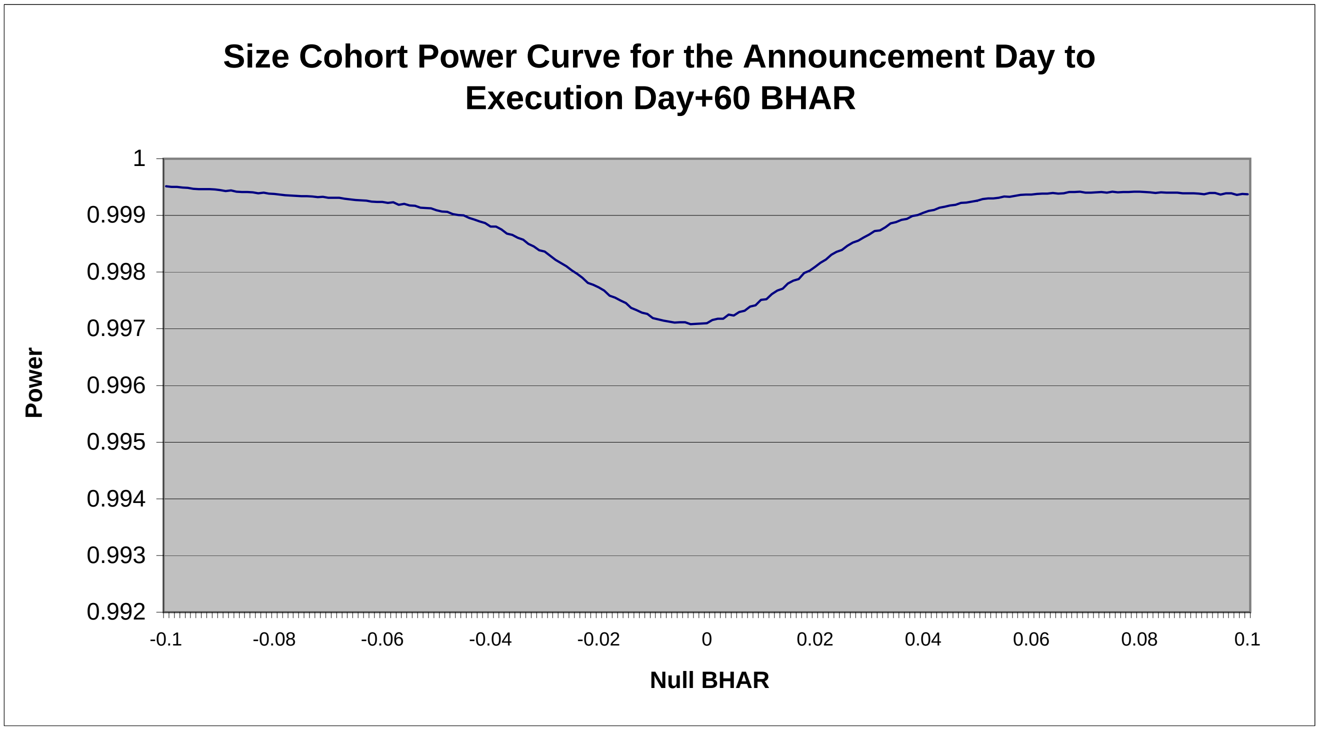


Figure 15. Bootstrapped Size Cohort Power Curve for Announcement Day to Execution Day+60 Interval
Figure 15 displays the power curve for the announcement day to execution day interval + 60 generated by the size cohort of the S&P sample using bootstrapped contemporaneous data. The daily firm and market returns were collected from the CRSP daily stock and index files. The size cohort was formed using the contemporaneous cohort of all firms traded on the NYSE which had between 80% and 120% of the market value for a given sample firm one week before that sample firm was added to the S&P 500. Returns over the interval [0,100] after announcement were collected for each of the size cohort's 15, 452 firms. For a given sample firm, the matrix of all such firms formed an Nx101 matrix of stock and index returns from which the means and standard errors of the buy and hold returns (BHARs) were calculated. The power curve was constructed in the following way: From the addition day (Day 0) to the execution day + 60, the firm and market returns for the cohort were selected. For these days, the firm and market returns were resampled 1000 times with replacement and a BHAR mean and standard error were calculated. The last step was repeated 1000 times to get a collection of 1000 means and standard errors. An interval of null means ranging from -10% (-0.10) to 10% (0.10) and an acceptance region of 1.96 (a chosen significance level of 5% for a two tailed test) were then used to calculate the probability of Type II error. For example, one would subtract null value, say -0.10, from the range of the BHAR means and then divide by the respective standard error and see how any observations out of the 1000 were less than -1.96 or greater than 1.96. This amount is the probability of Type II error. 1-Prob(Type II error | null=-0.10) is the power at the null value of -0.10. The same procedure is conducted over the range of null values [-0.10, 0.10] with an incremental increase of 0.001 to create a power curve. This procedure was averaged over all of the sample firms to form the graph above. The symmetrical nature of the above graph suggests that the size cohort has good power against the null BHAR of zero.

Chapter 3

S&P Dynamic Liquidity Analysis

Introduction

In the second chapter, I examine changes in static and dynamic liquidity measures around the time of addition. Although earlier studies such as BW find no permanent liquidity gains, subsequent studies such as Hedge and McDermott [10] (HM) that cover later years and eliminate Nasdaq firms find perdurable liquidity improvements. HM find statistically and economically meaningful reductions in the quoted, relative quoted spread and effective spreads. As well, they performed tests on ratios of the fixed cost and asymmetric information spread components pre and post addition and found that the fixed costs fell significantly. Due to the spurious results that can arise from static tests of the bid ask spread (see Van Ness et. al. [14]), I complement those tests with a dynamical liquidity measure whose results augment the previous literature with the conclusion that the spread reductions arise in an environment with less private information in the trade flow.

I decompose the effective spread using the methods described by Glosten and Harris[8] and Madhavan, Richardson and Roomans[13]. I confirm Hedge and McDermott's finding that the reduction in the spread (which is noticeable only in NYSE stocks) arises from a decrease in the fixed cost component. This may arise as institutional managers or index arbitrageurs would cause the number of daily trades to increase permanently, thereby allowing specialists to reduce the spread as their fixed costs are covered sooner in the day. On the other hand, the adverse selection component remains constant, indicating no perceptible change in the amount of asymmetric information.

Relying on static liquidity measures is problematic to two reasons. First, it ignores an important dimension of liquidity, resilience of the market. Second, as Van Ness et. al. (VN) demonstrate, one problem with spread decomposition analysis is that nonsensical estimates of the fixed cost and adverse selection components of the spread often arise (both extremely large positive numbers and/or negative numbers can confound coefficient aggregation). Lastly, the results from VN suggest that the models' choice of explanatory variables may do a poor job of capturing a firm's degree of asymmetric information.

Rather than selecting a model that best fits the sample or picking firms with economically reasonable results, I compare the results of the static analysis with those of a dynamic liquidity analysis. I analyze the dynamical changes in liquidity by adapting Hasbrouck's bi-variate VAR(p) model of quote revisions and order flow[5] to allow structural change in the model through admitting thresholds in the signed order flow. I then use a generalized impulse response function to measure the permanent price impact of an order flow shock to the buy and sell sides pre and post addition. To decompose the price impact into its component parts, I use Hasbrouck's model for trade informativeness [6] to identify whether the reduction in price impact occurs because of a smaller shock to the system, a smaller quote revision (caused by less private information in the trade flow), or some combination of the two. I find a significant reduction in the size of the shocks before and after addition, as well as smaller quote revisions that arise from trades that contain less private information. Both the dynamic and static measures indicate an improvement in liquidity.

The outline of the chapter is as follows. In Section 2, I describe the data. Section 3 features the various methodologies used. In Section 4, I discuss the results and conclude in Section 5.

Data

The raw sample consists of 249 firms. Seventy seven firms were deleted because of confounding events around the announcement or addition dates. Firm and index returns were drawn from the CRSP daily stock and index files. The index returns are value weighted and both returns files include distributions.[1] All of the intraday data are from the Trade and Quote (TAQ) database. All data were filtered for recording errors and were constrained to normal trading hours. Following previous studies, I selected the major exchange for each stock by summing the volume of each of the markets on which it was traded and chose the primary exchange as the one with the largest volume. This designation coincides with the CRSP index indicator. Lastly, all of the 'Day 0' event days mentioned in the paper are actually the trading day after the event as the announcement always takes place after trading hours.

[1] The results were similar for estimations carried out with the equally weighted index.

Methodology

Liquidity Measures

Spread Decomposition

S&P addition potentially affects stock's spreads in two ways. First, if the spread narrows permanently, as Hedge and McDermott note, then investors may benefit from lower transaction costs. In addition, they may also enjoy trading in a less volatile environment (with lower price impacts from order flow shocks). I first explore the former to determine if this sample shares a like reduction in transaction costs as measured by the effective spread.

To ascertain the source of the spread reduction, I decompose the spread into adverse selection and fixed costs components pre and post addition. I use a modified version of the Glosten and Harris (GH) (to accommodate the Lee and Ready algorithm[15]) and Madhavan, Richardson, and Roomans (MRR) models to decompose the spread before and after addition. Glosten and Harris' model is unique in that it features trade indicator variables and signed order flow to explain observed changes in price.

$$\Delta P = Q_t C_t - Q_{t-1} C_{t-1} + Q_t Z_t + e_t \tag{3.1}$$

where ΔP is the observed price change and Q_t is the trade indicator derived from Lee and Ready's algorithm. The fixed and variable cost spread components are assumed to take a linear functional form $C_t = c_0 + c_1 V_t$, $Z_t = z_0 + z_1 V_t$. Plugging in the values for C_t and Z_t, the fixed and variable costs, respectively, we have

$$\Delta P = c_0(Q_t - Q_{t-1}) + c_1(Q_t V_t - Q_{t-1} V_{t-1}) + z_0 Q_t + z_1 V_t + e_t \tag{3.2}$$

After running a test to see which of the four coefficients are significant in the cross section, I reran the estimation with the significant variables, constraining the fixed cost C_t to be fixed (not a function of volume), and the adverse selection component to be proportional to volume with a constant of proportionality z_1. Consequently, the specification used in this study is

$$\Delta P = c_0(Q_t - Q_{t-1}) + z_1 Q_t V_t + e_t \tag{3.3}$$

is the specification used in this paper.

Madhavan, Richardson, and Roomans (MRR) follow a different approach and model changes in observed prices as a function of trade indicator variables. MRR has a five moment GMM framework to identify a constant α, the fixed cost, ϕ, and adverse selection component of the spread, θ, the correlation in order flow trade signs, ρ, and the probability of trades crossing inside the spread, λ.

$$\Delta P = \alpha + (\phi + \theta)x_t - (\phi + \rho\theta)x_{t-1} + e_t \tag{3.4}$$

with moment conditions

$$E \begin{pmatrix} (u_t - \alpha) \\ (u_t - \alpha)x_t \\ (u_t - \alpha)x_{t-1} \\ x_t x_{t-1} - x_t^2 \\ |x_t| - (1 - \lambda) \end{pmatrix} = 0 \tag{3.5}$$

where u_t represents the expected price change minus the intercept (the intercept was found to be zero in their study as it was in mine). Table I contains the results from the effective spread estimation. The results of each model imply that the 15% decrease in the effective spread was due to a reduction in the fixed cost component of the spread.

Measuring Price Impact from an Order Flow Shock

The Structural Model for Quote Revisions and Signed Order Flow

In addition to testing for changes in static measures of liquidity around the time of addition, I also test for changes in the stock's ability to absorb order flow shocks, and hence provide additional evidence of an overall improvement to liquidity. I implement a variant of the Hasbrouck (1991) bi-variate structural VAR(p) model for quote revisions and signed order flow. I wish to examine the rate and magnitude of the quote revision in response to a trade *innovation* i.e. the portion of a trade shock which arises from private information in the context of a vector autoregression. However, I also want to take into consideration any threshold effects that may arise as signed order flow exceeds critical levels. In this sense, I try to capture any nonlinearities in the VAR through a piecewise linear approximation.

Assuming the stock's mid-point is an unbiased proxy for true price(3.6), we find the logged quote revision to be the difference between the present and past logged consensus prices (3.7).

$$As\, s \longrightarrow \tau, E[(q_s^b + q_s^a)/2 - P_\tau|\Phi_t] \longrightarrow 0 \tag{3.6}$$

$$r_t = ln\left(\frac{\frac{q_t^b+q_t^a}{2}}{\frac{q_{t-1}^b+q_{t-1}^a}{2}}\right) \tag{3.7}$$

From (3.7), q_t^b and q_t^a are the bid and ask quotes at time t and the fractions denote the current and one period lagged midpoint, respectively. Logged returns are used instead of dollar revisions to make comparisons across stocks easier. I utilize a generalized variant of Hasbrouck's bi-variate structural VAR(p) model of trade flow and quote revisions to measure the speed and magnitude of cumulative quote revisions from trade innovations of one and five standard deviations.[2]

In the original VAR specification, one models the revision process, r_t, as a linear function of past revisions and of past *and* present order flow. The structure in the VAR comes from this equation as the model assumes that the specialist knows all lagged price changes as well as the contemporaneous and past trade flow available to her at time t. In this sense we assume that r_t contains all of the publicly available information at time t, and that the specialist acts primarily on this information set. Any disturbances to this process should arrive randomly, thus the error vector for this equation should be white noise. One can then select the lag length of the VAR which coincides with a sufficiently white residual stream. The signed order flow equation, x_t, will contain private information if any exists in the system.

$$r_t = \sum_{i=1}^{p} a_i r_{t-i} + \sum_{i=0}^{p} b_i x_{t-i} + v_{1,t} \tag{3.8}$$

$$x_t = \sum_{i=1}^{p} c_i r_{t-i} + \sum_{i=1}^{p} d_i x_{t-i} + v_{2,t} \tag{3.9}$$

This system (3.8,3.9) has a vector moving average representation if the

[2]I will determine the lag length, p, of the bivariate VAR by successive iteration of the quote revision process. Starting at ten lags, one picks the lag, p, with the highest AIC from the set of lagged equations that have a Q-statistic (or Portmanteau test) higher than 98.5%.

two series are covariance stationary. Assuming this to be the case, by the Wold representation theorem, the two series are invertible and we can write infinite order moving average expressions that relate the errors in each model to quote revisions and signed order flow.

$$r_t = \sum_{i=1}^{\infty} \gamma_i v_{1,t-i} + \sum_{i=0}^{\infty} \delta_i v_{2,t-i} + \eta_{1,t} \tag{3.10}$$

$$x_t = \sum_{i=1}^{\infty} \eta_i v_{1,t-i} + \sum_{i=1}^{\infty} \theta_i v_{2,t-i} + \eta_{2,t} \tag{3.11}$$

In many instances, studying the impulse response functions that arise from such a VMA are instructive, but several problems may arise. First, one might expect asymmetric quote responses from buy and sell shocks. Hence a positive shock of one standard deviation could illicit different responses than a negative shock and the impact of a two or a five standard deviation shock should not be constrained to a two or five-fold multiple of the unit shock. Nevertheless, the scale invariance of the VMA does not permit such flexibility. As a result, it would not be possible to examine threshold effects in the order flow using the standard VMA/IRF technique.

Self-Exciting Thresholds If we can think about the bi-variate structural VAR(p) expressed in vector form as $Y_t = (r_t, x_t)^T$, then the VAR for

(3.8) and (3.9) can be expressed as

$$Y_t = \sum_{i=1}^{p} \boldsymbol{\alpha_i} Y_{t-i} + \boldsymbol{\epsilon_t} \tag{3.12}$$

where $\boldsymbol{\alpha_i}$ is the 2×p coefficient matrix. If no adverse selection were present in the order flow and no liquidity constraint applied, then a compact representation for the linear relation would be

$$Y_t = \vec{\alpha_1} X_{t-1} + \epsilon_{1,t} \tag{3.13}$$

where $X_{t-1} = (Y_{t-1} \ldots Y_{t-p})$, and $\vec{\alpha_1}$ is the 2×p vector of coefficients.

However, assuming that the size of an order proxies for the amount of adverse selection in a trade or that liquidity constraints from finite depth bind, the system should exhibit thresholds with respect to signed order flow. If thresholds exist, then the specification would be a function of the threshold value(s) (a nuisance parameter). The new m-threshold VAR (which I term T_m-VAR(p)) would cause (4.4) to behave as

$$Y_t = \vec{\alpha}_1 X_{t-1} I_{1,t}(\vec{\gamma}, d) + \ldots + \vec{\alpha}_m X_{t-1} I_{m,t}(\vec{\gamma}, d) + \vec{\epsilon}_{1,t} \tag{3.14}$$

where $X_{t-1} = (Y_{t-1} \ldots Y_{t-p})$, $\vec{\gamma} = (\gamma_1, \ldots, \gamma_m)$ which is the vector of threshold values, and $I_{j,t}(\vec{\gamma}, d) = I(\gamma_{j-1} < Y_{t-j} \leq \gamma_j)$. Hence, this is a model with m thresholds, or m-1 regimes, and a delay lag of d trades. It is assumed that d is zero or the threshold variable is the contemporaneous signed trade variable. Given $d = 0$, the indicator $I_{j,t}(\vec{\gamma}, d)$ will drop the d dependence and shall be expressed as a function of the threshold values, $I_{j,t}(\vec{\gamma})$. I implement Hansen's[17] method for testing whether there are one, two or three regimes to determine the proper specification of the T_m-VAR(p).

Testing a null of one regime versus an alternative of two (or zero thresholds versus one) would compare

$$Y_t = \sum_{i=1}^{p} \boldsymbol{\alpha_i} Y_{t-i} + \boldsymbol{\epsilon_t} \tag{3.15}$$

to

$$Y_t = \vec{\alpha} X_{t-1}(\gamma_1) + \epsilon_{1,t} \tag{3.16}$$

where

$$X_{t-}(\gamma_1) = \begin{pmatrix} X_{t-1} I_{1,t}(\gamma_1) \\ X_{t-1} I_{2,t}(\gamma_1) \end{pmatrix} \tag{3.17}$$

where γ_1 is the value of the first threshold, $\vec{\alpha} = (\vec{\alpha}_1, \vec{\alpha}_2)$, $I_{1,t}(\gamma_1) = I(x_t \leq \gamma_1)$, and $I_{2,t}(\gamma_1) = I(x_t > \gamma_1)$, where x_t is the contemporaneous order flow.

Both regression and nuisance parameters can be consistently estimated with OLS. In the one regime or linear model (4.24), OLS determines the $\boldsymbol{\alpha}_i$ that minimizes the sum of squared errors. Analogously, in the two regime case (4.25), one must pick the γ_1 and the $\vec{\alpha}$ such that the squared errors are minimized.

$$S_2 = \min_{(\gamma_1,\vec{\alpha})} (Y - \vec{\alpha}X(\gamma_1))'(Y - \vec{\alpha}X(\gamma_1)) \tag{3.18}$$

where S_2 is the sum of squared errors that is minimized over a grid search of the data.

To find the threshold value and coefficients that create S_2, first rank the data by the threshold variable and then divided it into deciles. A minimum grid size of a decile of the data was selected to insure that a reasonable fraction of the total number of observations for each stock would be available to form the alternative to the null. Once the γ_1 and $\vec{\alpha}$ are selected that minimize S_2, a further dissection of the data is attempted. One collects the γ_1 which minimizes S_2 and the γ with the second lowest sum of squared errors and divides the data between the two threshold values inclusive into another 10 deciles. So if the γ_1 is 2000 shares and the threshold value with the next lowest sum of squared errors is 1000, then a new set of regressions is performed on the data with 2000, 1900, etc. as the new breakpoints. Of course, if the difference between the threshold values only cover 400 shares, then only five points can be tested and hence a courser grid must be used.

The test statistic to test the null of zero thresholds (one regime) against the alternate of one threshold (two regimes) is

$$F_{0,1} = T\left(\frac{S_1 - S_2}{S_2}\right) \tag{3.19}$$

where T is the total number of observations, S_2 is defined in (4.27) and S_1 is the sum of squared errors of the linear model. The asymptotic distribution of $F_{0,1}$ is non standard because of the presence of the threshold variable. Hence, the statistic would not be distributed χ^2, and therefore one needs a method to approximate the asymptotic distribution for $F_{0,1}$.

Hansen (1996) provides an excellent overview of two approximation methods. The first method uses asymptotic distribution theory pertinent for random functions (so called empirical process theory). Because $F_{0,1}(\gamma)$ is a random function of γ, the $F_{0,1}$ test statistic is a random maximum of $F_{0,1}(\gamma)$ over the nuisance parameter space,

$$F_{0,1} = \max_{\gamma \in \Gamma} F_{0,1}(\gamma) \tag{3.20}$$

In order to ascertain significance levels of the $F_{0,1}$ statistic, one must calculate the asymptotic distribution of the empirical process $F_{0,1}(\gamma)$. Hansen calls this random limit function $T(\gamma)$. Hansen (1996) and (1999) cover the algorithm to calculate the approximation to the asymptotic distribution, T_n, using sample moments. The significance of $F_{0,1}$ is gauged by how many points of T_n exceed the observed value of $F_{0,1}$. Computationally, this is a low cost method.

The second method for determining the significance of $F_{0,1}$ centers on using a bootstrap approximation. Although the cost of using the bootstrap is high in that it requires placing structure on the residuals i.e. that the residuals are independent over time, as Hansen (1999) points out, several studies conclude that the bootstrap outperforms first order asymptotic theory as an approximation to finite sample distributions. Accordingly, I opt for the bootstrap approximation in this study.

In bootstrapping the significance of the observed $F_{0,1}$ or $F_{1,2}$ see below (4.32), one must determine whether heteroscedasticity is present in the data. This may involve having homoscedastic data unconditionally, regime specific homoscedasticity, or heteroscedasticity in each regime. I tested each firm for heteroscedasticity using White's test and then aggregated the results. The results were not significant over the entire data set in testing $F_{0,1}$ nor was it apparent in the regimes in testing $F_{1,2}$, therefore no heteroscedasticity adjustment was made in the bootstrap.

The steps for calculating the bootstrap approximation under the homoscedasticity assumption are as follows for $F_{0,1}$: once the lag length (p) of the VAR has been determined, estimate the coefficients for the model using OLS under the null of linearity and collect the residuals. Now take a data series of length p and sample with replacement from the residual vector to create generate a new sample of size n (n=1000 in this study). Calculate the value of $F_{0,1}$ from these data and then repeat 1000 times. Calculate the percentage of times that the simulated values of $F_{0,1}$ exceeds the the observed $F_{0,1}$. This is the bootstrapped p-value. Aggregate p-values over all of the sample firms according the Gibbons and Shanken algorithm (1987) to arrive at a sample significance level.

Testing for three regimes versus two follows a similar procedure where the (4.26) becomes

$$X_{t-1}(\vec{\gamma}, d) = \begin{pmatrix} X_{t-1} I_{1,t}(\vec{\gamma}, d) \\ X_{t-1} I_{2,t}(\vec{\gamma}, d) \\ X_{t-1} I_{3,t}(\vec{\gamma}, d) \end{pmatrix} \tag{3.21}$$

with $\vec{\gamma} = (\gamma_1, \gamma_2)$, $\vec{\alpha} = (\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3)$, $I_{1,t}(\gamma, d) = I(x_t \leq \gamma_1)$, $I_{2,t}(\gamma, d) = I(\gamma_2 \geq x_t \geq \gamma_1)$, and $I_{3,t}(\gamma_1, d) = I(x_t > \gamma_2)$. The least squares minimization is much akin to (4.27) only now the regression coefficients are subject to two threshold values.

$$S_3 = \min_{(\gamma_1, \gamma_2, \vec{\alpha})} (Y - \vec{\alpha}X(\gamma,d))'(Y - \vec{\alpha}X(\gamma,d)) \tag{3.22}$$

where S_3 is the minimized sum of squared errors over the data divided into

3 regimes, γ_1 is the first signed order flow threshold and γ_2 is the second.

Once (4.31) is performed, then the relevant test statistic for two regimes versus three is

$$F_{1,2} = T\left(\frac{S_2 - S_3}{S_3}\right) \tag{3.23}$$

where $F_{1,2}$ is the test statistic for the test of one versus two thresholds. The recipe for calculating the bootstrap approximation under the regime specific homoscedasticity assumption is as follows for $F_{1,2}$: estimate the coefficients for the model using OLS under the null of two regimes (one threshold) and collect the residuals from each regime. Now, take a data series of length p and sample with replacement from the residual vector germane to the size of the contemporaneous order flow (i.e. if the trade size meets or exceeds the threshold value, use a residual corresponding to that regime). Generate a new sample of size n (n=1000). Calculate the value of $F_{1,2}$ from these data and then repeat 1000 times. Calculate the percentage of times that the simulated values of $F_{1,2}$ exceeds the the observed $F_{1,2}$. This is the bootstrapped p-value. Aggregate p-values as before and arrive at a sample significance level.

Any further iteration to a higher dimensional threshold vector would result in a highly problematic test statistic. This statistic would be a function of nuisance parameter which is a function of another nuisance parameter which is itself a function of the third nuisance parameter. Confidence in this degree of nested estimation is not supported by the change point or SETAR literature.

In order to carry out all of the regressions in a timely manner (to avoid the N^2 regressions for each of the 110 stocks), Bai and Perron[18] show that because the least squares estimates are consistent for $\vec{\gamma}$, a two stage iterative minimization may be utilized to estimate the threshold vector. So if $\vec{\gamma} = (\gamma_1, \gamma_2)$ then the estimate of the first threshold will be a consistent estimate of either γ_1 or γ_2 via the above minimization in (4.27). The data are then divided by $\hat{\gamma}_1$ another grid search takes place to calculate the value of $\hat{\gamma}_2$. This second threshold estimate will be a consistent estimate of the remaining threshold value in the pair (γ_1, γ_2).

Once the values for $\hat{\gamma}_1$ and $\hat{\gamma}_2$ have been determined, one should perform further iterations on the data to assure consistency of the threshold estimates. For example, the algorithm for the iterations for a stock with trade deciles at 100, 200,. . ., 1000 (only buys in the data set) would be to estimate $\hat{\gamma}_1$ and $\hat{\gamma}_2$. Suppose that these values were 200 and 800 respectively. Next, to iterate on the first threshold, take the data from 100 to 800 and rerun the estimation for $\hat{\gamma}_1$ and then take all of the data that have trades from that value (the refined guess for γ_1) and above and rerun the estimation of $\hat{\gamma}_2$. Then do this iteration one more time over both $\hat{\gamma}_1$ and $\hat{\gamma}_2$. Third round estimates (two iterations) were used in this study.

Generalized Impulse Response Functions If the $F_{0,1}$ or $F_{1,2}$ values are significant for the sample, then (3.8) becomes

$$r_t = \sum_{i=1}^{p} \vec{a}_i r_{t-i} + \sum_{i=0}^{p} \vec{b}_i x_{t-i} + \varsigma_{1,t} \tag{3.24}$$

Thus, we will abandon the typical IRF/VMA in favor of a method that can accommodate the thresholds in the signed order flow in measuring the magnitude of the cumulative quote revision resulting from a shock from the order flow equation.

Accordingly, I opt to use a generalized impulse response function to measure the magnitude of quote revisions as a proxy for the trade informativeness of buy and sell trades before and after the index addition. A generalized impulse response function (GIRF) can be viewed as the answer to the conceptual question of what happens to the time profile of the effect of a shock on a series over and above what would have happened to the series without the shock. The GIRF, like the IRF, is a function of the length of the series, the magnitude of the shock, but also of the history of the system. It must contain an accurate benchmark with which one can compare the shocked series.

$$GIRF_{Rev}(n, v_t, \omega_{t-1}) = E[Rev_{t+n}|v_t, \omega_{t-1}] - E[Rev_{t+n}|\omega_{t-1}] \tag{3.25}$$

where n is the number of trades after the shock, v_t is an arbitrary shock to the quote revision equation (3.8), and ω_{t-1} is the particular history examined before the shock arrives at time t. The choice of n ranges from 0 to 20 throughout the paper although all of the figures are truncated to 15 lags for aesthetic purposes.

A shock of an arbitrary magnitude would not give a meaningful measure to the question at hand. I standardize this input to be a one standard deviation shock to the signed order flow equation in the *pre* addition data, thus facilitating comparisons of responses across time as well as between buys and sells. Hence (3.25) now becomes

$$GIRF_{Rev}(n, \delta, \omega_{t-1}) = E[Rev_{t+n} \,|\, \delta,\, \omega_{t-1}] - E[Rev_{t+n} \,|\, \omega_{t-1}] \tag{3.26}$$

One runs simulations (described in the Appendix) using orthogonalized errors to create averages over the different 'histories' possible at time t-1. In this sense, mean responses in the $E[Rev_{t+n} \,|\, \delta,\, \omega_{t-1}]$ term and benchmark term $E[Rev_{t+n} \,|\, \omega_{t-1}]$ are calculated. The average of this difference is still a function of time t-1 (as it is a function of the same quote and flow data), so the GI_{Rev} now can be thought of as a variable approximating the expectation over states $\omega_{t-1} \in \Omega_{t-1}$.

$$GIRF_{Rev}(n, \delta, \Omega_{t-1}) \longrightarrow \int_{\omega_{t-1}} GIRF_{Rev}(n, \delta, u_{t-1}) du \tag{3.27}$$

Running simulations over all histories and repeating that 20 times provides an average that approximates the mean response over all histories.

$$GIRF_{Rev}(n, \delta) \longrightarrow \int_{\Omega_t} GIRF_{Rev}(n, \delta, s) ds \tag{3.28}$$

which yields an approximation to the conditional expectation of interest.

$$GIRF_{Rev}(n, \delta) \; = \; E[Rev_{t+n} \,|\, \delta\,] - E[Rev_{t+n}] \tag{3.29}$$

Figure 1 shows the sample's GIRF pre and post addition. The asymmetry of buy and sell shocks on prices and the reduction of price impacts post addition (liquidity gains) is predicted by theory. The GIRF post addition is much smaller than that of the GIRF beforehand. To test whether the cumulative quote revision (height of the GIRF sans the original shock) is statistically different, I use a simple t test with unequal variances. With 20 observations of the mean response for each firm before and after, the Central Limit Theorem should allow inferences that we can aggregate over the sample. The results are in Table II. The rest of the figures plot the cumulative quote revision to two sizes of shocks (one and five standard deviation shocks) from the trade flow equation on the revision equation. The data are examined as a whole, divided up by market value and by spread (each measured pre addition). The data are also examined by the aforementioned criteria while noting the differences arising between pre and post addition data.

Results

The spread decomposition investigation using the modified GH and the MRR models yields interesting results in Table I. Although the GH model's effective spread estimate is larger than the sample average effective spread, and the MRR model's is smaller, each model did estimate an accurate two cent decrease in spread. Panels B and C report that this translates to an average 15% permanent decrease in the estimated effective spread. These results agree with the recent work on the liquidity issues surrounding S&P additions by Hedge and McDermott.

Panels B and C also show that the GH and MRR models each indicate that the cost savings comes strictly from the fixed cost spread component. Addition into the S&P 500 represents a permanent regime change, at least for stocks traded on the NYSE. Finding that the cost decrease comes from the fixed cost component agrees with economic theory two fold. First, as the number of trades permanently increases after the addition, fixed costs are covered earlier in the day and savings can be passed on to the trader. Second, the stocks picked for inclusion are medium to large, monitored stocks (approximately 95% of which were members of the MidCap 400). The adverse selection component of the spread of such stocks would be small to begin with, leaving little possible room for improvement.

Table II reports the findings from the dynamical liquidity analysis. Order flow shocks decrease over time. Panel A of Table II reports an aggregate p-value of 0.0000005. Therefore any reduction in the price impacts of order flow shock comes from smaller shocks to the revision equation as well as less private information in a trade innovation. In Panel B, the overall volatility of the quote revision process(the model's proxy for public information volatility) wanes after addition. Table IV shows strong rejection of the one and the two regime null hypotheses. 92% of the firms with a negative value for their first threshold (45% of the firms in the sample) had a positive value for their second. 93% of the firms with a positive first threshold value had a negative second value. This symmetry is consistent with theory as large orders in either direction would signal a larger amount of adverse selection and inspire a risk neutral specialist to enact larger than average quote revisions.

Table III reports the findings for the generalized impulse response analysis. Panel A shows the results from the unpaired t-tests run on pre and post buy side quote revisions. The low p-value corresponds to rejecting that the pre and post addition price impacts come from distributions with the same means. The reduction in quote revisions to buy and sell shocks are approximately equal post addition (at 25%). The sell side revisions are larger in magnitude and the sell side reductions are slightly larger at 27% than those of the buys at 25%. Panel B shows the analogous results from the larger five standard deviation shock. The results are qualitatively similar to those of the smaller shock.

The figures relay an interesting story that for the most part fall in line with theory. Figure 1 shows cumulative quote revisions by shock size pre and post addition. Quote revisions from the one standard deviation shocks (one sigma) to the buy side were slightly smaller but the price impact reductions were larger on the sell side. In the bottom graph (five sigma shocks), the pre and post buy side shocks were about equal, whereas the sell side shocks were much lower post addition. The buy side shocks pre and post were also smaller than their sell side counterparts. This figure is exemplary in that features key economic factors that the conventional linear model (IRF/VMA) representation. First, buy and sell side shocks have different permanent price impacts. Second, a five fold shock does not have a five fold effect on price impacts. Hence, the data appear to have many non linearities captured by this type of piece wise linear representation.

Figure 2 shows the same breakdown as Figure 1 but here only the bottom half of the sample as divided by effective spread (high spread firms) is examined. The buy and sell price impacts are smaller for the one sigma shock (top graph), and roughly symmetrical, but with the greater liquidity gain (lower price impacts) on the sell side. The graph for the five sigma shock looks very similar to the overall graph for the sample and has most of the reductions in price impacts going to the sell side with little improvement on the buy side. Figure 3 contains the results from the low spread firms. The gains for this sample were much less pronounced. The quote revisions to the one sigma shock pre and post looked similar with small reductions on both sides while those to the larger shock were larger to each side with the approximately the same shape as the one sigma cumulative quote revisions.

Figures 4 and 5 show the results of the firms in the top and bottom effective spread quartiles as calculated before inclusion. The firms with the smallest spreads had the greatest gains which seems counterintuitive given that the greater marginal gain should come from those firms with the larger initial spread. The price impacts in each figure appear roughly symmetrical pre and post addition but with the responses to the five sigma shock all appearing less than a five fold version of the smaller shock.

Figures 6 and 7 feature the results of the high and low market value halves of the sample pre and post addition (market values were calculated before addition). Although one normally associates low spreads with high market value, some differences arise in the two sets of samples. The high market value firms have most of their price impact reductions on the sell side for both the large and small shocks pre and post. The low market value firms have larger quote revisions pre and post. The impacts to the smaller shock have a symmetrical effect pre and post with the post addition impacts strictly smaller on both sides. For the larger shock, however, the pre and post shocks appear about equal with little liquidity gain arising post addition.

Figures 8 and 9 contain the results of the top versus bottom market value quartiles, respectively. As predicted, the high market value firms show smaller price impacts on both sides to both shocks than those of the smaller firms. Figure 8 shows an odd revision behavior. The sell side has a much *lower* price impact to both the small and large price shocks. The smaller firms have symmetrical patterns on both sides pre and post addition. In each case, for both sets of firms, the post addition impacts are smaller, with the sell side having the greater reduction. The high market value firms have a larger buy side impact, the smaller firms have a larger sell side impact (for each size shock).

Figures 10 and 11 contain the pre and post addition impacts of the high versus low market value firms. The pre addition impact patterns look similar for the two sets of firms. The high market value firms have smaller price impacts for each size shock, but the magnitude of the impact to the larger sell side shock appears equal for each set. The post addition data tell a different story. The post addition data have the high market value firms having the same size cumulative revision to the *smaller* buy side shock, but strictly smaller impacts to the larger shocks.

Figures 12 and 13 contain the results of comparing the top versus the bottom market value quartiles pre and post addition, respectively. Pre addition the price impacts are noticeably larger for the small and large firms. For the high value firms, a strange pattern appears as the impacts to sell side shocks are much less than those to the buy side. The same behavior is present for the impacts to small and large shocks. In the post addition data, a more symmetrical behavior is present for both classes of firms. For both sets of firms and for both sets of shocks, the sell side impacts were larger than those on the buy side and the high value firms had smaller impacts than those of the smaller firms.

Figures 14 and 15 show results from small and large spread firms pre and post addition, respectively. Similar to the pattern of based on market value, the small spread firms had smaller price impacts pre addition with the exception of the buy side impacts to one sigma shocks - the high spread firms had slightly smaller impacts. Both sets of patterns were symmetrical with the sell side impacts larger than those on the buy side. Post addition the differences in impacts were much smaller. The impacts to the large and small shocks were almost identical for both sets of firms post addition.

Figures 16 and 17 contain the results from comparing the top and bottom spread quartiles in the pre and post addition data. Pre addition, both high and low spread firms had fairly symmetrical price impacts. The sell side responding to the five sigma shock for the low spread firms was much smaller than the buy side, a result at variance with theory. Post addition, both sets of firms had symmetrical impacts with the low spread firms having very small impacts to the one sigma shock. The impacts to the five sigma shock very highly symmetric but not five fold larger than the those of the one sigma shock.

Conclusion

Analysis of the spread revealed that the quoted and relative spread fell by an average of 15% post addition. The reduction in trading cost, as measured by the effective spread, arises from a reduced fixed cost component of the spread. This reduction presumably arises from the presence of institutional traders greatly increasing the number of trades each day, enabling specialists to reduce the per share cost of transacting.

The dynamic facet of liquidity showed a significant increase after addition to the index. The analysis concludes that trade shocks cause a 25% smaller amount of quote revision after addition because of a reduced amount of private information in the order flow. This aligns itself well with the conclusion from the static analysis. The inference on the reduction in the amount of private information post addition yields new information about the amount of resilience in the market for NYSE S&P 500 stocks. Including information beyond the spread gives a more complete picture of the liquidity enhancement resulting from addition to the flagship S&P index.

Table I

Effective Spread Estimation of NYSE Traded Firms

The data for Table I were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. Panel A shows the average quoted or 'inside' spread and the average effective spread (the absolute value of the difference between the price and quote midpoint) for the set of 110 firms before and after addition to the S&P 500. The p-value arises from a difference of means t-test between the pre and post averages (with unequal variances) of each firm which is then aggregated over all firms. All numbers are expressed in dollar terms. Panel B reports the results from the Glosten and Harris ('88) spread decomposition model. The model breaks down the effective spread into fixed and variable cost components. The variable cost component is a linear function of signed order flow and is considered a proxy for the adverse selection component of the spread. The effective spread estimate was calculated for each firm by adding the adverse selection component per share*average number of shares per trade to the fixed cost component. Panel C contains the estimates from the Madhavan, Richardson, Roomans' ('97) spread decomposition model. The various components of the spread were calculated in a similar fashion to those of the Glosten and Harris model. All components in each model have aggregate p-values significant at or below the 5% level.

Panel A: Empirical Analysis of Spreads

	Pre	Post	p-value
Effective Spread	0.133	0.117	0.104
Quoted Spread	0.190	0.167	0.013

Panel B: Glosten Harris Decomposition

	Fixed Cost*	Adverse Selection*	Effective Spread Estimate
Before	0.112	0.00204	0.115
After	0.093	0.00147	0.094

Panel C: Madhavan, Richardson, Roomans Decomposition

	Fixed Cost*	Adverse Selection*	Effective Spread Estimate
Before	0.076	-0.00042	0.151
After	0.064	0.00121	0.130

* p value < 0.05

Table II

Volatility Analysis

The data for Table II were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. Panel A contains the results from an F test of the ratio of the standard deviations from the signed order flow equation from Hasbrouck's (1991) structural bi-variate VAR model. The ratio is composed of the pre addition signed order flow volatility over the post addition volatility for each of the 110 firms in the sample. The p-values from each F test are then aggregated by the Gibbons Shanken ('87) algorithm to an aggregated p value shown below. The transformation involves taking the natural log of each p-value and multiplying it by –2. Each summed variable follows a chi-square distribution with twice the degrees of freedom as the number of firms. The results show a statistically significant reduction between pre and post addition order flow volatilities. No adjustment was made for possible dependence across firms. Panel B contains the aggregate p-value for the test of whether the total public information volatility (as proxied by the volatility in the logged quote revision equation) has changed pre to post addition. The same methodology was applied here as in Panel A. The low p-value indicates logged quote revision volatility decreases post addition.

Panel A: F test on Order Flow Shock Pre and Post Addition

Aggregate p-value: 0.00000005

Panel B: F test on Quote Revision Volatility Pre and Post Addition

Aggregate p-value: 0.001

Table III

Test of Average Price Impact Pre and Post Addition

The data for Table III were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. Table III contains information derived from a generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow. The original model was modified to allow for m thresholds (or m+1 regimes) in the signed order flow. The sample value for m in this study was two. Given the specification of the two threshold VAR(p), a generalized impulse response function (GIRF) was created. Cumulative logged mid-quote revisions to a one and a five standard deviation shock were averaged from the GIRF over 1000 simulations for each history (set of *p* sequential data points) and then averaged over all histories (the entire data set) for 14 subsequent trades after the shock to create mean cumulative logged midquote revisions for each stock. 20 such averages were formed to create a distribution of means for the buys and the sells (for the one and five standard deviation shocks) pre and post addition for each of the 110 stocks in the sample. The p-value (for each shock on the buy and sell side) comes from a difference of means (the average cumulative quote revision for each stock) t-test (from pre and post addition distributions) with unequal variances. All t-statistics were over 5. Each price impact is the percentage of the midquote return to a given size shock. Logged quotes were chosen to afford better comparability across stocks. Panel A shows the cumulative quote revisions corresponding to shocking the flow equation with a one standard deviation shock to the buy side (a positive shock) and sell side (a negative shock), respectively. Panel B contains the same information for a five standard deviation shock to each side. The results indicate that statistically significant differences arise between buy and sell side shocks. Also, the five standard deviation shock results in less than a five fold effect of the one sigma shock.

Panel A: Cumulative Quote Revision (as a % of Mid-Quote) from a One Standard Deviation Shock

Buy	Pre	Post	p-value
	0.020	0.014	0

Sell	Pre	Post	p-value
	-0.026	-0.018	0

Panel B: Cumulative Quote Revision (as a % of Mid-Quote) from a Five Standard Deviation Shock

Buy	Pre	Post	p-value
	0.032	0.026	0

Sell	Pre	Post	p-value
	-0.036	-0.028	0

Table IV

Signed Order Flow Thresholds from Pre and Post S&P 500 Addition

The data for Table IV were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. Table IV contains information derived from a generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged midquote revisions and signed order flow. The original model was modified to allow for m thresholds (or m+1 regimes) in the signed order flow. The sample value for m in this study was two. Two signed order flow threshold parameters were calculated for each of the 110 NYSE firms for the pre and post addition periods. For Panel A and B, row one contains the average thresholds. For each firm, trade size averages on the same side as their first threshold parameter are reported in row two and percentages appear in row three. Row four reports that 55% of the 110 sample firms have first thresholds that were positive signed (buys) and 43% have positive second thresholds. The last row has the aggregated p-value for each threshold. Because the size of the threshold is not identified until the alternative, simulations are run to determine the significance level of each threshold. Each firm has a bootstrapped p-value arising from the non-standard F-test testing zero versus one threshold, F(0,1) and then for the non-standard F-test testing one versus two thresholds, F(1,2). The set of bootstrapped F-test p-values for each firm and each threshold were transformed to a sample aggregate p-value using the Gibbons and Shanken ('87) algorithm. The transformation involves taking the natural log of each p-value and multiplying it by -2, then summing all of the transformed p-values. Each summed variable follows a chi-square distribution with twice the degrees of freedom as the number of firms. Panel A contains the pre addition results. Both thresholds are significant and are smaller than the average trade. Panel B contains the post addition results. As with Panel A, both thresholds are significant and are smaller than the average trade. The average trade size and average first threshold both decreased after addition, though the second threshold increased in size. For both the pre and post samples, approximately 95% of the second thresholds are of the opposite sign of that of the first threshold.

Panel A: Pre-Addition

	First Threshold		Second Threshold	
AVERAGE THRESHOLD:	**BUY**	**SELL**	**BUY**	**SELL**
	1972	-1779	-1495	1400
AVERAGE TRADE SIZE:	2457	-2579		
PERCENT OF AVERAGE:	0.84	0.71		
PERCENT OF BUYS/TOTAL:	0.55		0.43	

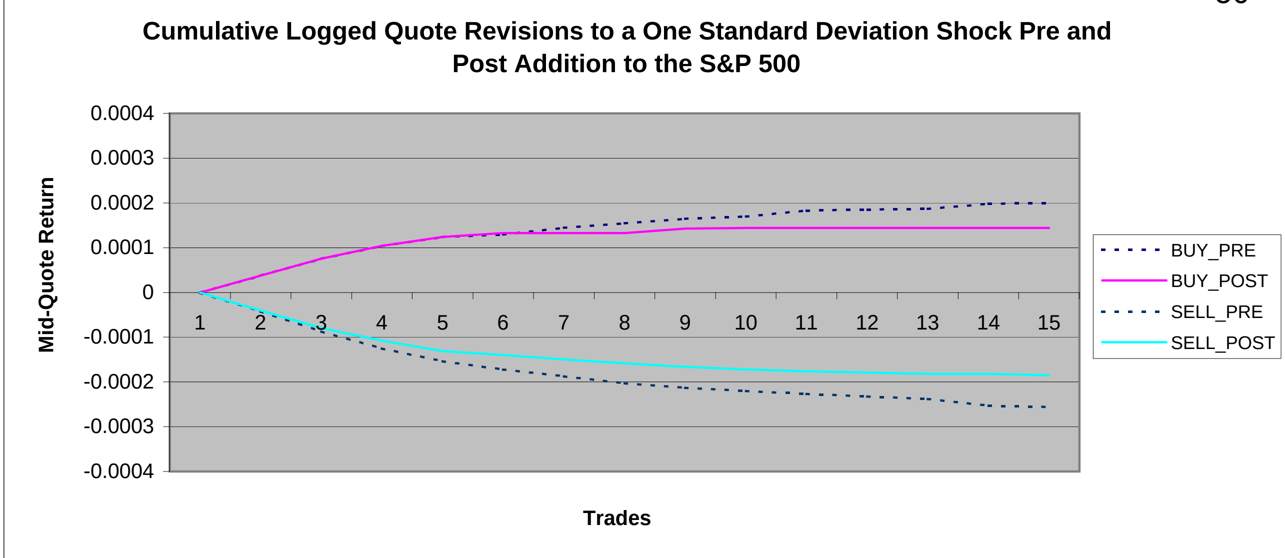
F(0,1) = 0.00 **F(1,2) =** 0.023

Panel B: Post Addition

	First Threshold		Second Threshold	
AVERAGE THRESHOLD:	**BUY** 1773	**SELL** -1491	**BUY** -1593	**SELL** 1489
AVERAGE TRADE SIZE:	2152	-2318		
PERCENT OF AVERAGE:	0.80			
PERCENT OF BUYS/TOTAL:	0.53		0.45	

F(0,1) 0.00

F(1,2) = 0.045



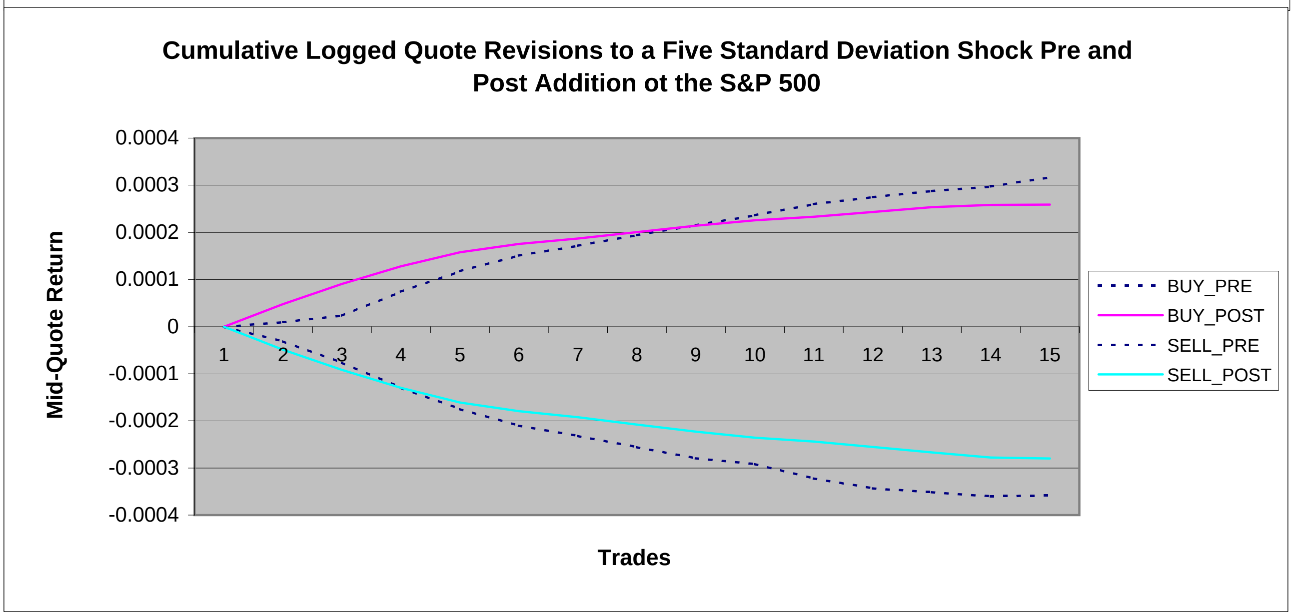
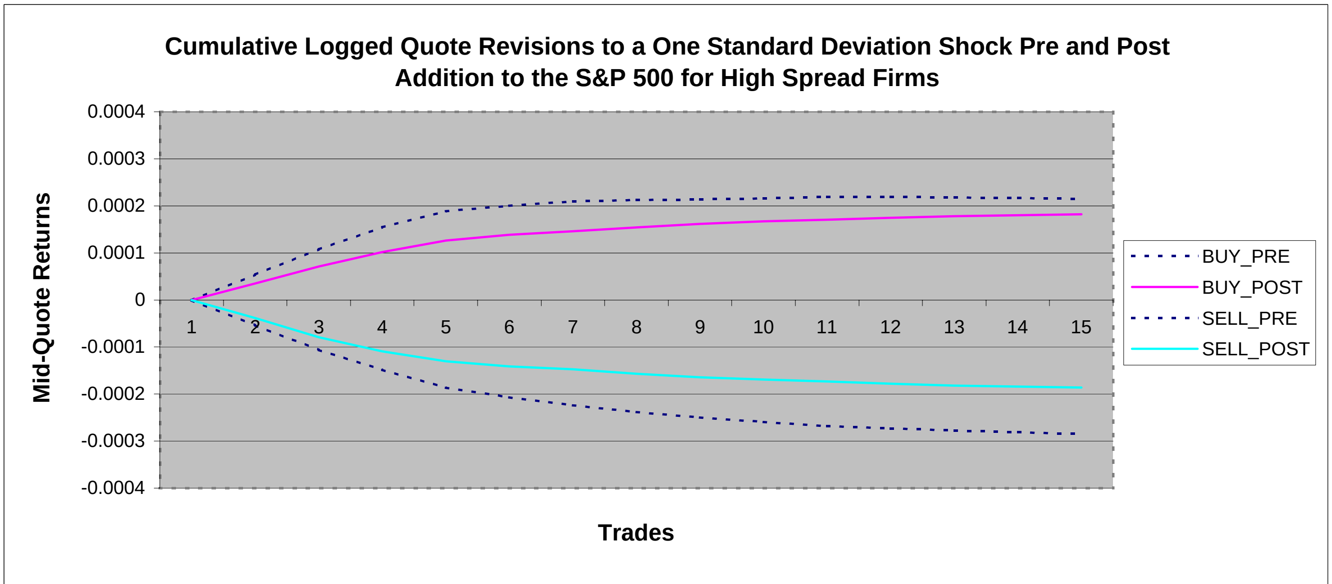


Figure 1. Sample Cumulative Generalized Impulse Response Functions (GIRFs).

Figure 1 contains results from the entire sample of firms (110) pre and post addition to the S&P 500. The data for Figure 1 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 1 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, the speed of quote revision is about equal for both the pre and post data sets, but smaller and asymmetrical quote revisions to both the buy and sell side obtain after addition. The thresholds in the data suggest that the sell side shocks have a larger effect on quote revisions. In the bottom graph, the five sigma shock on the buy side are almost identical pre and post addition while the sell side shows smaller and faster quote revisions post addition. Both graphs illustrate the effect of thresholds (non-linearities) and support the conventional wisdom that sell side shocks have a greater effect on quotes. Also, except for the buy side shock to the 5 sigma shock, the data suggest that less private information content is contained in the order flow as smaller permanent price impacts arise post addition.



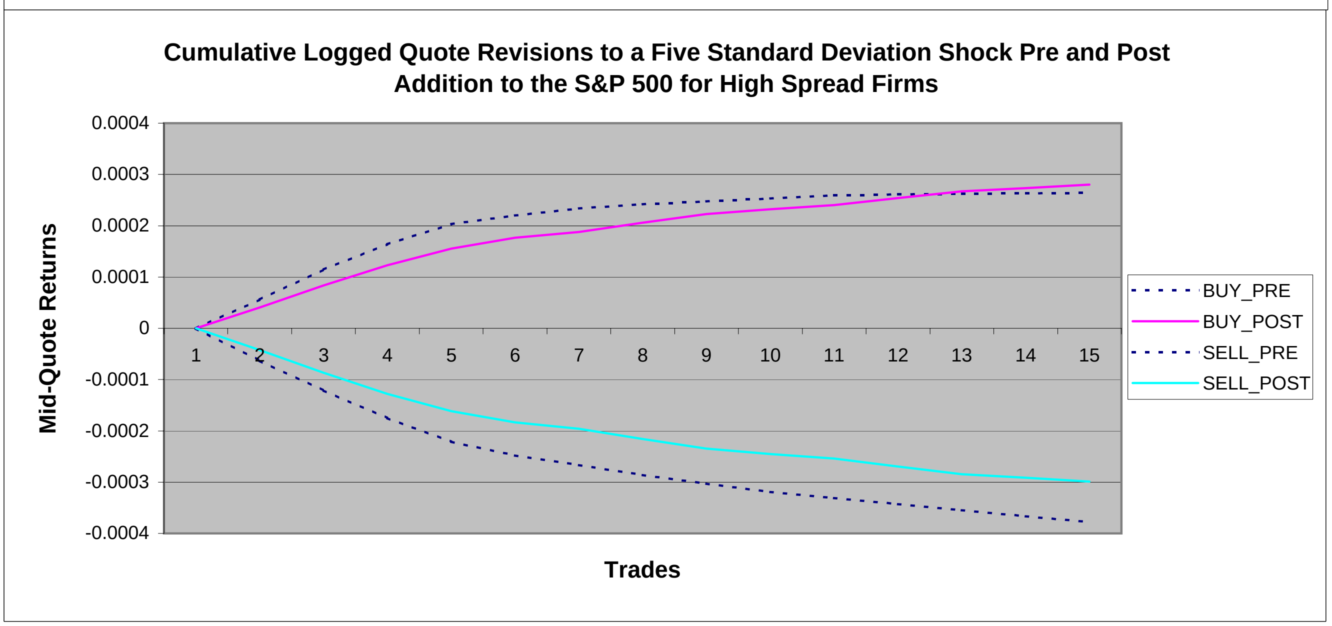
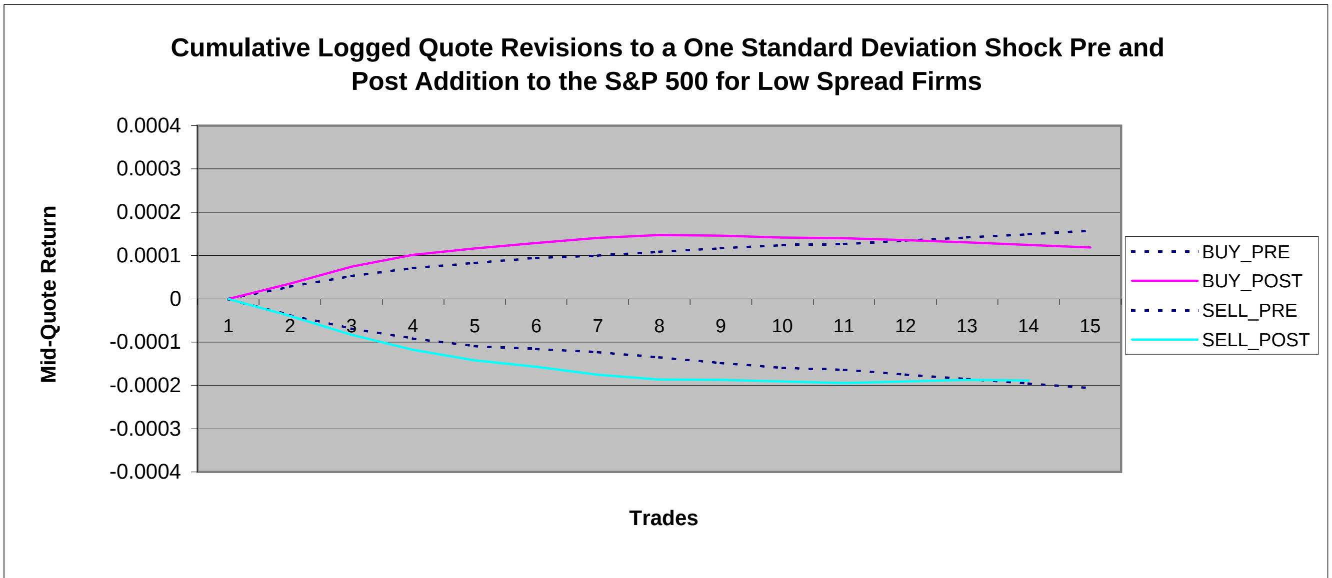


Figure 2. GIRFs of the High Spread Half of the Sample

Figure 2 contains results from the bottom half of the sample (55) with regards to the effective spread calculated pre addition to the S&P 500. The data for Figure 2 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 2 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. The top graph indicates that post addition price impacts (to a one sigma shock) are smaller with the greater decrease going to sell side shocks. While the price impacts to sell side shocks were markedly greater pre addition, the magnitudes of the buy and sell price impacts post addition are approximately equal. A different pattern emerges on the bottom graph (price impacts to a five sigma shock). Post addition price impacts on the buy side are slightly larger than the pre addition impacts. The post addition sell side impacts are each larger in size than their buy side counterparts with the post addition impact 22% smaller than its pre addition counterpart.



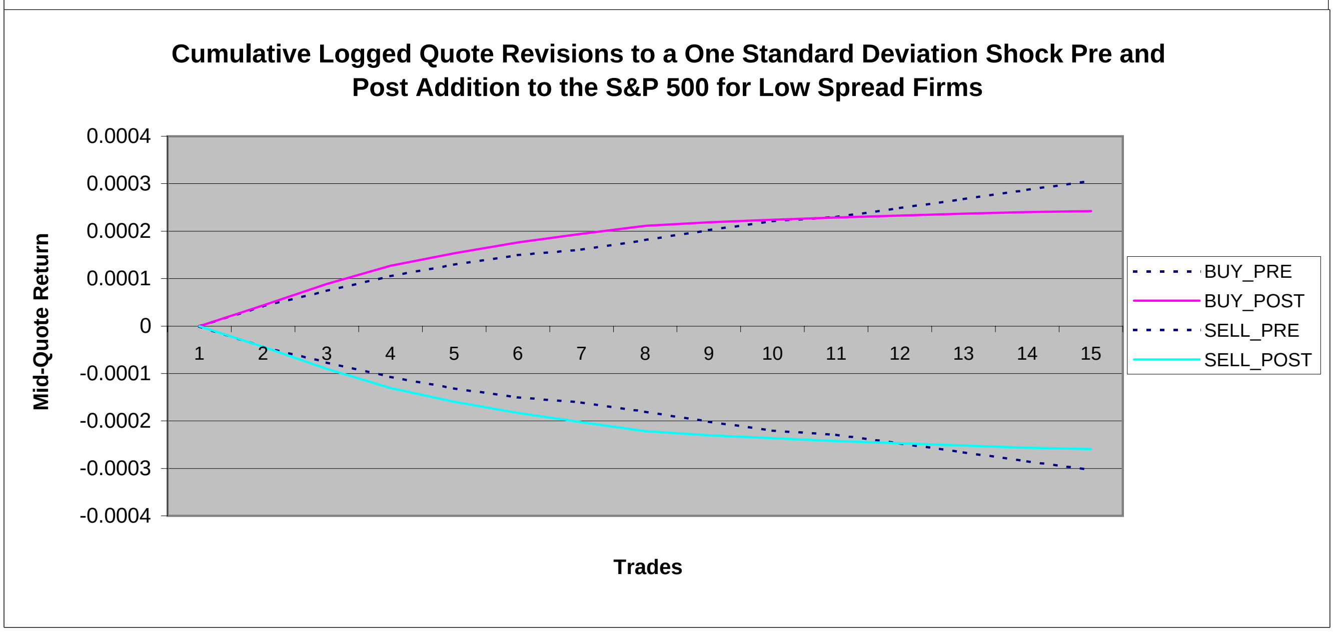
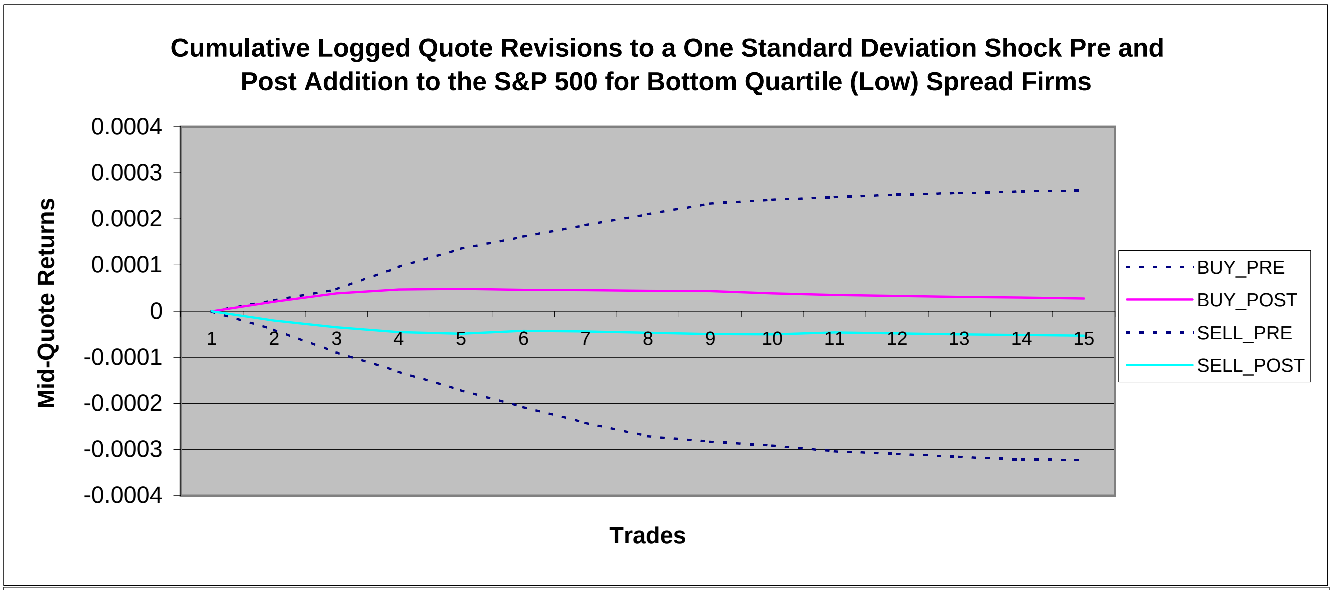


Figure 3. GIRFs of the Low Spread Half of the Sample

Figure 3 contains results from the top half of the sample (55) with regard to the effective spread calculated pre addition to the S&P 500. The data for Figure 3 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 3 contains cumulative GIRFs of a one(top) and a five(bottom) standard deviation shock from the signed order flow equation on the revision equation. The top graph indicates that post addition price impacts (to a one sigma shock) are approximately the same size as the pre addition impacts. The post addition buy side shock is slightly smaller in size and reaches its equilibrium level faster than its pre addition counterpart. Post addition price impacts to five sigma shocks on both sides are smaller and they reach equilibrium levels faster than their pre addition counterparts.



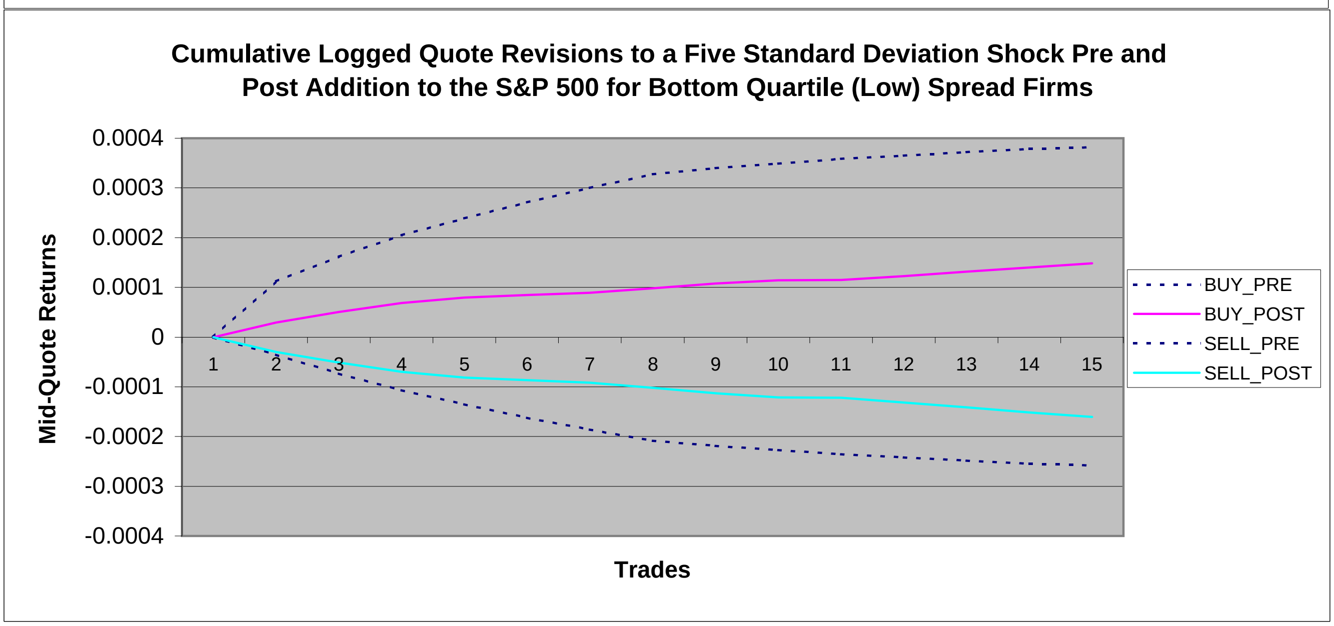
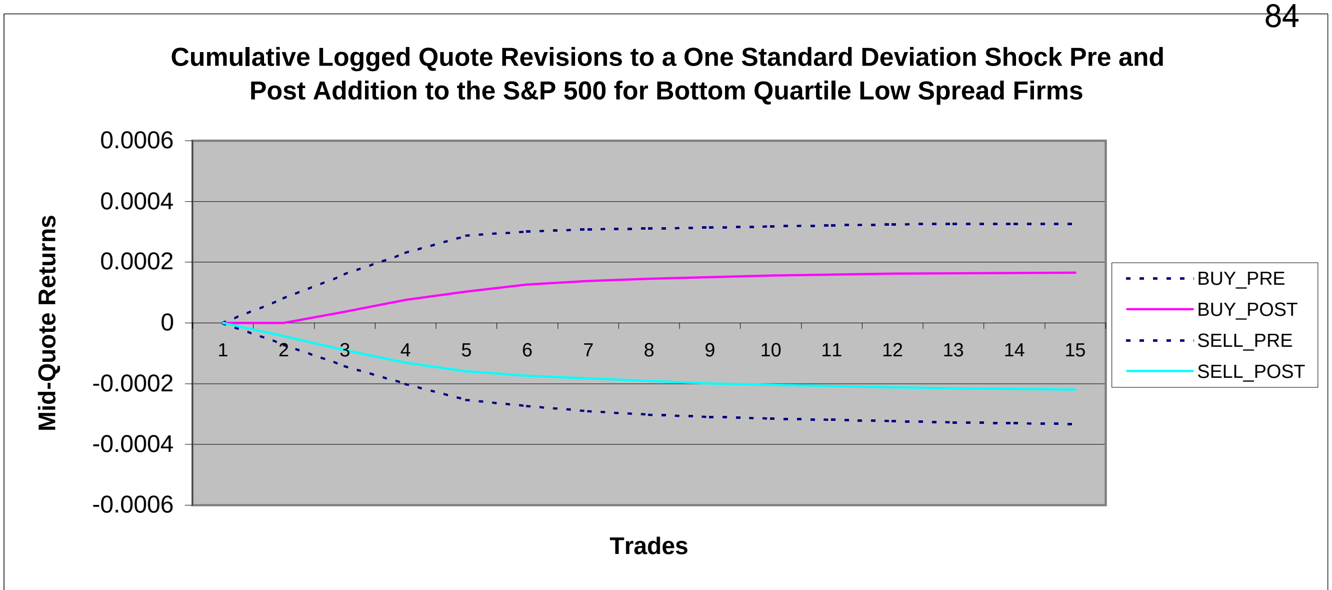


Figure 4. GIRFs of the Firms with Spreads in the Bottom Quartile of the Sample
Figure 4 contains results from bottom quartile, or smallest spread, (27) of the sample based on effective spread calculated pre addition. The data for Figure 4 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 4 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. Both graphs show evidence of great reductions in price impacts post addition. The top graph indicates that one sigma shocks post addition have about 1/6th the price impact that they did pre addition to the index. Both sides appear to have similar cumulative quote revisions pre and post addition. In each case the sell side impact is slightly larger than its buy side counterpart. In the bottom graph, the pre addition buy side impacts are surprisingly larger than their sell side counterparts. Post addition, the buy and sell side impacts are similar in shape and magnitude.



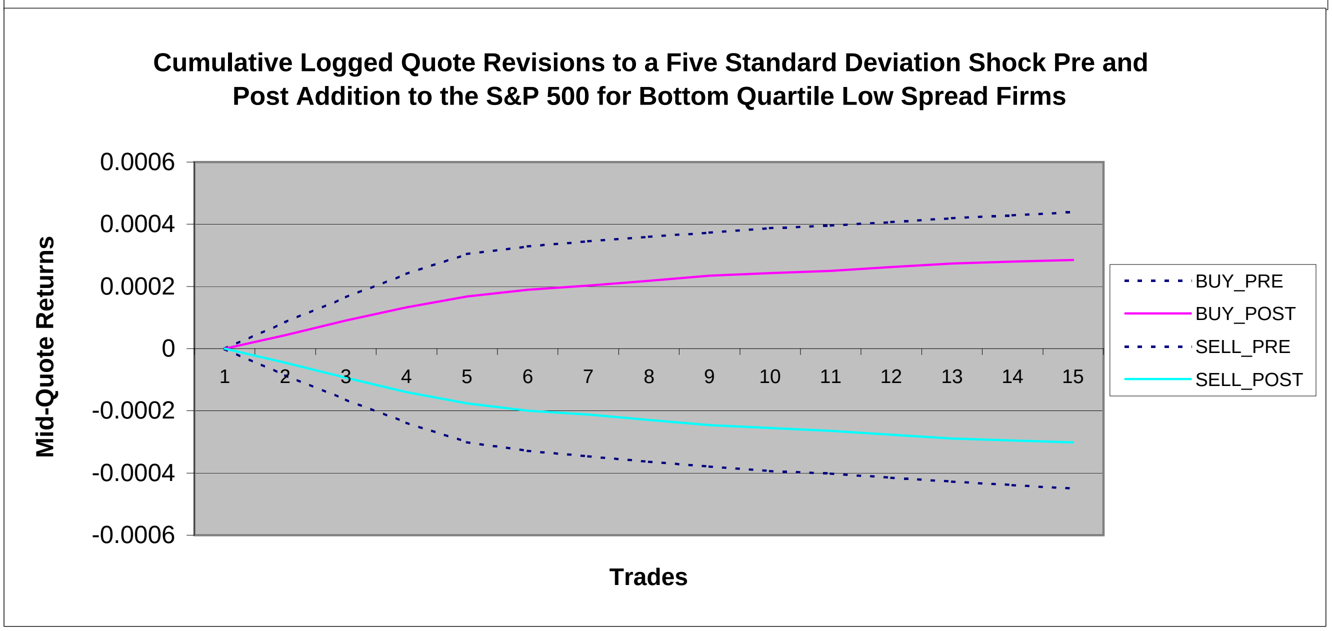
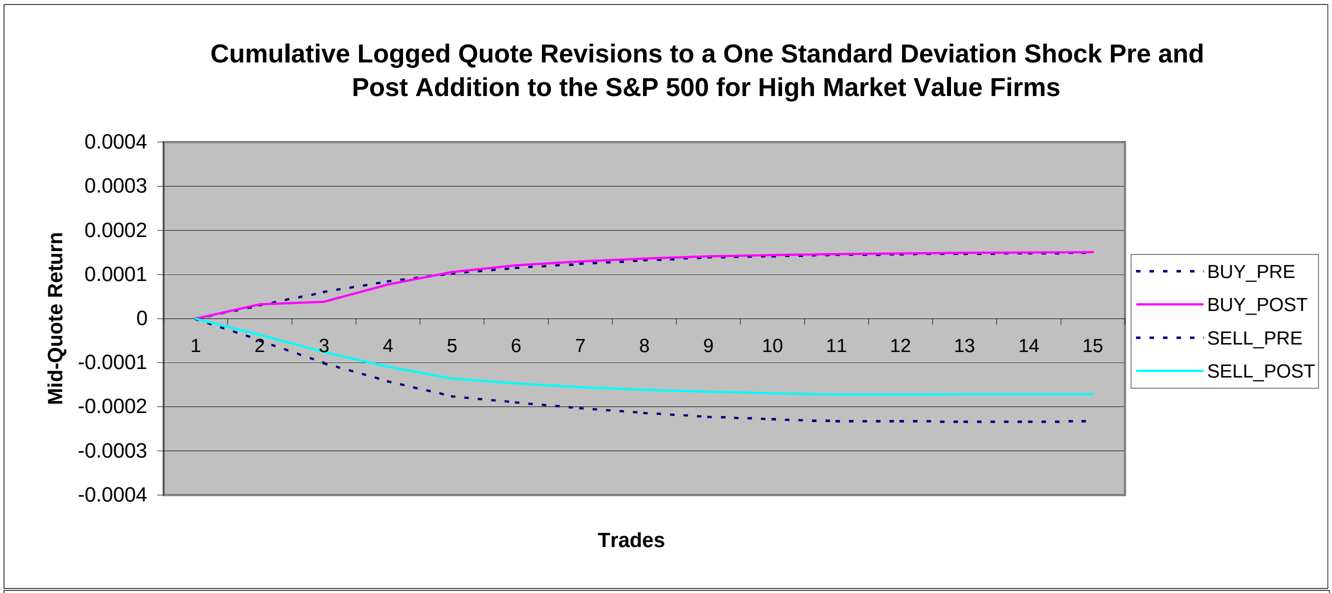


Figure 5. GIRFs of the Firms with Spreads in the Bottom Quartile of the Sample

Figure 5 contains results from the bottom quartile, or largest spread, (27) of the sample based on effective spread calculated pre addition to the S&P 500. The data for Figure 5 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 5 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. Consistent with theory, all price impacts featured in the above figure have a greater price impact than its counterpart in the low spread figure. In the top graph, pre addition shocks are similar in size while different in shape. The buy side quote revisions appear to reach equilibrium much faster than the sell side. Post addition, the size and shapes are quite similar with the sell side price impacts slightly larger. In the bottom graph, although the price impacts are much larger than their one sigma counterparts above, they are still smaller than a five fold increase expected from a IRF/VMA analysis. However, the buy and sell side price impacts are close in size and shape pre and post addition. Hence one form of linearity is rejected (a five fold shock has less than a five fold impact) while another aspect of linearity (symmetry) appears to be preserved for this set of firms.



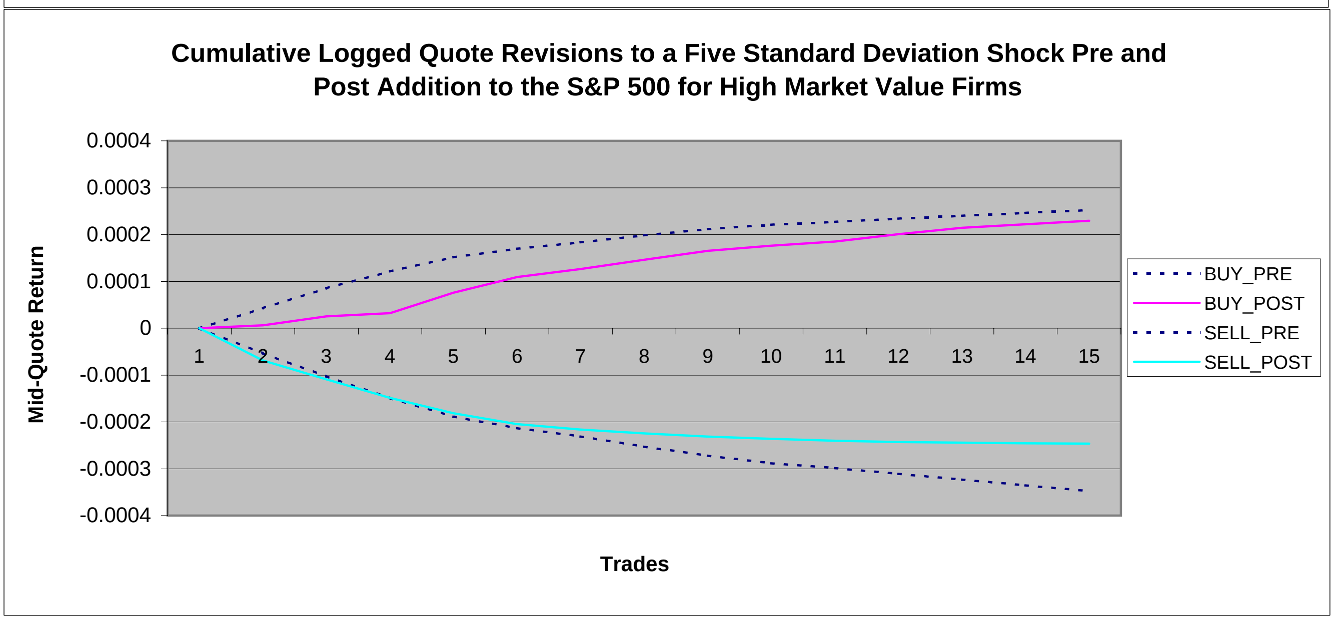
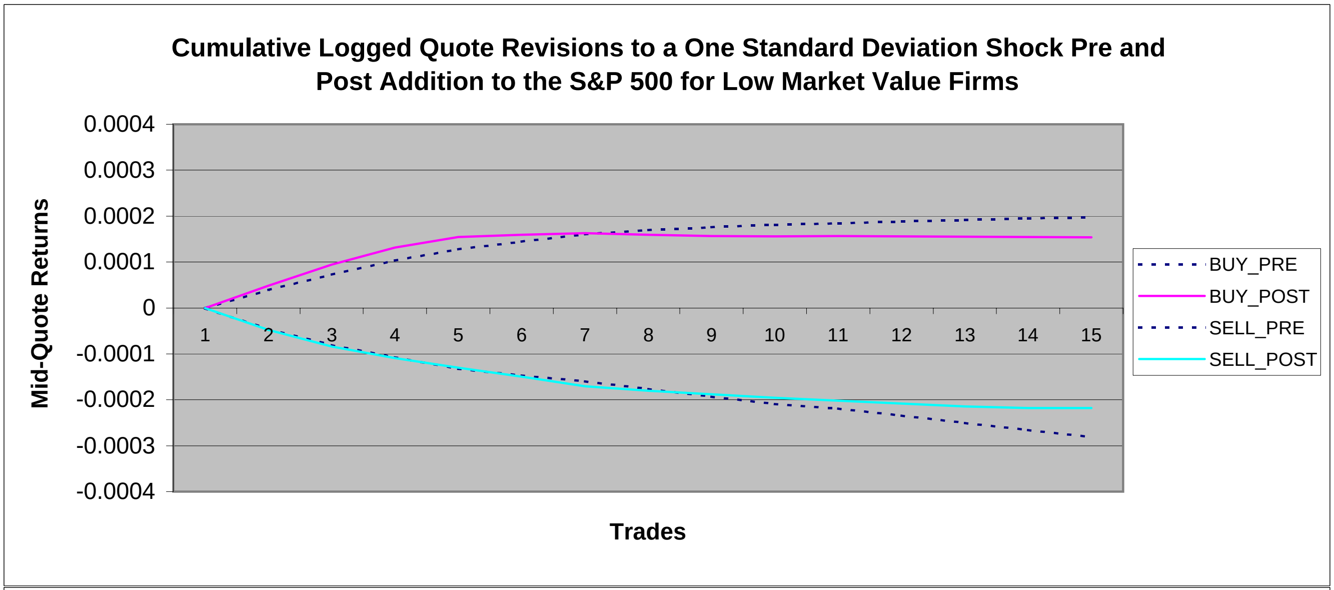


Figure 6. GIRFs of the Firms with Market Values in the Top Half of the Sample

Figure 6 contains results from the top half of the sample of firms (55) based on market values pre addition to the S&P 500. The data for Figure 6 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 6 contains cumulative GIRFs of a one(top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation.The top graph shows that for high market value firms the price impacts changed very little with respect to size or shape pre and post addition on the buy side. On the sell side, post impacts are smaller with a similar shape. On the bottom graph, both sides pre addition act strangely compared to their post addition counterparts. The post buy side price impacts are smaller for the first few trades, but then even out to slightly smaller than their pre addition counterpart while the post sell side is very close to its pre addition counterpart for a few trades and then diverges to a much smaller cumulative amount.



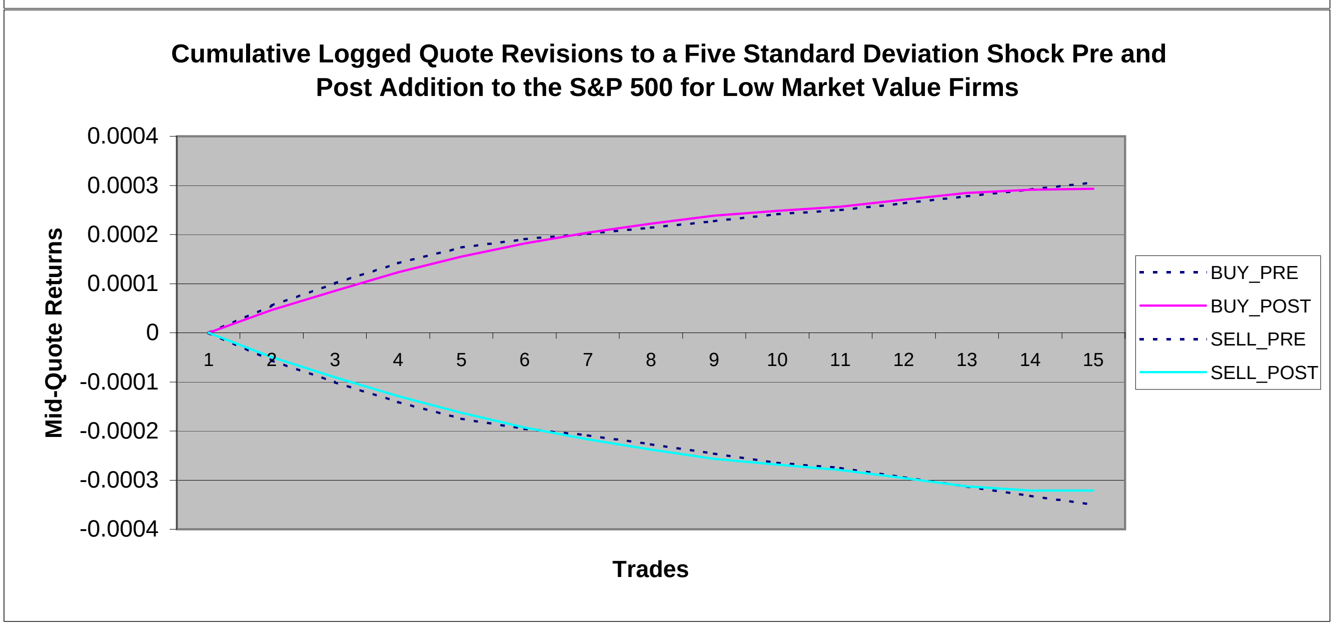
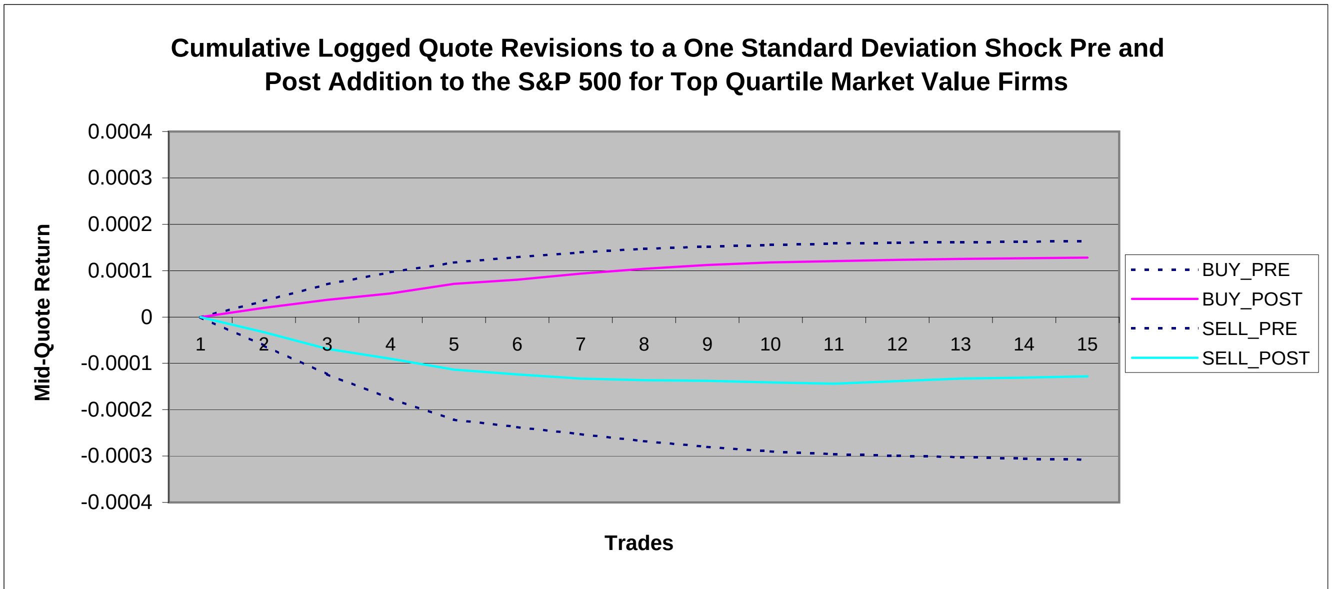


Figure 7. GIRFs of the Firms with Market Values in the Bottom Half of the Sample

Figure 7 contains results from the bottom half of the sample of firms (55) based on market values pre addition to the S&P 500. The data for Figure 7 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 7 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, both sides have smaller price impacts post addition and are roughly symmetrical in shape. Each side post addition also appears to reach an equilibrium price impact faster than their pre addition counterparts. In the bottom graph, both sides behave similarly in shape and size pre and post addition. Very little evidence of structural change is apparent in this sample with the five sigma shock to the revision equation.



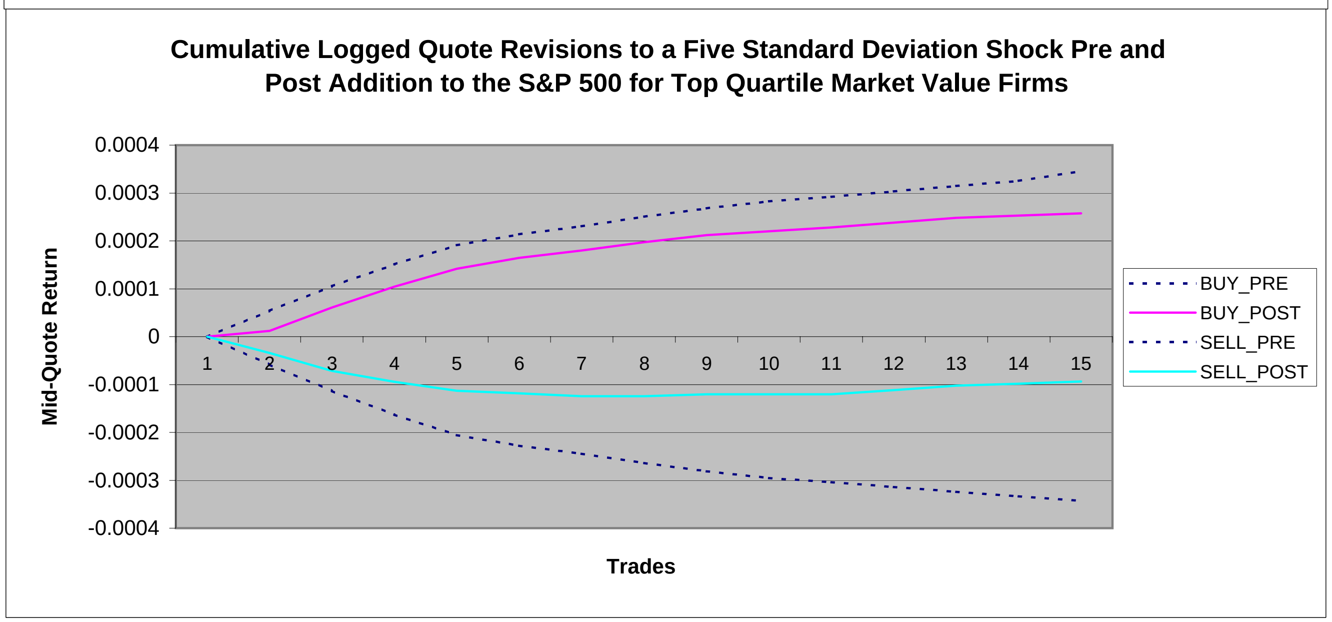
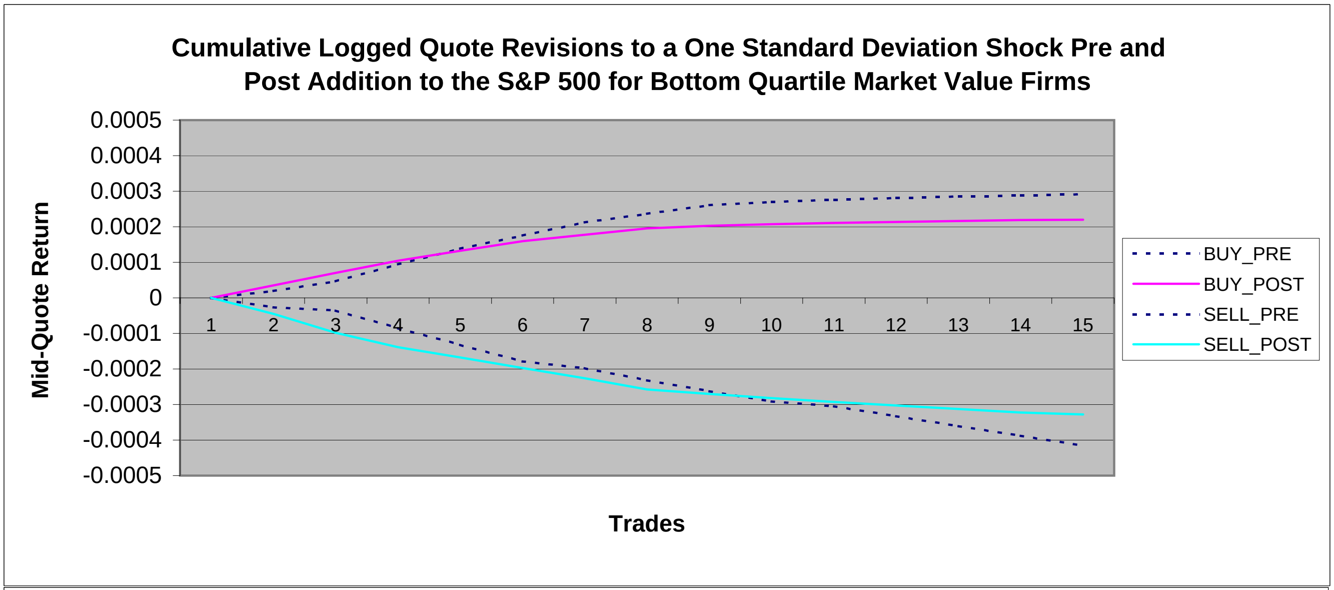


Figure 8. GIRFs of the Firms with Market Values in the Top Quartile of the Sample
Figure 8 contains results from top quartile, or largest firms, (27) of the sample based on market values calculated pre addition. The data for Figure 8 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 8 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, some atypical behavior appears. While the sell side impacts pre addition dominate the buy side pre addition, the pattern switches post addition as the buy side impacts are greater than the sells. In each case, the post addition impacts are smaller than their pre addition counterparts, with the sell side having the greater decrease in magnitude. In the bottom graph, a similar pattern as the above emerges and appears exacerbated due to the larger shock. The buy and sell side shocks are about equal in size and shape pre addition, but post addition the buy side shocks remain much larger than the sell side.



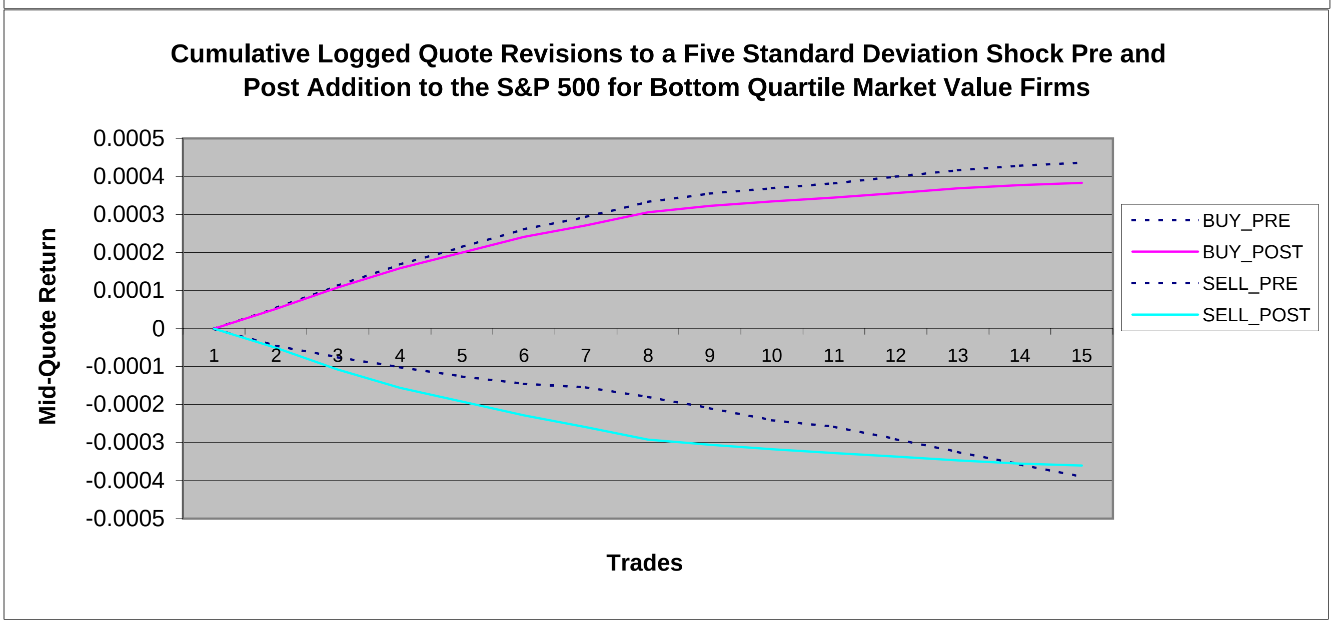
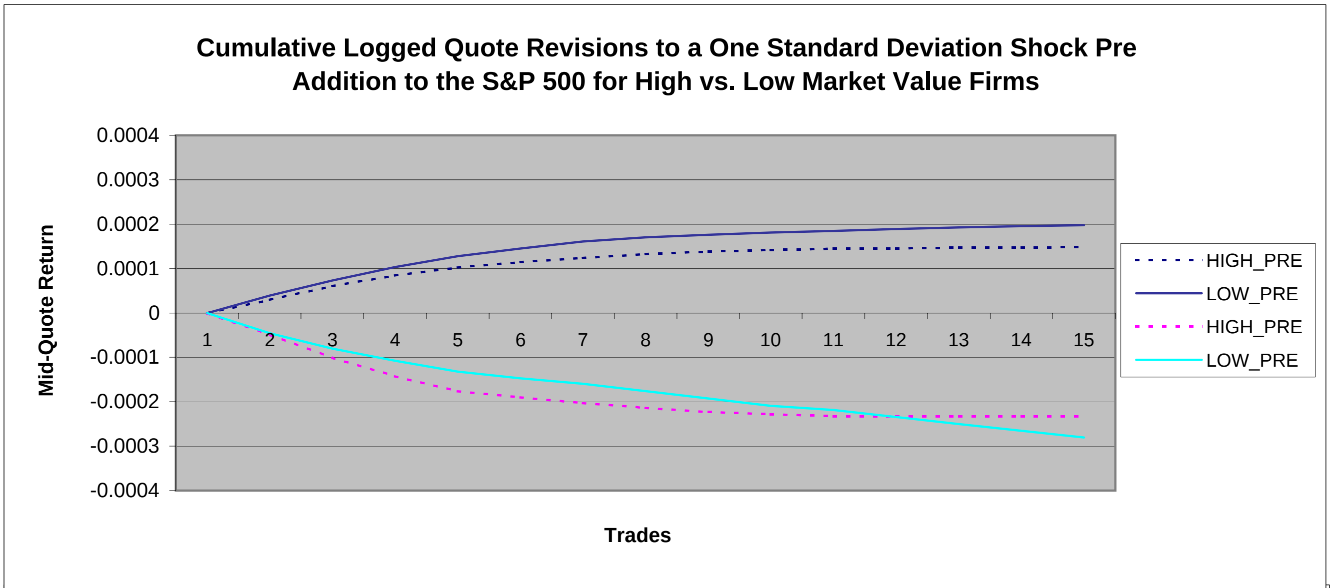


Figure 9. GIRFs of the Firms with Market Values in the Bottom Quartile of the Sample

Figure 9 contains results from bottom quartile, or smallest firms, (27) of the sample based on market values calculated pre addition. The data for Figure 9 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 9 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. The top graph show revision equation. The top graph shows a typical set of price impacts pre and post addition. Each side has a smaller price impact after addition and the sell side impacts are smaller than those of the buy side. In the bottom graph, different results obtain. On both sides, the post addition impacts are only slightly smaller than those pre addition. At variance with theory, the buy side impacts are slightly larger than the sell side both pre and post addition.



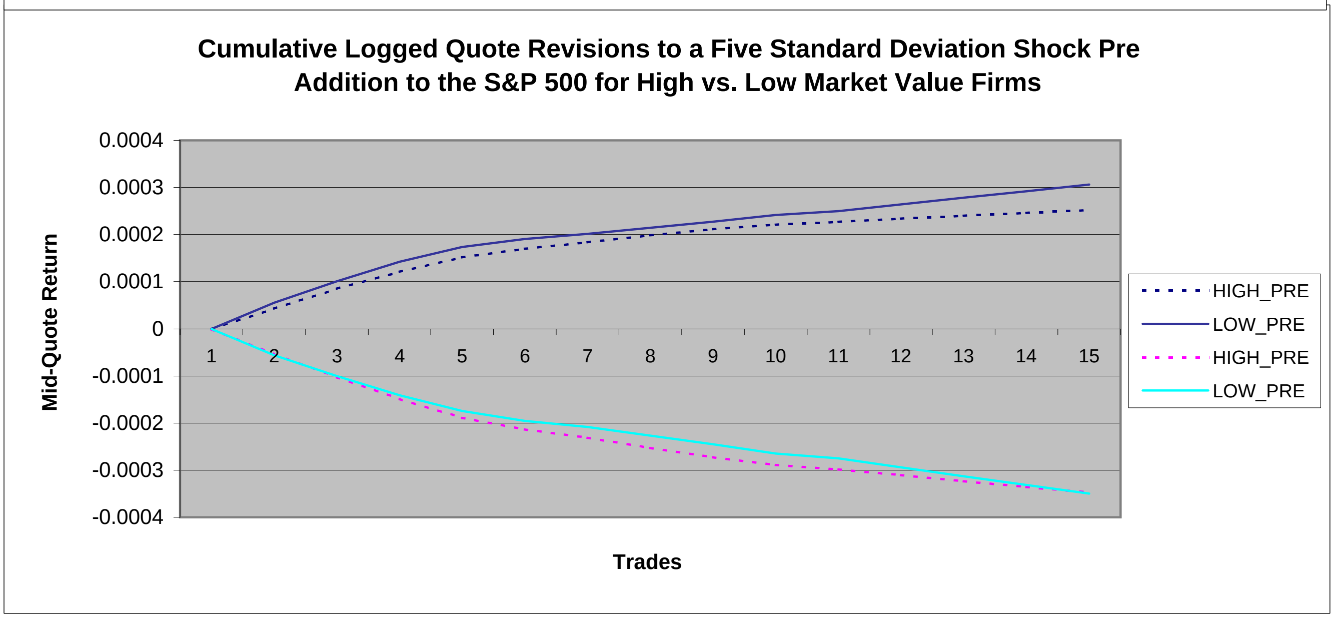
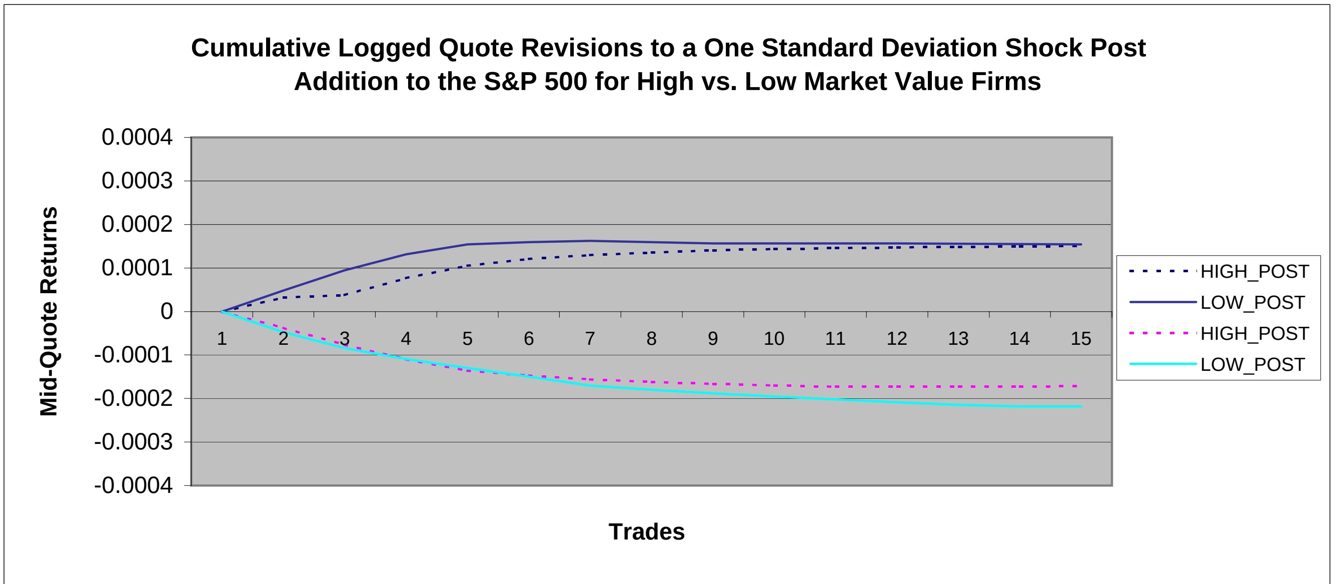


Figure 10. GIRFs of the Firms with Market Values in the Top and Bottom Halves of the Sample Pre Addition

Figure 10 contains results from the entire sample of firms (110) pre addition to the S&P 500 delineated by market value. The data for Figure 10 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 10 contains cumulative GIRFs of a one (top) and a five(bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, the high market value firms have smaller price impacts than those of smaller cap firms pre addition. As market value is a rough proxy for liquidity, this result is consistent with theory. In the bottom graph, although the buy side impacts are smaller for the higher cap firms, the sell side has both sets of firms having similar price impacts pre addition. Each set of sell side impacts is larger than the counterpart buy side impact.



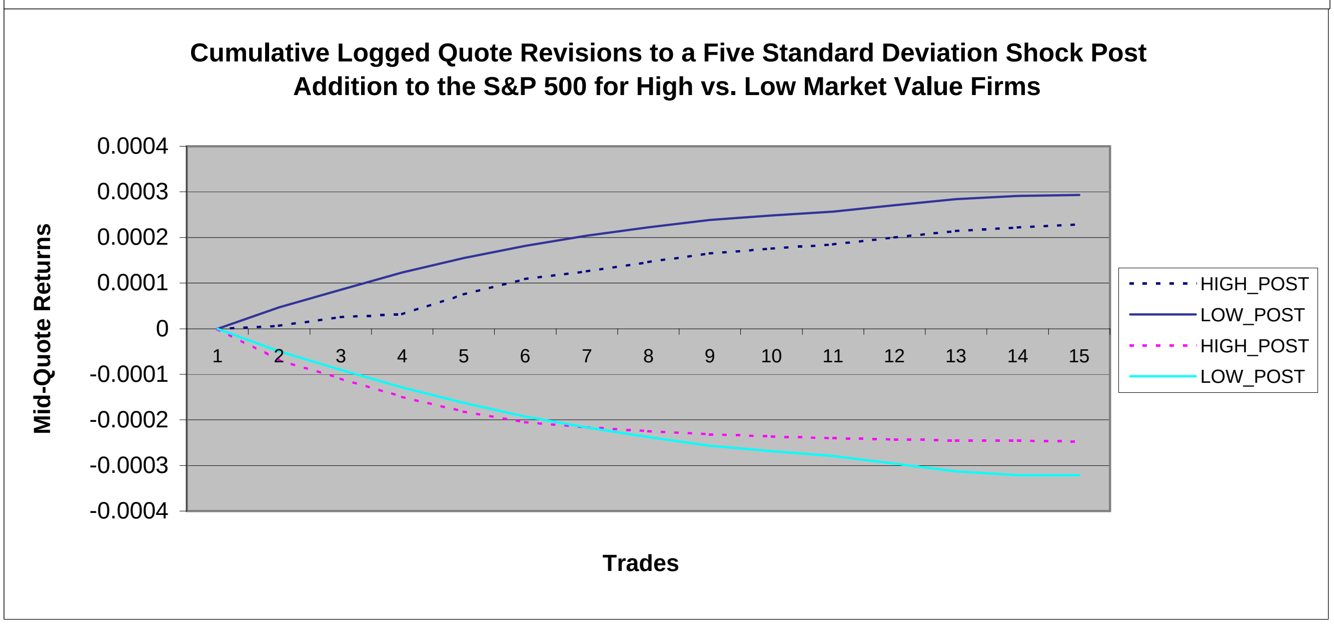
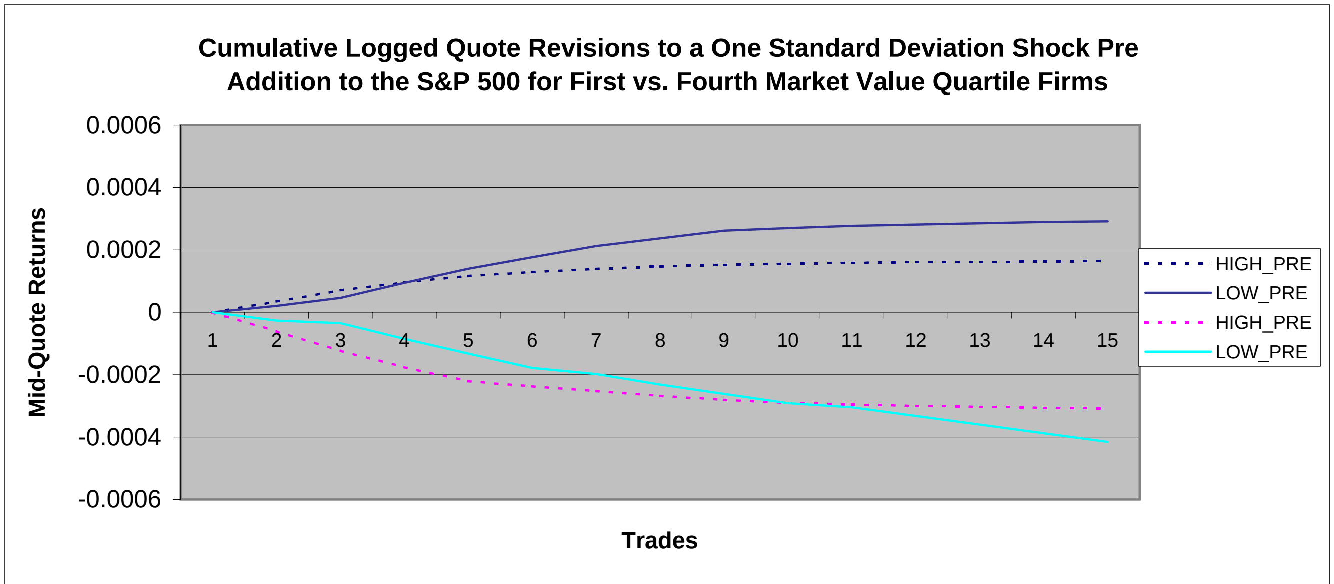


Figure 11. GIRFs of the Firms with Market Values in the Top and Bottom Halves of the Sample Post Addition
Figure 11 contains results from the entire sample of firms (110) post addition to the S&P 500 delineated by pre addition market value. The data for Figure 11 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow.Figure 11 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, the buy side impacts are similar in size for the high and low cap firms. Their behavior indicates that the high cap firms reach equilibrium faster but with larger revisions in the first few trades. On the sell side, the price impacts of both sets of firms have similar size and shapes for the first few trades, but the high cap firms have a steady flattening of their price impact curve resulting in smaller price impacts on that side. In the bottom graph, the high cap firms have smaller quote revisions to the larger shock with a nearly symmetrical reduction on the buy and sell sides.



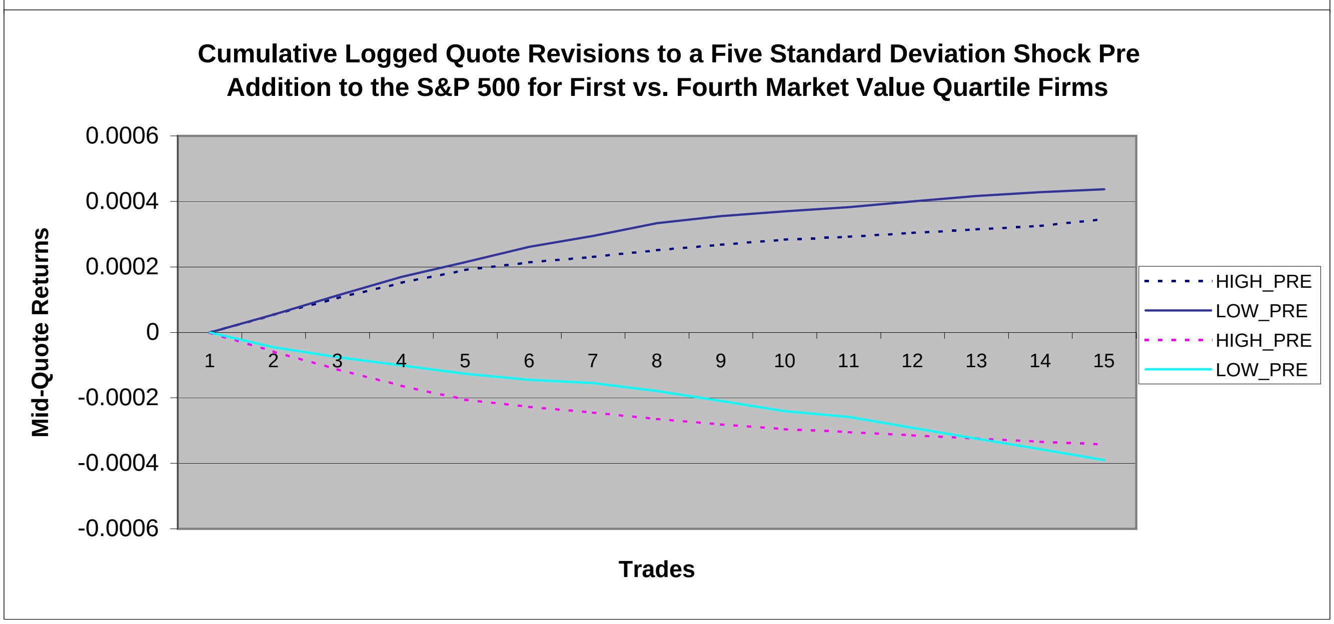
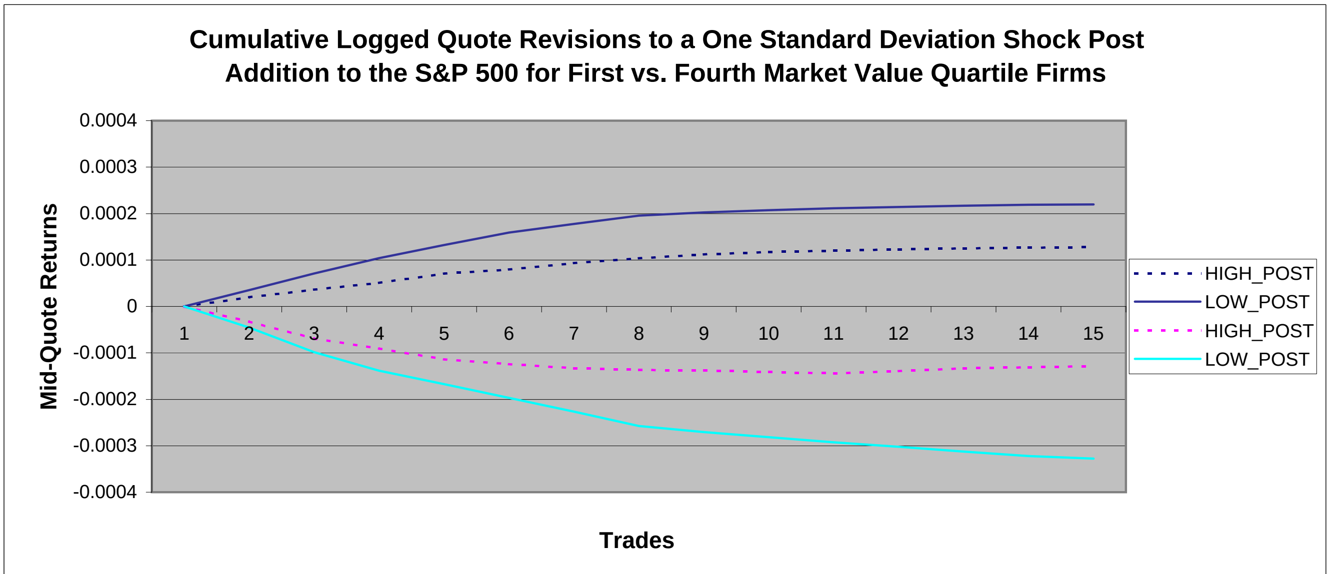


Figure 12. GIRFs of the Firms with Market Values in the Top and Bottom Quartiles of the Sample Pre Addition

Figure 12 contains results from the top (27) and bottom (27) market value quartiles based on pre addition market value. The data for Figure 12 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 12 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, smaller buy and sell side price impacts emerged for the high cap firms. The gains over the small cap firms were asymmetrical, though, as sell side impacts were larger for both capitalization sets. In the bottom graph, a less clear cut result is present. Although the high cap firms have smaller price impacts, the difference is very small for the sell side and emerges very slowly.



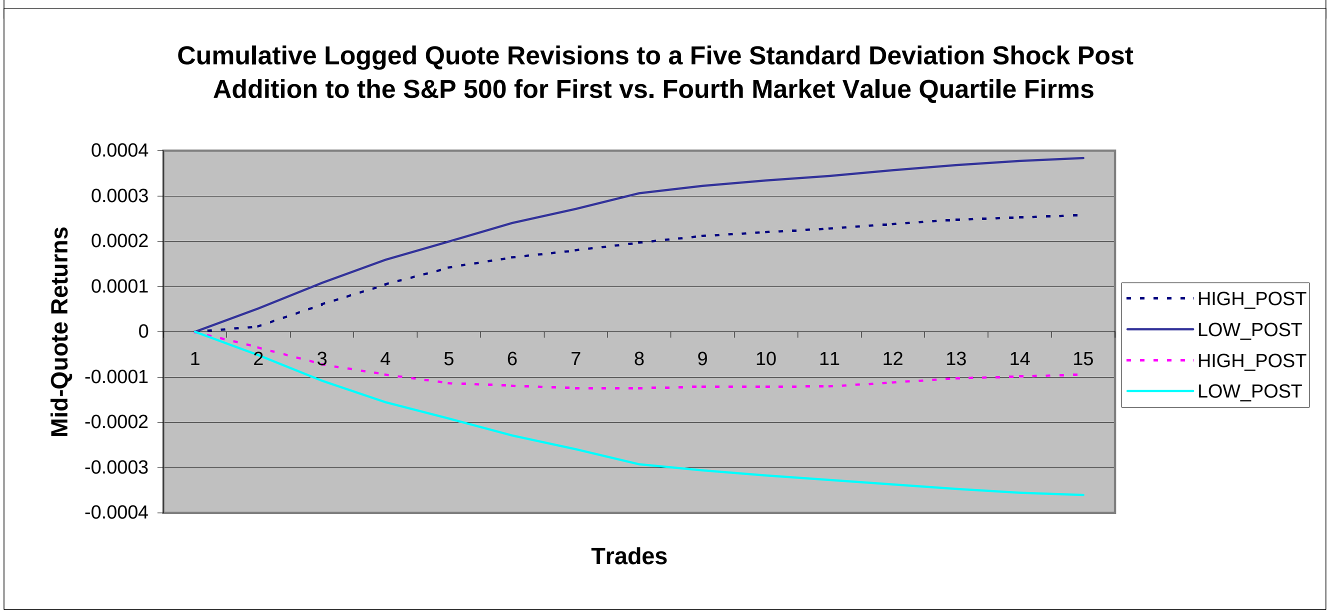
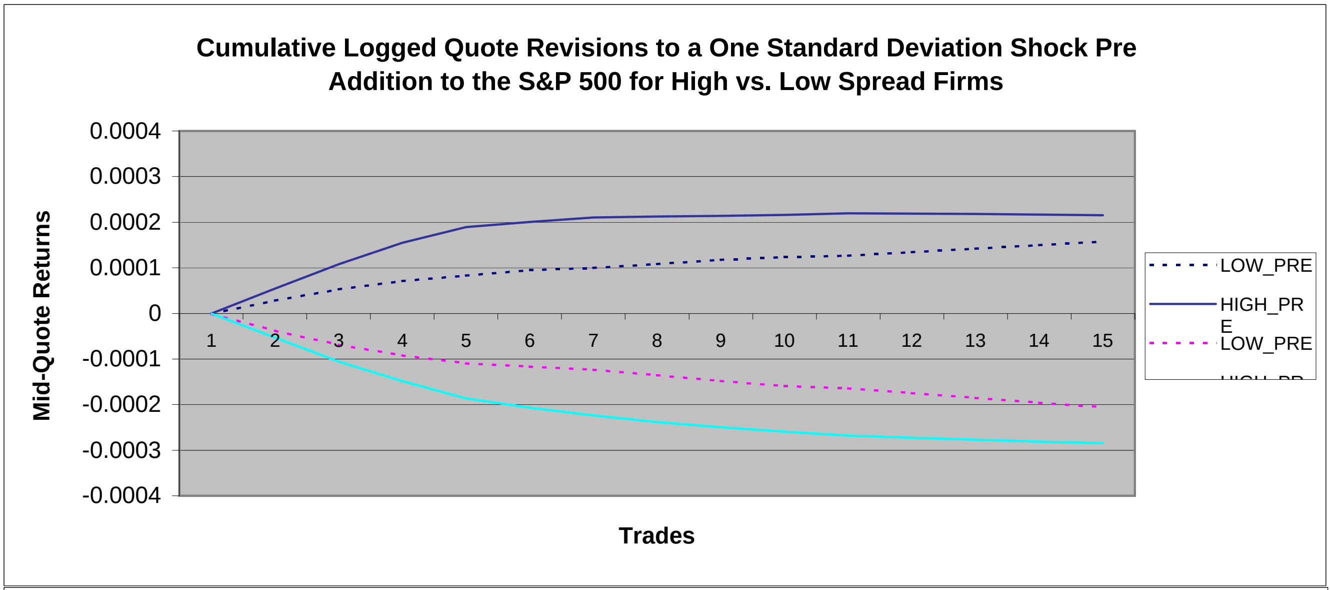


Figure 13. GIRFs of the Firms with Market Values in the Top and Bottom Quartiles of the Sample Post Addition

Figure 13 contains results from the top (27) and bottom (27) market value quartiles based on pre addition market value. The data for Figure 13 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 13 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation.In the top graph, the low cap firms have symmetrical price impacts while the high cap firms show signs of asymmetry. The buy side has a larger impact than the sell side but they both have similar shapes. On the bottom graph, again the smaller cap firms have symmetrical price impact behavior on the buy and sell side. The high cap firms have a much larger buy side impact pre addition, a counterintuitive result.



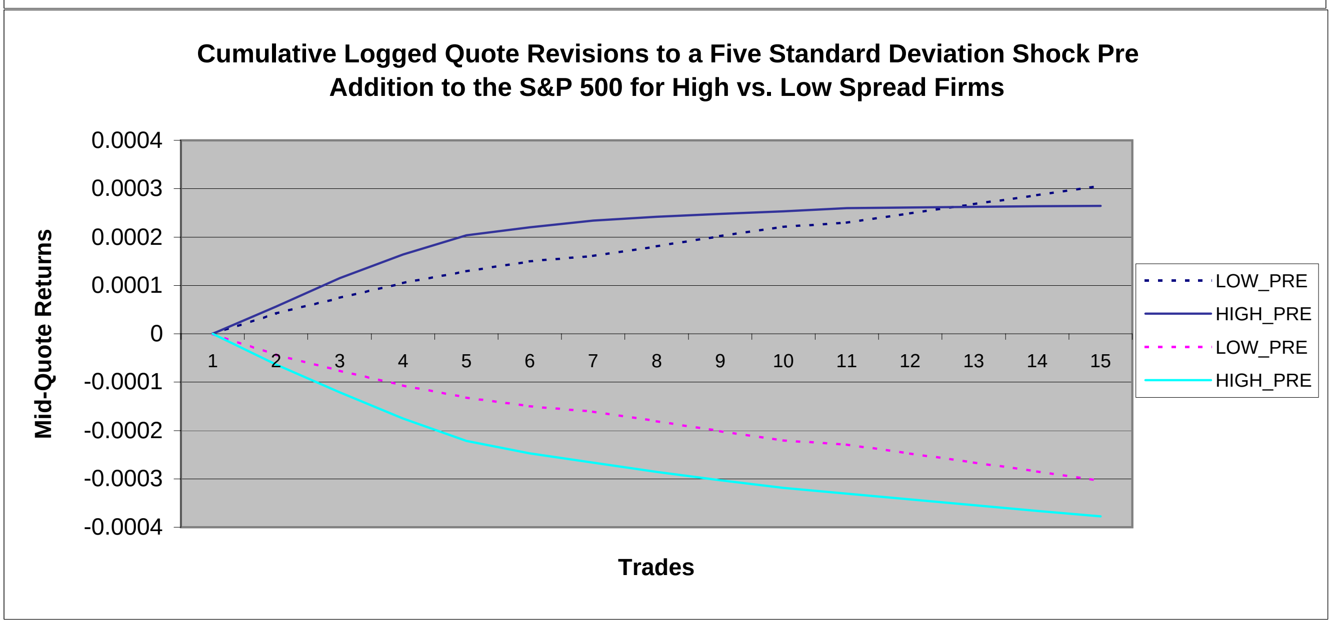
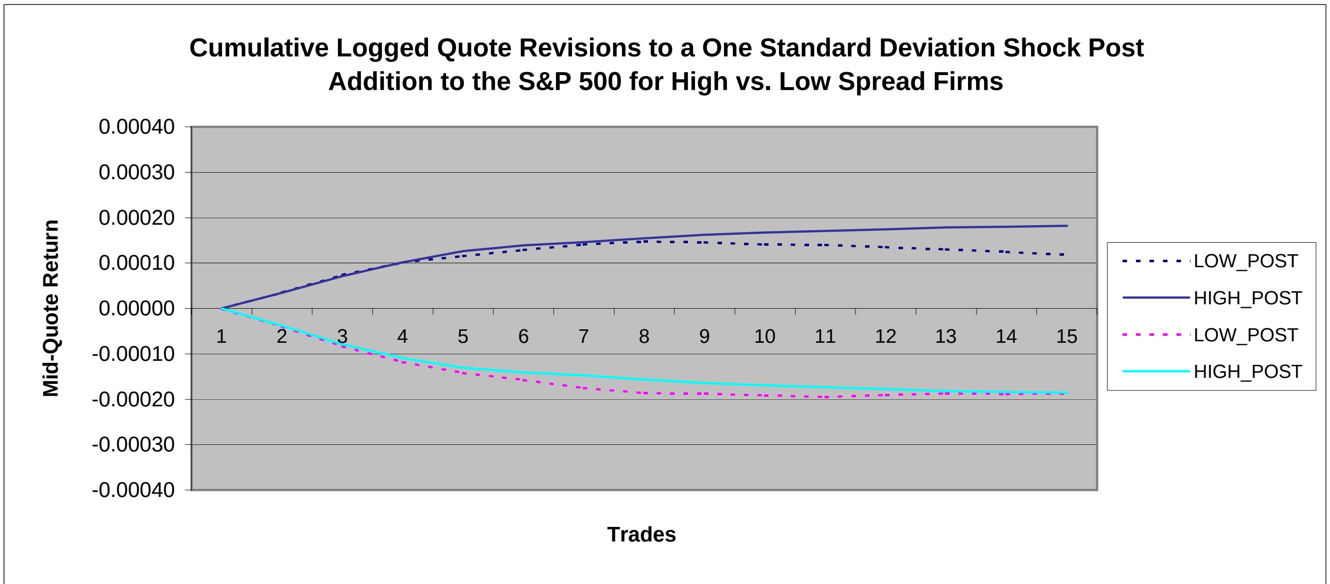


Figure 14. GIRFs of the Firms with Spreads in the Top and Bottom Halves of the Sample Pre Additio

Figure 14 contains results from the entire sample of firms (110) pre addition to the S&P 500 delineated by effective spread calculated pre addition. The data for Figure 14 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 14 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, the low spread firms demonstrate asymmetrical price impact on the buy and sell side while the high spread firms seem to have a symmetrical response on each side. In each set of firms has sell side price impacts dominating those on the buy side. In the bottom graph, the same basic pattern emerges with one exception -- the buy side impacts for the low spread firms exceeds those of the high spread firms 13 trades after the shock.



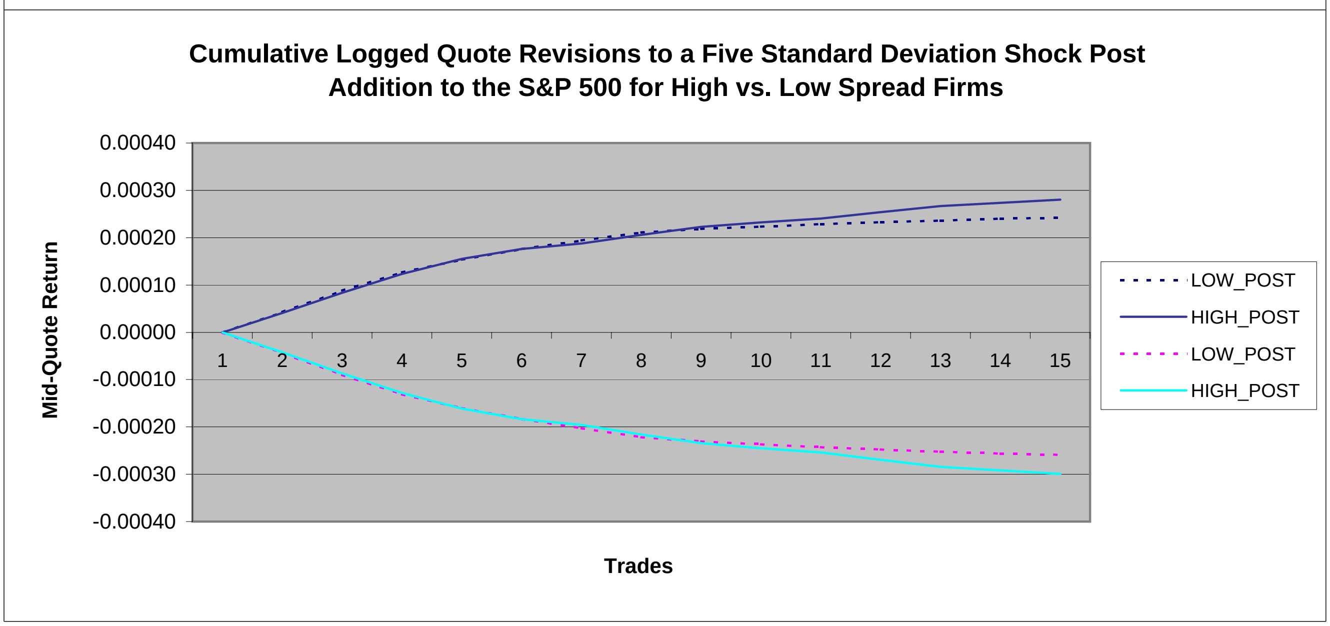
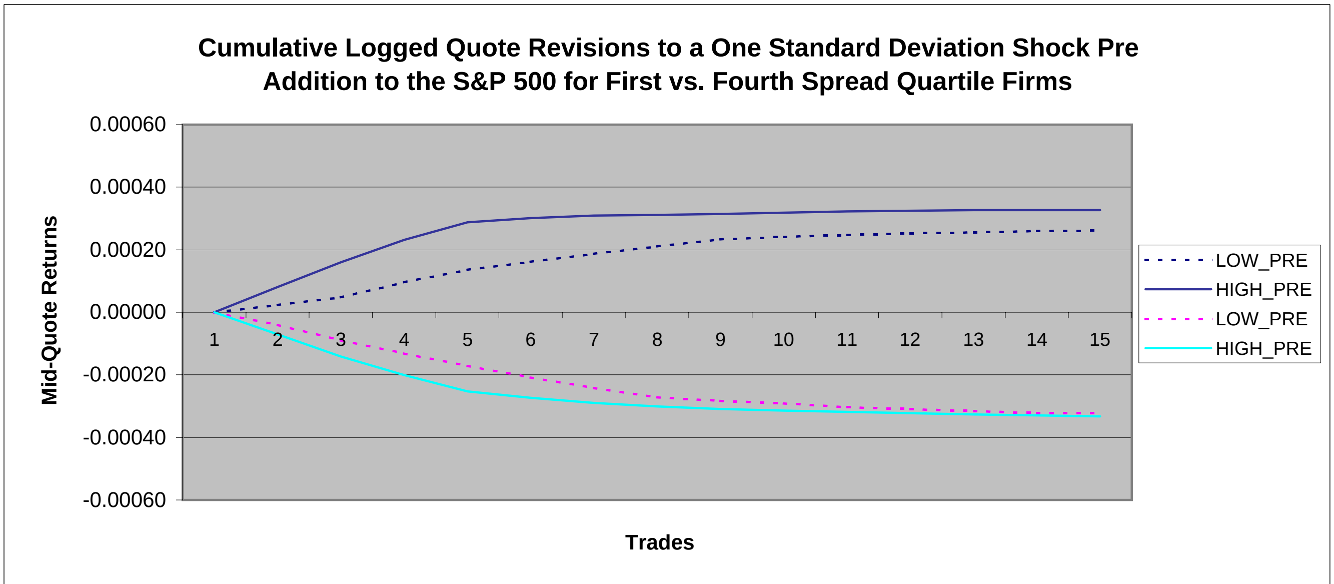


Figure 15. GIRFs of the Firms with Spreads in the Top and Bottom Halves of the Sample Post Addition
Figure 15 contains results from the entire sample of firms (110) post addition to the S&P 500 delineated by effective spread calculated pre addition. The data for Figure 15 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow.Figure 15 contains cumulative GIRFs of a one (top) and a five(bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, the high and low spread firms have similar shapes and sizes in price impacts to sell side shocks. On the buy side the low spread firms have a similar size and shape up to 8 or 9 trades after the shock. After that, the low spread firms price impact curve lies beneath the high spread firms'. In the bottom graph, each set of firms have a symmetrical response to the buy and sell side shocks with the low spread firms having smaller price impacts on both sides.



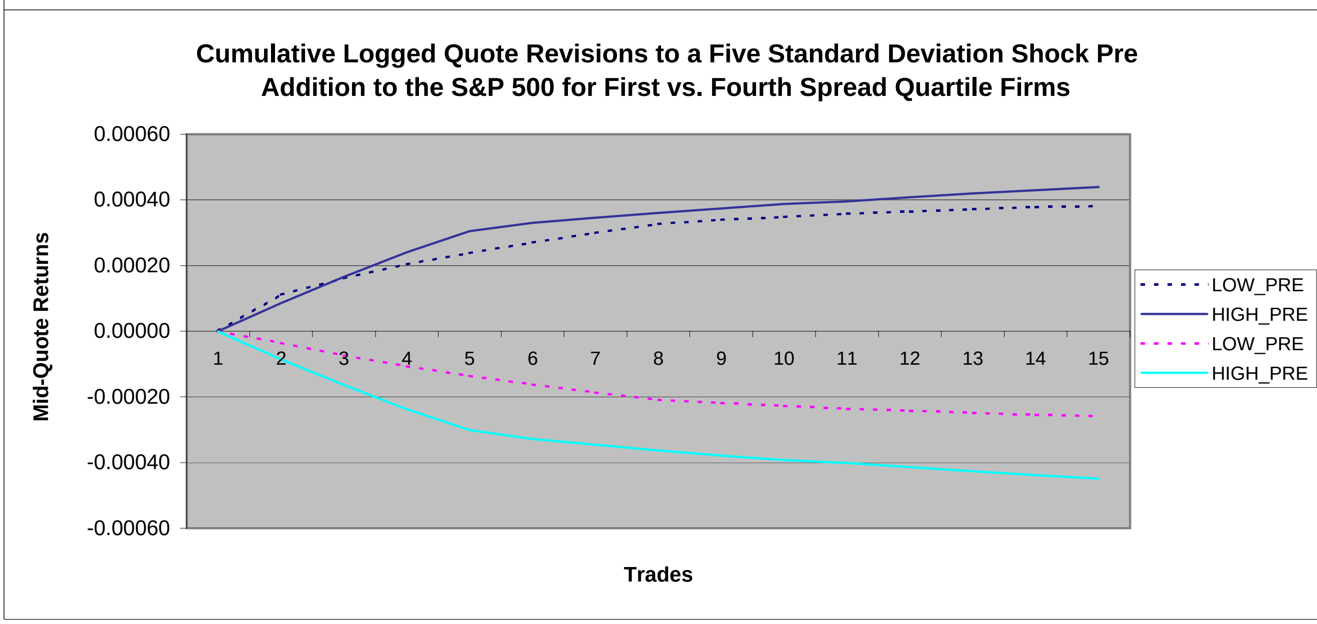
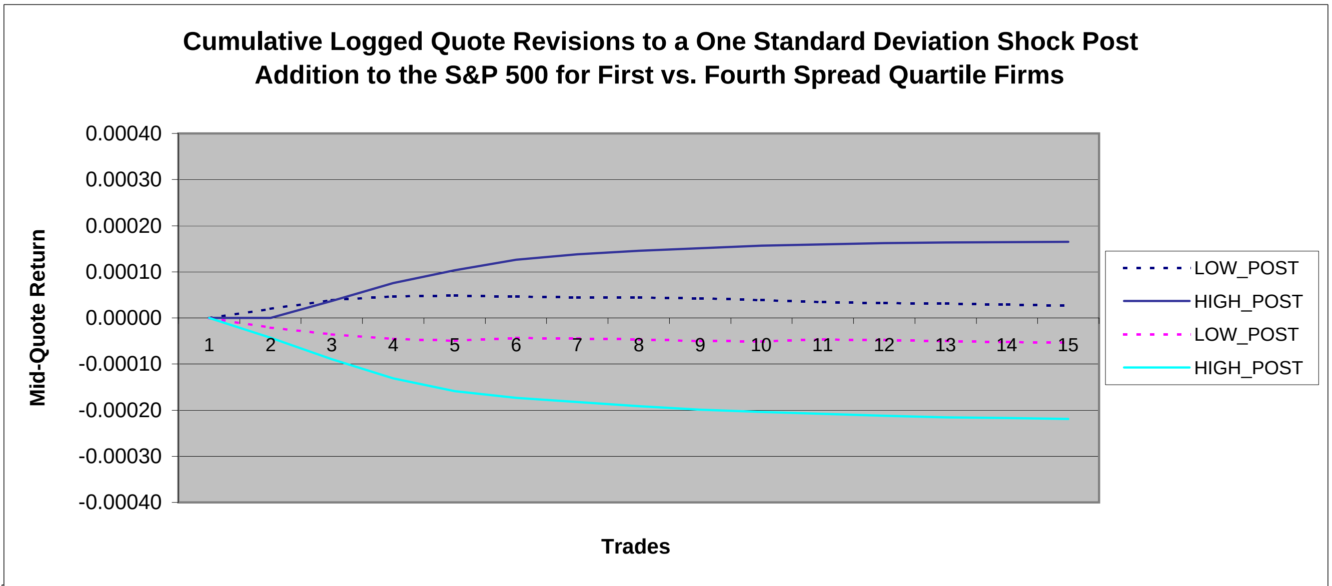


Figure 16. GIRFs of the Firms with Spreads in the Top and Bottom Quartiles of the Sample Pre Addition
Figure 16 contains results from the top (27) and bottom (27) spread quartiles based on pre addition effective spreads. The data for Figure 16 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 16 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, although the high and low spread firms' sell side price impacts have different shapes, the size of the impacts are approximately equal. The high spread firms' buy side impacts strictly dominate the low spread firms' pre addition. In the bottom graph, the buy side price impacts closely match each other while the sell side impacts of the low spread firms are noticeably smaller. The low spread firms appear more symmetrical in the top graph, whereas the high spread firms are more symmetrical in the bottom graph.



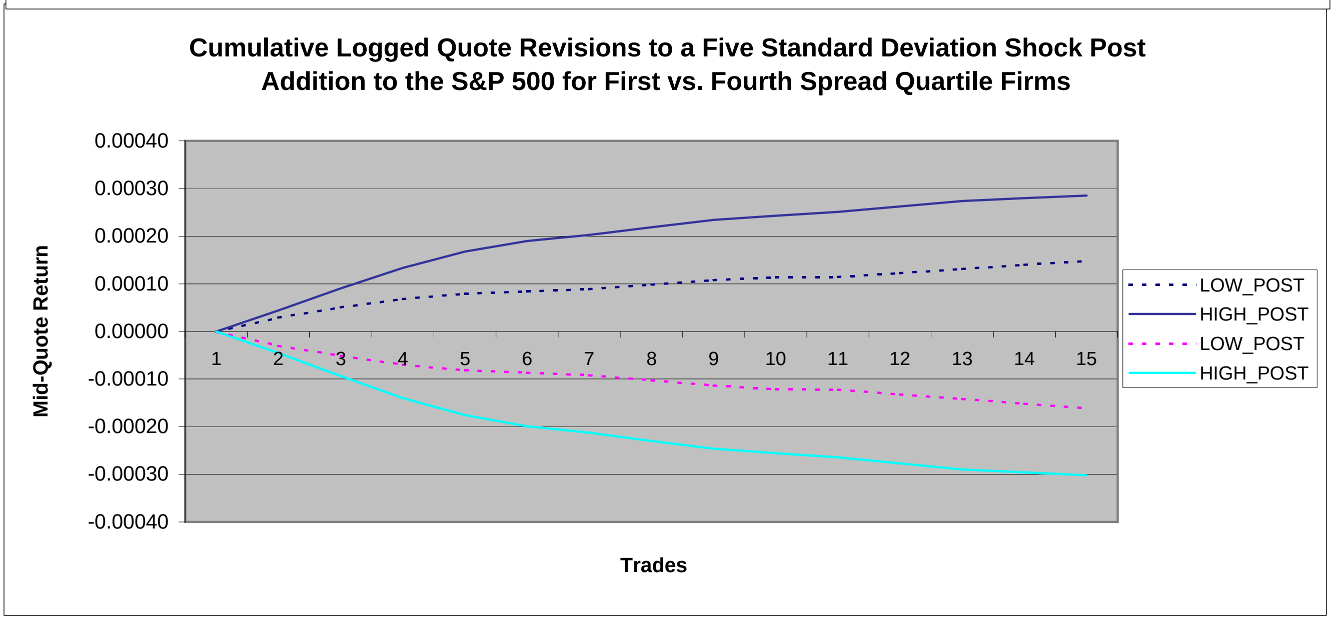


Figure 17. GIRFs of the Firms with Spreads in the Top and Bottom Quartiles of the Sample Post Addition
Figure 17 contains results from the top (27) and bottom (27) spread quartiles based on post addition data from firms effective spreads calculated pre addition. The data for Figure 17 were pulled from the Trade and Quote (TAQ) database. Pre addition data were pulled from the 10 trading day interval [-60,-51] before announcement. Post addition data were pulled from the 10 trading day interval [+51, +60] after announcement. A generalization of Hasbrouck's 1991 structural bi-variate VAR(p) of logged quote revisions and signed order flow that allows m thresholds to arise in the signed order flow was used to form the GIRFs. The threshold model allows asymmetries in buy and sell shocks that conventional IRF/VMA representations do not. The GIRF is an expectation of the quote revision response to a shock (of fixed magnitude) over and above what one would expect to happen in the absence of the shock. The cumulative quote revisions are interpreted as the permanent price impacts from a trade innovation and are used to measure the average amount of private information in the order flow. Figure 17 contains cumulative GIRFs of a one (top) and a five (bottom) standard deviation shock from the signed order flow equation on the revision equation. In the top graph, both sets of firms appear to have symmetrical responses to buy and sell side shocks. The low spread firms have a much smaller response post addition to each shock. In the bottom graph, each set of firms still maintains symmetry of responses while the magnitude of each a greater (but not five fold greater) than the one sigma shock in the top graph.

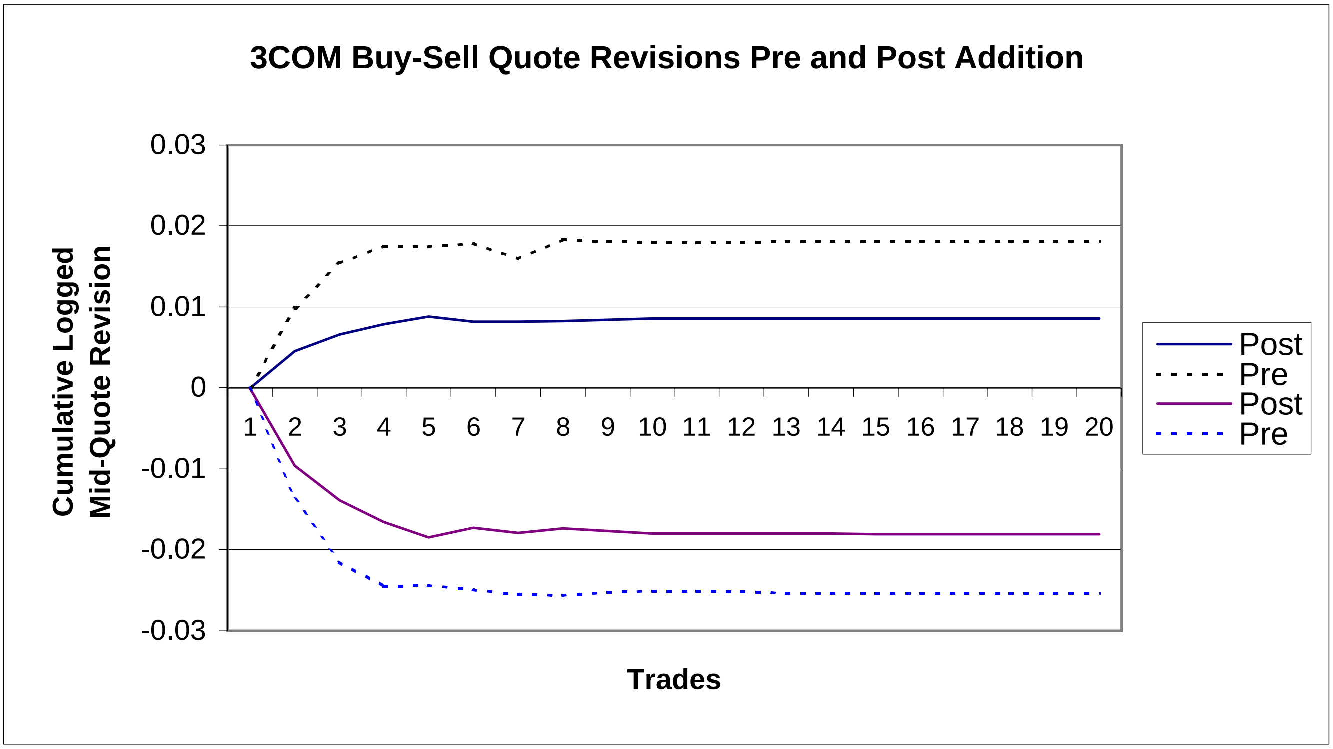


Figure 18. Example GIRFs of a Sample Firm, 3COM, from a One Standard Deviation Shock Pre and Post Addition

Figure 18. Cumulative generalized impulse response functions of a unit trade innovation to the buy and sell side in the quote revision equation. The GIRFs were formed using 3COM Corp data as an example of the price impact reduction arising post addition to the S&P 500. The GIRFs were formed by averaging 1000 simulations for each observation then sampling through the entire history that covered two trading weeks. 20 such averages were taken to construct a mean distribution over each of the 20 trades after the shock. The plot of the means for each trade appears above. The graph indicates that the private information content of a trade innovation is more slowly absorbed into the quote process after the addition to the S&P 500 Index, but with a smaller price impact on quotes. Although the stock featured here (3COM) is liquid before addition (its permanent price impact before addition is approximately 0.02% of the firm's stock price for a buy and about 0.025 % for a sell shock), it does have a significant marginal improvement after addition as each new price impact is approximately half of the previous level.

Chapter 4

Generalization of Glosten Harris (1988) Model

Introduction

Spread decomposition models strive to inform the researcher of how transaction costs are allocated between fixed and variable costs. Two approaches to this topic have been proposed. One method models the percentage of the spread each component maintains (see Lin, Sanger Booth[23]). The other attempts to estimate the dollar levels of each of the spread components such as Glosten and Harris[8], Madhavan, Richardson and Roomans[13] and Huang and Stoll[22]. This paper will extend the work of Glosten and Harris and will attempt to generalize their spread level model by admitting data driven thresholds and thus formulates a piece-wise linear model that approximates the nonlinear price behavior in auction markets.

The questions that I want to answer in this paper are: given the average effective spread for a stock, i.e. the difference between the transaction price and the midquote, what available and quantifiable economic variables contribute to the effective spread; what is their marginal contribution in conditional and unconditional settings, and does a nonlinear specification outperform the traditional linear specification when forecasting spreads? In developing a model that addresses these questions, one should be able to obtain ex ante trading strategies that minimize transaction costs.

Trading costs are broken down formally into the fixed costs of processing trades, those of keeping an inventory and the variable costs associated with trading against agents with private information. Research in the late 1980s and 1990s (Madhavan and Smidt [24] and Stoll[25]) suggests that the inventory holding costs are on average small relative to the other components of the effective spread. Consensus in the literature has the fixed cost component as the dominant fraction of the spread. This is supported in this study by regression estimates which indicate that a large fraction of the cost of trading is fixed and economies of scale obtain from this in the form of lower effective spreads for medium sized orders. The variable cost component enters into a model to proxy for the amount of adverse selection present in a trade. As in Glosten and Harris (1988), signed order flow will be a proxy for adverse selection in this paper[1].

Microstructure theory predicts that a price premium exists when trading larger amounts of stock or when one is trading in an unexpectedly brisk market. If the underlying dynamics of a market were linear, then one would not observe trading frictions such as those arising from finite depth or the added costs of supplying liquidity to the market. Identifying the magnitude and significance of each of these changepoints or thresholds not only allows one to explain a greater degree of the observed price changes, but also should provide better forecasts and hence enable one to identify conditions that minimize trading costs.

Unfortunately for the empiricist, theory does not place any bounds on the magnitude of either of these two threshold variables and as such their magnitude remains an empirical question[26]. Attempts to formalize a model with ad hoc threshold levels involves pre-assigning values to a nuisance parameter. Without verification of whether the data support such values, one will perform a misspecified test as the threshold is known only under the alternative hypothesis. I use Hansen's (1999) technique for identifying thresholds to test for both order flow and intensity threshold levels and their respective significance.

[1] Lee and Ready's[15] algorithm for determining trade sign was used to classify buys and sells

Testing for intensity thresholds also allows one to make some inference about the informational role of time in trading. Following Easley and O'Hara[27], informed traders can take both sides of a transaction and therefore can capitalize on private information at all times. A lack of trading would imply a lack of news and consequently longer trade durations or periods of no trading should precipitate waning spreads. On the other hand, Diamond and Verecchia[29] posit that informed traders may not always have an inventory to trade from and they are further restricted from trading by the uptick rule on short selling. Thus longer trade durations may be a sign of bad news and thus one would expect spreads to widen during periods of longer trade duration. Analyzing the intensity price premia (if any) on a given volume will indicate which line of reasoning the data support.

Data

To test for the presence of threshold effects on trading costs, I chose 30 common stocks from each of three market value deciles. Market value decile information was pulled from the CRSP daily stock returns for the year 1996. Firms from the New York Stock exchange were selected to minimize problems from arising due to inter-dealer trades and double counting on NASDAQ and to reduce zero duration observations from combining NYSE data with those of smaller trading venues.

After the data were filtered to control for any exchange effects, firms with SICs corresponding to public administration, museums, zoos, financial services, real estate, insurance, forestry, and the U.S. Post Office were dropped. The data were then partitioned into deciles on the arbitrary date of April 1, 1996. Each decile was sorted alphabetically on the ticker symbol and the top 30 firms from the tenth (largest), eighth and sixth decile were chosen. To allow for a reasonable number of observations in each group, three time horizons were chosen. The largest firms had a one month horizon, the middle group had a three month period and the smallest had a six month sample period. The out of sample data were collected in a similar manner. Starting on November 1, 1996, the largest firms had a one trading week period, the middle set had a two week period and the smallest firms had a trading month. None of the original firms were missing from the out of sample period.

The intra daily data from the New York Stock Exchange were pulled from the Trade and Quote (TAQ) data set. All observations before 9:45 and after 4:00 EST were dropped and the data were filtered to eliminate any observation with a price level change greater than 10% of the former observation. Zero duration trades were aggregated. Summaries of the order flow and spread data for each set of stocks in the sample are featured in Table I.

Methodology

Spread Decomposition

Two goals of this paper are to explain and to forecast trading costs conditional on the trading environment. Traditionally, these costs have been estimated through modelling the direct order processing costs, the costs of maintaining an inventory, and the costs arising from trading against privately informed traders. The convention used in this paper is that trading costs will have a fixed and a variable component. Fixed costs include those directly accrued in trading as well as those of maintaining an inventory. Any price effects arising from the fixed cost component are assumed to be transient. The variable cost component proxies for the amount of asymmetric information in a trade. Historically, the proxy for asymmetric information has centered on trade volume. This paper generalizes the Glosten and Harris model in two ways. First, another proxy for asymmetrical information, unexpected trade intensity, is included. Second, I abandon the linear structure in the price/volume and price/time relation and capture nonlinearities between these variables with a piecewise linear specification.

Several econometric assumptions from previous works are adhered to in this study. First as in Hasbrouck (1991), I assume that the order flow process is exogenous and ceteris paribus unexpected from a market participant's point of view. Also, quotes are set after trades. The causality runs from order flow to quotes and not the other way around. As in Engle and Dufour(2000), I assume that the time process is exogenous to both price and trade processes. They concede "...that the trade arrival rate might depend on market variables such as prices." However, they continue "We do not have knowledge, to this date, of any theoretical model that shapes the reciprocal interactions or price, trade, and time...In particular, it might be interesting to study the effects, if any, of market makers quote revision on trading flow."

Glosten and Harris' model (GH) is unique among most spread decomposition models in that it uses trade indicator variables and signed order flow to explain observed changes in price. Their full model is

$$\Delta P_t = Q_t C_t - Q_{t-1} C_{t-1} + Q_t Z_t + e_t \tag{4.1}$$

where ΔP_t is the observed price change at time t and Q_t is the trade indicator derived from Lee and Ready's (1991) algorithm. The fixed and variable cost spread components are assumed to take a linear functional form $C_t = c_0 + c_1 V_t$, $Z_t = z_0 + z_1 V_t$. Plugging in the values for C_t and Z_t, the fixed and variable costs, respectively, we have

$$\Delta P_t = c_0(Q_t - Q_{t-1}) + c_1(Q_t V_t - Q_{t-1} V_{t-1}) + z_0 Q_t + z_1 V_t + e_t. \tag{4.2}$$

Following the methodology in GH, I first ran unconstrained regressions to see which of the four coefficients were significant in the cross section. Then, I reran the estimation with the significant variables, constraining the adverse selection component to be proportional to volume with a constant of proportionality, z_1. Consequently, the base specification used in this paper is

$$\Delta P_t = c_0(Q_t - Q_{t-1}) + c_1(Q_t V_t - Q_{t-1} V_{t-1}) + z_1 Q_t V_t + e_t \tag{4.3}$$

where c_0 captures the fixed cost of trading, c_1 represents any economies or diseconomies of scale, and z_1 is marginal impact on price changes given an increase in order size. The marginal cost of trading, c_1, will be zero for a constant cost structure, positive for increasing costs and negative for decreasing costs. Inventory models predict that c_1 will be positive. If there is a large fixed cost component to trading, then c_1 will be negative reflecting a volume discount. Table II indicates that the marginal costs are negative for all three sets of firms in this study. This supports Madhavan and Schmidt's (1991)[24] conclusion that inventory costs are a small fraction of costs of trading.

Trade Duration Measure If we can think about the above specification in matrix form as $Y_t = (\Delta P_t)$, then (4.3) can be expressed as

$$Y_t = \vec{\alpha} X_t + \epsilon_t \tag{4.4}$$

where $\vec{\alpha}$ is the 1×3 coefficient vector and X_t is the matrix of significant right hand side variables. If no adverse selection were present in the order flow and no liquidity constraint applied, then the above model would most probably suffice. In the presence of adverse selection and liquidity constraints, the dynamical nature of the market may allow different regimes to arise in the face of large volume or unexpectedly high trading intensity.

Much has been written on the effect of varying the time between trades on spreads and liquidity (see Easley and O'Hara (1987)[28], Diamond and Verrecchia (1987)[29], Engle and Dufour (2000)[21], and Easley, Engle, O'Hara, Wu (2001)[31]). This paper addresses a point of contention between the results of Easley and O'Hara (EO) and those of Diamond and Verrecchia (DV). EO and DV have similar constructs in that one information event can occur per day and informed and uninformed agents trade according to their expectations. Different conclusions evolve in each paper as each makes a critical assumption about whether informed agents can trade on their information. In EO, informed traders can take both sides of a trade and as a result, longer trade durations (times between trades) imply no new information and spreads narrow while shorter durations imply an information event has occurred eliciting the opposite spread response. If short sellers are constrained by the uptick rule, i.e. one cannot short sell a stock until an uptick in prices is recorded, then informed traders who do not have an inventory cannot trade. Consequently, DV interpret the longer trade duration as bad news and hence, spreads should widen. Importantly, there is no overriding theory about how trade duration must affect spreads as market design, differential tax treatment, and ability to trade derivatives may all affect how time affects spreads. Hence, it's an empirical question as to what extent time between trade affects costs (see O'Hara (1995)[26]).

An empirical examination of the data indicates that time between trades, specifically *unexpected* time between trades has a secondary, though significant affect on spreads. The largest trades cost the most money, by up to a 22% premium, but changing the timing of a trade may increase the spread up to 8%. These effects may be seen in Panel B of Table IV as well as Tables V, VI, VII. To incorporate a time element into the model and to allow the empirically observed size premia to dominate, a specification such as

$$\Delta P_t = \vec{c}_0(Q_t - Q_{t-1}) + \vec{c}_1(Q_t V_t - Q_{t-1} V_{t-1}) + \vec{z}_1 Q_t V_t (1 + \varsigma_t) + e_t \qquad (4.5)$$

where $\vec{x}_i$ allows x_i to have m regimes due to $m-1$ order flow thresholds and n regimes from $n-1$ intensity thresholds, and ς_t is the proxy for unexpected trade intensity at time t.

Three different types of proxies were tested in the paper. All are functions of diurnally adjusted times between trades. The diurnal adjustment was made by regressing a cubic spline on unadjusted the times between trades during the day. 15 minute intervals were used as the knots for the spline regression.

$$\begin{aligned}\phi(\Delta t_{i-1}) &= \sum_{j=1}^{K} I_j[c_j + d_{1,j}(\Delta t_{i-1} - k_{j-1}) + d_{2,j}(\Delta t_{i-1} - k_{j-1})^2 + \\ &\quad d_{3,j}(\Delta t_{i-1} - k_{j-1})^3] \end{aligned} \tag{4.6}$$

where Δt_i is the ith unadjusted time between trades, I_j is a dummy variable with the value of 1 on the jth segment of the spline i.e. I_j is one where $k_{j-1} \leq \Delta t_{i-1} < k_j$ and 0 otherwise.

(4.6) may also be interpreted as the expected time between trades conditional on the time of day,

$$\phi(\Delta t_i) = E[\Delta t_i | \Phi(\tau)] \tag{4.7}$$

where Φ is the information set used to form expectations at time (τ) during the trading day.

The output from the spline model gives the average time between trades for every instant during the trading day. Deflate every observation by the appropriate average to get the normalized or diurnally adjusted trade duration. A simple moving average and an exponential moving average of the $\phi(\Delta t_i)$ are the first two proxies for unexpected trade durations and the last is Engle and Russell's WACD(p,q) (1998)[19]. The WACD(p,q) variable is estimated with MLE using the BHHH algorithm. This estimation involves optimizing the log likelihood of the Weibull distribution[2]

$$L(\xi, \alpha, \beta, \gamma) = \sum_{i=1}^{T} ln\left(\frac{\gamma}{\xi_i}\right) + \gamma\, ln\left(\frac{\Gamma(1+\frac{1}{\gamma})\xi_i}{\psi_i}\right) - \left(\frac{\Gamma(1+\frac{1}{\gamma})\xi_i}{\psi_i}\right)^{\gamma} \quad (4.8)$$

where γ is the Weibull parameter, $\Gamma(x)$ is the value of the Gamma distribution at the value x, and $\xi_i \equiv \frac{\Delta t_i}{\phi(\Delta t_i)}$.

(4.8) coupled with the GARCHesque expression

$$\psi_j = \omega + \sum_{j=1}^{p} \alpha_j \xi_{j-1} + \sum_{j=1}^{q} \beta_j \psi_{j-1} \quad (4.9)$$

subject to

$$\sum_{j=1}^{p} \alpha_j + \sum_{j=1}^{q} \beta_j < 1.0 \quad (4.10)$$

[2] The Weibull distribution was selected because it has a monotonic hazard function. The slope of the hazard function is governed by the so called Weibull parameter, γ. If γ is less than 1 then the hazard function is downward sloping. This is interpreted by saying as the duration from the last trade increases, the probability that a trade will come in the next instant increases. An analogous interpretation applies the to case where γ is greater than 1. Further generalization occurs if one uses the gamma distribution. With the gamma distribution, the hazard function is not constrained by monotonicity and thus may have parabolic shapes.

for the general case and subject to the following for the WACD(1,1),

$$\alpha + \beta < 1$$

$$\alpha , \beta > 0$$

Many different specifications arise for ψ_i which have all significant right hand side variables. The Bayesian Information Criterion was used to select the best specification from the collection of specifications that have all significant coefficients. The BIC takes the Weibull likelihood function

$$L(\xi|m, K, \Theta) = \prod_{i=1}^{N} L(\xi_i|m, K, \Theta) \tag{4.11}$$

where N is the number of observations for the model m, which has K parameters, and $\Theta_{m,K}$ is the parameter space for the model with $\Theta_m = \{\vec{\alpha}_m, \vec{\beta}_m, \vec{\gamma}_m\}$. Under the regularity conditions outlined in Schwartz(1978), an asymptotic approximation of the integrated log likelihood of L is

$$ln(L(\xi|m, K)) \approx ln(L(\xi|m, K, \hat{\Theta})) - \frac{\nu_{m,K}}{2} ln(N) \tag{4.12}$$

where $\hat{\Theta}$ is the maximum likelihood estimate of

$$\hat{\Theta} = \max_{\Theta} L(\xi|m, K, \Theta) \tag{4.13}$$

where $\nu_{m,K}$ is the number of free parameters for model m which has K parameters. Then, if one minimizes the Bayesian Information Criterion,

$$BIC_{m,K} = -2ln(L(\xi|m,K,\hat{\Theta})) + \nu_{m,K}ln(N) \tag{4.14}$$

where $ln(L(\xi|m,K,\hat{\Theta}))$ is the maximum log likelihood for the model m with K free parameters.

Two through 5 lags were examined for the SMA and the EMA and WACD(1,1) through WACD(3,3) inclusive were tested for the WACD. Tables IV and V indicate that the SMA(4), EMA(3) are the preferred lag lengths for the moving average proxies. The specification for the optimal lag length for the WACD was made by first searching for all specifications that had all variables significant at the 5% level. Then the specification with the smallest Bayesian Information Criterion (BIC) was used. The (1,2) specification minimized the BIC. Table VI shows the results of dissecting the data (all three data sets) using the WACD(1,2) specification. This is close to the WACD(2,1) used in Engle and Dufour (2000).

Threshold Analysis A compact representation for the nonlinear relation would be

$$Y_t = \boldsymbol{\alpha} X_t + \boldsymbol{\epsilon_t} \tag{4.15}$$

where $X_t = [(Q_t - Q_{t-1}),\ (Q_tV_t - Q_{t-1}V_{t-1}),\ Q_tV_t(1 + \varsigma_t)]$ over all regimes, and $\boldsymbol{\alpha}$ is the (m+n)×p matrix of coefficients, where p equals 3 or 4 depending on the significance of the duration threshold.

Assuming that the size of an order proxies for the amount of adverse selection in a trade or that liquidity constraints from finite depth bind, the system should exhibit thresholds with respect to signed order flow and/or trade duration. If thresholds exist, then the specification for a particular trade would be a function of the threshold vector (a vector of nuisance parameters). The generalized Glosten and Harris model would emerge as an (m,n) threshold model (which I term the $T_{m,n}$-GH) and would cause (4.3) to behave as

$$Y_t = \vec{\alpha}_1 X_t I_{1,t}(\vec{\gamma}, d) + \ldots + \vec{\alpha}_M X_t I_{M,t}(\vec{\gamma}, d) + \vec{\epsilon}_t \tag{4.16}$$

where $M = m + n$, $\vec{\gamma} = (\gamma_1, \ldots, \gamma_M)$ which is the vector of threshold values, and $I_{j,t}(\vec{\gamma}, d) = I(\gamma_{j-1} < X_t \leq \gamma_j)$. Hence, this is a model with M thresholds, or M-1 regimes, and a delay lag of d trades. It is assumed that d is zero or the threshold variable is the contemporaneous signed trade variable. Given $d = 0$, the indicator $I_{j,t}(\vec{\gamma}, d)$ will drop the d dependence and shall be expressed as a function of the threshold values, $I_{j,t}(\vec{\gamma})$. I implement Hansen's[17] method for testing whether there are one, two or three regimes both in order size and duration to determine the proper specification of the $T_{m,n}$-GH.

Testing for the significance of two threshold variables could be problematic at this point if the two variables are not independent. The Brock, Scheinkman, Dechert (BSD) model was used to test for dependence of an unspecified form. It calculates the significance level that the null of no dependence is rejected asymptotically. I adapted the test to see if the difference between the normalized trade intensity variable and the normalized signed order flow were independent. The results for each firm (each firm's p-value) were then aggregated according to the Gibbons/Shanken (1986) algorithm.

The BSD statistic uses a k dimension correlation integral to assess whether a process is random (for the purposes of this paper, k is one). To compare the signed order flow with the trade time process, I normalize each and examine the behavior of the difference between the normalized vectors. This difference, ζ_t, is

$$\zeta_t = \left(\frac{Q_t V_t - \mu_{VQ}}{\sigma_{QV}} - \frac{\xi_t - \mu_\xi}{\sigma_\xi} \right) \tag{4.17}$$

To begin the Brock, Dechert, and Scheinkman analysis, one must organize the ζ_t vector into n-histories which will be defined by the vector $\vec{\zeta}_t^n$

$$\vec{\zeta}_t^n = [\zeta_{t-n+1}, \ldots, \zeta_t] \tag{4.18}$$

where n is the so called embedding dimension. One then needs some proximity measure to ascertain which pairs of n-histories are close to each other.

The number k is the closeness metric where

$$k \text{ s.t. } \max_{i=0,\ldots,n-1} |x_{s-1} - x_{t-1}| < k. \tag{4.19}$$

This leads to defining a 'closeness indicator' function $K_{s,t}$ such that

$$K_{s,t} = \begin{cases} 1 & : \quad if\ max_{i=0,\ldots,n-1}\ i = 0, \ldots, n-1 \\ 0 & : \quad otherwise \end{cases}$$

so that $K_{s,t}$ is 1 if the n-histories are close and 0 otherwise. The fraction of pairs that are close in a sample of T such vectors $\vec{\zeta}_t^n$ will be defined to be

$$C_{n,T}(k) \equiv \frac{\sum_{s=1}^{T} \sum_{t=s}^{T} K_{s,t}}{T(T-1)/2} \tag{4.20}$$

Now let $C_n(k)$ equal the limit of this fraction as the sample size gets large. This limit is called the correlation integral, but may be interpreted as the probability that a random pair of ζ_t^n is close. Brock, Dechert, and Scheinkman (1987) show that if the data are IID, then for any n, the relationship

$$C_n(k) = C_1(k)^n \tag{4.21}$$

obtains. The usefulness of this is revealed when one interprets the ratio of $C_{n+1}(k)/C_n(k)$ as the probability of two points being close given that the previous n points are close. Thus, $C_{n+1}(k)/C_n(k)$ can be interpreted as the conditional probability

$$\frac{C_{n+1}(k)}{C_n(k)} = \mathbf{P}\left(\max_{i=0,\ldots,n-1} |x_{s-1} - x_{t-1}| < k \;\middle|\; \max_{i=1,\ldots,n-1} |x_{s-1} - x_{t-1}| < k\right)$$

$$\frac{C_{n+1}(k)}{C_n(k)} = \mathbf{P}\left(|x_s - x_t| < k \;\middle|\; \max_{i=1,\ldots,n-1} |x_{s-1} - x_{t-1}| < k\right) \tag{4.22}$$

So if the data are IID, then the conditional probability (4.22) must equal the unconditional probability that the two points are close. Brock, Dechert, and Scheinkman develop the test statistic

$$J_{n,T}(k) = \sqrt{T}\,\frac{C_{n,T}(k) - C_{1,T}(k)^n}{\hat{\sigma}_{n,T}(k)} \tag{4.23}$$

where $C_{n,T}(k)$ and $C_{1,T}(k)^n$ are the sample correlation integrals and $\hat{\sigma}_{n,T}(k)$ is the volatility of $C_{n,T}(k) - C_{1,T}(k)^n$. For large T, this statistic is standard normal under the null and has been found to have good power against common nonlinear models[32]. The results in Table VIII indicate that order flow and unexpected trade duration were asymptotically independent for all three sets of firms.

Given independence of the order flow and trade intensity, one may progress variable-wise to test for thresholds. Testing a null of one regime versus an alternative of two (or zero thresholds versus one) would compare

$$Y_t = \vec{\alpha} X_t + \epsilon_{1,t} \tag{4.24}$$

to

$$Y_t = \boldsymbol{\alpha} X_t(\gamma_1) + \boldsymbol{\epsilon_t} \tag{4.25}$$

where

$$X_t(\gamma_1) = \begin{pmatrix} X_t I_{1,t}(\gamma_1) \\ X_t I_{2,t}(\gamma_1) \end{pmatrix} \tag{4.26}$$

where $\vec{\alpha}$ is the $1\times$ p coefficient vector, $\boldsymbol{\alpha}$ is the (m+n)×p coefficient matrix, γ_1 is the value of the first threshold, $I_{1,t}(\gamma_1) = I(x_t \leq \gamma_1)$, and $I_{2,t}(\gamma_1) = I(x_t > \gamma_1)$, where x_t is the contemporaneous order flow or trade intensity.

Both regression and nuisance parameters can be consistently estimated with OLS. In the one regime or linear model (4.24), OLS determines the $\vec{\alpha}$ that minimizes the sum of squared errors. Analogously, in the two regime case (4.25), one must pick the γ_1 and the $\vec{\alpha}$ such that the squared errors are minimized.

$$S_2 = \min_{(\gamma_1, \vec{\alpha})} (Y - \vec{\alpha} X(\gamma_1))'(Y - \vec{\alpha} X(\gamma_1)) \tag{4.27}$$

where S_2 is the sum of squared errors that is minimized over a grid search of the data.

To find the threshold value and coefficients that create S_2, first rank the data by the threshold variable and then divided it into deciles. A minimum grid size of a decile of the data was selected to insure that a reasonable fraction of the total number of observations for each stock would be available to form the alternative to the null. Once the γ_1 and $\vec{\alpha}$ are selected that minimize S_2, a further dissection of the data is attempted. One collects the γ_1 which minimizes S_2 and the γ with the second lowest sum of squared errors and divides the data between the two threshold values inclusive into another 10 deciles. So if the γ_1 is 2000 shares and the threshold value with the next lowest sum of squared errors is 1000, then a new set of regressions is performed on the data with 2000, 1900, . . . , 1000 as the new breakpoints. Of course, if the difference between the threshold values only cover 400 shares, then only five points can be tested.

The test statistic to test the null of zero thresholds (one regime) against the alternate of one threshold (two regimes) is

$$F_{0,1} = T\left(\frac{S_1 - S_2}{S_2}\right) \tag{4.28}$$

where T is the total number of observations, S_2 is defined in (4.27) and S_1 is the sum of squared errors of the linear model. The asymptotic distribution of

$F_{0,1}$ is non standard because of the presence of the threshold variable. Hence, the statistic would not be distributed χ^2, and therefore one needs a method to approximate the asymptotic distribution for $F_{0,1}$.

Hansen (1996) provides an excellent overview of two approximation methods. The first method uses asymptotic distribution theory pertinent for random functions (also known as empirical process theory). Because $F_{0,1}(\gamma)$ is a random function of γ, the $F_{0,1}$ test statistic is a random maximum of $F_{0,1}(\gamma)$ over the nuisance parameter space,

$$F_{0,1} = \max_{\gamma \in \Gamma} F_{0,1}(\gamma). \tag{4.29}$$

In order to ascertain significance levels of the $F_{0,1}$ statistic, one must calculate the asymptotic distribution of the empirical process $F_{0,1}(\gamma)$. Hansen calls this random limit function $T(\gamma)$. Hansen (1996) and (1999) cover the algorithm to calculate the approximation to the asymptotic distribution, T_n, using sample moments. The significance of $F_{0,1}$ is gauged by how many points of T_n exceed the observed value of $F_{0,1}$. Computationally, this is a low cost method.

The second method for determining the significance of $F_{0,1}$ centers on a bootstrap approximation. Although the cost of using the bootstrap is high in that it requires placing structure on the residuals i.e. that the residuals are independent over time, as Hansen (1999) points out, several studies conclude that the bootstrap outperforms first order asymptotic theory as an approximation to finite sample distributions. Accordingly, I opt for the bootstrap approximation in this study.

The steps for calculating the bootstrap approximation are as follows for $F_{0,1}$: once the specification has been determined, estimate the coefficients for the model using OLS under the null of linearity and collect the residuals. Now take a data series of length p and sample with replacement from the residual vector to create generate a new sample of size n (n=10000 in this study). Calculate the value of $F_{0,1}$ from these data and then repeat 10000 times. Calculate the percentage of times that the simulated values of $F_{0,1}$ exceeds the observed $F_{0,1}$. This is the bootstrapped p-value. Aggregate p-values over all of the sample firms according the Gibbons and Shanken algorithm (1987) to arrive at a sample significance level.

Testing for three regimes versus two follows a similar procedure where the (4.26) becomes

$$X_t(\vec{\gamma}, d) = \begin{pmatrix} X_t I_{1,t}(\vec{\gamma}, d) \\ X_t I_{2,t}(\vec{\gamma}, d) \\ X_t I_{3,t}(\vec{\gamma}, d) \end{pmatrix} \tag{4.30}$$

with $\vec{\gamma} = (\gamma_1, \gamma_2)$, $\vec{\alpha} = (\vec{\alpha}_1, \vec{\alpha}_2, \vec{\alpha}_3)$, $I_{1,t}(\gamma, d) = I(x_t \leq \gamma_1)$, $I_{2,t}(\gamma, d) = I(\gamma_2 \geq x_t > \gamma_1)$, and $I_{3,t}(\gamma_1, d) = I(x_t > \gamma_2)$. The least squares minimization is much akin to (4.27) only now the regression coefficients are subject to two threshold values.

$$S_3 = \min_{(\gamma_1,\gamma_2,\vec{\alpha})} (Y - \vec{\alpha}X(\gamma,d))'(Y - \vec{\alpha}X(\gamma,d)) \tag{4.31}$$

where S_3 is the minimized sum of squared errors over the data divided into 3 regimes, γ_1 is the first signed order flow threshold and γ_2 is the second.

Once (4.31) is performed, then the relevant test statistic for two regimes versus three is

$$F_{1,2} = T\left(\frac{S_2 - S_3}{S_3}\right) \tag{4.32}$$

where $F_{1,2}$ is the test statistic for the test of one versus two thresholds. The recipe for calculating the bootstrap approximation under the regime specific homoscedasticity assumption is as follows for $F_{1,2}$: estimate the coefficients for the model using OLS under the null of two regimes (one threshold) and collect the residuals from each regime. Now, form a data matrix of length n and sample with replacement from the residual vector germane to the threshold variable regime (e.g. if the trade size meets or exceeds the threshold value, use a residual corresponding to the upper regime). Generate a new sample (y vector) of size n (n=10000) by adding the resampled error vector to the data matrix. Calculate the percentage of times that the simulated values of $F_{1,2}$ exceeds the observed $F_{1,2}$. This is the bootstrapped p-value.

Aggregate p-values as before and arrive at a sample significance level.

Any further iteration to a higher dimensional threshold vector would result in a highly problematic test statistic. This statistic would be a function of nuisance parameter which is a function of another nuisance parameter which is itself a function of the third nuisance parameter. Confidence in this degree of nested estimation is not supported by the change point or SETAR literature.

In order to carry out all of the regressions in a timely manner (to avoid the N^2 regressions for each of the 90 stocks), Bai and Perron[18] show that because the least squares estimates are consistent for $\vec{\gamma}$, a two stage iterative minimization may be utilized to estimate the threshold vector. So if $\vec{\gamma} = (\gamma_1, \gamma_2)$ then the estimate of the first threshold will be a consistent estimate of either γ_1 or γ_2 via the above minimization in (4.27). The data are then partitioned by $\hat{\gamma}_1$ and another grid search takes place to calculate the value of $\hat{\gamma}_2$. This second threshold estimate will be a consistent estimate of the remaining threshold value in the pair (γ_1, γ_2).

Once the values for $\hat{\gamma}_1$ and $\hat{\gamma}_2$ have been determined, one should perform further iterations on the data to assure consistency of the threshold estimates. For example, the algorithm for calculating the threshold values for a stock with trade deciles at 100, 200,. . ., 1000 (only buys in the data set) would estimate $\hat{\gamma}_1$ and $\hat{\gamma}_2$. Suppose that these values were 200 and 800 respectively.

Next, iterate on the first threshold, take the data from 100 to 800 and rerun the estimation for $\hat{\gamma}_1$ and then take all of the data that have trades from that value (the refined guess for γ_1) and above and rerun the estimation of $\hat{\gamma}_2$. Then do this iteration one more time over both $\hat{\gamma}_1$ and $\hat{\gamma}_2$. Third round estimates (two iterations) were used in this study.

Results

The average trade size, quoted and effective spreads for all three sets of firms are presented in Panel A of Table I. These averages (and all subsequent analysis) were calculated over different horizons to adjust for the differing frequency of trading over the three market value deciles. Average trade size differed by 400-500 shares from small to medium and then from the medium to large firms. The quoted spread differed by about 1.5 cents in a similar manner while the effective spread changed by only 0.2 cents. Panel B shows the paradoxical result of larger firms having larger amounts of adverse selection, an unfortunate consequence of having only one quantifiable variable for adverse selection in a trade. As Van Ness et. al. ([14]) point out, adverse selection is a very problematic element to identify and capture.

Table II shows the full and reduced Glosten and Harris model. Results from all three data sets on the full model suggest that the intercept term on the variable cost component is not significant and was dropped from all further analysis. The regression with the reduced specification indicates that fixed cost and variable costs increase with firm size and that there is evidence of some economies of scale in the order processing mechanism (the negative value on the c_1 coefficient indicates that the trades have decreasing costs to scale).

Three proxies were selected to capture the temporal adverse selection present in the order flow. All three are functions of the diurnally adjusted time between trades. Little formal structure exists in selecting the simple and exponential moving averages used in this paper, but the WACD specification was deduced in a straightforward manner. Table III contains all of the aggregated estimation results for all possible specification from WACD(1,1) to WACD(3,3) inclusive. The small firms had five specifications that had each variable statistically significant, the medium and large firms had three and five, respectively. The WACD(1,2) specification minimized the Bayesian Information Criterion for all three sets of stocks. The specification used by Engle and Dufour (2000)(WACD(2,1)) placed second in this analysis.

In Glosten and Harris' original paper, expected price changes were the sum of fixed costs, proxied by the coefficient on the change in the trade sign indicator, and by variable costs measured by the coefficient on the signed order volume. Table IV clearly shows a trend in the data that was captured in the modified functional form used in this note. For all three sets of stocks, the quoted spread monotonically increased as trade size increases. The effective spread forms a general "U" pattern, heavily suggesting that lower marginal costs reduce the spread until adverse selection concerns drive the spread back up. A breakdown by trade category confirms the robustness of the results.

Microstructure theory suggests that a temporal premium exists for trades. This could arise because of finite depth at any point in time, or because of concern that a privately informed agent is trading. Table V shows the relationship between one proxy for unexpected trade intensity, the simple moving average (SMA) and quoted and effective spreads. Only the simple moving average specification with lags 2 through 5 were featured in Table V in the interest of parsimony. The SMA(4) appears to fit the data the best. The criterion of establishing a good fit is one of observing a monotonic relationship, conditional on a given volume, between trade intensity and effective spreads. Table VI displays the same information for the exponential moving average proxy. A lag length of 3 was selected along the same criterion. Table VII has the results of the same type of analysis for the WACD(1,2) variable. It, like the moving average variables, has the same general monotonic relationship between quoted and effective spreads.

Table VIII has the results from the Brock, Scheinkman, Dechert (BSD) test for dependence between the trade intensity variables and signed order flow. This model calculates the significance level that the null of no dependence is rejected asymptotically. It uses the sample correlation integral and an estimator of the asymptotic standard deviation to ascertain whether the observations are asymptotically IID. I adapted the test to see if the difference between the normalized trade intensity variable and the normalized signed order flow were independent i.e. to see whether the threshold covariance kernel is diagonal. If cross correlations were present, then extra simulations would have to be performed to establish conditional threshold values. The results for each firm (each firm's p-value) were aggregated according to the Gibbons/Shanken (1986) algorithm. The results from the BSD analysis indicate no dependence between the order flow and duration for each proxy for all three sets of firms.

Before the threshold analysis can attach a significance to the threshold values for each firm, one must adjust the bootstrap F-test to account for heteroscedasticity in the data. The data were tested for heteroscedasticity over the entire data set and for regime specific heteroscedasticity. Table IX has the results from White's test over all of the data. The p-values from each Wald F-test were aggregated like the BSD statistics above. Table X features a similar test of the presence of heteroscedasticity in each order flow regime (as divided by the first order flow threshold). White's test confirms that regime specific heteroscedasticity of an unknown form exists in the data. A heteroscedasticity consistent bootstrap F-test must be used in establishing threshold significance.

A summary of the threshold analysis appears in Table XI. For the small firms, two thirds of the firms have first order flow thresholds that are negative. This falls in line with theory as risk averse agents would respond to perceived sell signals faster than to those of buy signals. Of those with negative first thresholds, about 2/3s of those have positive second thresholds, possibly reflecting a mildly symmetrical outlook that risk averse agents have in the market for smaller, less followed stocks. Two thirds of the small stocks also have first trade time thresholds that are less than one. Most of the second thresholds (of the firms with first thresholds less than one) are greater than one. Of firms with their first threshold greater than one, 90% of the second thresholds are less than unity. This would all imply that for small stocks (firms thought to have greater amounts of asymmetric information) price changes are more sensitive for sell orders and to orders demanding greater immediacy, conclusions in accordance with theory.

The medium sized firms tell a slightly different story. While 18 of the 30 stocks have negative first order thresholds, first time thresholds are split even between those larger and smaller than unity. About half of the firms with negative first order flow thresholds have positive second thresholds, but 10 of the 12 firms with positive first thresholds have negative second thresholds. The second time thresholds follow a similar pattern to those of the small firms: a third of the firms with their first threshold less than one have second thresholds greater than one, but of the 15 firms with a first thresholds greater than one, 14 have second thresholds less than one. Hence of the total number of temporal threshold parameters, most are less than one, which empirically demonstrates the data driven relationship between order urgency and price premium. The temporal threshold results support conclusions of Foster and Viswanathan (1990)[30].

Oddly, the results for the larger firms are very similar to those of the smaller firms. Over two thirds of the firms have first order thresholds that are negative, but most of those have second thresholds that are positive. 90% of the firms with positive first thresholds have negative second threshold values. Thus, the same symmetry over order flow thresholds levels exists in the larger, highly covered firms. The temporal thresholds hold a parity with the smaller firms. Most have first thresholds that are less than one and about 2/3s of those firms have second thresholds greater than one. Of the firms with first thresholds greater than one, over 90% have second thresholds less than one. The same pattern of order immediacy driven price premium found in the smaller firms is observed here also.

Table XII features the ratios of the root mean squared errors (RMSE) from the out of sample data for each firm. For each firm, coefficients and thresholds were calculated previously and applied to the new data. Like the initial data set, the horizon for the out of sample data was a function of the expected frequency of the size cohort. The small firms, having the lowest trading frequency, had the longest horizon of one month, the medium firms had two trading weeks and the large firms one trading week. The SMA had smaller RMSEs than the EMA and the WACD for the small and medium firms, but finished last in the large firm out of sample data. Although the SMA was closely edged out by the WACD, the EMA RMSE was over 2% smaller than that of the SMA. Consequently, overall the SMA appears to forecast best of the three variables that proxy for unexpected trade intensity, but it does not strictly dominate the others.

Conclusion

The data statistically support a nonlinear dynamical relationship between price changes and order size and intensity in an auction market. This aligns itself with microstructure theory insofar as it confirms that economically significant premia are demanded by liquidity providers under less than optimal market conditions. As well, it supports the notion that one would minimize trading costs if one were to trade medium size orders in a 1normal' trading environment for that particular stock. The threshold model also forecasted better than the linear counterpart. The results indicate that the informational role of time is such that no news or longer trade durations results in larger spreads as do smaller durations, a finding that supports Diamond and Verecchia (1987).

Signed order flow, on either side of the transaction, causes prices to move by additional amounts, ceteris paribus. This supports Kempf and Korn's[20] (1999) more general finding of a nonlinear relationship between the order size and its price impact. The magnitude of the order size needed to increase the price change was much smaller than initially thought (for both buys and sells). Several proxies for unexpected trade intensity were examined and the simple moving average forecasted best out of sample. All of the proxies had sizeable lag lengths which suggests that agents must have some confirmation that the unexpected trading is driven by something other than market noise. No evidence of dependence between order size and duration greatly simplified threshold calculations and their significance tests.

Table I

Summary of Trades from the 10th, 8th and 6th Deciles of the NYSE during 1996

The data in Table I were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. The TAQ stocks were filtered for trades only on the NYSE. Additionally, the sample was filtered for trades executed during regular trading hours and for price movements smaller than 10% of the previous share price. The entire sample contains 312,389 observations. Panel A contains the average trade size, the effective and quoted spread, percentage buys/sells as computed form the Lee and Ready (1991) algorithm. Panel B shows the effective spread estimates from the Glosten and Harris (1988) spread decomposition model. The model breaks down the effective spread into fixed and variable cost components. The full model specification presents the variable and fixed cost components as linear functions of signed order flow and the change in signed order flow, respectively. The variable cost component is considered a proxy for the adverse selection component of the spread. The fixed cost component may be interpreted as having a constant cost per share (if only the intercept is significant) or an increasing or decreasing cost per share (if the coefficient on the change in signed order flow is positive or negative, respectively). The effective spread estimate was calculated for each firm by adding the adverse selection component per share*average number of shares per trade to the fixed cost component. Panel B has the paradoxical result of larger adverse selection present in larger stocks. This finding is at variance with the notion that the increased analyst following of larger stocks would result in lower adverse selection.

Panel A: Average Trade Characteristics

		Overall	Buys	Sells
Trade Size	Small	1575	1627	1559
	Medium	2074	2220	2028
	Large	2396	2533	2289
Quoted Spread	Small	0.1969	0.1977	0.1966
	Medium	0.1805	0.1804	0.1809
	Large	0.1651	0.1650	0.1653
Effective Spread	Small	0.1242	0.1238	0.1246
	Medium	0.1225	0.1221	0.1228
	Large	0.1221	0.1221	0.1220
Percentage of Trades Buy:	Small	50.29%		
	Medium	50.64%		
	Large	48.97%		

Panel B : Estimates from Glosten Harris Spread Decomposition Model

GH		Effective Spread	Fixed Cost	Adverse Selection
	Small	0.0902	0.0891	0.0010
	Medium	0.0929	0.0918	0.0012
	Large	0.0956	0.0941	0.0015

Table II

Effective Spread Estimation of NYSE Stocks from the 10th, 8th and 6th Deciles

The data in Table II were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. Panel A shows the full GH model regression results that were aggregated over each size cohort. The 'c' and 'z' coefficients pertain to the fixed and variable components of the spread, respectively. A negative sign on c(1) would be interpreted as declining costs per share as the number of shares increase. The reduced form specification (the final specification for the paper) is shown in Panel B. The reduced form coefficients come from rerunning the Glosten and Harris model without the z(0) or fixed cost argument of the variable cost part of the effective spread.

Panel A: Full Model Glosten and Harris (1988) Decomposition

		c(0)	c(1)*	z(0)	z(1)*
SMALL	**Betas**	0.0455	-0.0005	-0.0018	0.0010
	t	8.0	-2.8	0.6	8.6
MEDIUM		**c(0)**	**c(1)***	**z(0)**	**z(1)***
	Betas	0.0464	-0.0002	-0.0008	0.0005
	t	7.2	-1.9	0.0	2.5
LARGE		**c(0)**	**c(1)***	**z(0)**	**z(1)***
	Betas	0.0468	-0.0003	0.0005	0.0006
	t	8.7	-5.3	1.0	6.3

Panel B: Reduced Model Glosten Harris (1988) Decomposition

		c(0)	c(1)*	z(1)*
SMALL	**Betas**	0.0446	-0.0003	0.0007
	t	75.6	-2.1	5.0

		c(0)	c(1)*	z(1)*
MEDIUM	**Betas**	0.0459	-0.0004	0.0003
	t	50.2	-2.0	2.3

		c(0)	c(1)*	z(1)*
LARGE	**Betas**	0.0470	-0.0003	0.0006
	t	89.8	-4.3	10.9

* indicates the coefficient is multiplied by 1000

Table III

Weibull Autoregressive Conditional Duration Specification

The data in Table III were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. Panel A shows the MLE estimates of the Engle's (1996) WACD model for the small stock cohort. Panel B and C show those of the medium and large stocks respectively. The specification examination was conducted over all specifications from WACD(1,1) to WACD(3,3) inclusive. The estimates for the (1,1) specification were constrained such that each coefficient had to be greater than zero and the sum had to be less than one. The other specifications were constrained to make sure that the total of the coefficients all added to less than unity (an arbitrary cutoff of 0.999998) was used). Specifications with all significant coefficients are boldfaced and the chosen specification for the paper is italicized. The chosen specification, WACD(1,2), was elected on the criteria that it has all significant coefficients and it must minimizes the Baynesian Information Criterion (BIC).

Panel A: Small Stock Sample

	intercept	**alpha(1)**	**alpha(2)**	**alpha(3)**	**beta(1)**	**beta(2)**	**beta(3)**
WACD(3,3)	0.091964	0.069568	0.048007	0.015888	0.272611	0.008653	0.494734
	2.5022545	8.12942	4.2256065	1.4167098	1.5144348	0.0474417	5.1634597
WACD(3,2)	0.062147	0.076963	0.022667	-0.01723	0.309514	0.546218	
	1.8454136	8.5974951	2.0149181	-2.5197282	2.8077781	5.3719927	
WACD(3,1)	0.054955	0.075587	-0.00573	-0.01579	0.890447		
	1.4638404	8.3939434	-0.7997369	-3.0757152	20.465109		
WACD(2,3)	0.031463	0.073389	-0.00217		0.665369	-0.06871	0.302238
	4.2320734	8.081479	-0.1535578		4.4521603	-0.5798248	4.5432612
WACD(2,2)	0.048495	0.063287	0.017117		0.443843	0.428168	
	3.1364808	7.4466091	1.3711582		3.8392321	3.939682	
WACD(2,1)	0.03128	0.075724	-0.0209		0.914708		
	2.5758695	8.1793442	-2.8965406		63.802521		
WACD(1,3)	0.177705	0.068379			0.732983	-0.22723	0.225255
	1.296849	8.6736429			6.4322883	-1.9108796	3.210444
WACD(1,2)	*0.027319*	*0.072228*			*0.658382*	*0.24396*	
	4.4726415	*8.6152559*			*10.186075*	*3.9543506*	
WACD(1,1)	0.02858	0.059637			0.912277		
	4.0757843	9.31954			86.296487		

Panel B: Medium Stock Sample

	intercept	alpha(1)	alpha(2)	alpha(3)	beta(1)	beta(2)	beta(3)
WACD(3,3)	0.057062	0.058019	0.017981	0.021889	0.342618	-0.13313	0.635962
	5.3603757	6.187772	1.5667462	2.9116828	1.9360703	-0.782637	7.7365929
WACD(3,2)	0.03444	0.068551	0.012566	-0.01354	0.321326	0.57701	
	5.6789619	7.2333131	1.0432649	-2.1908462	3.3739601	6.485217	
WACD(3,1)	0.07948	0.066901	-0.00388	-0.01333	0.871674		
	1.355884	7.0041731	-0.5231447	-2.5154345	13.497759		
WACD(2,3)	0.029672	0.067025	-0.01657		0.681987	-0.09854	0.336728
	4.5364739	7.1649308	-1.2383001		3.9789793	-0.7200708	5.1431597
WACD(2,2)	0.034182	0.059926	0.00179		0.452073	0.452502	
	6.1198385	6.4918911	0.1400282		4.3901329	4.7051211	
WACD(2,1)	0.050914	0.067529	-0.01961		0.901375		
	1.7147698	6.9146245	-2.7756152		27.961254		
WACD(1,3)	0.026395	0.061156			0.788029	-0.21262	0.338152
	6.4819744	7.5493606			7.4434239	-1.5889101	5.4674758
WACD(1,2)	*0.024638*	*0.056797*			*0.691537*	*0.228057*	
	6.287102	*7.3450913*			*9.8124127*	*3.3324754*	
WACD(1,1)	0.126693	0.049959			0.926689		
	1.7265233	7.1617887			125.43382		

Panel C: Large Stock Sample

	intercept	alpha(1)	alpha(2)	alpha(3)	beta(1)	beta(2)	beta(3)
WACD(3,3)	0.035345	0.073175	0.014378	-0.01449	0.436566	0.348707	0.105362
	4.1006645	18.499844	1.4060163	-1.4795418	3.1443061	2.61466	0.8806078
WACD(3,2)	0.022617	0.084297	-0.00911	-0.02594	0.554906	0.373308	
	6.2098358	21.39337	-0.9450415	-4.6590518	4.3600339	3.1041813	
WACD(3,1)	0.013343	0.084485	-0.02799	-0.02357	0.954002		
	9.2990118	20.894726	-7.5104547	-9.9006129	355.49001		
WACD(2,3)	0.02804	0.082114	-0.03984		1.061847	-0.23131	0.099202
	2.9073882	20.690964	-3.6034893		7.8784766	-2.2397404	2.1072063
WACD(2,2)	0.023274	0.070928	-0.02188		0.842756	0.084497	
	4.1880025	16.719159	-1.9075912		5.7377506	0.6270714	
WACD(2,1)	0.016746	0.08485	-0.04716		0.945626		
	9.9774759	20.936517	-11.677789		282.86542		
WACD(1,3)	0.030849	0.068963			0.612546	0.01057	0.277314
	9.5417536	19.089943			6.9093149	0.0879075	6.2277069
WACD(1,2)	*0.030445*	*0.063174*			*0.516592*	*0.389897*	
	9.3036121	*19.395934*			*14.268038*	*11.04317*	
WACD(1,1)	0.02253	0.046087			0.931334		
	8.7725383	18.160516			197.04142		

Table IV

Trade Size, Quoted and Effective Spread by Trade Quintile and Category

The data in Table IV were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. Panel A of Table IV contains trade size and effective and quoted spread statistics for the signed order flow quintiles. The sign of the trade was determined using Lee and Ready's (1991) algorithm. The quoted spread is the inside spread (best offer-best bid), and the effective spread is the average difference between execution price and inside quote midpoint. In Panel A, but for one observation, for each of the three sets of firms, the quoted spread increases by trade quintile. The effective spread summaries tell a much different story. The effective spread has a U-shaped pattern over the trade quintiles. For the small and medium firms, the effective spread *decreases* by quintile until the third and then increases. The smallest quintile for the small and medium firms has the largest effective spread, perhaps indicating decreasing fixed costs per share. The large firms have the same U-shaped pattern with the largest trade quintile having the largest effective spread, possibly indicating that the decreasing fixed costs were dominated by increased variable costs (adverse selection). In Panel B, the breakdown by trade category shows that each set of firms has a monotonic increase in the quoted spread as trade size increases. The smallest firms have a mildly monotonically decreasing effective spread until the largest category where the spread implies a 2.5 cent or 17% premium for larger trades. The medium sized firms have a mostly increasing effective spread. The largest firms' effective spreads have a near perfect U-shape from the 1st through 5th set. The trades over 10,000 have a penny or 7% premium over the previous trade category.

Panel A: Breakdown by Trade Quintile

I.	SMALL	MEDIUM	LARGE
	MEAN	MEAN	MEAN
TRADE SIZE	100	100	100
QUOTED SPR	0.1693	0.1587	0.1452
EFFECTIVE SPR	0.1534	0.1476	0.1414

II.	SMALL	MEDIUM	LARGE
	MEAN	MEAN	MEAN
TRADE SIZE	200	200	269
QUOTED SPR	0.1687	0.1592	0.1457
EFFECTIVE SPR	0.1497	0.1457	0.1407

III.	SMALL	MEDIUM	LARGE
	MEAN	MEAN	MEAN
TRADE SIZE	410	427	614
QUOTED SPR	0.1719	0.1614	0.1504
EFFECTIVE SPR	0.1477	0.1444	0.1414

IV.	SMALL	MEDIUM	LARGE
	MEAN	MEAN	MEAN
TRADE SIZE	903	1057	1452
QUOTED SPR	0.1757	0.1683	0.1545
EFFECTIVE SPR	0.1481	0.1459	0.1417

V.	SMALL	MEDIUM	LARGE
	MEAN	MEAN	MEAN
TRADE SIZE	5958	9769	9103
QUOTED SPR	0.1819	0.1673	0.1625
EFFECTIVE SPR	0.1475	0.1431	0.1439

Panel B: Breakdown by Trade Category

	SMALL	MEDIUM	LARGE
	MEAN	MEAN	MEAN
100-500 SHARES			
TRADE SIZE	264	255	253
QUOTED SPR	0.1805	0.1633	0.1486
EFFECTIVE SPR	0.1544	0.1474	0.1414
501-1,000			
TRADE SIZE	848	844	867
QUOTED SPR	0.1911	0.1725	0.1529
EFFECTIVE SPR	0.1535	0.1471	0.1400
1,001-2,500			
TRADE SIZE	1703	1713	1773
QUOTED SPR	0.1985	0.1766	0.1552
EFFECTIVE SPR	0.1548	0.1489	0.1392
2,501-5,000			
TRADE SIZE	3740	3752	3861
QUOTED SPR	0.2045	0.1879	0.1625
EFFECTIVE SPR	0.1528	0.1491	0.1400
5,001-10,000			
TRADE SIZE	7501	7626	7760
QUOTED SPR	0.2147	0.1905	0.1701
EFFECTIVE SPR	0.1508	0.1528	0.1413
>10,000			
TRADE SIZE	40877	43498	30162
QUOTED SPR	0.2169	0.1942	0.1769
EFFECTIVE SPR	0.1763	0.1633	0.1513

Table V

Order Flow Deciles Dissected by Simple Moving Average Intensity Quintiles

The data in Table V were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. Table V contains each set of stocks broken down first by trade deciles and then further by unexpected trade intensity quintiles. For the buy and the sell orders (as assigned by the Lee and Ready (1991) algorithm) in each decile, mean trade size, quoted and effective spreads were calculated. The measure of unexpected trade intensity is captured by a simple moving average, SMA, (of the present trade and the previous 2-5 trades) of the diurnally adjusted (where diurnally adjusting involves taking the time between trades and then deflating that by the expected time between trades conditioned on the time of day) times between trades. The simple moving average lag of 4 captured the sharpest results in the data while remaining parsimonious (only lags 1-5 were considered). The moving average variable (MA QUINTILE) quintile breakpoints are shown to the left, with the averages of the quoted or inside spread (SPREAD), trade size (SIZE), and effective spread (EFFSPR) to the right. For the small, medium and large sized stocks, there exists a general increase in effective spread as trade intensity (as proxied by the SMA intensity) increases, potentially indicating an increase in the adverse selection component of the spread. Panels A through D display the trade size / trade intensity relationship for all sets of stocks. Each Panel features the results from the small, medium, followed by those of the large stocks. Microstructure theory predicts that, all else equal, order processing has a temporal premium. Consequently, a pattern should arise such that the quoted and effective spreads should increase as unexpected time between trades decreases. This general pattern appears over all sizes of stocks in the sample using 4 specifications of simple moving average (SMA). The optimal lag length was chosen as the one with the most consistent, monotonic pattern across all stocks in the sample.

Panel A: SMA(2)

SMALL

BUYS					SELLS				
MEAN TRADE	6754	MEAN SPR	0.1979		MEAN TRADE	-6950	MEAN SPR	0.2014	
		MEAN EFF	0.1070				MEAN EFF	0.1065	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.0885		0.1992	6119	0.1153	0.1032		0.2120	-6641	0.1216
0.2969		0.1977	7240	0.1083	0.3410		0.2031	-7074	0.1081
0.5976		0.1990	7017	0.1073	0.6641		0.2006	-7153	0.1046
1.0718		0.1990	6855	0.1018	1.1585		0.1960	-7096	0.0990
2.5014		0.1945	6542	0.1021	2.6406		0.1956	-6776	0.0992

MEAN TRADE	1021	MEAN SPR	0.1896		MEAN TRADE	-1021	MEAN SPR	0.1882	
		MEAN EFF	0.1199				MEAN EFF	0.1193	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.0865		0.1986	1023	0.1246	0.1026		0.1983	-1024	0.1261
0.3080		0.1897	1027	0.1201	0.3410		0.1885	-1018	0.1197
0.6165		0.1888	1020	0.1175	0.6616		0.1861	-1020	0.1202
1.0905		0.1876	1018	0.1204	1.1635		0.1848	-1019	0.1158
2.5025		0.1834	1016	0.1168	2.6126		0.1834	-1022	0.1150

MEAN TRADE 487 MEAN SPR 0.1847
MEAN EFF 0.1242

MA QUINTILE	SPREAD	SIZE	EFFSPR
0.1008	0.1908	490	0.1293
0.3373	0.1867	486	0.1269
0.6522	0.1842	487	0.1222
1.1222	0.1820	488	0.1225
2.5302	0.1799	487	0.1199

MEAN TRADE -522 MEAN SPR 0.1901
MEAN EFF 0.1229

MA QUINTILE	SPREAD	SIZE	EFFSPR
0.1027	0.2027	-520	0.1320
0.3432	0.1927	-521	0.1271
0.6596	0.1856	-522	0.1198
1.1581	0.1888	-522	0.1179
2.5573	0.1823	-523	0.1186

MEAN TRADE 239 MEAN SPR 0.1765
MEAN EFF 0.1325

MA QUINTILE	SPREAD	SIZE	EFFSPR
0.1155	0.1852	241	0.1364
0.3706	0.1755	238	0.1329
0.6993	0.1737	239	0.1322
1.1736	0.1742	240	0.1300
2.5286	0.1738	238	0.1311

MEAN TRADE -341 MEAN SPR 0.1741
MEAN EFF 0.1278

MA QUINTILE	SPREAD	SIZE	EFFSPR
0.1046	0.1809	-341	0.1325
0.3440	0.1733	-341	0.1271
0.6654	0.1759	-341	0.1295
1.1671	0.1718	-342	0.1272
2.5795	0.1698	-341	0.1236

MEAN TRADE 100 MEAN SPR 0.1788
MEAN EFF 0.1371

MA QUINTILE	SPREAD	SIZE	EFFSPR
0.1201	0.1826	100	0.1403
0.3895	0.1761	100	0.1352
0.7167	0.1842	100	0.1437
1.2026	0.1768	100	0.1353
2.5901	0.1742	100	0.1310

MEAN TRADE -144 MEAN SPR 0.1745
MEAN EFF 0.1323

MA QUINTILE	SPREAD	SIZE	EFFSPR
0.1051	0.1816	-145	0.1362
0.3441	0.1755	-144	0.1348
0.6669	0.1746	-145	0.1327
1.1641	0.1735	-144	0.1308
2.5516	0.1690	-143	0.1286

MEDIUM

BUYS					SELLS				
MEAN TRADE	9855	MEAN SPR	0.1815		MEAN TRADE	-9697	MEAN SPR	0.1876	
		MEAN EFF	0.1104				MEAN EFF	0.1113	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1033		0.1745	8015	0.1188	0.1091		0.1845	-7937	0.1234
0.3112		0.1801	11227	0.1120	0.3312		0.1877	-9548	0.1145
0.6086		0.1842	10717	0.1105	0.6404		0.1875	-9893	0.1094
1.0896		0.1858	9233	0.1043	1.1277		0.1911	-11218	0.1045
2.5245		0.1829	10084	0.1065	2.6170		0.1870	-9890	0.1046

MEAN TRADE	1141	MEAN SPR	0.1822		MEAN TRADE	-1119	MEAN SPR	0.1827	
		MEAN EFF	0.1200				MEAN EFF	0.1189	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1120		0.1863	1139	0.1265	0.1067		0.1943	-1123	0.1309
0.3407		0.1813	1135	0.1223	0.3327		0.1825	-1119	0.1210
0.6449		0.1836	1143	0.1178	0.6400		0.1824	-1113	0.1182
1.1211		0.1770	1153	0.1164	1.1341		0.1788	-1123	0.1127
2.5472		0.1827	1135	0.1171	2.5681		0.1769	-1120	0.1135

MEAN TRADE	513	MEAN SPR	0.1771		MEAN TRADE	-512	MEAN SPR	0.1743	
		MEAN EFF	0.1238				MEAN EFF	0.1230	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1231		0.1788	513	0.1296	0.1112		0.1808	-515	0.1316
0.3632		0.1787	516	0.1254	0.3298		0.1749	-509	0.1242
0.6808		0.1763	512	0.1198	0.6443		0.1737	-511	0.1207
1.1607		0.1785	511	0.1196	1.1321		0.1731	-511	0.1209
2.5135		0.1734	512	0.1244	2.5130		0.1707	-513	0.1202

MEAN TRADE	239	MEAN SPR	0.1675		MEAN TRADE	-300	MEAN SPR	0.1655	
		MEAN EFF	0.1298				MEAN EFF	0.1289	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1405		0.1669	241	0.1322	0.1150		0.1722	-300	0.1335
0.4047		0.1667	238	0.1287	0.3333		0.1671	-300	0.1295
0.7290		0.1706	239	0.1296	0.6451		0.1643	-300	0.1280
1.2026		0.1659	238	0.1289	1.1282		0.1658	-300	0.1274
2.5275		0.1675	237	0.1295	2.5038		0.1603	-300	0.1275

MEAN TRADE	100	MEAN SPR	0.1681		MEAN TRADE	-144	MEAN SPR	0.1658	
		MEAN EFF	0.1342				MEAN EFF	0.1314	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1540		0.1697	100	0.1361	0.1149		0.1741	-147	0.1382
0.4249		0.1684	100	0.1335	0.3342		0.1684	-144	0.1334
0.7521		0.1650	100	0.1327	0.6408		0.1648	-144	0.1301
1.2289		0.1694	100	0.1331	1.1338		0.1627	-143	0.1282
2.6097		0.1682	100	0.1354	2.4908		0.1628	-143	0.1301

LARGE

BUYS					SELLS				
MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1688		0.1747	7844	0.1291	0.1828		0.1758	-7572	0.1328
0.3911		0.1818	8599	0.1164	0.4189		0.1797	-8247	0.1203
0.6708		0.1832	8935	0.1127	0.7087		0.1811	-8962	0.1137
1.1072		0.1853	9378	0.1081	1.1590		0.1823	-9002	0.1074
2.4407		0.1823	10160	0.1031	2.5288		0.1811	-10080	0.1032

MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1545		0.1651	1553	0.1298	0.1770		0.1646	-1781	0.1299
0.3661		0.1695	1570	0.1240	0.4175		0.1673	-1789	0.1235
0.6318		0.1698	1559	0.1229	0.7090		0.1683	-1788	0.1203
1.0581		0.1702	1573	0.1197	1.1597		0.1685	-1776	0.1157
2.3754		0.1688	1572	0.1184	2.5073		0.1659	-1773	0.1145

MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1549		0.1605	611	0.1328	0.1744		0.1606	-877	0.1309
0.3812		0.1644	608	0.1285	0.4175		0.1636	-874	0.1265
0.6659		0.1623	610	0.1264	0.7061		0.1635	-871	0.1235
1.1138		0.1629	615	0.1270	1.1618		0.1635	-868	0.1211
2.4547		0.1589	617	0.1257	2.5277		0.1612	-863	0.1201

MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1797		0.1536	272	0.1343	0.1738		0.1569	-415	0.1335
0.4377		0.1549	270	0.1318	0.4207		0.1564	-412	0.1297
0.7461		0.1537	268	0.1299	0.7094		0.1555	-409	0.1272
1.2136		0.1531	267	0.1306	1.1616		0.1565	-409	0.1261
2.5835		0.1545	268	0.1307	2.5248		0.1559	-406	0.1241

MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1891		0.1497	100	0.1359	0.1748		0.1503	-140	0.1355
0.4628		0.1512	100	0.1332	0.4215		0.1514	-140	0.1323
0.7865		0.1516	100	0.1326	0.7110		0.1510	-139	0.1315
1.2679		0.1516	100	0.1329	1.1655		0.1519	-138	0.1308
2.6688		0.1540	100	0.1341	2.5294		0.1517	-138	0.1301

Panel B: SMA(3)

SMALL

BUYS					SELLS				
MEAN TRADE	6754	MEAN SPR	0.1979		MEAN TRADE	-6950	MEAN SPR	0.2014	
		MEAN EFF	0.1070				MEAN EFF	0.1065	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1373		0.1997	6990	0.1135	0.1601		0.2102	-6617	0.1186
0.3840		0.1996	6749	0.1079	0.4293		0.2037	-7369	0.1093
0.6800		0.1973	6401	0.1060	0.7363		0.1994	-6492	0.1048
1.1143		0.1970	7173	0.1045	1.1799		0.1981	-6928	0.1002
2.3074		0.1957	6462	0.1029	2.4059		0.1958	-7334	0.0998

MEAN TRADE	1021	MEAN SPR	0.1896		MEAN TRADE	-1021	MEAN SPR	0.1882	
		MEAN EFF	0.1199				MEAN EFF	0.1193	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1374		0.1980	1025	0.1246	0.1584		0.1964	-1022	0.1247
0.3971		0.1916	1022	0.1189	0.4302		0.1898	-1023	0.1241
0.6971		0.1888	1020	0.1197	0.7377		0.1845	-1016	0.1170
1.1163		0.1850	1019	0.1175	1.1849		0.1879	-1019	0.1154
2.3063		0.1848	1019	0.1185	2.3504		0.1827	-1024	0.1159

MEAN TRADE	487	MEAN SPR	0.1847		MEAN TRADE	-522	MEAN SPR	0.1901	
		MEAN EFF	0.1242				MEAN EFF	0.1229	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1537		0.1900	488	0.1305	0.1603		0.2013	-520	0.1305
0.4250		0.1879	487	0.1251	0.4291		0.1913	-521	0.1248
0.7299		0.1836	487	0.1226	0.7410		0.1886	-522	0.1209
1.1501		0.1812	487	0.1230	1.1803		0.1891	-522	0.1191
2.3080		0.1808	488	0.1197	2.3209		0.1812	-523	0.1195

MEAN TRADE	239	MEAN SPR	0.1765		MEAN TRADE	-341	MEAN SPR	0.1741	
		MEAN EFF	0.1325				MEAN EFF	0.1278	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1708		0.1829	240	0.1357	0.1620		0.1812	-341	0.1311
0.4579		0.1754	240	0.1332	0.4310		0.1745	-339	0.1271
0.7678		0.1768	239	0.1311	0.7417		0.1733	-342	0.1302
1.1863		0.1744	239	0.1312	1.1848		0.1737	-341	0.1274
2.3103		0.1730	238	0.1313	2.3228		0.1692	-341	0.1238

MEAN TRADE	100	MEAN SPR	0.1788		MEAN TRADE	-144	MEAN SPR	0.1745	
		MEAN EFF	0.1371				MEAN EFF	0.1323	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1805		0.1831	100	0.1395	0.1654		0.1824	-144	0.1365
0.4694		0.1758	100	0.1366	0.4321		0.1761	-145	0.1339
0.7750		0.1842	100	0.1429	0.7397		0.1735	-144	0.1320
1.2064		0.1757	100	0.1340	1.1801		0.1737	-144	0.1316
2.3504		0.1752	100	0.1324	2.3393		0.1688	-144	0.1289

MEDIUM

BUYS					SELLS				
MEAN TRADE	9855	MEAN SPR	0.1815		MEAN TRADE	-9697	MEAN SPR	0.1876	
		MEAN EFF	0.1104				MEAN EFF	0.1113	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1463		0.1728	8572	0.1162	0.1597		0.1831	-8939	0.1230
0.3895		0.1832	10397	0.1135	0.4163		0.1864	-9831	0.1139
0.6848		0.1834	10719	0.1107	0.7229		0.1913	-9517	0.1093
1.1204		0.1854	9326	0.1047	1.1571		0.1910	-9776	0.1058
2.3295		0.1826	10261	0.1071	2.3849		0.1859	-10426	0.1044
MEAN TRADE	1141	MEAN SPR	0.1822		MEAN TRADE	-1119	MEAN SPR	0.1827	
		MEAN EFF	0.1200				MEAN EFF	0.1189	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1665		0.1852	1139	0.1285	0.1619		0.1912	-1121	0.1319
0.4225		0.1835	1133	0.1216	0.4196		0.1834	-1123	0.1207
0.7198		0.1798	1150	0.1168	0.7225		0.1813	-1117	0.1174
1.1373		0.1808	1144	0.1170	1.1604		0.1812	-1119	0.1125
2.3093		0.1816	1139	0.1162	2.3371		0.1777	-1117	0.1144
MEAN TRADE	513	MEAN SPR	0.1771		MEAN TRADE	-512	MEAN SPR	0.1743	
		MEAN EFF	0.1238				MEAN EFF	0.1230	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1815		0.1795	514	0.1306	0.1678		0.1797	-513	0.1312
0.4490		0.1786	515	0.1214	0.4212		0.1770	-510	0.1234
0.7531		0.1766	510	0.1215	0.7260		0.1751	-511	0.1215
1.1754		0.1771	513	0.1228	1.1562		0.1730	-511	0.1212
2.3001		0.1737	512	0.1226	2.3074		0.1683	-513	0.1203
MEAN TRADE	239	MEAN SPR	0.1675		MEAN TRADE	-300	MEAN SPR	0.1655	
		MEAN EFF	0.1298				MEAN EFF	0.1289	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2049		0.1663	241	0.1324	0.1675		0.1727	-300	0.1347
0.4886		0.1692	239	0.1293	0.4234		0.1683	-300	0.1309
0.7936		0.1677	238	0.1280	0.7260		0.1654	-300	0.1281
1.2116		0.1668	237	0.1289	1.1621		0.1634	-300	0.1255
2.2886		0.1676	238	0.1304	2.2728		0.1605	-300	0.1276
MEAN TRADE	100	MEAN SPR	0.1681		MEAN TRADE	-144	MEAN SPR	0.1658	
		MEAN EFF	0.1342				MEAN EFF	0.1314	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2203		0.1701	100	0.1376	0.1713		0.1760	-146	0.1372
0.5097		0.1654	100	0.1340	0.4229		0.1676	-144	0.1340
0.8176		0.1680	100	0.1302	0.7256		0.1641	-144	0.1293
1.2322		0.1682	100	0.1337	1.1611		0.1626	-143	0.1291
2.3643		0.1689	100	0.1353	2.2857		0.1631	-144	0.1300

LARGE

BUYS					SELLS				
MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2245		0.1776	8038	0.1263	0.2410		0.1772	-7836	0.1315
0.4663		0.1821	8510	0.1176	0.4895		0.1802	-8443	0.1208
0.7332		0.1839	8987	0.1132	0.7649		0.1820	-8833	0.1137
1.1312		0.1826	9291	0.1080	1.1696		0.1816	-8994	0.1077
2.2395		0.1811	10084	0.1044	2.3053		0.1791	-9757	0.1038

MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2087		0.1665	1553	0.1286	0.2354		0.1668	-1783	0.1290
0.4416		0.1691	1564	0.1249	0.4894		0.1673	-1789	0.1241
0.7058		0.1701	1565	0.1224	0.7647		0.1681	-1784	0.1203
1.0966		0.1705	1573	0.1209	1.1694		0.1663	-1779	0.1175
2.1958		0.1671	1572	0.1181	2.2874		0.1659	-1773	0.1134

MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2153		0.1628	610	0.1322	0.2307		0.1623	-877	0.1316
0.4626		0.1621	607	0.1293	0.4898		0.1641	-875	0.1265
0.7395		0.1624	613	0.1267	0.7646		0.1631	-872	0.1224
1.1415		0.1621	612	0.1268	1.1687		0.1615	-867	0.1218
2.2461		0.1596	619	0.1253	2.2930		0.1613	-862	0.1202

MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2428		0.1548	272	0.1335	0.2330		0.1574	-415	0.1325
0.5137		0.1540	269	0.1315	0.4911		0.1558	-411	0.1308
0.8066		0.1538	269	0.1311	0.7656		0.1557	-410	0.1273
1.2189		0.1531	268	0.1305	1.1690		0.1569	-408	0.1256
2.3407		0.1541	268	0.1307	2.2899		0.1552	-405	0.1247

MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2595		0.1509	100	0.1350	0.2361		0.1512	-140	0.1352
0.5465		0.1511	100	0.1337	0.4916		0.1516	-140	0.1323
0.8422		0.1512	100	0.1330	0.7680		0.1512	-138	0.1323
1.2630		0.1523	100	0.1331	1.1704		0.1514	-139	0.1305
2.4080		0.1526	100	0.1339	2.2985		0.1512	-138	0.1301

Panel C: SMA(4)

SMALL

BUYS					**SELLS**				
MEAN TRADE	6754	MEAN SPR	0.1979		MEAN TRADE	-6950	MEAN SPR	0.2014	
		MEAN EFF	0.1070				MEAN EFF	0.1065	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1712		0.2003	6387	0.1125	0.1991		0.2101	-7186	0.1156
0.4379		0.1999	6765	0.1085	0.4811		0.2024	-6840	0.1108
0.7234		0.1938	7139	0.1065	0.7787		0.2009	-6173	0.1055
1.1243		0.1996	7429	0.1040	1.1826		0.1980	-7250	0.1004
2.1673		0.1959	6059	0.1034	2.2508		0.1957	-7294	0.1004
MEAN TRADE	1021	MEAN SPR	0.1896		MEAN TRADE	-1021	MEAN SPR	0.1882	
		MEAN EFF	0.1199				MEAN EFF	0.1193	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1763		0.1955	1026	0.1237	0.1971		0.1966	-1025	0.1242
0.4539		0.1918	1020	0.1181	0.4833		0.1895	-1020	0.1209
0.7428		0.1898	1018	0.1206	0.7794		0.1865	-1020	0.1189
1.1319		0.1867	1023	0.1189	1.1859		0.1856	-1015	0.1171
2.1856		0.1843	1016	0.1179	2.2216		0.1832	-1023	0.1158
MEAN TRADE	487	MEAN SPR	0.1847		MEAN TRADE	-522	MEAN SPR	0.1901	
		MEAN EFF	0.1242				MEAN EFF	0.1229	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1947		0.1910	488	0.1293	0.1985		0.2025	-519	0.1307
0.4793		0.1847	487	0.1242	0.4835		0.1915	-521	0.1243
0.7727		0.1834	486	0.1231	0.7814		0.1876	-523	0.1209
1.1590		0.1844	487	0.1231	1.1847		0.1896	-522	0.1196
2.1772		0.1800	488	0.1212	2.1900		0.1806	-523	0.1195
MEAN TRADE	239	MEAN SPR	0.1765		MEAN TRADE	-341	MEAN SPR	0.1741	
		MEAN EFF	0.1325				MEAN EFF	0.1278	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2148		0.1824	240	0.1365	0.2014		0.1804	-340	0.1311
0.5090		0.1774	239	0.1317	0.4823		0.1758	-341	0.1273
0.7988		0.1736	239	0.1316	0.7812		0.1741	-342	0.1307
1.1844		0.1763	239	0.1317	1.1851		0.1729	-342	0.1267
2.1748		0.1728	238	0.1309	2.1892		0.1685	-341	0.1237
MEAN TRADE	100	MEAN SPR	0.1788		MEAN TRADE	-144	MEAN SPR	0.1745	
		MEAN EFF	0.1371				MEAN EFF	0.1323	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2250		0.1822	100	0.1396	0.2066		0.1825	-145	0.1368
0.5154		0.1843	100	0.1429	0.4828		0.1760	-144	0.1336
0.8123		0.1763	100	0.1355	0.7792		0.1734	-144	0.1322
1.2046		0.1757	100	0.1358	1.1823		0.1736	-144	0.1308
2.2127		0.1755	100	0.1318	2.1945		0.1691	-144	0.1297

MEDIUM

BUYS					SELLS				
MEAN TRADE	9855	MEAN SPR	0.1815		MEAN TRADE	-9697	MEAN SPR	0.1876	
		MEAN EFF	0.1104				MEAN EFF	0.1113	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1777		0.1739	8913	0.1162	0.1947		0.1825	-9116	0.1220
0.4402		0.1833	10563	0.1138	0.4707		0.1885	-9749	0.1134
0.7312		0.1831	9149	0.1088	0.7664		0.1909	-9024	0.1077
1.1313		0.1860	10741	0.1071	1.1695		0.1897	-10049	0.1074
2.2098		0.1811	9907	0.1063	2.2669		0.1862	-10551	0.1058

MEAN TRADE	1141	MEAN SPR	0.1822		MEAN TRADE	-1119	MEAN SPR	0.1827	
		MEAN EFF	0.1200				MEAN EFF	0.1189	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2042		0.1847	1135	0.1280	0.2022		0.1909	-1122	0.1301
0.4741		0.1842	1141	0.1206	0.4724		0.1853	-1124	0.1215
0.7640		0.1799	1142	0.1175	0.7657		0.1791	-1113	0.1166
1.1467		0.1814	1144	0.1165	1.1669		0.1849	-1120	0.1132
2.1827		0.1808	1142	0.1175	2.1907		0.1747	-1118	0.1153

MEAN TRADE	513	MEAN SPR	0.1771		MEAN TRADE	-512	MEAN SPR	0.1743	
		MEAN EFF	0.1238				MEAN EFF	0.1230	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2267		0.1813	514	0.1293	0.2083		0.1794	-512	0.1325
0.5110		0.1767	514	0.1214	0.4719		0.1773	-511	0.1212
0.7973		0.1769	511	0.1220	0.7686		0.1752	-511	0.1226
1.1770		0.1770	513	0.1221	1.1667		0.1716	-510	0.1202
2.1532		0.1737	512	0.1242	2.1638		0.1695	-514	0.1214

MEAN TRADE	239	MEAN SPR	0.1675		MEAN TRADE	-300	MEAN SPR	0.1655	
		MEAN EFF	0.1298				MEAN EFF	0.1289	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2515		0.1663	240	0.1309	0.2056		0.1726	-300	0.1347
0.5413		0.1688	238	0.1303	0.4738		0.1687	-300	0.1306
0.8274		0.1677	239	0.1281	0.7694		0.1658	-300	0.1272
1.2099		0.1667	238	0.1292	1.1703		0.1632	-300	0.1269
2.1437		0.1681	238	0.1305	2.1434		0.1599	-300	0.1275

MEAN TRADE	100	MEAN SPR	0.1681		MEAN TRADE	-144	MEAN SPR	0.1658	
		MEAN EFF	0.1342				MEAN EFF	0.1314	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2666		0.1696	100	0.1362	0.2120		0.1740	-146	0.1370
0.5647		0.1653	100	0.1338	0.4758		0.1695	-145	0.1323
0.8488		0.1679	100	0.1305	0.7688		0.1639	-143	0.1306
1.2253		0.1698	100	0.1345	1.1673		0.1637	-143	0.1296
2.2228		0.1678	100	0.1358	2.1491		0.1618	-144	0.1298

LARGE

BUYS					SELLS				
MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2641		0.1792	8127	0.1244	0.2786		0.1795	-8006	0.1300
0.5108		0.1831	8540	0.1189	0.5329		0.1806	-8442	0.1208
0.7682		0.1832	9040	0.1121	0.7991		0.1815	-8575	0.1138
1.1385		0.1823	9156	0.1094	1.1731		0.1813	-9207	0.1079
2.1200		0.1796	10044	0.1049	2.1669		0.1771	-9632	0.1049
MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2483		0.1672	1554	0.1271	0.2740		0.1673	-1784	0.1280
0.4910		0.1700	1559	0.1253	0.5329		0.1687	-1786	0.1248
0.7486		0.1697	1573	0.1231	0.8004		0.1675	-1786	0.1202
1.1129		0.1697	1570	0.1209	1.1749		0.1664	-1779	0.1172
2.0832		0.1666	1571	0.1184	2.1471		0.1643	-1773	0.1145
MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2586		0.1631	609	0.1318	0.2720		0.1630	-879	0.1311
0.5136		0.1637	608	0.1292	0.5330		0.1648	-874	0.1261
0.7810		0.1618	612	0.1271	0.7985		0.1626	-871	0.1230
1.1529		0.1606	614	0.1261	1.1739		0.1616	-868	0.1222
2.1230		0.1596	619	0.1260	2.1572		0.1603	-861	0.1202
MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2881		0.1558	272	0.1331	0.2727		0.1585	-414	0.1325
0.5622		0.1536	269	0.1316	0.5342		0.1556	-411	0.1294
0.8394		0.1536	268	0.1309	0.8007		0.1571	-410	0.1282
1.2165		0.1529	269	0.1307	1.1723		0.1556	-408	0.1259
2.2054		0.1538	268	0.1310	2.1522		0.1544	-407	0.1249
MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.3091		0.1524	100	0.1345	0.2787		0.1527	-140	0.1352
0.5962		0.1504	100	0.1333	0.5354		0.1515	-139	0.1320
0.8779		0.1507	100	0.1332	0.8028		0.1506	-138	0.1325
1.2605		0.1525	100	0.1344	1.1756		0.1511	-139	0.1304
2.2668		0.1520	100	0.1333	2.1642		0.1510	-138	0.1302

Panel D: SMA(5)

SMALL

BUYS					SELLS				
MEAN TRADE	6754	MEAN SPR	0.1979		MEAN TRADE	-6950	MEAN SPR	0.2014	
		MEAN EFF	0.1070				MEAN EFF	0.1065	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1982		0.2003	6189	0.1112	0.2286		0.2098	-7294	0.1158
0.4746		0.2000	6997	0.1093	0.5184		0.2019	-6569	0.1085
0.7537		0.1948	7392	0.1071	0.8019		0.1994	-6456	0.1064
1.1286		0.1989	7273	0.1035	1.1851		0.2004	-6565	0.1022
2.0628		0.1955	5931	0.1039	2.1413		0.1958	-7860	0.0998
MEAN TRADE	1021	MEAN SPR	0.1896		MEAN TRADE	-1021	MEAN SPR	0.1882	
		MEAN EFF	0.1199				MEAN EFF	0.1193	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2084		0.1952	1025	0.1224	0.2272		0.1952	-1025	0.1224
0.4908		0.1912	1019	0.1200	0.5163		0.1912	-1018	0.1228
0.7697		0.1888	1022	0.1190	0.8020		0.1852	-1019	0.1178
1.1323		0.1879	1019	0.1199	1.1858		0.1867	-1019	0.1180
2.1066		0.1851	1019	0.1179	2.1399		0.1829	-1022	0.1160
MEAN TRADE	487	MEAN SPR	0.1847		MEAN TRADE	-522	MEAN SPR	0.1901	
		MEAN EFF	0.1242				MEAN EFF	0.1229	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2263		0.1895	488	0.1290	0.2313		0.2012	-519	0.1299
0.5176		0.1846	487	0.1238	0.5167		0.1904	-522	0.1259
0.7987		0.1841	486	0.1254	0.8006		0.1898	-522	0.1204
1.1638		0.1846	488	0.1229	1.1865		0.1901	-522	0.1210
2.0864		0.1808	488	0.1198	2.1276		0.1798	-523	0.1177
MEAN TRADE	239	MEAN SPR	0.1765		MEAN TRADE	-341	MEAN SPR	0.1741	
		MEAN EFF	0.1325				MEAN EFF	0.1278	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2466		0.1822	239	0.1369	0.2335		0.1783	-340	0.1304
0.5428		0.1777	240	0.1311	0.5193		0.1745	-341	0.1284
0.8232		0.1731	239	0.1318	0.8019		0.1757	-341	0.1293
1.1835		0.1762	239	0.1317	1.1843		0.1741	-342	0.1275
2.0942		0.1732	238	0.1309	2.1126		0.1686	-341	0.1239
MEAN TRADE	100	MEAN SPR	0.1788		MEAN TRADE	-144	MEAN SPR	0.1745	
		MEAN EFF	0.1371				MEAN EFF	0.1323	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2586		0.1831	100	0.1396	0.2408		0.1822	-145	0.1365
0.5544		0.1768	100	0.1355	0.5207		0.1756	-145	0.1331
0.8329		0.1822	100	0.1428	0.8043		0.1746	-144	0.1328
1.2106		0.1766	100	0.1356	1.1827		0.1731	-144	0.1310
2.1212		0.1752	100	0.1320	2.1108		0.1690	-144	0.1294

MEDIUM

BUYS				SELLS			
MEAN TRADE 9855	MEAN SPR 0.1815			MEAN TRADE -9697	MEAN SPR 0.1876		
	MEAN EFF 0.1104				MEAN EFF 0.1113		
MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.2013	0.1754	8862	0.1167	0.2208	0.1825	-8750	0.1225
0.4757	0.1823	10233	0.1123	0.5060	0.1894	-10573	0.1129
0.7607	0.1825	9859	0.1093	0.7914	0.1907	-8927	0.1086
1.1386	0.1859	10082	0.1063	1.1731	0.1891	-10039	0.1059
2.1205	0.1812	10242	0.1076	2.1769	0.1859	-10199	0.1064
MEAN TRADE 1141	MEAN SPR 0.1822			MEAN TRADE -1119	MEAN SPR 0.1827		
	MEAN EFF 0.1200				MEAN EFF 0.1189		
MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.2356	0.1862	1135	0.1279	0.2310	0.1908	-1125	0.1305
0.5171	0.1816	1140	0.1196	0.5099	0.1859	-1117	0.1195
0.7910	0.1809	1145	0.1174	0.7887	0.1803	-1120	0.1173
1.1454	0.1824	1141	0.1173	1.1703	0.1834	-1120	0.1144
2.0909	0.1799	1144	0.1179	2.0826	0.1749	-1115	0.1151
MEAN TRADE 513	MEAN SPR 0.1771			MEAN TRADE -512	MEAN SPR 0.1743		
	MEAN EFF 0.1238				MEAN EFF 0.1230		
MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.2603	0.1814	513	0.1292	0.2381	0.1790	-512	0.1300
0.5496	0.1772	514	0.1218	0.5089	0.1769	-512	0.1227
0.8214	0.1774	512	0.1209	0.7911	0.1750	-510	0.1230
1.1802	0.1759	513	0.1234	1.1725	0.1719	-509	0.1198
2.0766	0.1736	512	0.1236	2.0747	0.1700	-515	0.1221
MEAN TRADE 239	MEAN SPR 0.1675			MEAN TRADE -300	MEAN SPR 0.1655		
	MEAN EFF 0.1298				MEAN EFF 0.1289		
MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.2875	0.1687	240	0.1321	0.2391	0.1708	-300	0.1372
0.5774	0.1660	238	0.1282	0.5103	0.1685	-300	0.1293
0.8519	0.1671	239	0.1280	0.7925	0.1659	-300	0.1271
1.2092	0.1694	239	0.1295	1.1664	0.1632	-300	0.1265
2.0508	0.1662	237	0.1312	2.0447	0.1611	-300	0.1277
MEAN TRADE 100	MEAN SPR 0.1681			MEAN TRADE -144	MEAN SPR 0.1658		
	MEAN EFF 0.1342				MEAN EFF 0.1314		
MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.3042	0.1695	100	0.1359	0.2435	0.1745	-145	0.1385
0.6023	0.1661	100	0.1322	0.5099	0.1685	-145	0.1323
0.8700	0.1673	100	0.1318	0.7963	0.1640	-144	0.1295
1.2246	0.1685	100	0.1349	1.1675	0.1638	-143	0.1297
2.1136	0.1690	100	0.1359	2.0639	0.1623	-144	0.1302

LARGE

BUYS					**SELLS**				
MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2945		0.1811	8311	0.1238	0.3053		0.1799	-8160	0.1300
0.5426		0.1835	8581	0.1177	0.5620		0.1819	-8341	0.1205
0.7935		0.1832	8841	0.1132	0.8203		0.1813	-8566	0.1129
1.1388		0.1815	9271	0.1100	1.1712		0.1802	-9255	0.1081
2.0337		0.1781	9904	0.1050	2.0713		0.1767	-9540	0.1059
MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2793		0.1681	1555	0.1272	0.3007		0.1674	-1782	0.1286
0.5280		0.1701	1565	0.1247	0.5607		0.1695	-1788	0.1239
0.7785		0.1694	1567	0.1234	0.8207		0.1676	-1787	0.1209
1.1242		0.1695	1566	0.1211	1.1701		0.1661	-1776	0.1173
2.0108		0.1661	1572	0.1183	2.0504		0.1638	-1774	0.1143
MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2928		0.1643	609	0.1324	0.3007		0.1638	-878	0.1304
0.5521		0.1642	609	0.1292	0.5628		0.1647	-875	0.1274
0.8081		0.1609	610	0.1263	0.8202		0.1625	-871	0.1229
1.1577		0.1605	616	0.1268	1.1725		0.1612	-867	0.1222
2.0355		0.1590	617	0.1256	2.0604		0.1599	-862	0.1200
MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.3209		0.1555	272	0.1332	0.3024		0.1586	-415	0.1325
0.5963		0.1548	268	0.1317	0.5631		0.1565	-412	0.1297
0.8604		0.1526	269	0.1309	0.8223		0.1568	-410	0.1281
1.2109		0.1531	268	0.1307	1.1741		0.1552	-407	0.1260
2.1126		0.1537	268	0.1309	2.0592		0.1542	-407	0.1247
MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.3442		0.1523	100	0.1343	0.3070		0.1526	-141	0.1345
0.6299		0.1508	100	0.1335	0.5647		0.1521	-139	0.1327
0.8978		0.1510	100	0.1338	0.8224		0.1513	-139	0.1324
1.2555		0.1512	100	0.1332	1.1729		0.1503	-139	0.1301
2.1667		0.1528	100	0.1339	2.0684		0.1509	-138	0.1306

Table VI

Order Flow Deciles Dissected by Exponential Moving Average Intensity Quintiles

The data in Table VI were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. Table VI contains each set of stocks broken down first by trade decile and then further by unexpected trade intensity quintiles. For the buys and the sells orders (as assigned by the Lee and Ready (1991) algorithm) in each decile, mean trade size, quoted and effective spreads were calculated. The measure of trade intensity is captured by an exponential moving average, EMA, (of the present trade and the previous 2-5 trades) of the diurnally adjusted (where diurnally adjusting involves taking the time between trades and then deflating that by the expected time between trades conditioned on the time of day) times between trades. The exponential moving average lag of 3 captured the sharpest results in the data while remaining parsimonious (only lags 1-5 were considered). The moving average variable (MA QUINTILE) quintile breakpoints are shown to the left, with the averages of the quoted or inside spread (SPREAD), trade size (SIZE), and effective spread (EFFSPR) to the right. For the small, medium and large sized stocks, there exists a general increase in effective spread as trade intensity (as proxied by the EMA intensity) increases, potentially indicating an increase in the adverse selection component of the spread. Panels A through D display the trade size / trade intensity relationship for all sets of stocks. Each Panel features the results from the small then medium, followed by those of the large stocks. Microstructure theory predicts that, all else equal, order processing has a temporal premium. Consequently, a pattern should arise such that the quoted and effective spreads should increase as unexpected time between trades decreases. This general pattern appears over all sizes of stocks in the sample using 4 specifications of exponential moving average (EMA). The optimal lag length was chosen as the one with the most consistent, monotonic pattern across all stocks in the sample.

Panel A: EMA(2)

SMALL

BUYS						SELLS					
MEAN TRADE	6754	MEAN SPR	0.1979			MEAN TRADE	-6950	MEAN SPR	0.2014		
		MEAN EFF	0.1070					MEAN EFF	0.1065		
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**		MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	
0.0852		0.1989	6128	0.1157		0.0986		0.2123	-6721	0.1212	
0.2889		0.1995	6824	0.1073		0.3293		0.2021	-7342	0.1083	
0.5872		0.1974	7477	0.1061		0.6491		0.1993	-6891	0.1043	
1.0637		0.1981	6696	0.1047		1.1506		0.1968	-7003	0.0991	
2.5738		0.1955	6648	0.1010		2.7043		0.1967	-6783	0.0996	

MEAN TRADE	1021	MEAN SPR	0.1896			MEAN TRADE	-1021	MEAN SPR	0.1882		
		MEAN EFF	0.1199					MEAN EFF	0.1193		
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**		MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	
0.0831		0.1992	1024	0.1256		0.0983		0.1976	-1022	0.1261	
0.2961		0.1867	1025	0.1188		0.3309		0.1893	-1022	0.1205	
0.6018		0.1903	1019	0.1176		0.6481		0.1872	-1018	0.1176	
1.0764		0.1864	1019	0.1194		1.1518		0.1848	-1017	0.1177	
2.5644		0.1856	1017	0.1180		2.6702		0.1822	-1023	0.1149	

MEAN TRADE 487 MEAN SPR 0.1847
MEAN EFF 0.1242

MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.0958	0.1914	490	0.1292
0.3230	0.1860	486	0.1273
0.6346	0.1846	487	0.1224
1.1053	0.1818	488	0.1223
2.5940	0.1799	487	0.1197

MEAN TRADE -522 MEAN SPR 0.1901
MEAN EFF 0.1229

MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.0970	0.2036	-520	0.1323
0.3313	0.1899	-522	0.1273
0.6506	0.1889	-521	0.1186
1.1506	0.1855	-522	0.1192
2.6340	0.1838	-523	0.1179

MEAN TRADE 239 MEAN SPR 0.1765
MEAN EFF 0.1325

MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.1092	0.1858	240	0.1365
0.3545	0.1753	239	0.1330
0.6734	0.1743	239	0.1317
1.1540	0.1735	239	0.1300
2.5882	0.1735	239	0.1314

MEAN TRADE -341 MEAN SPR 0.1741
MEAN EFF 0.1278

MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.0993	0.1808	-341	0.1336
0.3313	0.1732	-341	0.1272
0.6473	0.1750	-340	0.1283
1.1490	0.1734	-341	0.1273
2.6071	0.1692	-341	0.1237

MEAN TRADE 100 MEAN SPR 0.1788
MEAN EFF 0.1371

MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.1128	0.1825	100	0.1402
0.3659	0.1773	100	0.1365
0.6870	0.1817	100	0.1421
1.1792	0.1773	100	0.1353
2.6309	0.1752	100	0.1314

MEAN TRADE -144 MEAN SPR 0.1745
MEAN EFF 0.1323

MA QUINTILE	**SPREAD**	**SIZE**	**EFFSPR**
0.1015	0.1812	-144	0.1366
0.3335	0.1757	-144	0.1340
0.6522	0.1747	-145	0.1339
1.1507	0.1734	-144	0.1302
2.6043	0.1690	-143	0.1282

MEDIUM

BUYS					SELLS				
MEAN TRADE	9855	MEAN SPR	0.1815		MEAN TRADE	-9697	MEAN SPR	0.1876	
		MEAN EFF	0.1104				MEAN EFF	0.1113	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.0991		0.1739	8716	0.1193	0.1050		0.1855	-7755	0.1229
0.3009		0.1825	10573	0.1114	0.3203		0.1850	-9613	0.1139
0.5940		0.1813	10535	0.1093	0.6268		0.1885	-10357	0.1100
1.0751		0.1867	9664	0.1064	1.1197		0.1914	-11383	0.1042
2.6036		0.1831	9788	0.1057	2.6848		0.1874	-9377	0.1055
MEAN TRADE	1141	MEAN SPR	0.1822		MEAN TRADE	-1119	MEAN SPR	0.1827	
		MEAN EFF	0.1200				MEAN EFF	0.1189	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1079		0.1854	1137	0.1271	0.1021		0.1930	-1124	0.1303
0.3291		0.1828	1140	0.1220	0.3224		0.1845	-1117	0.1211
0.6289		0.1825	1136	0.1168	0.6266		0.1812	-1113	0.1168
1.1163		0.1788	1152	0.1161	1.1125		0.1788	-1128	0.1150
2.6039		0.1814	1140	0.1181	2.6385		0.1776	-1115	0.1131
MEAN TRADE	513	MEAN SPR	0.1771		MEAN TRADE	-512	MEAN SPR	0.1743	
		MEAN EFF	0.1238				MEAN EFF	0.1230	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1173		0.1794	513	0.1300	0.1066		0.1807	-513	0.1313
0.3503		0.1771	515	0.1245	0.3238		0.1760	-510	0.1231
0.6629		0.1777	513	0.1197	0.6271		0.1718	-511	0.1217
1.1416		0.1771	511	0.1208	1.1212		0.1735	-511	0.1200
2.5901		0.1742	512	0.1238	2.5704		0.1709	-512	0.1212
MEAN TRADE	239	MEAN SPR	0.1675		MEAN TRADE	-300	MEAN SPR	0.1655	
		MEAN EFF	0.1298				MEAN EFF	0.1289	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1339		0.1671	241	0.1323	0.1108		0.1728	-300	0.1338
0.3877		0.1673	239	0.1285	0.3272		0.1667	-300	0.1302
0.7016		0.1692	239	0.1300	0.6282		0.1635	-300	0.1258
1.1778		0.1667	238	0.1288	1.1173		0.1647	-300	0.1282
2.5705		0.1672	238	0.1294	2.5558		0.1618	-300	0.1280
MEAN TRADE	100	MEAN SPR	0.1681		MEAN TRADE	-144	MEAN SPR	0.1658	
		MEAN EFF	0.1342				MEAN EFF	0.1314	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1445		0.1701	100	0.1361	0.1091		0.1743	-147	0.1380
0.4060		0.1671	100	0.1343	0.3242		0.1673	-144	0.1333
0.7221		0.1654	100	0.1315	0.6294		0.1657	-143	0.1304
1.2060		0.1694	100	0.1335	1.1234		0.1623	-144	0.1285
2.6588		0.1687	100	0.1355	2.5444		0.1628	-143	0.1295

LARGE

BUYS					**SELLS**				
MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1596		0.1760	8006	0.1278	0.1725		0.1762	-7721	0.1307
0.3752		0.1822	8619	0.1155	0.4015		0.1802	-8301	0.1197
0.6523		0.1833	9025	0.1123	0.6864		0.1819	-8762	0.1134
1.0946		0.1845	9344	0.1083	1.1428		0.1822	-9271	0.1080
2.5047		0.1814	9923	0.1056	2.5818		0.1796	-9807	0.1057

MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1485		0.1657	1554	0.1292	0.1670		0.1650	-1781	0.1295
0.3545		0.1700	1566	0.1238	0.3996		0.1674	-1792	0.1217
0.6190		0.1693	1567	0.1224	0.6879		0.1689	-1783	0.1212
1.0515		0.1701	1568	0.1199	1.1415		0.1677	-1779	0.1169
2.4407		0.1682	1571	0.1195	2.5602		0.1654	-1772	0.1150

MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1493		0.1610	611	0.1318	0.1648		0.1615	-876	0.1306
0.3689		0.1649	609	0.1282	0.4011		0.1633	-875	0.1254
0.6515		0.1620	612	0.1273	0.6888		0.1642	-870	0.1236
1.1009		0.1621	612	0.1269	1.1394		0.1626	-869	0.1219
2.5087		0.1589	617	0.1261	2.5737		0.1608	-862	0.1208

MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1726		0.1537	272	0.1343	0.1651		0.1568	-414	0.1335
0.4224		0.1554	269	0.1312	0.4031		0.1567	-412	0.1295
0.7255		0.1534	268	0.1306	0.6870		0.1554	-409	0.1273
1.1988		0.1530	268	0.1301	1.1422		0.1565	-408	0.1252
2.6312		0.1543	268	0.1311	2.5721		0.1556	-406	0.1251

MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1808		0.1503	100	0.1362	0.1668		0.1508	-140	0.1354
0.4477		0.1515	100	0.1328	0.4038		0.1512	-139	0.1319
0.7651		0.1512	100	0.1323	0.6906		0.1518	-139	0.1318
1.2483		0.1527	100	0.1333	1.1467		0.1515	-138	0.1308
2.7128		0.1525	100	0.1341	2.5673		0.1512	-138	0.1302

Panel B: EMA(3)

SMALL

BUYS					SELLS				
MEAN TRADE	6754	MEAN SPR	0.1979		MEAN TRADE	-6950	MEAN SPR	0.2014	
		MEAN EFF	0.1070				MEAN EFF	0.1065	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1142		0.2003	7061	0.1126	0.1309		0.2096	-6901	0.1159
0.3325		0.1991	6845	0.1060	0.3672		0.2017	-7156	0.1077
0.6215		0.1952	6258	0.1061	0.6629		0.1971	-6449	0.1034
1.0836		0.1989	7220	0.1045	1.1303		0.2005	-7258	0.1039
2.5518		0.1959	6391	0.1056	2.6325		0.1984	-6976	0.1017

MEAN TRADE	1021	MEAN SPR	0.1896		MEAN TRADE	-1021	MEAN SPR	0.1882	
		MEAN EFF	0.1199				MEAN EFF	0.1193	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1141		0.1976	1024	0.1232	0.1309		0.1962	-1022	0.1252
0.3404		0.1892	1019	0.1196	0.3671		0.1873	-1018	0.1207
0.6312		0.1893	1021	0.1200	0.6622		0.1873	-1022	0.1185
1.0873		0.1867	1019	0.1182	1.1263		0.1846	-1020	0.1177
2.5376		0.1852	1022	0.1183	2.5439		0.1859	-1022	0.1148

MEAN TRADE	487	MEAN SPR	0.1847		MEAN TRADE	-522	MEAN SPR	0.1901	
		MEAN EFF	0.1242				MEAN EFF	0.1229	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1255		0.1923	488	0.1299	0.1331		0.2003	-520	0.1296
0.3605		0.1843	487	0.1245	0.3666		0.1910	-521	0.1225
0.6541		0.1830	488	0.1235	0.6586		0.1872	-522	0.1209
1.1187		0.1815	487	0.1228	1.1352		0.1880	-522	0.1226
2.5353		0.1826	488	0.1202	2.5353		0.1846	-523	0.1190

MEAN TRADE	239	MEAN SPR	0.1765		MEAN TRADE	-341	MEAN SPR	0.1741	
		MEAN EFF	0.1325				MEAN EFF	0.1278	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1393		0.1824	239	0.1357	0.1345		0.1802	-341	0.1293
0.3862		0.1760	240	0.1335	0.3663		0.1735	-340	0.1267
0.6905		0.1745	240	0.1314	0.6649		0.1751	-341	0.1305
1.1470		0.1746	238	0.1298	1.1310		0.1724	-342	0.1277
2.5332		0.1749	238	0.1321	2.5290		0.1703	-341	0.1250

MEAN TRADE	100	MEAN SPR	0.1788		MEAN TRADE	-144	MEAN SPR	0.1745	
		MEAN EFF	0.1371				MEAN EFF	0.1323	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1444		0.1837	100	0.1412	0.1353		0.1836	-144	0.1361
0.3942		0.1779	100	0.1357	0.3678		0.1757	-144	0.1335
0.6903		0.1734	100	0.1335	0.6627		0.1734	-145	0.1326
1.1557		0.1822	100	0.1416	1.1239		0.1709	-143	0.1297
2.5454		0.1767	100	0.1334	2.5559		0.1709	-145	0.1306

MEDIUM

BUYS					SELLS				
MEAN TRADE	9855	MEAN SPR	0.1815		MEAN TRADE	-9697	MEAN SPR	0.1876	
		MEAN EFF	0.1104				MEAN EFF	0.1113	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1248		0.1753	8703	0.1141	0.1354		0.1833	-10102	0.1205
0.3376		0.1823	10292	0.1126	0.3616		0.1881	-9373	0.1129
0.6200		0.1829	10040	0.1097	0.6564		0.1885	-9468	0.1082
1.0778		0.1844	10162	0.1086	1.1145		0.1908	-9188	0.1075
2.5548		0.1826	10077	0.1071	2.5983		0.1870	-10356	0.1072
MEAN TRADE	1141	MEAN SPR	0.1822		MEAN TRADE	-1119	MEAN SPR	0.1827	
		MEAN EFF	0.1200				MEAN EFF	0.1189	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1391		0.1862	1137	0.1271	0.1368		0.1896	-1118	0.1292
0.3651		0.1827	1143	0.1211	0.3632		0.1830	-1124	0.1204
0.6484		0.1797	1141	0.1179	0.6581		0.1813	-1121	0.1158
1.1001		0.1795	1143	0.1177	1.1107		0.1815	-1120	0.1141
2.4978		0.1829	1140	0.1165	2.5377		0.1791	-1114	0.1168
MEAN TRADE	513	MEAN SPR	0.1771		MEAN TRADE	-512	MEAN SPR	0.1743	
		MEAN EFF	0.1238				MEAN EFF	0.1230	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1504		0.1805	513	0.1285	0.1405		0.1821	-512	0.1283
0.3889		0.1792	515	0.1226	0.3673		0.1746	-511	0.1238
0.6763		0.1770	513	0.1217	0.6571		0.1747	-511	0.1221
1.1286		0.1733	510	0.1242	1.1087		0.1720	-511	0.1214
2.5403		0.1755	512	0.1220	2.4949		0.1698	-513	0.1211
MEAN TRADE	239	MEAN SPR	0.1675		MEAN TRADE	-300	MEAN SPR	0.1655	
		MEAN EFF	0.1298				MEAN EFF	0.1289	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1676		0.1674	240	0.1314	0.1417		0.1740	-300	0.1327
0.4159		0.1687	239	0.1299	0.3668		0.1642	-300	0.1312
0.7090		0.1660	238	0.1290	0.6559		0.1665	-300	0.1284
1.1590		0.1670	238	0.1296	1.1108		0.1623	-300	0.1264
2.5082		0.1684	238	0.1292	2.4743		0.1628	-300	0.1271
MEAN TRADE	100	MEAN SPR	0.1681		MEAN TRADE	-144	MEAN SPR	0.1658	
		MEAN EFF	0.1342				MEAN EFF	0.1314	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1777		0.1681	100	0.1365	0.1424		0.1737	-146	0.1374
0.4322		0.1677	100	0.1353	0.3663		0.1672	-144	0.1317
0.7316		0.1685	100	0.1328	0.6581		0.1653	-144	0.1303
1.1930		0.1661	100	0.1309	1.1125		0.1627	-144	0.1293
2.5735		0.1703	100	0.1353	2.4768		0.1631	-143	0.1304

LARGE

BUYS | **SELLS**

MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1878		0.1798	8204	0.1216	0.2004		0.1787	-8416	0.1270
0.4062		0.1835	8917	0.1161	0.4277		0.1821	-8658	0.1194
0.6713		0.1839	8964	0.1131	0.6954		0.1814	-8680	0.1130
1.1000		0.1805	9173	0.1113	1.1196		0.1790	-9033	0.1105
2.4482		0.1797	9652	0.1075	2.4948		0.1789	-9075	0.1075

MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1781		0.1684	1554	0.1265	0.1967		0.1679	-1780	0.1268
0.3914		0.1688	1568	0.1241	0.4277		0.1685	-1788	0.1228
0.6537		0.1701	1565	0.1233	0.6977		0.1664	-1786	0.1220
1.0747		0.1689	1568	0.1215	1.1228		0.1665	-1776	0.1185
2.4225		0.1671	1571	0.1194	2.4801		0.1648	-1777	0.1149

MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1851		0.1637	609	0.1306	0.1948		0.1639	-877	0.1294
0.4126		0.1624	610	0.1298	0.4273		0.1634	-874	0.1256
0.6838		0.1619	611	0.1280	0.6982		0.1634	-872	0.1243
1.1124		0.1612	613	0.1262	1.1193		0.1610	-867	0.1230
2.4488		0.1597	618	0.1258	2.4808		0.1605	-864	0.1206

MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2041		0.1557	271	0.1324	0.1963		0.1574	-414	0.1309
0.4486		0.1541	269	0.1313	0.4290		0.1570	-411	0.1299
0.7322		0.1538	268	0.1319	0.6986		0.1561	-410	0.1280
1.1667		0.1529	269	0.1312	1.1222		0.1563	-409	0.1268
2.5298		0.1534	268	0.1306	2.4639		0.1544	-406	0.1251

MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2173		0.1521	100	0.1338	0.1995		0.1519	-139	0.1338
0.4761		0.1516	100	0.1332	0.4295		0.1527	-140	0.1330
0.7672		0.1513	100	0.1343	0.7000		0.1507	-139	0.1309
1.2130		0.1514	100	0.1337	1.1242		0.1512	-138	0.1316
2.5827		0.1517	100	0.1337	2.4961		0.1504	-138	0.1305

Panel C: EMA(4)

SMALL

BUYS					SELLS				
MEAN TRADE	6754	MEAN SPR	0.1979		MEAN TRADE	-6950	MEAN SPR	0.2014	
		MEAN EFF	0.1070				MEAN EFF	0.1065	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1219		0.2002	6724	0.1097	0.1382		0.2098	-7225	0.1123
0.3399		0.1991	6813	0.1090	0.3671		0.1997	-7116	0.1074
0.6213		0.1971	6785	0.1053	0.6603		0.2000	-6681	0.1045
1.0692		0.1967	7199	0.1065	1.1234		0.1989	-6074	0.1052
2.5003		0.1963	6258	0.1043	2.5770		0.1988	-7647	0.1033
MEAN TRADE	1021	MEAN SPR	0.1896		MEAN TRADE	-1021	MEAN SPR	0.1882	
		MEAN EFF	0.1199				MEAN EFF	0.1193	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1255		0.1942	1023	0.1220	0.1361		0.1944	-1021	0.1236
0.3522		0.1903	1022	0.1192	0.3667		0.1893	-1021	0.1177
0.6415		0.1883	1021	0.1191	0.6587		0.1873	-1024	0.1174
1.1022		0.1897	1021	0.1220	1.1252		0.1866	-1019	0.1193
2.5481		0.1856	1018	0.1169	2.5635		0.1837	-1018	0.1189
MEAN TRADE	487	MEAN SPR	0.1847		MEAN TRADE	-522	MEAN SPR	0.1901	
		MEAN EFF	0.1242				MEAN EFF	0.1229	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1370		0.1894	488	0.1267	0.1384		0.2015	-519	0.1263
0.3701		0.1855	487	0.1247	0.3658		0.1880	-522	0.1234
0.6673		0.1815	486	0.1244	0.6617		0.1912	-523	0.1216
1.1202		0.1820	487	0.1216	1.1203		0.1862	-522	0.1214
2.5246		0.1852	489	0.1235	2.5608		0.1843	-522	0.1218
MEAN TRADE	239	MEAN SPR	0.1765		MEAN TRADE	-341	MEAN SPR	0.1741	
		MEAN EFF	0.1325				MEAN EFF	0.1278	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1480		0.1821	239	0.1353	0.1393		0.1786	-341	0.1300
0.3885		0.1765	240	0.1320	0.3673		0.1758	-340	0.1276
0.6818		0.1747	239	0.1317	0.6605		0.1735	-340	0.1287
1.1328		0.1737	239	0.1311	1.1238		0.1736	-343	0.1275
2.5124		0.1753	239	0.1323	2.5391		0.1698	-341	0.1258
MEAN TRADE	100	MEAN SPR	0.1788		MEAN TRADE	-144	MEAN SPR	0.1745	
		MEAN EFF	0.1371				MEAN EFF	0.1323	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1521		0.1820	100	0.1387	0.1432		0.1814	-144	0.1345
0.3908		0.1841	100	0.1433	0.3678		0.1740	-145	0.1330
0.6858		0.1761	100	0.1339	0.6629		0.1741	-144	0.1324
1.1461		0.1756	100	0.1361	1.1271		0.1735	-144	0.1324
2.5585		0.1761	100	0.1334	2.5309		0.1708	-144	0.1299

MEDIUM

BUYS					SELLS				
MEAN TRADE	9855	MEAN SPR	0.1815		MEAN TRADE	-9697	MEAN SPR	0.1876	
		MEAN EFF	0.1104				MEAN EFF	0.1113	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1322		0.1782	9054	0.1147	0.1417		0.1841	-9761	0.1181
0.3415		0.1818	9968	0.1108	0.3668		0.1905	-9361	0.1105
0.6236		0.1822	9846	0.1101	0.6588		0.1870	-9222	0.1097
1.0738		0.1829	9778	0.1091	1.1172		0.1886	-10072	0.1080
2.5660		0.1822	10627	0.1074	2.6382		0.1876	-10073	0.1100
MEAN TRADE	1141	MEAN SPR	0.1822		MEAN TRADE	-1119	MEAN SPR	0.1827	
		MEAN EFF	0.1200				MEAN EFF	0.1189	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1492		0.1844	1135	0.1259	0.1442		0.1864	-1120	0.1253
0.3737		0.1830	1143	0.1199	0.3692		0.1866	-1120	0.1188
0.6510		0.1819	1142	0.1183	0.6592		0.1812	-1121	0.1168
1.1044		0.1809	1141	0.1160	1.1150		0.1807	-1117	0.1172
2.5207		0.1809	1143	0.1200	2.5304		0.1793	-1120	0.1178
MEAN TRADE	513	MEAN SPR	0.1771		MEAN TRADE	-512	MEAN SPR	0.1743	
		MEAN EFF	0.1238				MEAN EFF	0.1230	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1642		0.1815	514	0.1257	0.1509		0.1822	-513	0.1288
0.4000		0.1780	512	0.1231	0.3706		0.1742	-509	0.1237
0.6893		0.1769	513	0.1234	0.6581		0.1727	-513	0.1210
1.1377		0.1751	512	0.1223	1.1161		0.1718	-509	0.1209
2.4923		0.1741	513	0.1244	2.5195		0.1725	-514	0.1225
MEAN TRADE	239	MEAN SPR	0.1675		MEAN TRADE	-300	MEAN SPR	0.1655	
		MEAN EFF	0.1298				MEAN EFF	0.1289	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1738		0.1673	239	0.1304	0.1496		0.1708	-300	0.1363
0.4181		0.1678	239	0.1300	0.3717		0.1662	-300	0.1269
0.7052		0.1679	239	0.1279	0.6573		0.1676	-300	0.1282
1.1479		0.1657	237	0.1302	1.1137		0.1623	-300	0.1259
2.4966		0.1688	239	0.1305	2.4781		0.1619	-300	0.1290
MEAN TRADE	100	MEAN SPR	0.1681		MEAN TRADE	-144	MEAN SPR	0.1658	
		MEAN EFF	0.1342				MEAN EFF	0.1314	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1851		0.1698	100	0.1351	0.1516		0.1712	-145	0.1339
0.4346		0.1674	100	0.1324	0.3713		0.1680	-145	0.1322
0.7295		0.1666	100	0.1328	0.6604		0.1662	-144	0.1312
1.1801		0.1669	100	0.1351	1.1208		0.1638	-144	0.1307
2.5599		0.1698	100	0.1354	2.4812		0.1618	-143	0.1299

LARGE

BUYS					SELLS				
MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1958		0.1821	8671	0.1192	0.2041		0.1814	-8464	0.1236
0.4112		0.1843	8577	0.1140	0.4256		0.1829	-8762	0.1180
0.6722		0.1818	8987	0.1158	0.6911		0.1803	-8692	0.1151
1.0890		0.1809	9099	0.1114	1.1130		0.1793	-8687	0.1108
2.4515		0.1782	9574	0.1092	2.4732		0.1762	-9257	0.1100

MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1884		0.1693	1557	0.1251	0.2021		0.1687	-1784	0.1249
0.4033		0.1696	1562	0.1234	0.4271		0.1682	-1780	0.1235
0.6643		0.1700	1570	0.1241	0.6908		0.1680	-1792	0.1194
1.0826		0.1679	1571	0.1221	1.1145		0.1662	-1777	0.1198
2.4286		0.1665	1567	0.1201	2.4548		0.1631	-1774	0.1175

MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1983		0.1645	610	0.1302	0.2011		0.1644	-877	0.1277
0.4207		0.1635	609	0.1287	0.4265		0.1637	-872	0.1258
0.6890		0.1615	612	0.1276	0.6917		0.1628	-870	0.1242
1.1167		0.1602	614	0.1278	1.1124		0.1615	-867	0.1239
2.4570		0.1592	617	0.1260	2.4550		0.1599	-866	0.1216

MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2125		0.1566	271	0.1317	0.2033		0.1595	-413	0.1300
0.4489		0.1543	269	0.1322	0.4267		0.1573	-410	0.1296
0.7259		0.1533	269	0.1315	0.6921		0.1560	-409	0.1271
1.1610		0.1524	267	0.1303	1.1106		0.1552	-409	0.1273
2.5312		0.1532	269	0.1315	2.4612		0.1534	-409	0.1266

MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2264		0.1527	100	0.1332	0.2052		0.1533	-140	0.1332
0.4744		0.1522	100	0.1333	0.4285		0.1523	-139	0.1325
0.7605		0.1518	100	0.1343	0.6932		0.1510	-138	0.1315
1.2110		0.1501	100	0.1341	1.1167		0.1505	-138	0.1320
2.6022		0.1512	100	0.1338	2.4791		0.1501	-139	0.1305

Panel D: EMA(5)

SMALL

BUYS					SELLS				
MEAN TRADE	6754	MEAN SPR	0.1979		MEAN TRADE	-6950	MEAN SPR	0.2014	
		MEAN EFF	0.1070				MEAN EFF	0.1065	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1312		0.1996	6470	0.1098	0.1478		0.2099	-7371	0.1107
0.3496		0.2005	6878	0.1072	0.3790		0.1990	-6670	0.1086
0.6320		0.1955	6507	0.1059	0.6708		0.1999	-6139	0.1062
1.0691		0.1964	7248	0.1064	1.1242		0.2011	-7306	0.1030
2.4305		0.1975	6678	0.1056	2.5075		0.1972	-7257	0.1042

MEAN TRADE	1021	MEAN SPR	0.1896		MEAN TRADE	-1021	MEAN SPR	0.1882	
		MEAN EFF	0.1199				MEAN EFF	0.1193	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1394		0.1943	1022	0.1214	0.1477		0.1947	-1025	0.1219
0.3670		0.1887	1022	0.1211	0.3806		0.1866	-1016	0.1177
0.6516		0.1875	1018	0.1170	0.6728		0.1873	-1024	0.1214
1.0905		0.1888	1020	0.1211	1.1253		0.1860	-1020	0.1179
2.5234		0.1889	1022	0.1187	2.5042		0.1865	-1019	0.1180

MEAN TRADE	487	MEAN SPR	0.1847		MEAN TRADE	-522	MEAN SPR	0.1901	
		MEAN EFF	0.1242				MEAN EFF	0.1229	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1469		0.1888	486	0.1276	0.1476		0.1974	-520	0.1269
0.3864		0.1833	487	0.1240	0.3812		0.1923	-523	0.1248
0.6788		0.1822	487	0.1230	0.6724		0.1904	-522	0.1235
1.1212		0.1849	489	0.1246	1.1215		0.1848	-521	0.1196
2.4920		0.1844	488	0.1217	2.5478		0.1862	-522	0.1199

MEAN TRADE	239	MEAN SPR	0.1765		MEAN TRADE	-341	MEAN SPR	0.1741	
		MEAN EFF	0.1325				MEAN EFF	0.1278	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1584		0.1827	239	0.1361	0.1521		0.1772	-341	0.1290
0.3991		0.1764	240	0.1315	0.3800		0.1735	-341	0.1283
0.6843		0.1749	239	0.1317	0.6728		0.1750	-340	0.1267
1.1307		0.1730	238	0.1322	1.1204		0.1713	-342	0.1289
2.5296		0.1754	239	0.1310	2.4802		0.1739	-340	0.1264

MEAN TRADE	100	MEAN SPR	0.1788		MEAN TRADE	-144	MEAN SPR	0.1745	
		MEAN EFF	0.1371				MEAN EFF	0.1323	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1621		0.1839	100	0.1399	0.1534		0.1797	-145	0.1338
0.4045		0.1754	100	0.1357	0.3834		0.1762	-144	0.1334
0.6986		0.1761	100	0.1327	0.6741		0.1733	-145	0.1318
1.1581		0.1821	100	0.1421	1.1220		0.1730	-143	0.1323
2.5600		0.1763	100	0.1349	2.4751		0.1712	-143	0.1308

MEDIUM

BUYS					SELLS				
MEAN TRADE	9855	MEAN SPR	0.1815		MEAN TRADE	-9697	MEAN SPR	0.1876	
		MEAN EFF	0.1104				MEAN EFF	0.1113	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1404		0.1788	8916	0.1141	0.1498		0.1854	-8989	0.1185
0.3572		0.1833	9253	0.1102	0.3786		0.1888	-10405	0.1103
0.6366		0.1821	10190	0.1086	0.6627		0.1879	-9911	0.1096
1.0859		0.1825	10188	0.1091	1.1169		0.1879	-9479	0.1092
2.5063		0.1807	10730	0.1102	2.5778		0.1878	-9704	0.1087
MEAN TRADE	1141	MEAN SPR	0.1822		MEAN TRADE	-1119	MEAN SPR	0.1827	
		MEAN EFF	0.1200				MEAN EFF	0.1189	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1583		0.1857	1138	0.1258	0.1560		0.1877	-1124	0.1233
0.3900		0.1828	1137	0.1178	0.3802		0.1868	-1116	0.1209
0.6676		0.1821	1143	0.1176	0.6634		0.1818	-1117	0.1158
1.1013		0.1795	1140	0.1199	1.1204		0.1795	-1126	0.1172
2.4907		0.1810	1146	0.1191	2.4685		0.1788	-1115	0.1184
MEAN TRADE	513	MEAN SPR	0.1771		MEAN TRADE	-512	MEAN SPR	0.1743	
		MEAN EFF	0.1238				MEAN EFF	0.1230	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1747		0.1798	513	0.1267	0.1602		0.1797	-515	0.1262
0.4128		0.1787	513	0.1253	0.3792		0.1737	-510	0.1231
0.6963		0.1767	511	0.1204	0.6652		0.1729	-510	0.1218
1.1358		0.1775	514	0.1221	1.1158		0.1738	-511	0.1219
2.5255		0.1728	514	0.1244	2.4971		0.1724	-512	0.1232
MEAN TRADE	239	MEAN SPR	0.1675		MEAN TRADE	-300	MEAN SPR	0.1655	
		MEAN EFF	0.1298				MEAN EFF	0.1289	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1869		0.1691	240	0.1315	0.1615		0.1694	-300	0.1321
0.4289		0.1670	239	0.1286	0.3796		0.1657	-300	0.1301
0.7104		0.1669	238	0.1292	0.6660		0.1620	-300	0.1287
1.1519		0.1664	239	0.1302	1.1232		0.1645	-300	0.1272
2.4942		0.1680	237	0.1295	2.4838		0.1666	-300	0.1276
MEAN TRADE	100	MEAN SPR	0.1681		MEAN TRADE	-144	MEAN SPR	0.1658	
		MEAN EFF	0.1342				MEAN EFF	0.1314	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.1991		0.1711	100	0.1334	0.1623		0.1719	-145	0.1358
0.4479		0.1676	100	0.1336	0.3825		0.1681	-145	0.1322
0.7340		0.1665	100	0.1341	0.6666		0.1647	-144	0.1308
1.1736		0.1659	100	0.1340	1.1185		0.1626	-144	0.1292
2.5165		0.1694	100	0.1357	2.4797		0.1641	-143	0.1306

LARGE

BUYS					SELLS				
MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2065		0.1843	8563	0.1183	0.2110		0.1829	-8675	0.1233
0.4241		0.1840	9041	0.1158	0.4314		0.1823	-8479	0.1182
0.6817		0.1823	8911	0.1147	0.6945		0.1807	-8746	0.1138
1.0889		0.1804	9240	0.1110	1.1031		0.1788	-8826	0.1127
2.4111		0.1764	9153	0.1098	2.4305		0.1754	-9138	0.1094

MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2014		0.1696	1561	0.1242	0.2089		0.1691	-1784	0.1246
0.4192		0.1703	1568	0.1238	0.4309		0.1687	-1781	0.1227
0.6798		0.1699	1567	0.1240	0.6923		0.1677	-1785	0.1213
1.0967		0.1672	1564	0.1224	1.1064		0.1650	-1782	0.1191
2.4324		0.1664	1566	0.1204	2.4278		0.1638	-1775	0.1178

MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2102		0.1663	609	0.1302	0.2086		0.1646	-874	0.1276
0.4359		0.1628	612	0.1282	0.4322		0.1640	-874	0.1261
0.7019		0.1611	615	0.1283	0.6904		0.1619	-870	0.1240
1.1197		0.1602	611	0.1277	1.1056		0.1613	-871	0.1234
2.4340		0.1585	614	0.1259	2.4148		0.1604	-865	0.1221

MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2226		0.1570	270	0.1322	0.2110		0.1588	-412	0.1304
0.4600		0.1537	269	0.1317	0.4310		0.1580	-411	0.1293
0.7323		0.1533	268	0.1311	0.6942		0.1556	-411	0.1273
1.1582		0.1532	269	0.1316	1.1047		0.1547	-408	0.1274
2.4947		0.1525	268	0.1307	2.4186		0.1543	-408	0.1263

MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.2370		0.1530	100	0.1332	0.2119		0.1536	-140	0.1333
0.4821		0.1515	100	0.1344	0.4321		0.1517	-138	0.1317
0.7643		0.1509	100	0.1347	0.6926		0.1508	-139	0.1317
1.2040		0.1508	100	0.1333	1.1053		0.1512	-139	0.1316
2.5590		0.1518	100	0.1331	2.4185		0.1499	-138	0.1313

Table VII

Order Flow Deciles Dissected by WACD Quintiles

The data in Table VII were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. Table VII contains each set of stocks broken down first by trade decile and then further by unexpected trade intensity quintiles. For the buys and the sells orders (as assigned by the Lee and Ready (1991) algorithm) in each decile, mean trade size, quoted and effective spreads were calculated. The measure of trade intensity is captured by Engle's (1996) Weibull distributed autoregressive conditional duration model. The lag structure for the model was determined to be a WACD(1,2). The optimal lag structure was determined minimizing the Bayesian Information Criterion (BIC) statistic over the specifications that had each variable significant at the 5% level. The universe of specifications contained all possible permutations of (1,1) through (3,3) inclusive. The unexpected intensity variable (MA QUINTILE) quintile breakpoints are shown to the left, with the averages of the quoted or inside spread (SPREAD), trade size (SIZE), and effective spread (EFFSPR) to the right. For the small, medium and large sized stocks, there exists a general increase in effective spread as trade intensity (as proxied by the WACD(1,2) variable) increases, potentially indicating an increase in the adverse selection component of the spread. Panel A contains the results of the small stock portfolio, Panels B and C contain those of the medium and large market value stocks, respectively.

Panel A: Small

BUYS					**SELLS**				
MEAN TRADE	6754	MEAN SPR	0.1979		MEAN TRADE	-6950	MEAN SPR	0.2014	
		MEAN EFF	0.1070				MEAN EFF	0.1065	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.4846		0.1926	6272	0.1081	0.5139		0.2049	-7085	0.1115
0.7508		0.2005	7315	0.1110	0.7795		0.1993	-6434	0.1090
0.9147		0.2019	7283	0.1051	0.9392		0.2050	-6651	0.1063
1.0970		0.1991	6712	0.1061	1.1253		0.2023	-7577	0.1043
1.5924		0.1953	6188	0.1046	1.6150		0.1956	-7004	0.1015
MEAN TRADE	1021	MEAN SPR	0.1896		MEAN TRADE	-1021	MEAN SPR	0.1882	
		MEAN EFF	0.1199				MEAN EFF	0.1193	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.5101		0.1859	1026	0.1215	0.5166		0.1843	-1032	0.1211
0.7709		0.1924	1020	0.1224	0.7787		0.1879	-1016	0.1204
0.9246		0.1940	1020	0.1187	0.9384		0.1976	-1017	0.1203
1.1003		0.1923	1016	0.1182	1.1265		0.1885	-1014	0.1204
1.6017		0.1835	1023	0.1185	1.6220		0.1810	-1026	0.1144
MEAN TRADE	487	MEAN SPR	0.1847		MEAN TRADE	-522	MEAN SPR	0.1901	
		MEAN EFF	0.1242				MEAN EFF	0.1229	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.5439		0.1797	488	0.1275	0.5234		0.1930	-520	0.1274
0.7899		0.1865	487	0.1280	0.7835		0.1910	-522	0.1246
0.9387		0.1910	488	0.1221	0.9395		0.1965	-522	0.1221
1.1080		0.1884	487	0.1223	1.1226		0.1913	-521	0.1226
1.5849		0.1781	488	0.1210	1.5977		0.1780	-522	0.1182

MEAN TRADE	239	MEAN SPR	0.1765	
		MEAN EFF	0.1325	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.5582		0.1729	242	0.1350
0.7978		0.1744	238	0.1321
0.9395		0.1834	238	0.1317
1.1056		0.1802	238	0.1320
1.5587		0.1715	239	0.1317

MEAN TRADE	-341	MEAN SPR	0.1741	
		MEAN EFF	0.1278	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.5315		0.1704	-341	0.1285
0.7829		0.1727	-341	0.1282
0.9379		0.1785	-342	0.1302
1.1241		0.1778	-341	0.1276
1.6074		0.1686	-340	0.1241

MEAN TRADE	100	MEAN SPR	0.1788	
		MEAN EFF	0.1371	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.5882		0.1764	100	0.1392
0.8172		0.1861	100	0.1425
0.9532		0.1816	100	0.1364
1.1142		0.1775	100	0.1365
1.5613		0.1724	100	0.1308

MEAN TRADE	-144	MEAN SPR	0.1745	
		MEAN EFF	0.1323	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.5450		0.1737	-147	0.1353
0.7849		0.1738	-143	0.1332
0.9388		0.1807	-144	0.1332
1.1235		0.1758	-144	0.1325
1.5944		0.1660	-143	0.1280

Panel B: Medium

BUYS					SELLS				
MEAN TRADE	9855	MEAN SPR	0.1815		MEAN TRADE	-9697	MEAN SPR	0.1876	
		MEAN EFF	0.1104				MEAN EFF	0.1113	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.4542		0.1632	9620	0.1175	0.4986		0.1723	-9104	0.1208
0.7637		0.1876	8808	0.1115	0.7956		0.1967	-8686	0.1096
0.9314		0.1927	9829	0.1068	0.9498		0.1964	-9488	0.1098
1.0940		0.1887	10762	0.1057	1.1034		0.1911	-9698	0.1065
1.6149		0.1753	10257	0.1106	1.6402		0.1814	-11506	0.1097
MEAN TRADE	1141	MEAN SPR	0.1822		MEAN TRADE	-1119	MEAN SPR	0.1827	
		MEAN EFF	0.1200				MEAN EFF	0.1189	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.5401		0.1757	1127	0.1288	0.5329		0.1804	-1122	0.1282
0.8039		0.1908	1150	0.1206	0.7976		0.1895	-1121	0.1186
0.9454		0.1860	1146	0.1145	0.9489		0.1858	-1123	0.1167
1.0892		0.1841	1135	0.1163	1.1018		0.1822	-1121	0.1160
1.5616		0.1745	1147	0.1200	1.5760		0.1744	-1110	0.1180
MEAN TRADE	513	MEAN SPR	0.1771		MEAN TRADE	-512	MEAN SPR	0.1743	
		MEAN EFF	0.1238				MEAN EFF	0.1230	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.5848		0.1727	514	0.1272	0.5461		0.1733	-512	0.1306
0.8296		0.1838	513	0.1245	0.7973		0.1800	-511	0.1209
0.9631		0.1807	514	0.1201	0.9487		0.1788	-509	0.1228
1.1033		0.1789	509	0.1214	1.1031		0.1715	-512	0.1220
1.5205		0.1696	513	0.1257	1.5442		0.1673	-515	0.1218
MEAN TRADE	239	MEAN SPR	0.1675		MEAN TRADE	-300	MEAN SPR	0.1655	
		MEAN EFF	0.1298				MEAN EFF	0.1289	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.6264		0.1612	240	0.1315	0.5547		0.1648	-300	0.1347
0.8456		0.1703	240	0.1274	0.8001		0.1712	-300	0.1319
0.9707		0.1718	238	0.1292	0.9494		0.1661	-300	0.1276
1.1033		0.1682	238	0.1301	1.1022		0.1617	-300	0.1252
1.4677		0.1661	238	0.1307	1.4922		0.1641	-300	0.1289
MEAN TRADE	100	MEAN SPR	0.1681		MEAN TRADE	-144	MEAN SPR	0.1658	
		MEAN EFF	0.1342				MEAN EFF	0.1314	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.6532		0.1644	100	0.1356	0.5696		0.1705	-146	0.1366
0.8606		0.1698	100	0.1319	0.7983		0.1707	-144	0.1337
0.9768		0.1694	100	0.1334	0.9499		0.1657	-143	0.1302
1.1099		0.1661	100	0.1342	1.1018		0.1620	-144	0.1293
1.4734		0.1710	100	0.1359	1.5015		0.1634	-144	0.1305

Panel C: Large

BUYS					SELLS				
MEAN TRADE	8983	MEAN SPR	0.1815		MEAN TRADE	-8772	MEAN SPR	0.1800	
		MEAN EFF	0.1139				MEAN EFF	0.1155	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.6438		0.1879	8970	0.1252	0.6462		0.1879	-8759	0.1296
0.8072		0.1828	8836	0.1180	0.8146		0.1808	-8373	0.1190
0.9335		0.1801	8477	0.1112	0.9421		0.1779	-8729	0.1135
1.0842		0.1801	9008	0.1093	1.0957		0.1783	-8787	0.1097
1.4270		0.1766	9624	0.1057	1.4447		0.1752	-9213	0.1056
MEAN TRADE	1565	MEAN SPR	0.1687		MEAN TRADE	-1781	MEAN SPR	0.1668	
		MEAN EFF	0.1230				MEAN EFF	0.1210	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.6496		0.1706	1561	0.1300	0.6469		0.1697	-1786	0.1303
0.8128		0.1694	1565	0.1234	0.8156		0.1686	-1793	0.1237
0.9352		0.1694	1558	0.1215	0.9421		0.1677	-1783	0.1196
1.0835		0.1688	1570	0.1196	1.0950		0.1659	-1781	0.1170
1.4226		0.1652	1572	0.1203	1.4385		0.1624	-1765	0.1153
MEAN TRADE	612	MEAN SPR	0.1618		MEAN TRADE	-871	MEAN SPR	0.1624	
		MEAN EFF	0.1281				MEAN EFF	0.1246	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.6546		0.1643	607	0.1329	0.6446		0.1641	-879	0.1313
0.8243		0.1625	608	0.1288	0.8157		0.1649	-875	0.1261
0.9466		0.1633	615	0.1266	0.9423		0.1637	-869	0.1234
1.0955		0.1612	617	0.1255	1.0942		0.1608	-866	0.1225
1.4348		0.1576	613	0.1265	1.4411		0.1588	-865	0.1204
MEAN TRADE	269	MEAN SPR	0.1540		MEAN TRADE	-410	MEAN SPR	0.1562	
		MEAN EFF	0.1315				MEAN EFF	0.1281	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.6597		0.1542	271	0.1335	0.6396		0.1588	-415	0.1324
0.8378		0.1546	269	0.1308	0.8157		0.1573	-413	0.1299
0.9634		0.1550	270	0.1310	0.9426		0.1573	-408	0.1278
1.1125		0.1538	269	0.1309	1.0958		0.1555	-408	0.1266
1.4594		0.1521	267	0.1311	1.4421		0.1528	-407	0.1245
MEAN TRADE	100	MEAN SPR	0.1516		MEAN TRADE	-139	MEAN SPR	0.1513	
		MEAN EFF	0.1337				MEAN EFF	0.1319	
MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**	MA QUINTILE		**SPREAD**	**SIZE**	**EFFSPR**
0.6873		0.1512	100	0.1343	0.6469		0.1534	-141	0.1346
0.8572		0.1502	100	0.1330	0.8171		0.1522	-139	0.1327
0.9807		0.1529	100	0.1340	0.9439		0.1513	-139	0.1315
1.1275		0.1524	100	0.1344	1.0955		0.1516	-138	0.1313
1.4752		0.1514	100	0.1330	1.4441		0.1491	-138	0.1304

Table VIII

BSD Test for Dependence of Signed Order Flow and Trade Intensity

The data in Table VIII were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. The Brock, Scheinkman, Dechert (BSD) model calculates the significance level that the null of no dependence is rejected asymptotically. I adapted the test to see whether the difference between the normalized trade intensity variable and the normalized signed order flow were independent. The results for each firm (each firm's p-value) were then aggregated according to the Gibbons and Shanken (1986) algorithm. The transformation involves taking the natural log of each p-value and multiplying it by –2. Each summed variable follows a chi-square distribution with twice the degrees of freedom as the number of firms. Each size cohort and each unexpected trade intensity variable is presented below.

Panel A: p-Value from Aggregated BDS Statistic for Large Firms

SMA Aggregate p-value: 0.0014
EMA Aggregate p-value: 0.0012
WACD Aggregate p-value: 0.0000

Panel B: p-Value from Aggregated BDS Statistic for Medium Firms

SMA Aggregate p-value: 0.0000
EMA Aggregate p-value: 0.0020
WACD Aggregate p-value: 0.0015

Panel C: p-Value from Aggregated BDS Statistic for Small Firms

SMA Aggregate p-value: 0.0006
EMA Aggregate p-value: 0.0008
WACD Aggregate p-value: 0.0000

Table IX

Heteroscedasticity Tests for Threshold Glosten and Harris Model

The data in Table IX were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. White's test for heteroscedasticity was performed on the entire sample for each of the stocks in the small, medium and large cohorts. This regression consists of squaring the residuals from the regression of the observed price changes projected on the change in sign, the change in signed order flow, the signed order flow, and the signed order flow crossed with one of three unexpected trade intensity variables (where unexpected trade intensity is a proxy for temporal adverse selection). The three unexpected trade intensity variables are a simple moving average with lag length four, and exponential moving average of lag length three and Engle's WACD(1,2) model. Next, one regresses these squared residuals on a vector of ones and the squares of the right hand side variables in the original regression. High condition numbers prohibited adding cross products and higher order terms as exogenous variables. The p-value from each firm's Wald F-test was stored and then aggregated according to the Gibbons and Shanken (1986) algorithm. The transformation involves taking the natural log of each p-value and multiplying it by –2. Each summed variable follows a chi-square distribution with twice the degrees of freedom as the number of firms. Each size cohort is presented below. Panel A has the aggregated estimates, their t-statistic and their aggregate significance level in the form of a cohort p-value (from each firm's Wald F-test) from White's test for the three small firm sets. Each small firm data set is identical except for the choice of unexpected trade intensity variable (the same holds true for the medium and large data sets). Panels B and C have similar results for the medium and large firms, respectively.

Panel A: Small Firm Estimates and Aggregated p-values from White's Heteroscedasticity Test

Simple Moving Average(4)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald-p
Betas	4.50E-03	7.22E-04	-1.14E-12	3.13E-12	-5.20E-13	0.00
t	7.16	12.25	-0.90	1.38	-0.35	

Exponential Moving Average(3)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	4.49E-03	7.25E-04	-1.14E-12	4.16E-12	-6.79E-13	0.00
t	7.16	12.22	-0.88	1.56	-0.76	

WACD(1,2)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	4.48E-03	7.24E-04	-1.14E-12	-5.38E-12	1.12E-11	0.00
t	7.18	12.21	-0.91	-1.44	1.74	

Panel B: Medium Firm Estimates and Aggregated p-values from White's Heteroscedasticity Test

Simple Moving Average(4)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	2.70E-03	6.73E-04	4.24E-13	-2.57E-13	7.46E-13	0.00
t	10.06	19.16	1.54	-0.63	2.38	

Exponential Moving Average(3)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	2.70E-03	6.73E-04	4.10E-13	6.90E-14	6.23E-13	0.00
t	10.07	19.27	1.50	0.13	1.15	

WACD(1,2)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	2.70E-03	6.74E-04	4.23E-13	-3.53E-13	7.93E-13	0.00
t	10.10	19.12	1.54	-0.38	0.88	

Panel C: Large Firm Estimates and Aggregated p-values from White's Heteroscedasticity Test

Simple Moving Average(4)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	1.72E-03	6.59E-04	6.02E-13	4.80E-13	3.58E-14	0.00
t	7.42	15.79	3.71	1.07	0.48	

Exponential Moving Average(3)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	1.72E-03	6.59E-04	5.98E-13	2.39E-13	4.57E-14	0.00
t	7.42	15.81	3.68	0.87	0.86	

WACD(1,2)

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	1.72E-03	6.58E-04	6.09E-13	3.43E-14	5.08E-13	0.00
t	7.42	15.86	3.67	0.06	1.14	

Table X

Regime Specific Heteroscedasticity Tests for Threshold Glosten and Harris Model

The data in Table X were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. White's test for heteroscedasticity was performed on the each regime (as delineated by the first signed order flow threshold) for each of the stocks in the small, medium and large cohorts. This regression consists dividing each firm's data two sets, one with signed order flow less than or equal to the first threshold, and the second with the remainder of the original data set. Then, one squares the residuals from the regime specific regression of the observed price changes projected on the change in sign, the change in signed order flow, the signed order flow, and the signed order flow crossed with one of three unexpected trade intensity variables (where unexpected trade intensity is a proxy for temporal adverse selection). The three unexpected trade intensity variables are a simple moving average with lag length four, and exponential moving average of lag length three and Engle's WACD(1,2) model. Next, one regresses these squared residuals on a vector of ones and the regime specific squares of the right hand side variables in the original regression. High condition numbers prohibited adding cross products and higher order terms as exogenous variables. The p-value from each firm's Wald F-test was stored and then aggregated according to the Gibbons and Shanken (1986) algorithm. The transformation involves taking the natural log of each p-value and multiplying it by –2. Each summed variable follows a chi-square distribution with twice the degrees of freedom as the number of firms. Each size cohort is presented below. Panel A has the estimates from White's test for the small firm upper and lower regimes as divided by first signed order flow threshold. Each small firm data set is identical expect for the choice of the unexpected trade intensity variable. Panels B and C have similar results for the upper and lower regimes of medium and large firms, respectively.

Panel A: Small Firm Estimates and Aggregated p-values from White's Heteroscedasticity Test

Simple Moving Average(4)

Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald-p
Betas	5.09E-03	8.00E-04	-1.36E-11	2.27E-11	-4.07E-12	0.00
t	7.42	6.44	-1.10	1.32	-0.79	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald-p
Betas	4.37E-03	7.35E-04	-8.71E-13	5.00E-12	-5.61E-13	0.00
t	7.13	7.42	-0.96	1.62	-0.24	

Exponential Moving Average(3)

Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	5.08E-03	7.67E-04	-1.31E-11	2.08E-11	-1.28E-12	0.00
t	7.47	9.56	-0.99	1.22	-0.73	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	4.37E-03	7.12E-04	-8.24E-13	4.98E-12	-5.57E-13	0.00
t	7.15	8.13	-0.94	2.45	-0.57	

WACD(1,2)

Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	5.02E-03	7.23E-04	-1.33E-11	1.91E-11	3.36E-12	0.00
t	7.54	11.23	-0.91	1.23	0.55	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	4.46E-03	7.19E-04	-7.95E-13	-1.06E-12	9.04E-12	0.00
t	7.16	8.07	-0.93	-0.10	0.74	

Panel B: Medium Firm Estimates and Aggregated p-values from White's Heteroscedasticity Test

Simple Moving Average(4)
Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	3.08E-03	7.22E-04	4.40E-14	1.66E-12	3.25E-13	0.00
t	10.27	14.92	0.61	1.34	0.84	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	3.26E-03	6.36E-04	1.94E-12	-1.52E-12	4.29E-14	0.00
t	7.85	15.12	1.19	-0.94	0.07	

Exponential Moving Average(3)
Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	3.01E-03	7.42E-04	6.45E-14	2.09E-12	-3.52E-13	0.00
t	10.00	17.14	1.07	1.55	-0.68	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	3.22E-03	6.19E-04	2.03E-12	-1.65E-12	-1.67E-13	0.00
t	8.03	15.21	1.23	-1.00	-0.29	

WACD(1,2)
Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	3.19E-03	7.35E-04	1.26E-13	2.32E-12	-1.50E-12	0.00
t	9.67	13.68	0.50	0.58	-0.44	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	3.24E-03	6.47E-04	1.92E-12	-4.87E-12	4.88E-12	0.00
t	8.11	15.02	1.15	-1.63	1.37	

Panel C: Large Firm Estimates and Aggregated p-values from White's Heteroscedasticity Test

Simple Moving Average(4)

Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	2.23E-03	7.23E-04	6.72E-13	7.06E-13	-1.13E-13	0.00
t	7.11	19.22	2.82	1.55	-1.28	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	1.69E-03	6.37E-04	7.18E-13	7.54E-14	7.49E-14	0.00
t	6.78	14.01	4.03	0.20	0.92	

Exponential Moving Average(3)

Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	2.05E-03	7.21E-04	7.15E-13	4.02E-13	-1.57E-14	0.00
t	8.20	18.02	2.92	0.79	-0.33	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	1.78E-03	6.37E-04	6.84E-13	1.16E-13	6.74E-14	0.00
t	6.48	13.58	3.84	0.31	0.93	

WACD(1,2)

Upper

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	2.27E-03	7.42E-04	7.75E-13	1.14E-13	1.74E-13	0.00
t	7.21	14.78	3.06	0.23	0.43	

Lower

	const	c(0)^2	c(1)^2*	z(1)^2*	z(2)^2*	Wald
Betas	1.71E-03	6.31E-04	6.69E-13	-1.23E-13	3.50E-13	0.00
t	6.87	14.39	4.07	-0.20	0.64	

Table XI

Signed Order Flow and Trade Intensity Thresholds under Heteroscedasticity

The data in Table XI were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. Table XI contains information derived from a generalization of Glosten and Harris' spread decomposition model. The original model was modified to allow for m thresholds (or m+1 regimes) in the signed order flow and n thresholds (or n+1 regimes) in the trade intensity variable. The sample value for m and n in this study was 2 and 2. Two signed order flow and 2 trade intensity threshold parameters were calculated for each of the three sets of firms. For Panel A, section one contains the average thresholds for both variables for all three sets of firms. For each firm, trade size (on the same side as their first threshold parameter) and trade intensity averages are reported in section two and percentages appear in section three. Panel B has the results of the significance tests for the first and second thresholds. For the trade volume followed by the trade intensity variable, aggregated p-values for the first and second threshold are presented as an F-statistic. Because the size of the threshold is not identified until the alternative, simulations are run to determine the significance level of each threshold. Each firm has a bootstrapped p-value arising from the non-standard F-test testing zero versus one threshold, F(0,1) and then for the non-standard F-test testing one versus two thresholds, F(1,2). The set of bootstrapped F-test p-values for each firm and each threshold were transformed to a sample aggregate p-value using the Gibbons and Shanken ('87) algorithm. The transformation involves taking the natural log of each p-value and multiplying it by -2, then summing all of the transformed p-values. Each summed variable follows a chi-square distribution with twice the degrees of freedom as the number of firms. Surprisingly, threshold values hover around 35-45% of the average order flow. Statistical test indicate that three regimes (two thresholds) exist for both order flow and unexpected trade intensity. Although the results from Tables IX and X suggest that heteroscedasticity is present in the data, adjusting the bootstrapped F-test to account for heteroscedasticity did not alter the results of the aggregated p-values. The first and second volume thresholds are broken down into average of those that are positive (Pos) and those that are negative (Neg). The first and second trade intensity thresholds are broken down into those that are less than one (occur faster than normal) and those that are greater than one (occur slower than normal).

Panel A: Threshold Summary: Small Firms

		Volume				**Trade Intensity**			
		Pos	**Neg**			**< 1**	**> 1**		
		First		Second		First		Second	
AVERAGE THRESHOLD	SMA	467	-619	653	-731	0.520	1.860	0.629	1.5477
	EMA	589	-667	672	-900	0.571	1.499	0.365	1.9064
	WACD	625	-711	587	-793	0.797	1.179	0.800	1.1906
AVERAGE TRADE SIZE	SMA	1631	-1528						
	EMA	1614	-1564						
	WACD	1624	-1593						
PERCENT OF AVERAGE	SMA	0.286	0.405	0.400	0.478				
	EMA	0.365	0.426	0.417	0.575				
	WACD	0.385	0.446	0.361	0.498				

VOLUME

SMA	**F(0,1) =**	0.000*
EMA	**F(0,1) =**	0.000*
WACD	**F(0,1) =**	0.000*

VOLUME

SMA	**F(1,2) =**	0.000*
EMA	**F(1,2) =**	0.000*
WACD	**F(1,2) =**	0.000*

TRADE INTENSITY

SMA	**F(0,1) =**	0.000*
EMA	**F(0,1) =**	0.000*
WACD	**F(0,1) =**	0.000*

TRADE INTENSITY

SMA	**F(1,2) =**	0.002
EMA	**F(1,2) =**	0.000*
WACD	**F(1,2) =**	0.000*

Panel B: Threshold Summary: Medium Firms

		Volume				Trade Intensity			
		Pos	Neg			< 1	> 1		
		First		Second		First		Second	
AVERAGE THRESHOLD	SMA	1233	-978	669	-576	0.590	1.461	0.641	1.272
	EMA	938	-942	777	-606	0.583	1.921	0.509	1.255
	WACD	1233	-1267	708	-506	0.845	1.152	0.827	1.117
AVERAGE TRADE SIZE	SMA	1908	-2250						
	EMA	1862	-2277						
	WACD	1908	-2250						
PERCENT OF AVERAGE	SMA	0.646	0.435	0.351	0.256				
	EMA	0.504	0.412	0.417	0.266				
	WACD	0.646	0.563	0.371	0.225				

VOLUME		
SMA	**F(0,1) =**	0.000*
EMA	**F(0,1) =**	0.000*
WACD	**F(0,1) =**	0.000*
VOLUME		
SMA	**F(1,2) =**	0.000*
EMA	**F(1,2) =**	0.000*
WACD	**F(1,2) =**	0.000*
TRADE INTENSITY		
SMA	**F(0,1) =**	0.000*
EMA	**F(0,1) =**	0.000*
WACD	**F(0,1) =**	0.000*
TRADE INTENSITY		
SMA	**F(1,2) =**	0.000*
EMA	**F(1,2) =**	0.001
WACD	**F(1,2) =**	0.000*

Panel C: Threshold Summary: Large Firms

		Volume				Trade Intensity			
		Pos	**Neg**			**< 1**	**> 1**		
		First		Second		First		Second	
AVERAGE THRESHOLD	SMA	1029	-1000	810	-720	0.792	1.168	0.629	1.178
	EMA	1014	-1122	795	-1145	0.603	1.603	0.465	1.373
	WACD	1500	-1225	853	-736	0.792	1.168	0.906	1.178
AVERAGE TRADE SIZE	SMA	2296	-2389						
	EMA	1862	-2277						
	WACD	2519	-2328						
PERCENT OF AVERAGE	SMA	0.448	0.419	0.353	0.301				
	EMA	0.545	0.493	0.427	0.503				
	WACD	0.596	0.526	0.339	0.316				

VOLUME		
SMA	**F(0,1) =**	0.000*
EMA	**F(0,1) =**	0.000*
WACD	**F(0,1) =**	0.000*

VOLUME		
SMA	**F(1,2) =**	0.000*
EMA	**F(1,2) =**	0.005
WACD	**F(1,2) =**	0.000*

TRADE INTENSITY		
SMA	**F(0,1) =**	0.000*
EMA	**F(0,1) =**	0.000*
WACD	**F(0,1) =**	0.000*

TRADE INTENSITY		
SMA	**F(1,2) =**	0.001
EMA	**F(1,2) =**	0.000*
WACD	**F(1,2) =**	0.000*

*values smaller than 0.0001

Table XII

Relative Root Mean Squared Errors of Effective Spread Forecasts

The data in Table XII were pulled from the Trade and Quote (TAQ) database during 1996. Three samples of 30 stocks were taken at random from the 10th (large), 8th (medium) and 6th (small) deciles. The large sample has trade data for 20 trading days (twenty trading days after April 1st), the medium sample for 60 trading days (three trading months after April 1st) and the small for 120 trading days after April 1st. Table XII contains information derived from a generalization of Glosten and Harris' spread decomposition model. The original model was modified to allow for m thresholds (or m+1 regimes) in the signed order flow and n thresholds (or n+1 regimes) in the trade intensity variable. The sample value for m and n in this study was 2 and 2. Two signed order flow and 2 trade intensity threshold parameters were calculated for each of the three sets of firms. In Panel A, the relative root mean square errors appear for the small, medium and large stocks using out of sample data sets for each of the three sets of stocks. For each firm in each cohort, the square residuals are collected from the out of sample data. The squared residuals are collected over all stocks in the cohort and the square root of the mean of the cohort vector is calculated. Ratios of the cohort RMSE are formed using the RMSE for the simple moving average (SMA) as the numerator for each of the ratios i.e. ratios less than one indicate that the SMA RSME was less than the RMSE of the model with another trade intensity variable. The SMA works best in the small and medium stock cohorts. With the large cohort of stocks, the SMA intensity proxy approximately ties the RMSE with the WACD trade intensity and lags the EMA by 2.4%.

Relative Root Mean Square Errors

	SMALL	MEDIUM	LARGE
SMA	1	1	1
EMA	0.977042	0.947153	1.024261
WACD	0.962874	0.917799	1.009846

Appendix A

Parametric Estimation of Abnormal Returns

The choice of which method one should use to identify the presence and the magnitude of abnormal returns is central to many papers. Typically one thinks of a market model type analysis where the errors are interpreted as abnormal returns for firm i during an event window T.

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + e_{i,t} \tag{A.1}$$

so the cumulative abnormal returns are the summation of $e_{i,t}$ from day 0 to T.

$$CAR_i = \sum_{t=0}^{T} e_{i,t} \tag{A.2}$$

What happens to the CAR estimates when there exists structural change at the time of the event? The parameter shift would cause the residuals to take on a different behavior. This should cause the CARs (using the old parameters) to trend in one direction or another during the event window. This is exactly what we see in Table IV Panel B. The farther the event window is carried out, the more negative the CARs become. Clearly, this statistical result does not wash with our economic intuition. A test of parameter instability is used to test whether structural change has occurred in the market model after the addition to the S&P 500.

Panel A of Table VI shows the results from the Wald tests of changes in the intercept and slope coefficients. The results of the Wald test of whether the slope *and* the intercept change from pre to post addition indicate parameter instability arises around the event. Aggregate p-values were created from each firm's test statistic. The Wald tests show p-values close to zero indicating a high probability of structural change around the time of index addition. This seems to imply that parametric estimation of abnormal returns is a much more complicated issue that it first appears.

In Panel C we see that when the CARs are calculated with a set of parameters estimated from an interval of time beginning 100 days after Day 0, the price effects of addition are fleeting. This qualitatively agrees with the bootstrapped BHAR result proposed by Efron and Tibshirani. Also, the inferences from the parametric model are problematic because of survivorship bias. 16 firms dropped out of the sample one year after the event. This was most likely due to mergers and acquisitions which were in turn a function of their abnormal performance. This leads me to advocate side stepping the parametric estimation issue altogether and relying on a BHAR measure with a bootstrapped null test statistic.

Appendix B

The Generalized Impulse Response Function Algorithm

1. Calculate the appropriate lag length for each firm as described in the text and collect the residuals from the VAR(p) regressions.

2. If V is the KxT matrix of residuals from the regressions,

$$V = \begin{pmatrix} V_{1,1}, V_{1,2}, \ldots, V_{1,T} \\ V_{2,1}, V_{2,2}, \ldots, V_{2,T} \end{pmatrix}$$

or

$$V = (\overrightarrow{V_1}, \overrightarrow{V_2})^T \tag{B.1}$$

which is TxK. Let V have a mean zero, normal distribution with variance Σ. We assume that Σ is constant and may be estimated from the data. Given that Σ is the constant covariance matrix for V, then C is the Cholesky Decomposition of $var(V)$ such that $\Sigma = CC^T$.

3. Premultiply V by C^{-1}, we can compute the underlying mean 0 standard deviation 1 shocks, e_t $N(0, I)$, where the dimension of I is KxK.

4. Premultiply this diagonal matrix e_t by C to give each normalized error vector the same standard deviation as the original errors. This returns the original variance structure while eliminating the dependence between revision and order flow shocks. Now only the diagonal elements of C appear in the covariance matrix.

5. Select a history consisting of a series of p observations from the revision and signed order flow series, ω_{t-1}.

6. Sample N+1 elements from the new orthogonal errors.

7. At N=0, include the error of the magnitude of one standard deviation in the order flow equation and an error collected from the step 4 in the revision equation. This will form the left hand side of the difference in the $GIRF_{Rev}$ definition.

8. The benchmark case will use the same history as 5, but with an N+1 vector of orthogonal shocks in both equations.

9. Now apply the same remaining N+20 shocks to each set of equations (the experiment or left hand side and the benchmark case).

10. Repeat 4-7 1000 times to create averages of the GI over the same

ω_{t-1} and then sample over all histories to get the sample average.

11. Repeat 8 100 times to get 100 draws from the mean distribution of the GIRF. Collect the mean and standard deviation to compare to other cases.

Appendix C

Wall Street Journal Article

From the Wall Street Journal:

The effect may be accentuated by some funds' trading techniques. Robert Laible, director of sales and trading at ITG Inc. which operates an electronic stock-trading system for institutions, says index-fund managers are paid to match the index; so as to minimize their "tracking error", they want to buy new additions with the price guaranteed at the closing price on the addition date. But that often means buying it at an artificially high price.

To avoid that, some index managers buy before the addition date, but that is risky if the stock does not rally. So instead, some fund managers strike profit-participation agreements with brokers. The fund contracts with the broker to buy the needed stock prior to the addition date, with the purchase price guaranteed at the close the day the stock enters the index. But if the *broker makes a profit by buying the stock at less than the price at which he sells it to the fund manager, he splits the profit with the fund, giving the fund a slight extra return on the index.*

Mr. Laible says index managers, realizing they are buying the stock at an artificial high, figure that "might as well try to recoup some of that" by entering into some sort of profit-sharing arrangement to earn back some of the excess mark-up in the stock's price.

Some of brokers fret that profit-sharing may give traders the incentive to drive up the price of a stock being added to the index as high as possible to maximize the profit to be shared.

Bibliography

[1] Amihud, Y., Mendelson, H., 1986, *Asset Pricing and the Bid-Ask Spread*, Journal of Financial Economics, 17, 223-249.

[2] Beneish, M., Whaley, R., 1996, *An Anatomy of the "S&P Game": The Effects of Changing the Game*, Journal of Finance, 51, 1909-1930.

[3] Harris, L., Gurel, E., 1986, *Price and Volume Asscociated with Changes in the S&P500 List: New Evidence for the Existence of Price Pressure*, Journal of Finance, 41, 815-829.

[4] Ghysels, E., Hall, A., 1991, *A Test for Structural Stability of Euler Conditions Parameters Estimated Via the Generalized Method of Moments Estimator*, International Economic Review, 31, ,2, 355-365.

[5] Hasbrouck, J., 1991, *The Summary Informativeness of Stock Trades: An Econometric Analysis*, Review of Financial Studies, 4, s, 571-595.

[6] Hasbrouck, J., 1991, *Measuring the Information Content of Stock Trades*, Journal of Finance,44, 1, 179-207.

[7] Lynch, A. Mendenhall, R., 1997, *New Evidence on Stock Price Effects Associated with Changes in the S&P500 Index*, Journal of Business, 70, 351-383.

[8] Glosten, M., Harris L.E., 1988, *Estimating the Components of the Bid/Ask Spread*, Journal of Financial Economics, 21, 1, 123-142.

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

[10] Hedge, S., McDermott, J., 2001, *The Liquidity Effects of Revisions to the S&P500 Index*, Working Paper, University of Connecticut.

[11] Fama, E., 1998, *Market Efficiency, Long-Term Returns and Behavioral Finance*, Journal of Financial Economics, 49, 3, 283-306.

[12] Efron, B., Tibshirani, R., 1998, *An Introduction to the Bootstrap*, Monographs on Statistics and Applied Probability 57.

[13] Madhavan, A., Richardson, M., Roomans, M., 1997, *Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks*, Review of Financial Studies, 10, 4, 1035-1064.

[14] Van Ness, B., Van Ness, R., Warr, R., 2001, *How Well Do Adverse Selection Components Measure Adverse Selection?*, Financial Management, Vol. 30, Iss. 3,

[15] Lee, C. M. C., Ready, M., 1991, *Inferring Trade Direction from Intraday data*, Journal of Finance, 46, 2, 733-746.

[16] Gibbons, M., Shanken, J., 1987, *Subperiod Aggregation and the Power of Multivariate Tests of Portfolio Efficiency*, Journal of Financial Economics, 19, 389-394.

[17] Hansen, B., 1999 *Testing for Linearity*, Journal of Economic Surveys, 13, 551-576.

[18] Bai, J., Perron, P. 1998 *Estimating and Testing Linear Model with Multiple Structural Changes*, Econometrica, 66, 47-78.

[19] Engle, R., Russell, J., 1998, *Autoregressive Conditional Durations:A New Model for Irregularly Spaced Transction Data*, Econometrica, 66, 1127-1162.

[20] Kempf, A., Korn, O., 1999, *Market Depth and Order Size*, Journal of Financial Markets, Vol. 2, 29-48.

[21] Dufour, A., Engle, R., 2000, *Time and the Price Impact of a Trade*, Journal of Finance, Vol. 2, 29-48.

[22] Huang, R. D., Stoll, H. R., 1997, *The Components of the Bid-Ask Spread: A General Approach*, Review of Financial Studies, 10, 4, 995-1034.

[23] Lin, J., Sanger, G., Booth, G.,G., 1995, *Trade Size and Components of the Bid-Ask Spread*, Review of Financial Studies, 8, 4, 1153-1183.

[24] Madhavan, A., Smidt, S, 1991 *A Bayesian Model of Intraday Specialist Pricing*, Journal of Financial Economics, 30, 99-134.

[25] Stoll, H., 1989 *Inferring the Componenets of the Bid-Ask Spread: Theory and Empirical Tests*, Journal of Finance, 44, 115-134.

[26] O'Hara, M., 1995 *Market Microstructure Theory*, Oxford: Basil Blackwell Inc.

[27] Easley, D., O'Hara, M. 1992 *Time and the Process of Security Price Adjustment*, Journal of Finance, 47, 577-605.

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

[29] Diamond, D., Verrecchia, R. 1987 *Constraints on Short Selling and Asset Price Adjustment to Private Information*, Journal of Financial Economics, 18, 277-311.

[30] Foster, F., Viswanathan, R. 1993 *Variations in Trading Volume, Return Volatility and Trading Costs: Evidence on Recent Price Formation Models*, Journal of Finance, 48, 187-212.

[31] Easley, D., Engle, R., O'Hara, M., Wu, L. 2001 *Time Varying Arrival Rates of Informed and Uninformed Trades*, Working Paper.

[32] Campbell, J., Lo, A., MacKinlay, A., 1997 *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

Edward Glidewell

940 Cricklewood Drive #137
State College, Pa. 16803
ebg110@psu.edu

EDUCATION

The Pennsylvania State University
Ph.D. Finance August 2004

Texas A&M University
M.S. Mathematics August 1998

ACADEMIC AND PROFESSIONAL EXPERIENCE

Academic

Teaching Penn State University
1/02 - 6/04 Finance 410, *Speculative Markets*

Research Penn State University
9/00 - 12/01 Assisted Ian Domowitz

Real

1/97 - 8/98 Law and Economics Consulting Group (LECG) Assisted in various price fixing/ collusion cases as well as the State of Washington tobacco settlement

1/01 - 9/01 NO2 Inc.
Performed research in the pollution credit market to price forwards and option in the absence of a liquid spot market market

HONORS, ACADEMIC AND PROFESSIONAL MEMBERSHIP

Doctoral Fellowship 2000-2002
AFA, FMA, AMS
Passed CFA Level III exam and am currently working toward the work experience requirement for the CFA® designation