1989

Essays on the Components of the Bid-Ask Spread.

Pei-hwang Wei

Louisiana State University and Agricultural & Mechanical College

Order Number 9017306

Essays on the components of the bid-ask spread

Wei, Pei-Hwang, Ph.D.

The Louisiana State University and Agricultural and Mechanical Col., 1989

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Essays on the Components of the Bid-Ask Spread

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Interdepartmental Program in Business Administration

by
Pei-Hwang Wei
B.S., National Taiwan University, 1980
M.S., Louisiana State University, 1984

August 1989

ACKNOWLEDGEMENTS

I wish to thank members of my dissertation committee, Gary C. Sanger, G. Geoffrey Booth, R. Carter Hill, John S. Howe, and William R. Lane, for their effort and guidance. With their outstanding guidance, preparation of this dissertation is a tremendous learning experience to me. I especially like to thank my chairman, Gary C. Sanger for numerous helpful comments and suggestions. Also, thanks to many other people who are not in my committee but provide helpful comments. Special thanks is to Mel Jameson, who taught me the foundation of the literature related to my dissertation and motivates me to pursue in this area. Christopher Lamoureux provides some stimulating suggestions. Stephen Looney gives some very useful thoughts.

Fellow students also give me helpful suggestions. I cannot overemphasize the importance of the friendly and supportive working environment that is shared by many students in the finance department. Finally, I would be like to thank my parents who give me mental support and strength in preparing this dissertation.

TABLES OF CONTENTS

LIST OF TABLES

Abstract

The presence of the bid-ask spread causes equilibrium prices to deviate from transaction prices. More importantly, the components of the spread - the order component, the inventory component, and the information component - have different impacts on transaction prices. In general, the order and the inventory components induce negative correlation in successive transaction price changes. On the other hand, transaction prices will form a martingale if only the information component exists. This dissertation is composed of three related essays that utilize this general relationship between transaction prices and the components of the spread. The first essay employs a cross-sectional analysis that relates components of the spread (costs of specialists) to measures of market activity, risk, and information risk. Primary results indicate that information component is positively related to transaction size and insider holdings. The order component is positively associated with number of transactions. Institutional holdings is negatively related to the order and the inventory components. The second essay gives empirical evidence that trading activity, price variability, information component are higher in the opening

transactions. On the other hand, order component, inventory component, and a measure of the transaction cost are estimated to be lower in the opening. These results give support to the importance of liquidity trading and the role of transaction costs. For the last trading hour, there are significant changes in trading activity and price variability without significant changes in components of spread, which is consistent with the previous finding of a positive relationship between volume and price changes. Some problems remain in the estimation of components of the spread. Therefore, caution must be taken with respect to results provided in essay one and two. A model that relates components of order imbalances (net volumes) and components of the spread is established in the third essay. The main result highlights the potentially important role of order imbalances in the information process. That is, variations in order imbalances give information concerning the degree of informational asymmetry in financial markets.

Chapter I

Introduction

In classical economic analysis, the existence of a Walrasian auctioneer assures that prices instantaneously converge to equilibrium prices. In reality, not all markets have such a perfect setting. Often traders have to rely on some imperfect substitutes such as a dealership market. To maintain a continuous and liquid market, a dealer stands ready to trade with incoming orders for his/her own account. In assuming this task, however, dealers bear some costs and incur some risk that must be compensated for. Therefore, the dealer buys low and sells high in order to cover the cost. The difference between dealer's purchase and sale prices is termed the dealer's bid-ask spread. Demsetz (1968) pioneers the study of the bid-ask spread and hypothesizes that the spread is a compensation to the dealer for providing immediacy to the traders. What is still unclear is the complete nature of the cost(s) of providing such services.

Stoll (1978) states that there are three main costs that a dealer might face : orders processing cost, inventory cost and information cost. Earlier studies focus on the inventory implications of trades to the dealer. The dealer, as many other investors, seeks diversification and has a desired inventory level for the stock the dealer specializes in. This means inventory imbalances present a cost to the dealer. Ordering costs are those incurred in handling orders such as communication and clerical costs. Surprisingly, not many studies have been devoted solely to the investigation of ordering costs. One natural question is whether the dealer industry is characterized by economies of scale or economies of scope in ordering costs. Answer to this question is still unclear at this stage.

Bagehot (1971) first points out another type of cost that is potentially important. Knowing the possibility that the trade may be originated by a trader with superior information over the dealer, the dealer will adjust the quoted price so as to make up the loss to informed traders. Eventually, the dealer loses to informed traders and gains from uninformed traders.

The studies in the bid-ask spread and the components of the spread have important implications for many other areas in finance. First, the spread introduces noise into transaction prices. The earliest study goes back to Neiderhoffer and Osborne (1966). These authors show that there is a general tendency for price reversal between trades and suggest the spread as being one of the causes for price reversal. In a theoretical paper, Glosten and Milgrom

(1985) show that different types of costs affect transaction prices differently. More explicitly, Glosten (1987) shows that the ordering/inventory cost produces biases in mean return, variance and serial covariance while the information cost attenuates the biases. This indicates that knowledge of the magnitude of the spread as well as the composition of the spread is needed to correct biases induced by the spread. Second, the information cost model ia strongly related to market operations and market efficiency. Being a central figure in the market, how does a dealer react to trades that may or may not originate from an informed trader? Easley and O'Hara (1985) give an intriguing discussion of market efficiency and the role of dealer's trading activity. Third, these studies are relevant for making decisions about the regulation of the security industry. For example, a large order cost suggests the security industry is not operated efficiently. Fourth, since the spread represents a form of transaction cost to investors, the bid-ask spread can be used as an indicator of market liquidity. Amihud and Mendelson (1986) hypothesize there is a transactions cost clientele and show that expected return is an increasing and concave function of the spread. Increasingly, other types of applications can also be found. Barclay and Smith (1988), for example, explore the information implication of spread adjustment to corporate repurchase announcements.

The primary purpose of this dissertation is to further extend the analysis of the components of the bid-ask spread, with particular emphasis on the application and extension of the concept of the components of the spread. The dissertation is composed of three separate essays, although they are strongly related. Also, although the dealer function is frequently observed on other markets (e.g., used car market), we focus on dealers on the organized stock exchanges (i.e., specialists).

To empirically estimate the magnitude of the three components of the spread has been shown to be difficult. For one thing, the precise nature of inventory and information costs is subjective and therefore an objective measure of these costs appears to be lacking. Ho and Macris (1984) and Glosten and Harris (1988) utilize the intuition provided by Glosten and Milgrom to estimate the components of the spread. Nevertheless, each estimation approach assumes the absence of a certain component of spread. A more general approach can be found in Stoll (1989) that extends Roll's (1984) analysis in measuring the effective spread. In testing his model, Stoll (1989) uses daily closing data. The first essay applies Stoll's analysis to transaction-by-transaction data, which is more appropriate for his model. More importantly, we apply a slight variant of his approach so that the complete nature of dealer's costs can be studied.

Recent studies (e.g., Foster and Viswanathan (1988)) find that volume and price variability are higher in the opening and the closing hours. Satisfactory explanation for these results is still lacking. These results are consistent with investors' minimization of the transaction cost if the transaction cost is lower in the opening and the closing periods. Admati and Pfleiderer (1988) present a more restricted model based on the notion of minimization of the information cost. Their model, as well as other explanations that may be consistent with the concentration of volume and price variability, are tested in the second essay.

The trading volume can be broadly decomposed into two components : the information-motivated and the liquidity -motivated trading. Observed trading activity then reflects the interaction between informed and uninformed traders. In the third essay, an attempt is made to link together the concept of the components of the spread and the concept of the components of trading volume. The intuition behind this is that the components of the spread reflect the degree of informational asymmetry between the dealer and traders while the components of trading volume reflect the degree of informational asymmetry between informed and uninformed traders. Combination of these two pieces of information should be valuable in inferring the degree of informational asymmetry in financial markets.

Since these three essays are strongly related, Chapter II gives a comprehensive literature review so that the reader can more easily grasp the relationship among these essays. Chapters III, IV, and V present the results of essay one, two and three, respectively. Chapter VI concludes this dissertation.

Chapter II

Literature Review

II.A. The Components of the Bid-Ask Spread

The bid-ask spread is generally viewed as a composition of three types of costs faced by the dealer[1] : the ordering cost, the inventory cost and the information cost components, ignoring the potential monopoly rent that the dealer may enjoy. Absence of monopoly power may be approximately true for specialists in NYSE because they face competition such as traders submitting limit orders, floor traders, large block brokers and dealers on other exchanges. On the other hand, monopoly power may be also absent in the NASDAQ market where a system of multiple market makers is employed. The degree of competition is usually assumed to be great so that the spread does not contain a monopoly component.

The inventory cost arises from the diversification need of the dealer. In this context, a desired level of the inventory for the dealer is usually assumed. Moving away

[1]Dealer is a more general terminology. The specialists on organized exchanges and the market makers on the options market and NASDAQ market are performing the dealer function.

from the desired inventory level is undesirable to the dealer and the dealer will attempt to adjust the quoted price in order to restore his/her desired inventory level. The inventory cost is modeled by Garman (1976), Bradfield (1979), Amihud and Mendleson (1980), Ho and Stoll (1981), Mildenstein and Schleef (1983) and O'Hara and Oldfield (1986). These studies differ with respect to the dealer's utility function, the degree of competition among dealers, and the stochastic arrival of orders. Generally, they show that optimal quoted prices are function of the dealer's desired inventory position.

Since Bagehot (1971), many studies model information cost based on the asymmetric information between the dealer and the traders. Nevertheless, these models often follow diverse approaches. Copeland and Galai (1983) model spread being equivalent to writing a combination of put and call options to informed traders. According to the option pricing theory, the spread is then a function of stock return variances. Other models focus on the role of the informational asymmetry in a dynamic trading environment. Kyle (1985) develops a sequential equilibrium model in which a single informed trader can choose the timing of his/her trading. His/her information is gradually revealed to the public through a number of trades. Further, the speed of revelation of information through trades depends on several market activity variables such as the depth of the market (i.e., amount of uninformed trading or noise trading). The basic intuition is that informed trader will prefer to trade in a deeper market where the cost of trade is lower. Easley and O'Hara (1987) hypothesize that the trade size is an important factor since informed traders will prefer to trade large quantities to maximize their profit. The impact of the information component on transaction prices will be discussed in section C. More discussion of the role of informational asymmetry in the trading dynamics can be found in section E.

One shortcoming of the previous research is that the potential monopoly power of specialists is not explicitly considered. Glosten (1989) provides a rationale for the coexistence of the specialist system and the multiple market makers system. If the information problem is severe in the market, a competitive market making system would likely result in a market shut-down. Under the specialist system, the market will likely remain open even in the presence of severe information asymmetry, since the specialist likely enjoys some monopoly power and can make up today's losses in the future. However, specialists do face some competition, and the issue of the degree of competition is clearly an empirical one.

II. B. Determinants of the (Total) Spread

Another class of literature is less model-specific; instead it presents the more general notion that the spread is a function of several important variables such as market activity, risk, and the degree of competition.

Demsetz (1968) provides the theoretical foundation for the bid-ask spread as well as the first empirical test of the determinants of the spread. In his test, the magnitude of the spread is used as the dependent variable. Independent variables that proxy for the level of trading activity and the degree of competition are used to explain the variations in the magnitude of the spread.

Tinic (1972), Tinic and West (1972), Benston and Hagerman (1974), Hamilton (1976), Branch and Freed (1977) and Stoll (1978a) generally follow Demsetz' approach with some refinement in the measurement of the market activity, risk, and competition. The intuition of their studies is simple. Since the magnitude of the spread represents the costs of the dealer, it should be related to the scale and the riskiness of his/her business. Generally speaking, these empirical studies find that the level of stock price, volume, number of dealers, number of transactions, and number of shareholders to be negatively related to the magnitude of the bid-ask spread. Measures of risk (either unsystematic or total risk) and measures of insider risk are positively associated with the magnitude of the spread.

In addition to the measurement problem that is often encountered in this type of study, there is another problem that is often ignored. That is, some empirical findings can be consistent with any of the three views of spread and can be interpreted differently. For instance, a strong negative relationship between the total spread and volume can be consistent with the economies of scale in a purely order cost environment. The same relationship can also be consistent with the information model that says greater market depth should induce lower spread.

Furthermore, these studies do not consider adjustments in the location of the quoted prices. This point is fairly important and will be made clearer when the relationship between transaction prices and the components of the spread is addressed in the following section.

II.C. Bid-Ask Spread and Transaction Prices

A number of studies have noted that the spread influences transaction prices.[2] However, it is not until Roll (1984) that the relationship between the spread and transaction prices has presented itself with a practical use. Recognizing that some transactions are within the

---

[2]See, for example, Neiderhoffer and Osborne (1966), Cohen, Maier, Schwartz and Whitcomb (1978, 1979), Goldman and Beja (1979).

spread and that the quoted spread is not an accurate measure of the actual spread, Roll proposes a measure of the effective spread.

If no information arrives at the market and if the magnitude of the spread remains constant, all transactions will be conducted on either the ask price or the bid price. There is no possibility of successive price changes occurring in the same direction. Therefore, the covariance of price changes will be generally negative. Moreover, the absolute value of the covariance reflects the magnitude of the effective spread. Roll shows that the effective spread equals $2\sqrt{-covP}$, where covP is the covariance of transaction price changes.

Roll's model corresponds to the purely ordering cost environment in that the dealer, when knowing there is neither possibility of informed trading nor need to maintain a particular inventory level, will not adjust prices to trades. Therefore, it is a rather simplified model. A complete model requires specifying the relationship between transaction prices and the inventory cost, and between transaction prices and the information cost.

In their seminal paper, Glosten and Milgrom (1985) show that transaction prices will form a martingale with informed trading but exhibit negative serial correlation when trading is liquidity-motivated. This result lays the foundation for the analysis of the components of the spread.

The impact of the order cost on transaction prices has been explained. The inventory cost also produces negative covariance of price changes. If the inventory level is above (below) the desired level, the quoted price is adjusted downward (upward) in order to induce offsetting trades. This means trade types (purchase or sale) as well as successive price changes will be generally negatively correlated. Although negative covariance of price changes is also expected under the order cost environment, the order cost differs from the inventory cost in that no adjustment in quoted price is predicted.

Under the information cost model, the quoted price is adjusted downward (upward) if an investor sells to (purchases from) the dealer. This is because an investor's sale (purchase) implies the true price of the security is lower (higher). Therefore, inventory and information models predict similar price adjustment by the dealer. The key difference between the inventory and the information cost models lies in the changes in the perceived true price. Under the inventory cost model, the true price remains to be the same, but under the information cost model, the true price is conveyed to the dealer through types of trades. More importantly, since the perceived true price is revised accordingly, an upward (downward) adjustment in the quoted price does not necessarily induce a sale (purchase). Rather, the movement of quoted prices reflects the arriving of private information and no dependence in successive price changes is predicted.

Note that in the inventory model the impact of trades on quotes is transitory since the true price remains to be the same. On the other hand, the information model predicts that the impact is permanent since the dealer permanently revises the true price according to types of trades. Glosten and Harris (1988) thus refer to the information component as the permanent component and the inventory/order component as the transitory component.

Glosten (1987) further elaborates the difference between the permanent component and the transitory component. He shows that only the transitory component will induce statistical biases in mean return, variance and serial covariance but the information component attenuates the biases. Additionally, Glosten shows that, if there is an additional information component, the serial covariance of transaction price changes is equal to $-0.25\alpha S^2$, where α is the proportion of the sum of the order cost and the information cost due to the order cost, and S is the total proportional spread. This implies Roll's measure of effective spread, $2\sqrt{-covP}$, is an upward biased estimator of the order cost.

There are two related studies that point to factors influencing quote adjustment. Schultz, Gustavson and Reily (1985) provide an experimental test (rather than a statistical test) of factors influencing the setting of opening prices. With the assistance of some specialists, these authors design an experiment so as to determine the relative importance of the following factors in the setting of opening prices : current demand/supply, the number of limit orders, inventory position, price trend, and volume trend. Six NYSE specialists participate in this experiment and the test results show that current demand and supply is the most important factor; price trend and volume trend do not appear to be very important; limit orders and inventory position are modestly important. The shortcomings of their study are that the experimental design is limited to the opening quotations and that the sample is fairly small. Another study by Marsh and Rock (1986) finds a statistically positive relationship between percentage changes in quoted prices and measures of order imbalances (net number of trades, net volume).

II.D. Estimating the Components of the Bid-Ask Spread[3]

Ho and Macris (1984) and Glosten and Harris (1988) discern the components of the spread from transaction price

---

[3] I benefit greatly from Jameson (1988) with respect to some of the points made in this section. Of course, any error is mine alone.

changes. The two models utilize similar concept and can be summarized as follows.

The transaction price equals to the true price plus the price effect of three components of spread.

$$P_t = M_t + CQ_t - H(I_t - I^*) + ZQ_t$$

where P_t = transaction price at time t

M_t = true price at time t

Q_t = an indicator which = 1 if an investor buys
=-1 if an investor sells

I_t = dealer's inventory level at time t

I^* = dealer's desired inventory level

C = ordering cost component

H = inventory cost component

Z = information cost component

The dealer adjusts the true price downward (upward) if an investor sells (buys). The true price revision is therefore :

$$M_t = M_{t-1} + ZQ_{t-1} + e_t$$

where e_t = a random term describing the arrival of public information.

If all transactions are of unit trade, then

$$I_t = I_{t-1} - Q_{t-1}$$

Taking the first difference of prices and substituting the above two equations, the following obtains.

$$D_t = P_t - P_{t-1} = (C + Z)(Q_t - Q_{t-1}) + (H + Z)Q_{t-1} + e_t \quad .. \quad (I)$$

Equivalently :

$$D_t = P_t - P_{t-1} = C(Q_t - Q_{t-1}) + HQ_{t-1} + ZQ_t + e_t \quad .... \quad (II)$$

Given D_t and Q_t, the simultaneous estimation of C, H, and Z is still not possible unless any of C, H or Z is known. This is because an estimation of equation (I) provides estimates for (C+Z) and (H+Z), but these two estimates are insufficient to decompose the spread into three components. Ho and Macris ignore Z while Glosten and Harris basically treat H and C as one component; and therefore they are able to solve the estimation problem. Since theory predicts these three components of the spread affect transaction prices differently, both Ho/Macris and Glosten/Harris are not entirely satisfactory.

Additionally, Ho and Macris have the knowledge of Q_t while Glosten and Harris do not. In the latter case, the authors assume some prior distribution for Q_t and then use maximum likelihood estimation to estimate Q_t, Z and C. The possibility of size-dependent components is also considered in their test. Specifically, Glosten and Harris assume both the ordering/inventory and information components are linear in transaction size. That is :

$$D_t = P_t - P_{t-1} = (C_0 + C_1V_t)(Q_t - Q_{t-1}) + (Z_0 + Z_1V_t)Q_t + e_t$$

Where C is now the sum of the inventory and the order components and V_t is the transaction size at time t. They find, through a specification search, that the most useful model is one that assuming $Z_0 = C_1 = 0$, constant e_t mean and variance without accounting for the discreteness in prices. That is, the information component is dependent on the transaction size but the inventory/ordering component is not. Nevertheless, it is suspected that the specification of trade size dependent components does not necessarily outperform one that trade size does not play a crucial role, since it is possible that linear trade size function may be misspecified.

Stoll (1989) presents another type of approach. His basic insight is that the three views of spread predict different values for two summary parameters for price changes; the probability of price reversal (hereinafter, π) and the magnitude of price reversal as a proportion of spread (hereinafter, $1-\delta$). Price reversal here refers to consecutive trades of different type (e.g., a purchase is followed by a sale).

The probability of a sell or buy is equally likely under ordering cost and information cost models. In contrast, under the inventory model, if a dealer is satisfied with the current inventory position, an investor's sale (purchase) will always induce the dealer to adjust the spread downward (upward) to induce an offsetting trade. Once the desired inventory position is regained, there is no further incentive for the dealer to adjust the spread. Therefore, under the inventory model, π is within the range of 0.5 and 1. If the inventory effect is strong, the value of π should be close to 1.0.

Assuming that inventory cost is linear in the size of the inventory and symmetric with respect to a purchase or a sale, the magnitude of price change will be one half of the spread. Assume under the information model dealer's response to buy or sell is symmetric,[4] δ should also equal to one half of the spread. The ordering cost component differs from the information and the inventory components in that there is no adjustment in dealer's quoted prices. The magnitude of price change will be equal to the size of the spread (i.e., $\delta = 0$) if a reversal occurs.

His basic point can then be summarized as below. Each component predicts different pair of π and δ, and a

[4]If the short-selling constraint is binding, the information conveyed by a purchase or a sale may not be symmetric.

clearer distinction of the components of the spread is made possible.

	π	δ
ordering cost	0.5	0
information cost	0.5	0.5
inventory cost	> 0.5	0.5

Extending Roll's analysis, Stoll shows that the covariance of transaction price changes (covP) and the covariance of quoted price changes (hereinafter, covQ) are functions of π, δ and S (the quoted dollar spread) as follows.

$$covP = S^2(\delta^2(1-2\pi)-\pi^2(1-2\delta)) \quad \text{(III)}$$

$$covQ = S^2(\delta^2(1-2\pi)) \quad \text{(IV)}$$

There are several implicit assumptions in solving Equations (III) and (IV) to obtain estimates of π and δ. First, the market is efficient with respect to public information. If the market is inefficient, the covariance of transaction price changes will reflect some lagged adjustment in resolving information differences among traders. Second, the spread is assumed to be a constant over the estimation period. It then does not matter whether an ask price, a bid price or the average of bid and ask prices is used. Third, all transactions occur on the inside quote (i.e.,the highest bid or the lowest ask available in the market).

The same set of assumptions are also implicitly made in Glosten and Harris with the exception that some type of lagged adjustment in public information is allowed (e_t is some function of elapsed time between successive transactions). In addition, components of spread are allowed to be dependent on trade size, provided that the relationship between components of spread and trade size is correctly specified. On the other hand, Stoll's approach uses the additional information of probability of price reversal. Distinction of various components of the spread is made more clearly. Each approach has its merits and problems.

Implicitly, the above approaches assume that the dealer adjusts the quoted price in the next trade. It is possible that empirically, there might be some lagged adjustments in dealer's quotes. If this is the case, the empirical results will contain additional noises. Hasbrouck (1988) presents a general approach in that no restriction is imposed on the number of trades for the inventory or the information effects to be complete. His basic intuition is that, if the inventory effect is stationary, trade and quote revisions will be a moving average process in which the coefficients sum to zero while the information effect predicts the revisions are independent. Hasbrouck's primary results say that the information effect is present, inventory effect is more apparent in low volume stocks, but not in high volume stocks. Nevertheless, his approach cannot be used to estimate the magnitude of components of the spread. Rather, his primary objective is to ascertain the existence of inventory and information effects.

II.E. The Components of Trading Volume

The information problem faced by a dealer is essentially one type of the adverse selection problem that is elaborated by Akerloff (1970). In response to the adverse selection problem, dealers compensate themselves for the losses to the informed by the gains from the uninformed[5] through adjustment in quoted prices. The inference of the degree of the informational asymmetry from trade revision is nevertheless difficult. Trade type is a valuable piece of information from our discussion in section C. Some other potentially useful pieces of information exist. One of them

---

[5]Grossman and Stiglitz (1980) show the impossibility of informational efficiency in a world where informed and uninformed traders coexist. Bagehot (1971) introduces the concept of informational asymmetry into the area of market microstructure.

is the components of volume (i. e., the composition of trading by the informed and the uninformed).

We broadly use the term " uninformed " because the cause(s) for such trading is (are) unclear. Glosten and Milgrom (1985) model the variation of time preference across market participants as the cause for such trading. Therefore, in their model, uninformed trading is basically liquidity-motivated trading. In Black (1986), " Market participants trade on noise as if it were information. " Noise, then, results from human error. In this sense, liquidity trading may be different from noise trading since liquidity trading is not motivated by information. It is suspected that the property of noise trading can be different from that of the liquidity trading. This paper will use these terms loosely, although a distinction between them might be necessary.

Easley and O'Hara (1987) distinguish two types of trading equilibrium. In a "separating equilibrium", informed traders trade only large quantities. As a result, the presence of informed trading may not always result in a nontrivial spread. This is because the dealer can easily distinguish informed trading from uninformed trading so that the term for small trades is unaffected. This is not true under a "pooling equilibrium" in which the informed traders trade both small and large quantities. However, their results indicate that large trades will still be made at less favorable price. Glosten and Milgrom (1985) do not explicitly model the role of volume and therefore the presence of informed trading will always induce a nontrivial spread even when the dealer is risk-neutral and makes no profit.

In Easley and O'Hara (1987,1985), the speed of adjustment of price to its full information value is not instantaneous because the presence of liquidity trading introduces noise in the trades. Rather, the speed of adjustment varies with various market characteristics including market depth, size, volume and return variance. This is because the dealer's precision of inferences about the presence of informed trading from observing trades depends upon the amount of noise in the market. A similar yet different point is made in Kyle (1985). In his model, an informed trader prefers to trade a small (large) quantity now and large (small) quantity later when the current amount of uninformed trading is small (large). His model differs from Easley and O'Hara in that the informed trader has monopoly power of the private information (as a result of insider information rather than the effort to collect and analyze information) and therefore enjoys an additional choice variable : the timing of trading. Therefore, the trade size becomes a noisy signal since the informed trader will arrange his/her transactions so that these transactions are indistinguishable from those of uninformed traders.

Admati and Pfleiderer (1988) further separate the uninformed traders into 'discretionary' and 'nondiscretionary'. The discretionary trader has a certain degree of control over the timing of trades. These authors show that the discretionary traders prefer to trade when there is a large number of informed traders to the extent that the degree of competition among informed traders is great. Informed traders will also prefer to trade when there is a large number of uninformed traders in the market so that the informed traders will not be exposed to the adverse pricing by the dealer. As a result, Admati and Pfleiderer give a potential explanation for the concentration of trading in the opening and closing periods as documented by Jain and Joh (1988). Foster and Viswanathan (1989) further incorporate the possibility of an imperfect public signal. That is, the private information is imperfectly revealed to the market through some public signal (e.g., corporate news release). Since Monday is likely to be when the public signal is weak, the informational asymmetry between the dealer and informed traders is greater. The results are that trading costs are the highest, the price variability and volume are the lowest on Monday.

In summary, the trading activity may be thought of as a result of the interaction among various trading parties. Each party pursues its optimal trading strategy and thus both transaction price and volume will contain noises in the process.

II. F. The Quoted Spread, the Cost of Transacting and the Effective Spread

A related issue is whether the quoted spread represents the cost of transacting. If the quoted spread does not represent the cost of transacting, absolute estimates of the components of the spread is difficult to obtain even if there is an exact approach to decompose the spread.

First of all, the quoted spread is hardly meaningful to ordinary investors since it represents the difference between the bid and ask prices at one point in time (Grossman and Miller (1988) and Stoll (1989)). A careless investor may find himself/herself paying a very high transacting cost if he/she buys and sells stocks at points in time when the market is especially illiquid.

Smidt (1988) is particularly critical of the usefulness of the quoted spread. For the quoted spread to be representative of the true transacting cost, three conditions have to be met. First, the quoted spread has to be a constant over time. Second, all transactions should take place either at the bid or the ask. That is, there is no possibility of within quote trades. Third, the probability of purchase or sale is equally likely at any point in time. The third point is related to the effect of inventory and the information components on transaction prices. That is, inventory adjustment induces trades that restore the dealer's desired inventory level (i.e., unequal probability of purchase and sale). The arrival of new information also leads to an unequal probability of purchase or sale.

Furthermore, to give a correct measure of the transacting cost, the suppliers and consumers of immediacy have to be identified. This is an impossible task, since suppliers of immediacy are not necessarily professional dealers. Traders submitting limit orders can perform the dealer function. Using information from the SEC and Options Clearing Corporation (to obtain proportion of volume due to market makers), Smidt (1988) estimates that the approximate transacting cost is only 4% of the minimum quoted spread (i.e., one-eighth of a dollar). Stoll (1985) conducts a similar study and concludes the spread realized by the dealer is approximately half of the quoted spread. Although these two studies give very different estimates of the transacting cost, it suffices to say that the quoted spread is generally not equivalent to the transacting cost.

Similar arguments are not difficult to find. Grossman and Miller (1988) argue the quoted spread is not an adequate measure of liquidity except when the dealer simultaneously executes crossing orders. Roll's (1984) measure of the effective spread[6] is among one of the attempt to measure the transacting cost, although he is motivated by a different reason (the possibility of within quote trades). That is, he considers only the second reason given by Smidt (1988). Glosten and Harris (1988) also give a measure of the effective spread. They define the effective spread as (2*C+Z), the quoted spread as 2*(C+Z), where C and Z are estimated from Equation (I). The intuition is that informed traders do not pay the amount of Z at the time of initial transaction, but do so at the time of the offsetting transaction. Similarly, Stoll (1989) gives the realized spread as the sum of inventory and ordering components. It is clear that both Glosten/Harris and Stoll consider only the third reason given by Smidt. To conclude, the measurement of the transacting cost is difficult to say the least. Advancement in this area will likely require more information.

II. G. Conclusions

In reviewing the literature, we also point out many problems in the area of the components of the spread. Some

[6]Choi, Salandro and Shastri (1988) and Bhattacharya (1985) extend Roll's analysis to include considerations about the possibility of serial correlation in transaction type and the possibility of within quote transactions.

are not given solution (e. g., monopoly component), some are dealt with to some extent in this dissertation.

The discussion of the literature of the determinants of the spread points to a deficit in this type of study. That is, implications of these empirical findings are vague. Our first essay employs a revised version of Stoll's approach to decompose the spread, and then presents separate analysis of components of the spread to obtain a complete picture of dealers' costs.

The quote adjustment can take place in the form of adjustment in the quoted prices and/or the adjustment in the magnitude of the spread. This point is partly considered in essay two. Further, essay two gives estimates of the components of the spread using two different approaches; the estimation should be more reliable if two approaches give similar conclusions. Finally, essay three presents a model that relates the concept of the components of the spread and the concept of the components of trading volume, providing a framework for utilizing information about order imbalance.

Chapter III

A Cross-Sectional Analysis of Costs Faced by NYSE Specialists

III.A. Introduction

The information about the relative importance of various components of the spread is potentially useful for correcting biases in transaction prices (as discussed in Chapter II). Other potential uses exist. One of them is to analyze the cross-sectional variations of the components of the spread with respect to factors that are important to each of these components. Such analysis is an extension of the studies of the determinants of the (total) spread such as Benston and Hagerman (1974).

Such an extension is important in understanding the total picture of dealers' costs. For instance, a negative relationship between the total spread and volume can be consistent with any of three components of the spread. If the specialist industry is characterized by economies of scale, higher volume will imply a lower order cost. Under a purely inventory cost environment, higher volume suggests the possibility of permanent inventory imbalance is less, and the concern for inventory level is likely less. On the other hand, higher volume will also imply a lower spread under the information model (Kyle (1985)). This is the primary intuition of this essay. Stoll's (1989) model is utilized to decompose the spread to investigate this issue.

Glosten and Harris's (1988) cross-sectional analysis is similar to ours in spirit. However, ours is more extensive than theirs. Ours includes several additional variables that are not considered in their test, allowing some hypotheses concerning the cross-sectional variation of these variables to be made, drawing from an extensive body of literature. Furthermore, ours is based on a more general approach (Stoll(1989)) through which three components can be simultaneously estimated. Glosten and Harris treat the ordering/inventory component as one component, and therefore inference of their result is limited.

However, because Stoll's empirical test is not easily adapted to cross-sectional analysis, we employ a slight variant of Stoll's empirical approach, which will be discussed in details in Section C.

Stoll uses daily data for his empirical test. Jameson (1988) suggests that, with daily data, the estimated inventory and information effects may be biased. This is because quote and trade adjustments should take place fairly quickly, and daily closing prices are not likely to reflect these adjustments. Glosten and Harris (1988) also point out that daily closing quotes may be subject to some degree of manipulation by the specialist knowing that the possibility of trades occurring at these quotes is very low. Harris (1989) finds that, in NYSE, a large price change is associated with the last transaction, and the large price change may be related to the tendency for the last transaction to occur at the ask price. Based on these grounds, we use transaction-by-transaction data that should be more appropriate for Stoll's model.

This chapter is organized as follows. Section B gives a description of the data. Section C presents the empirical methodology. The results of the relative importance of components of the spread are shown in Section D. Section E presents hypotheses regarding the cross-sectional variations of components of the spread, and Section F gives the cross-sectional empirical results. The conclusions are made in Section G.

III.B. Data and Sample Selection

The available data consists of two months' - September and October, 1987 - information about every transaction for common stocks listed on NYSE and AMEX.[7] Price, trade size, time of trade, and the exchange where the trade is originated are recorded for every trade. Similarly, for every quote, information about the bid/ask price, time of

[7]This data is provided by the Institute of the Study of Security Markets (ISSM), Pennsylvania State University.

quote, and the size for which the quote is good for (the depth of the quote) is given.

The sample period is arbitrarily chosen to be the five trading days Friday September 18 through Thursday September 24, 1987. The sample stocks contain the first 100 stocks in the data base using four screening criteria.[8] First, stocks which go ex-dividend over the sample period are eliminated since the effect of dividends on prices is still a much debated issue. Second, low-volume stocks are excluded due to potentially greater measurement error. (In particular, covariance measure may be more unreliable in a small sample.) For a particular stock to be included in the sample, the minimum trades and the minimum quotes per day are arbitrarily set to be 15. Third, the primary exchange is NYSE. Fourth, dollar stocks (minimum price increment is one-sixteenth) are excluded, since the effect of price discreteness for these stocks is likely to be different from that of other stocks (price increment is one-eighth).

Table 1 shows some descriptive statistics for the sample of stocks[9]. The total volume and number of

---

[8]These stocks are selected in alphabetical order. Sample size is limited to one hundred for the ease of collecting additional information not provided by the data base.

[9]For the NYSE as a whole over these five days, the mean and median (in parentheses) across stocks and days for NT, NQ, vol, price are, respectively, 75 (33), 67 (34), 138,000 (44,300) shares, $26 ($20). Therefore, our sample stocks tend to be higher-volume and higher-priced stocks.

transactions vary considerably across days and across stocks. The skewness statistics indicate that, especially for the dollar spread and price, (and volume on the day of September 24,) distributions of these variables are highly skewed. This observation however does not contradict previous evidence and is not critical to the following analysis. The sample seems fairly evenly distributed among various industries (see Appendix A for a list of company names), and the possibility of industry bias is small.

Some unusual transactions are excluded. They are broadly categorized as : (1) large trades,[10] (2) unusual delivery (usual delivery is five business days), (3) trades reported out of sequence, (4) quotes during trading halts, (5) pre-opening quotes, and (6) transactions identified as potential errors.[11] These excluded transactions account for less than 10% of the total original transactions.

III.C. Estimation of the Components of the Spread

---

[10]Such large trades (larger than 32,760 shares) represent less than 0.001% of the total trades. The classification rule (32,760 shares) appears to be arbitrarily set by ISSM. These large trades are excluded primarily for convenience.

[11]For our selected sample, we find one case of abnormal price changes for Aetna life and Casualty. In this instance, the price jumps up about one-third and then return to its previous level. Other than that, no abnormal price changes were found.

For the convenience of readers, Stoll's basic equations are restated below.

$$covP = S^2(\delta^2(1-2\pi)-\pi^2(1-2\delta)) \quad \text{(III)}$$

$$covQ = S^2(\delta^2(1-2\pi)) \quad \text{(IV)}$$

For every stock and day, covariance of transaction price changes (covP), covariance of quoted price (the midpoint of the bid and ask prices) changes (covQ), and the average dollar quoted spread (S) are estimated.[12] These estimates are then used to solve for π (probability of reversal) and $1-\delta$ (magnitude of price change, as the proportion of the spread), using numerical methods[13] for Equations (III) and (IV). This is different from Stoll's empirical approach. In his approach, covP and covQ are regressed, separately, on the square of the spread to obtain average values of π and δ across stocks. In so doing, average values of the components of the spread can be

---

[12]Fuller (1976, p. 242) suggests an adjusted covariance for small sample. In a small sample, ordinary covariance estimates are biased downward. Intuitively, small sample may appear to have significant negative autocorrelation while in fact there is none (Hasbrouck(1988)). Our average number of transactions used to calculate the covariance of price changes is above 50 in most cases. The ordinary covariance estimator is used in this essay.

[13]IMSL subroutine NEQNF is employed to solve the system of equations. This routine uses a variation of Newton's method. Starting values for π and δ are 0.5 and 0.5. A solution is accepted if the difference between two successive approximations is less than 0.002.

obtained but cannot be readily applied to cross-sectional analysis. Therefore, this modification is important for our purpose. Moreover, it is unclear what the error term represents under Stoll's regression approach[14].

One of the implicit assumptions for Stoll's model is that the probability distribution of price changes is stationary over the sample period. Roll (1984) finds results of covP using daily data and five-day data are quite different and comments : "This implies informational inefficiency (although not necessarily profit opportunities) or else very short-term nonstationarity in expected return." If it can be argued that the return distribution is relatively more stationary over a short interval of time, it is more appropriate to perform the analysis using data within a shorter time span. French (1980) finds mean return is different across the day of the week, which suggests nonstationarity in daily returns. Also, because the opening trading procedure is different from the rest of the day, it seems undesirable to treat opening transactions as continuation of the last closing transaction. We therefore chose to apply the methodology to observations within a day. This means there are five pairs of (π, δ) for each stock in the sample, one pair each day.

---

[14]Roll (1984) infers the effective spread to be $2\sqrt{-covP}$. He too does not utilize a regression approach.

In Stoll's terminology, the realized spread, what the dealer earns on net, is equal to the sum of the inventory and ordering components. The information model says that the dealer loses to informed traders and gains from uninformed traders; the information component is therefore not a part of the realized spread and is the difference between the quoted spread and the realized spread. Stoll shows that the realized spread equals $2*(\pi-\delta)S$[15]. Realized spread is further decomposed into inventory and ordering components. The inventory component is shown to be $2*(\pi-0.5)$ and the ordering component is $2*(0.5-\delta)$.

Our interest here is to infer the expected value of π and δ over the estimation period. Since Equations (III) and (IV) are nonlinear in π and δ and since our covariances of price changes are also estimates, the numerical solutions would only give approximate expected values of π and δ, according to Jenson's inequality. Also, in Equations (III) and (IV), a constant spread is assumed. The Appendix B gives a revised version of Stoll's derivation of the covariance of transaction price changes (i.e., Equation (III)) allowing changes in the magnitude of the spread. If changes in the magnitude of the spread are random and small,

---

[15]Jameson (1988) is particularly critical of Stoll's decomposition. For a purely inventory model, the value of π has to be one to let realized spread equal S. Nevertheless, the general intuition of Stoll remains true. That is, a larger value of π indicates a stronger inventory effect.

the covariance of transaction price changes reduces exactly to Equation (III).

III.D. Relative Importance of the Components of the Spread

Table 2 contains the means and standard deviations for π, δ, covP, covQ, and estimated components of the spread. The average value of δ is greater than 0.4 while π is close to 0.5. Such estimates are most consistent with the information model. Accordingly, the estimated information component is extremely large. Over 80% of the spread is due to the information cost faced by the dealer. The inventory cost is small (about 2% of the spread) and insignificant. Both the t test and the sign rank test fail to reject the null hypothesis that the inventory cost is equal to zero in two out of four days, and the t test (but not the sign rank test) fails to reject the null hypothesis in two other days. The ordering cost accounts for about 15% of the spread. This pattern is fairly similar across days[16]. The covariance of quoted price changes is insignificantly different from zero in four out of five days (more consistent with the information cost model), and is generally less negative than

[16]Because the trading procedure for the opening market is quite different from that of the rest of the day, we also did a similar investigation excluding the first hour of transactions. The results are not much different. This issue will be looked at more carefully in Chapter Four.

the covariance of transaction price changes, which is what theory suggests[17].

Hasbrouck (1988) finds strong evidence for the information effect but mixed evidence for the inventory effect, although his test cannot indicate which component is the most important. Glosten and Harris (1988) find the estimated information component is not large using NYSE transaction data. However, since their estimation specifies a linear relationship between the information component and the transaction size, the information component can be substantial for a large trade. According to their specification, the quoted spread is $2*(C_0+Z_1V_t)$, where V_t is stated in thousands of shares. The quoted spread is estimated to be $2*(0.0242+0.0133*V_t)$. This means for a 10,000 share transaction, the estimated quoted spread is $0.31, which is close to the actual quoted spread. This suggests the information component is (2*0.0133*10) $0.266 that is approximately 86% of the estimated quoted spread. Smidt (1988) estimates that the return realized by specialists accounts for approximately 4% of the quoted spread. Therefore, our empirical evidence is partially consistent with some previous studies. Using daily NASDAQ data, Stoll (1989) estimates (through regression of covP, covQ on S^2) that 43%, 9%, 48% of the spread is due to the

[17]That is, covQ does not contain an order cost component. See also Equations (III) and (IV).

information cost, the inventory cost, the ordering cost, respectively (δ=0.265, π=0.550). His and our average estimates are quite different. We suspect the difference may be due to the difference in data. As pointed out in section A, the inventory and information components may not be captured using daily data[18]. According to Stoll's model, the specialist's price adjustment (the information and the inventory components) should immediately follow the previous trade. Daily data may thus reflect more of the order component. Moreover, NASDAQ market makers are usually less capitalized than NYSE specialists, which suggests higher order costs for NASDAQ stocks. Additionally, his sample consists of lower-priced, higher-spread stocks (on average, price = $14, percentage spread = 2.9%) than ours, which suggests discreteness problem may be more severe for his sample. A greater degree of discreteness problem is consistent with higher inventory and ordering costs.

Stoll and our results are consistent with respect to a small inventory component. Stoll (1985) argues that the inventory cost is unimportant because (1) specialists'

---

[18]We compare covP and covQ using every transaction prices and using only daily ending prices for the same set of stocks for the period of August 31 through October 15 (33 days). The mean (standard deviation) δ and π are 0.445 (0.133) and 0.526 (0.110) for transaction data, 0.544 (1.137) and 0.186 (1.769) for daily data. The estimated π and δ using daily data give negative order and inventory costs, which is very difficult to interpret. It seems covariance of price changes is sensitive to lags of prices.

portfolio of stocks may be only a fraction of the specialist's total wealth, and (2) unlike ordinary investors, the specialist can go long and short with little constraint. These may partially explain the relatively small magnitude of the inventory cost. The remaining part of this section is primarily devoted to examining whether or not the estimation is robust with respect to subsets of transactions or stocks.

Most stocks listed on NYSE are allowed to be traded on regional exchanges and/or the-over-counter market. It seems more reasonable to employ NYSE transactions since specialists do not participate in off-NYSE transactions. However, there might be a potential cross-market information effect since the specialist may infer information from transactions occurring on other markets. If cross-market effect exists, we are utilizing only partial information with NYSE transactions. To see whether off-NYSE transactions may seriously affect the results, we do a simple comparison of transactions including and excluding off-NYSE transactions.

Table 3 contains all transactions while Table 2 contains only NYSE transactions. Comparing these two tables, we find on average the estimated π (over 0.6) and δ (over 0.5) are all higher in Table 3. The information component remains about 80% of the spread but the inventory component now accounts for about 20% of the spread. The weaker inventory effect in Table 2 seems counter-intuitive, since excluding trades that occur outside NYSE should make the inventory effect more apparent while the result suggests otherwise. However, the spread on regional markets or OTC is often higher than that of NYSE for comparable stocks, which is also confirmed by comparing the proportional quoted spread in the two tables. Recall that the components of the spread are directly inferred from covP and covQ; it is therefore very likely that the result is an artifact of the higher absolute values of covP and covQ in Table 3, induced by larger spreads. In addition, mean order costs are negative for all days if all transactions are included. Given the above observations, we conclude that off-NYSE transactions are not directly comparable to NYSE transactions. The following analysis will use exclusively NYSE transactions, which appear to be a more reasonable set of data for the purpose of this chapter.

The options market may interact with the stock market. In particular, the specialist may sometimes trade in the options market to hedge his/her position in stocks. However, the specialist can buy and sell stocks in both long and short position while he/she has to pay additional transaction cost for trading in the option market. By the standard option valuation model (Black and Scholes (1973)), options are purely redundant securities since they can be replicated by some combination of the risk-free asset and the underlying stock. The mere existence of the options market then suggests there are transaction cost difference between the options market and the stock market. Also, some information effect may exist. In light of the greater degree of leverage that a trader enjoys in options market, we may suspect that informed traders prefer to trade in the options market. This argument however also implies segmentation of traders for the stock market and the options market. Therefore, the true effect is unclear.

The sample stocks are divided into two subsamples according to the presence or absence of options. A standard t test[19] (assuming unequal variances) is used to test the difference in means of the components of the spread between subsamples for each day. Only on September 23 does the t test indicate the inventory component for non-option stocks is significantly (t=2.38) different from and is higher than that of the option-stocks. No significant difference is observed in other components of the spread. The results for September 23 are shown in Table 4 Panel A. The average price, average transaction size and volume are higher for option-stocks than for non-option stocks. The π and the absolute value of covQ are higher for non-option stocks,

---

[19]$t = (x_1-x_2)/((var(x_1)/n_1)+(var(x_2)/n_2))^{1/2}$, where x_1, $var(x_1)$, and n_1 are the sample mean, variance, and sample size for group 1, similar notations apply to group 2. If variances are assumed to be unequal, the degree of freedom can be substantially less than the addition of n_1 and n_2 if sample variances turn out to be quite different.

which indicates a higher inventory cost. Panel B shows the results for September 24, the lightest trading day of the five days. For option stocks, mean covQ is large in absolute term relative to September 23 and other days, which suggests quote adjustments to be larger in light trading days. Another noteworthy observation is that mean inventory cost is especially high (0.04) for option stocks relative to those of other days (0.02 on average).[20]

Since the number of transactions used to estimate covP and covQ differs for each stock, mean covP and covQ among volume groups may differ. Furthermore, volume of trading should be a critical factor in the business of market making. Nevertheless, volume is a noisy indicator of the risk of dealer's business, based on the following observations that offer different interpretations of the role of volume. First, volume may be considered as a measure of competition if higher volume attracts more large block brokers. Second, volume can be considered as a measure of liquidity since higher volume usually suggests a more liquid market. Furthermore, since there are informed as well as uninformed trading, total volume alone does not mean much to the specialist.

[20]Results for other days are similar to those of September 23 and especially, September 24, in term of differences in components of spread.

To test the influence of volume, we arbitrarily separate the sample into three subsamples, based on cutoff points of 80,000 shares and 300,000 shares per day. Paired t tests indicate only on the day of September 18 is the difference in order costs between high and low volume stocks significant at the 10% level (t=2.10). Table 5 Panel A presents comparisons between volume subgroups for the day of September 23. The low-volume group appears to be less homogeneous than the other two groups, but there are no significant differences across volume subgroups with respect to means of δ, π, and the components of the spread. Panel B shows the results for September 24. Although the volume is the lowest on this day, mean components of the spread demonstrate no significant difference among volume subgroups.

There are some institutional factors that complicate the estimation of the components of the spread. The most important one perhaps is the presence of limit orders. Limit orders can be thought of as the competition to the dealer from traders who commit to trade at a particular price.[21] Limit orders may bias the estimation because some of the trades have no consequence to the dealer's inventory.

---

[21]An implication for the limit orders is that there is a need to distinguish dealer spread from market spread, where market spread is the better of dealer quote and quotes from traders submitting limit orders (Cohen, et al (1978,1981), Hamilton (1987)).

On the other hand, Conroy and Winkler (1981) argue that limit orders have some information effect to the dealer, since the information in the limit book is available to the dealer only. Thus, the effect of limit orders on our estimation is not an easily resolved issue. On a purely technical basis, if limit orders cluster around a certain price, an adjustment of quotes to that price will trigger a series of transactions occurring in the same direction (Choi, Salandro and Shastri (1988), Hasbrouck and Ho (1987)). Hasbrouck (1988) assumes limit orders arise purely from uninformed traders and are uncorrelated with the stochastic arrivals of market orders. In this case, the presence of limit orders simplifies into a simple measurement error. There are, however, reasons to suggest otherwise. In Kyle (1985), a monopolistic informed trader can time his/her trades and perhaps types of orders in order to avoid the adverse pricing by the dealer.

Other factors include broken orders and the discreteness in prices. Broken orders tend to introduce positive correlation in transaction types (Garbade and Lieber (1977)) while discreteness in prices tends to produce negative values of the covariance of price changes (Gottlieb and Kalay (1985), Cho and Frees (1988) and Ball (1988)). In Section F, our cross-sectional analysis uses prices to proxy and control the discreteness problem, since the discreteness problem is likely to be negatively related to the price of stocks.

Although the average values of π and δ are fairly stable across days, there are considerable variations in π and δ for individual stocks. This is not inconsistent with the theory but we need to point out that the estimated components of the spread are sensitive to the values of covP and covQ. This and factors discussed above (e. g., limit orders) indicate that we should be especially careful with measurement problems. Nevertheless, since our focus is on the cross-sectional variation of various components of the spread, measurement errors should not greatly affect the cross-sectional analysis if the errors are uniformly distributed across stocks. Additionally, there may be some reporting errors. In Section F, estimated components for individual stocks are averaged across days to reduce the effect of reporting errors and the effect of including an abnormal trading day.

In conclusion to this point, our empirical evidence indicates that dealers face a large information problem. The result is fairly robust. The aggregate estimates do not change substantially across days. The result is not likely to be due to sampling bias, since aggregate estimates are about the same whether the sample is separated into option groups or volume groups. We now turn to our focus of this chapter - the cross-sectional variation of the components of the spread.

III.E. Cross-Sectional Hypotheses

This section presents cross-sectional hypotheses of components with respect to important factors that influence the magnitude of these components. Empirical test of these hypotheses will be shown in Section F.

Number of transactions, measures of risk, number of shareholders, measure of insider risk, institutional holding, price, and volume have been found by previous empirical studies (see Chapter II) to be important determinant of the spread. Number of shareholders is generally used as a proxy for the degree of informational asymmetry. However, since not all shareholders will be active in trading, the number of shareholders does not seem to be an appropriate proxy for the degree of informational asymmetry.

The first important question concerning the order cost is whether the specialist industry is characterized by economies of scale. Some components of the order cost may be considered relatively fixed (e.g., clerical and administrative costs) while others may be considered as variable (e.g., interest expenses and communication cost). If a larger portion of the order cost is fixed (variable), then the relative importance of the order cost should decrease (increase) with the number of transactions, and the industry will be generally characterized by economies (diseconomies) of scale.

The average transaction size will be negatively correlated with the order cost if the cost per transaction is relatively fixed, will not be related to the order cost if cost per transaction varies. Easley and O'Hara (1987) hypothesize informed traders prefer to trade larger quantities. This predicts a positive relationship between the information cost and the transaction size. Kyle's (1985) model, however, indicates that transaction size is a noisy signal. The intuition is that informed traders will not necessarily always prefer to trade large quantities, since if they do, their private information will be revealed unambiguously to the market. In the empirical test, since a certain transaction size may be considered as large for a low volume stock while small for a high volume stock, rather than absolute transaction size, relative size is used instead (i. e., size/volume).

Benston and Hagerman (1974) argue that since the expected return of a security will compensate the holder for the systematic risk associated with it, only the unsystematic risk is relevant to the dealer.[22] On the other hand, Stoll (1978a) argues that the total risk is more relevant and states that "...market making causes the dealer to move away from his desired position on the efficient risk-return frontier to a point yielding a lower level of utility. The expected return of a stock is not sufficient compensation for the loss of utility." Either theory predicts a positive relationship between some measure of risk and the inventory cost.

Since the dealer may sometimes trade in the options market to hedge his/her stock position, the presence of the options for the stock the dealer deals in suggests the dealer will care less about the inventory cost. This predicts a negative relationship between the presence of options and the inventory cost.

Insider trading as well as unsystematic risk (Benston and Hagerman (1974)) are measures of information risk to the dealer, thus the information cost should be positively related to insider trading and/or unsystematic risk.

Smidt (1971) observes that the specialist seems to be more willing to absorb inventory imbalances if institutions account for a larger proportion of trades. This may be

[22]In equilibrium, investors require an expected return to compensate them for bearing the systematic risk associated with an asset. This is the Capital Asset Pricing Model (CAPM) developed by Sharpe (1964), Litner (1965), and Mossin (1966).

consistent with either or both of the following explanations : (1) a high institutional trading represents a more liquid market as competition for handling institutional transactions is greater, and/or (2) institutional traders are not informed traders. Explanation (1) implies institutional trading is negatively related to the ordering /inventory cost, and (2) suggests a negative relationship between the information cost and the amount of institutional trading.

Recent evidence shows that smaller firms' stocks have higher returns than those of larger firms' stocks, and the return differentials cannot be accounted for by risk differentials (Banz (1981) and Reinganum (1981)). Barry and Brown (1984) consider that the scarcity of information available for smaller firms may be a potential explanation. Glosten and Milgrom (1985) suggest that since smaller firms may have a greater degree of information asymmetry between shareholders, a larger information cost would be associated with smaller firms' stocks. However, Keim (1983) and Lamoureux and Sanger (1989) present evidence that most of the abnormal returns for smaller firms' stocks accrue in January. Our inference will be fairly limited because of the unavailability of January data.

III.F. Cross-Sectional Tests

The following notations will be used throughout this chapter.

C = order cost

H = inventory cost

Z = information cost

NT = number of transactions

AZ = average transaction size

RZ = relative transaction size

TR = total risk

UR = unsystematic risk

OP = option trading

IT = insider trading

IN = institutional trading

FZ = firm size

P = stock price

VOL = volume

The ordinary least squares (OLS) regression approach is employed to examine the cross-sectional variations of components of the spread with respect to the variables described in section E. The cross-sectional hypotheses are outlined below. The predicted signs of each of the coefficients are summarized in parentheses.

$$C = \alpha_1 + \underset{(\pm)}{\alpha_2 \, NT} + \underset{(\pm)}{\alpha_3 \, AZ} + \underset{(-)}{\alpha_4 \, IN} + \alpha_5 \, P + \alpha_6 \, VOL \quad \text{..........} \quad (A1)$$

$$H = \beta_1 + \underset{(+)}{\beta_2 \, UR(\text{or } TR)} + \underset{(-)}{\beta_3 \, OP} + \underset{(-)}{\beta_4 \, IN} + \beta_5 \, P + \beta_6 \, VOL \quad \text{...} \quad (A2)$$

$$Z = \theta_1 + \underset{(\pm)}{\theta_2 \, RZ} + \underset{(+)}{\theta_3 \, UR} + \underset{(-)}{\theta_4 \, IN} + \underset{(-)}{\theta_5 \, FZ} + \underset{(+)}{\theta_6 \, IT} + \theta_7 \, P + \theta_8 \, VOL \quad \text{.............} \quad (A3)$$

Price is included in the regression as a proxy of the degree of the discreteness problem and for other unknown risk (see Stoll (1978a) for a further discussion of the price factor). Discreteness may artificially produce larger inventory and ordering components (through more negative estimates of covariance of price changes) and is likely to be a more important problem for low-priced stocks. Kross (1985) relates prices to estimation risk : "Since more consistent parameter estimates can be generated for high priced stocks, individuals will invest relatively more in the high priced set of stocks than in the low priced set. Limited diversification results unless investors are paid, through higher returns, to include low priced securities in their portfolios." Both observations indicate that the level of price is negatively related to estimation risk. On the other hand, since the capital required for higher-priced stocks is higher given a fixed number of shares, it can be argued that price is positively associated with (specialist') business risk.

The role of volume is also rather vague. Volume may reflect the degree of competition. As such, higher volume represents a disadvantage for the dealers. On the other hand, high volume may be a blessing to the dealers since higher volume also suggests less risk. Further, volume is used extensively by previous studies and should be investigated. Since the role of neither price nor volume is clear, we present regression results for each of Equation (A1), (A2) and (A3) in three versions : (1) price and volume excluded, (2) price included, volume excluded, (3) price and volume included. Accordingly, regression (A1.3) refers to regression (A1) when price and volume are both included; regression (A2.1) refers to regression (A2) when price and volume are both excluded; etc.

IN, IT and number of shares are collected from Standard and Poor's Corporation Stock Reports (November, 1987). OP is an indicator variable and is equal to one if there are options traded for the stock and zero otherwise. The CRSP[23] daily file is used for the purpose of estimating TR (measured by return variance) and UR (measured by residual variance from the market model) for the period of 120 through 20 days prior to September, 1987 (i.e., approximately March through July). We use lagged values for TR and UR because dealers are uninformed (or else there will be no information cost) and presumably use lagged TR or UR as the estimated risk.[24] The NT, AZ, RZ, OP, FZ, VOL are contemporaneous estimates to reflect current market activity. Averages of five days of estimates of C, H, Z, NT, AZ, RZ and VOL for individual stocks are used to reduce

---

[23]The Center for Research in Securities Prices, University of Chicago.

[24]Benston and Hagerman (1974), Stoll (1978a), Glosten and Harris (1988) also use lagged estimates for risk.

the effect of including an abnormal day and the effect of potential reporting errors.

The final sample contains 91 stocks; 9 stocks are excluded because either data on previous prices is unavailable (five stocks are newly listed) or Stock Reports do not provide information about some of the variables investigated (two are ADRs, and two are limited partnerships).

Since some of the independent variables are highly correlated (see Table 6), OLS regression coefficients estimates are still unbiased but the t values may not be totally reliable. Therefore, for each regression, the value of the maximum eigenvalue divided by the minimum eigenvalue for the X'X matrix is given (denoted as cn). If cn is larger than 900, multicollinearity presents a serious problem to OLS estimates and ordinary t tests are unreliable (see Judge, Hill, Griffiths, Lütkepohl and Lee (1982)).

Regression results are given in Table 7 with Panels A, B, and C presenting the results for regression (A1), (A2), and (A3), respectively.[25]

---

[25]Prices may be correlated with some of the variables investigated. Instead of using contemporaneous prices, lagged prices are also used for the same set of regressions. The results are similar in terms of the signs and the p values of the coefficients. However, regressions using lagged prices have much higher R squares and a greater problem of multicollinearity than those using current prices in general. Results shown in this essay use current prices.

In both regressions (A1.1) and (A1.2), NT is significantly (at 10%) positively related to C. Such relationship is not supportive of economies of scale for the industry as a whole. AZ is insignificantly related to C, which indicates that cost per transaction is not fixed. IN is negatively related to C (although significant only at the 15% level in (A1.2-3)). It may be that a greater institutional holding represents a more competitive market. Alternatively, we may argue that institutional traders prefer to hold stocks handled by more efficient dealers. The reverse is unlikely to be true as previous evidence (Barnea(1976)) indicates more efficient specialists are not necessarily rewarded by a preferred allocation of stocks[26]). The negative sign for P is supportive of the conjecture that P is a proxy for discreteness problem and estimation risk. Regression (A1.3) is quite troublesome for cn is quite high (=344.6) relative to other regressions[27]. Furthermore, NT becomes insignificantly related to C. We may argue that NT is really a consequence of VOL, since NT is significant only when VOL is not included (correlation coefficient of NT and VOL is 0.854, which is significant at the 1% level). Since multicollinearity may be a problem in (A1.3), inference from

---

[26]Also, Stoll (1985) finds specialists' portfolios are not much different from a random portfolio.

[27]Further analysis indicates that the smallest eigenvalue affects the variation of the intercept term and IN, which says that the variation in IN is small.

(A1.3) is imprecise. Nevertheless, to the extent that NT is a consequence of VOL, the positive relationship between VOL and C is a piece of evidence against industry economies of scale.

In Panel B, all regressions use TR as the proxy for risk. Results using UR are not reported here because results using TR and UR are remarkably similar (correlation coefficient of TR and UR is 0.997 and is highly significant). Regressions (A2.1-3) suggest TR is not a factor in specialist's inventory decision, which contradicts previous evidence. Perhaps because the dealer cannot estimate precisely the degree of risk, he/she relies on more intuitive measures of risk (such as IN[28]) or perhaps our estimation of TR and UR is unreliable. Also, OP appears to be unrelated to H. It is consistent with the notion that options trading, although providing hedging opportunity, may involve additional costs that outweigh the benefit. IN is significantly (at 5%) negatively related to H in all regressions (A2.1-3), and this reinforces our conclusions regarding IN from regression (A1). VOL is unrelated to H. A possible explanation is that specialist' inventory position can accumulate fairly quickly if market volume is high but transactions predominantly occur at one side of the

[28]TR (UR) is significantly correlated with OP, IN, IT, FZ and P (see Table 6).

market. Thus we conclude a negative relationship between volume and spread is inconsistent with the inventory model.

The information cost regression (panel C) is generally subject to a greater multicollinearity problem as more variables are included. However, multicollinearity may not be very serious here as all three versions of regression (A3) generally show consistent signs and p values for estimated coefficients. Z is positively related to RZ, supporting Easley and O'Hara's conjecture that informed traders prefer to trade large quantities. According to their model, even in a pooling equilibrium, larger trades will still be made at less favorable term although trade size is a noisy signal of informed trading in this context. UR again is unrelated to Z, contradicting previous evidence. The result that θ_4 is significantly positive (at 1% level) suggests that the institutional traders are not totally uninformed. Although some theories (e.g., Admati and Pfleiderer (1988)) assume the institutional traders are uninformed, they are not totally convincing, as institutional traders presumably spend a lot of resources in collecting and analyzing information. However, since there may be more financial analysts working on stocks with a large amount of institutional holdings, this result may be an artifact of a positive relationship between information cost and the number of financial analysts. FZ is insignificantly associated with Z, although such result is not strong evidence against the hypothesis that smaller firms have a larger degree of informational asymmetry. It may be that informational asymmetry is more severe in January when annual results for most companies are revealed. Our data period does not include January, thus the result is fairly weak. IT, the measure of insider risk, is positively related to Z. It may be that the information asymmetry problem is more related to the insider information than to the superior analytic ability of investors.

This positive sign of P in regression (A3) does not contradict the notion that price is a proxy for the degree of the discreteness problem as higher (lower) price suggests less (more) degree of discreteness problem, which in turn suggests more positive (negative) covP and a higher (lower) Z. It is contradicting the conjecture that price is an appropriate proxy for risk of specialists' business.

In all regressions, the intercept terms are significantly positive (at 1%), which suggests a constant portion of C, H, and Z. Certainly, theory lacks a satisfactory explanation of what the intercept term really represents (except that a constant proportion of C may represent a fixed cost). The R-squares are generally not impressive, especially for regressions (A1) and (A2)[29].

---

[29]Studies on the determinants of the spread usually find a R-square value above 0.3. To see whether the error term is related to any of the variables in the regression, a test equivalent to Breusch and Pagan (1979) is employed to investigate whether the square of error terms divided by

III. G. Conclusions

For the sample of 100 stocks and for the period of September 18-24, 1987, our empirical results support the following : 1) a large portion of the cost faced by a NYSE specialist is attributed to the adverse selection problem; 2) the NYSE specialist industry is not characterized by economies of scale; 3) higher institutional holdings are associated with a more liquid market; 4) the information cost is positively related to the transaction size; 5) the information cost is positively associated with insider holdings. Overall, except for the result related to the measure of total risk, results show consistency with an extensive body of literature. This essay may then be considered as a first step in understanding a complete story of NYSE specialists' costs, since some measurement problems in the estimation of the components of the spread remain unresolved. These measurement problems include those arising from institutional factors such as limit orders and those arising from theoretical considerations such as the assumption of a constant spread. Little is known as to whether these measurement problems may vary across stocks.

---

mean of the square of the error terms are linear function of the independent variables in each of the regression. The result does not reject the null hypothesis that the error term is unrelated to these variables.

The level of price is used as a proxy to control the discreteness problem, since it is reasonable to expect discreteness problem may affect stocks differently. Finally, our results may be considered as weakly supportive of Stoll's model if the hypothesized cross-sectional variations detailed in section E are taken as given.

Our specification implicitly assumes the independence of specialists' trading behavior. In NYSE, a specialist may specialize in more than one stock. Often specialists form into a specialist unit (usually a firm).[30] These facts suggest some possible diversification effect across stocks. Moreover, many NYSE members participate in both specialist and brokerage businesses, although Smidt (1988) notes that "It is unlikely that specialists today can be thought of as dealing in stocks in order to attract brokerage business, although in the past, when competing specialists were common, dealing may have been a subsidiary by-product service offered to attract brokerage business." Another type of possible interaction among specialists may occur when the business of market making is especially good or especially bad. For example, specialists may maintain a moderate spread under extremely adverse market conditions in

---

[30]Using information from NYSE Unit of Trade Directory, we find a total of 52 specialist units responsible for the NYSE common stocks. Our 100 sample stocks are handled by 35 units, and the number of stocks each unit handles ranges from one to seven.

the hope that they will recover any loss incurred in an adverse market when the market becomes favorable again.

Chapter IV

The Intraday Variations in Trading Activity, Price Variability and The Transaction Cost : An Exploratory Investigation

IV. A. Introduction

Recent evidence (Wood, McInish, and Ord (1985), Jain and Joh (1988), and Mulherin and Gerety (1989)) document that volume and price changes are higher in the opening and the closing periods (i.e., exhibit a U-shaped pattern). Although this phenomenon is partly consistent with the finding of a positive relationship between volume and price variability (as summarized by Karpoff (1987)), the concentration of volume and price changes in the opening and the closing lacks a satisfactory explanation.

Admati and Pfleiderer (1988) present a model that is consistent with the U-shaped pattern in volume and price variability. Their primary innovation is to point out the critical role of the discretionary uninformed traders (i.e., uninformed traders that have the capacity to time their trades). The concentration of volume and price variability results from the interaction among and between discretionary uninformed traders and informed traders. Essentially, Admati and Pfleiderer argue the concentration of volume and price variability results from variations in transaction costs. This essay provides a test of their model, as well as tests for other explanations that are consistent with the concentration in volume and price variability.

Foster and Viswanathan (1988) conduct a similar study in that they also investigate intraday variations in the components of the spread. There are, however, a number of differences between their study and ours. In particular, competing hypotheses are presented and tests are more extensive in our study. Also, the measurement problem is addressed here. A brief description of Foster and Viswanathan's (1988) test design and empirical results will be given in Section B.

The plan of this chapter is as follows. Section B gives a brief review of Admati and Pfleiderer (1988) and Foster and Viswanathan (1988). The competing hypotheses are presented in Section C. The description of methodology and samples can be found in Section D. Section E gives the empirical results. Section F then summarizes and concludes this chapter.

IV. B. The Admati and Pfleiderer (1988) Model

The interaction between informed traders and uninformed traders has profound implications for security returns and volume. Kyle (1985) gives the insight that informed traders prefer to trade when the market is deeper (a greater amount of uninformed trading), so that they will not be subject to the adverse pricing by the dealer.

Admati and Pfleiderer (1988) (AP) extend Kyle's analysis by including discretionary uninformed traders. (Institutional traders, for example, may belong to this class of traders.) The intuition behind AP is as follows. Discretionary uninformed traders will prefer to trade when the transaction cost is lower, which usually occurs when the market is deeper. The concentration of volume can therefore occur because discretionary traders prefer to trade together. However, informed traders will also prefer to trade in a deeper market. To the extent that the degree of competition among informed traders is great, more informed traders will increase the welfare of uninformed traders. Consequently, this induces more discretionary uninformed traders, and the concentration of volume occurs. Prices will be more informative and price variability will be greater in periods of higher volume.

AP further argue that the opening and the closing periods are when trading is most likely to be concentrated. This is because overnight liquidity demands of traders will have be satisfied in these hours. This explanation, however, does not seem convincing for the higher volume that is often observed in the closing.

Under the AP model, the dealer responds to public information and private information conveyed through order flows by setting the price according to the following schedule :

$$P_t = F_t + \Sigma^t_{\tau=1} e_\tau + \lambda_t Q_t \quad \text{.........................} \quad (B1)$$

where P_t = transaction price at time t

F_t = expected value of the security at time t

e_τ = public information arriving at time τ

Q_t = order flow at time t

The key parameter is λ_t, the magnitude of price adjustment set by the dealer in response to order flows. The reciprocal of λ is a measure of the depth of the market. Their main result is that the dealer responses to a deeper (narrower) market by setting a lower (higher) spread (or magnitude of price adjustment).

According to Glosten and Harris's (1988) model, the revision in the expected value of the security is :

$$F_t = F_{t-1} + \lambda_t Q_{t-1}$$

Transform Equation (B1) to its price changes form, the resulting equation is :

$$P_t - P_{t-1} = \lambda_t Q_t + e_t$$

Note that this is similar to Glosten and Harris's basic equation (Equation (I)) :

$$P_t - P_{t-1} = (C + Z)(Q_t - Q_{t-1}) + (H + Z)Q_{t-1} + e_t$$

Where C, H, and Z represent, respectively, the order cost, the inventory cost and the information cost. If C = H = 0, two equations are exactly the same and $\lambda = Z$.

It is apparent that the AP model does not consider the inventory and the order components of the spread, which are found to be important components of the spread by several studies (Glosten and Harris (1988), Hasbrouck (1988), Stoll (1989), and the evidence provided in Chapter III).

Foster and Viswanathan (1988) (FV) investigate the intraday and interday variations in the components of the spread, volume and return variance for the years 1985-86. They use a revised version of Glosten and Harris to estimate the information component, the inventory component, and the fixed cost component.[31] Their main test appears to be (p. 13) : " ... an inverse relationship should exist between the adverse selection spread (the information component) and volume." The motivation for studying intraday variations in H and C is unclear.

The concentration of volume is tested by whether hourly volume accounts for one-sixth of a day's volume. For return variance, components of the spread, tests of zero differences in adjacent means are conducted. This

---

[31]As pointed out in Chapter II, Glosten and Harris's approach requires additional assumption(s) (e.g., H=0, or C and Z as some function of volume). FV does not assume that C and Z are independent but details of their decomposition are not given, and the solution to the identification problem is unclear.

estimation procedure is valid only for comparing adjacent means. The stated confidence level may be incorrect, since there are a family of hypotheses to be tested.

Generally, they find various components of the spread, volume, and return variance demonstrate U shaped patterns. These results should reject AP model. However, it seems the correlation between volume and price variability leads them to conclude : " ... results 'weakly' support" the AP model.

FV separate the sample into 10 dollar volume deciles. The motivation for using dollar volume appears to use dollar volume to proxy for the degree of attention given to the stock. However, no specific hypothesis is made regarding the expected differences among deciles. Moreover, dollar volume is influenced by price, which may reflect the degree of the discreteness problem.

Although there are 10 deciles available, only results for the first, the fifth, and the tenth deciles are shown. Less active firms show a more pronounced U shaped pattern. No explanation is given for this result. Moreover, the U shaped pattern in Z is not robust for larger volume stocks.

IV. C. Competing Hypotheses

In addition to the AP model, there are other explanations that may be consistent with the concentration of volume and price variability. This section is devoted to discuss these competing hypotheses.

The AP model predicts that when the market is deeper, λ (or Z) is lower; and this occurs with more informed traders coming into the market because of the competition factor. However, H and C can also be expected to be lower in a deeper market. In an environment characterized by the coexistence of H, C, and Z, lower Z (and the competition factor) is no longer a necessary condition for the concentration of volume and price variability.

A more general concept inspired by AP is that the concentration of volume and price variability should come with a lower total transaction cost[32]. Since informed traders prefer to trade in a deeper market, a higher Z may likely to be present if the degree of competition among informed traders is insufficient. It is clear that the concentration of volume will arise only if the total transaction cost is lower (hereinafter, the transaction cost hypothesis). The total transaction cost should be lower with lowering in Z, H, C or some combination of Z, H and C. Otherwise, discretionary uninformed traders will choose not to trade together and a concentration of volume is not possible.

---

[32]We concentrate on cost arising from trading with dealers and ignore commissions and other transaction costs.

A test of this simple concept is important, since it provides a test of the relative importance of discretionary uninformed traders and nondiscretionary uninformed traders. If all uninformed traders can time their trades, no trading will occur in some periods and heavy trading will be observed in some other periods. On the other hand, if all uninformed traders cannot time their trades, the concentration of volume can seldom occur.

It appears that the overnight liquidity demands alone can explain the opening concentration in volume and price variability. If all uninformed traders cannot time their trades, their overnight liquidity demands have to be satisfied at the opening. The accumulated demands can cause a greater depth in the opening and may result in a lowering in the transaction cost (hereinafter, the overnight demand hypothesis). This hypothesis predicts higher volume in the opening but the consequence on price changes and transaction cost is not so clearcut. If more informed traders are induced by greater liquidity trading in the opening, the information component may or may not be greater, depending on the degree of competition among informed traders. However, lower order/inventory cost can be expected in the opening. In contrast, the transaction cost hypothesis does not strongly predict a particular intraday pattern.

The literature of the relationship between volume and price variability is also relevant here, since a concentration of volume and price variability also implies a positive relationship between volume and price variability. Generally, according to the theory of volume and price variability, uncertainties in information will induce more trading and more variability in prices. If the rate of information flow is much faster in certain periods, greater volume and price variability will be observed in those periods (hereinafter, the rate of information flow hypothesis).

The AP model is distinguished from the literature of volume and price variability in two aspects : (1) the role of the dealer is explicitly considered, and (2) the interaction between and among informed and uninformed traders is modeled. Under AP, a concentration of volume and price variability can occur with a constant rate of information flow. More importantly, an important and testable hypothesis of the transaction cost model is that the transaction cost should be lower in periods of concentration of volume and price variability. This prediction is not a necessary consequence of the literature of volume and price variability. In the literature of volume and price variability, the liquidity demand does not enter into the rate of information flow argument. More importantly, the role of the transaction cost is not explicitly considered. We suspect that the information component may well be higher in periods of higher volume that suggests information is more uncertain. That is, the uncertain nature of the information precludes competition (and lower Z) being a major factor in trades. Also, the rate of information flow hypothesis does not predict a particular intraday pattern[33].

The following give important predictions that various hypotheses would prescribe.

Hypothesis (1) : The concentration of volume and price variability will occur if the information component is lower. The concentration of volume and price variability will not always occur; when it does not, the information component should be fairly constant. (The AP model.)

Hypothesis (2) : The concentration of volume and price variability will occur if the transaction cost is lower. The concentration of volume and price variability will not always occur; when it does not, the transaction cost should be fairly constant. (The transaction cost hypothesis.)

[33]Brock and Kleidon (1989) argue inelastic demands at the opening and closing may account for the higher volume in these periods. Demands are more inelastic in the opening if traders have greater divergences in belief. Even if this is true, there should be no intraday pattern in C.

Hypothesis (3) : Higher volume, price variability, and lower order/inventory cost should be observed in the opening (The overnight demands hypothesis.)

Hypothesis (4) : Higher (lower) volume occurs with higher (lower) price variability, and this likely occurs with higher (lower) information component and transaction cost. There should be no particular intraday pattern. (The rate of information flow hypothesis.)

It is clear that these four hypotheses are not mutually exclusive. Moreover, the relative explanatory power of various hypotheses may vary over time and across stocks.

We suspect that the degree of concentration of volume and price variability may vary across stocks, based on three reasons. The first reason relates to the competition factor that AP model heavily relies upon. The degree of competition is of course difficult to measure. However, since high-activity stocks are more likely to be closely watched, the degree of competition among informed traders is likely to be greater. This suggests the concentration of volume and price variability should be more pronounced for higher-activity stocks. Second, if the relative importance of the discretionary uninformed traders varies across stocks, the concentration of volume and price variability will be more obvious for stocks that discretionary traders play a more important role in. The relative importance of the discretionary traders across stocks is unclear. Institutional traders seem to prefer high-activity stocks. Nevertheless, institutional traders may not be equivalent to discretionary uninformed traders. Furthermore, although there may be fewer institutional traders interested in low-activity stocks, the relative importance of these traders may be actually higher if they assume more active trading than ordinary investors. No specific pattern across stocks is suspected at this point. Third, if liquidity trading is relatively more important for some stocks, the opening volume should be relatively high, according to the overnight liquidity demands hypothesis. The suspected cross-sectional differences motivate our analysis of different samples, separated by number of trades (to proxy for trading activity).

Although the data we use cover a short period (two months), it includes days of record volume and price variability (the so called 1987 October Crash). One obvious problem is that trading halts are more prevalent during these record-setting days. This may distort prices and volume. Also, there are reasons to expect that the post-crash behavior of specialists may differ from their pre-crash one. A report of the impact of the crash on Amex states[34] : "Amex specialists risked unprecedented amounts of capital to maintain orderly markets, buying six times the usual number of shares on October 19-20.", and "..., so on October 21, when the Amex Market Value Index shot up almost 9 percent, they (specialists) fulfilled their responsibilities by selling stocks, often at a substantial loss." For these reasons, we present separate results for the two periods.

There is still another complication. French (1980) and French and Roll (1986) document a day of the week effect in mean return and price variability. Theoretical explanation is nevertheless unclear. French and Roll argue the difference between weekend and weekday variances reveals the relative importance of public information, private information, and noise trading. Foster and Viswanathan (1989) hypothesize that the private information is revealed partially through some public signal. The public signal is weak during the weekend, and therefore discretionary uninformed traders prefer to trade later in the week. This predicts less volume and price variability, and higher information component on Monday. It appears that the day of the week effect is related to the interaction between the public and private information processes, and there is no

---

[34]American Stock Exchange. A Report to Listed Companies on the Performance of the American Stock Exchange During the October 1987 "Crash". January 30, 1988.

strong link between the time of the day effect and the day of the week effect. Nevertheless, we are concerned with the nonstationarities in return and volume distributions that are implied by the day of the week effect. The following analysis includes consideration of this factor.

IV. D. Methodology and Samples

Our main interest is the intraday variations of the trading activity, the price variability, the information component, the transitory component (the sum of inventory and order components), and the total transaction cost. Proportion of block trade (the number of block trades - 10,000 shares and above[35]- divided by the total number of trades, BLOCK) is also analyzed. BLOCK may represent either or both of the following : the amount of informed trading (if informed traders prefer to trade larger quantities (Easley and O'Hara (1987)), and the amount of institutional trading (if institutional traders tend to trade larger quantities). Therefore, BLOCK should provide additional information concerning cross-sectional variations in the concentration of volume and price variability.

What constitutes an adequate measure of trading activity is not as apparent as it seems. Number of trades

[35]Defined by New York Stock Exchange.

should represent the level of interest in the stock. However, trade size enters as an additional signal of trading interest (Easely and O'Hara (1987)). This suggests volume should also be investigated. Moreover, volume is more of the focus of previous studies (e.g., Jain and Joh (1988)). We will investigate both the intraday variations in number of trades (NT) and volume (VOL). The measurement of price variability also involves some problems. Previous studies in this area use different measures. Some (e.g., Wood, et al (1985)) use return variance (RETV) while others (Mulherin and Gerety (1989)) use the absolute values of returns (RET). We look at both RETV and RET in this chapter.

There are problems in the estimation of components of the spread, as discussed in Chapters II and III. However, if the Stoll's and the Glosten/Harris's approaches give similar results, it will increase our confidence in the estimation of the components of the spread. Stoll's approach is explained in details in Chapters II and III. The application of Glosten and Harris's approach requires some further discussion.

Glosten and Harris's empirical model can be specified in a number of different ways. To search for the most useful specification, they test the significance of the coefficients of the following regression :

$$P_t - P_{t-1} = C\ (Q_t - Q_{t-1}) + Z\ Q_t + e_t$$

where

Q_t = trade type (purchase or sale) at time t

C = transitory component = $C_0 + C_1 V_t$

Z = information component = $Z_0 + Z_1 V_t$

V_t = transaction size at time t

They find the most useful model is one that $Z_0 = C_1 = 0$, constant e_t mean and variance without accounting for the discreteness in prices, which indicates an OLS estimation of the form :

$$P_t - P_{t-1} = C_0 \ (Q_t - Q_{t-1}) + Z_1 \ Q_t \ V_t + e_t$$

provided that Q_t's are known[36]. Since our analysis requires a large number of estimations, specification search for every stock and every hour is simply not economical. Rather, we use the most useful model found in their study. (Here V_t is stated in hundreds of shares.) The information component is two times Z_1V_t and the transitory component (C2) is twice C_0, reflecting average cost for round-trip transactions. The information component is notated as Z1 if Stoll's approach is employed, Z2 if Glosten and Harris's is adopted. Likewise, C1 and C2 represent the transitory

---

[36]To identify the trade type Q_t (purchase or sale), the transaction price is compared to the most recent quote. Midpoint transactions (about 20% of all trades) are treated as neutral. To this end, it should be pointed out there are some classification rules to classify these midpoint transactions as discussed by Hasbrouck (1987). His classification rules are not utilized here.

component when employing Stoll and Glosten/Harris, respectively.

To measure the total transaction cost is much more difficult than it appears. The quoted spread seems to be a candidate for the measure of the transaction cost. However, as Grossman and Miller (1988) and Stoll (1989) point out, the quoted spread is the difference between the ask and the bid prices at one point in time. The quoted spread is therefore not representative of the true cost that is faced by an ordinary investor who seldom, if ever, simultaneously buys and sells a security at the same point in time. Furthermore, Stoll (1985) and Smidt (1988) find the return realized by dealers is substantially different from (in most cases, substantially lower than) the quoted spread. We include average proportional quoted spread (AVGS) nevertheless. The percentage changes in quoted prices ($\triangle Q$) (quoted price is given as the average of the bid and ask prices) is given as a measure of transaction cost. An astute investor would prefer to trade when dealers are less sensitive to trades, which should be captured by $\triangle Q$.

Two studies point to potential factors that may influence $\triangle Q$. Schultz, Gustavson and Reily (1985) provide an experimental study of factors influencing the opening quotes. They find that the price trend, the inventory position, the limit book, and especially, the order imbalance are important factors in the setting of opening prices. Marsh and Rock (1986) find a statistical relationship between ΔQ and net number of trades[37]. An empirical model of ΔQ is developed based partly on their findings.

Comparisons of hourly means for ΔQ are insufficient for our purpose. First of all, ΔQ are influenced by the factors discussed above. Second, the theory in the components of the spread indicates that order imbalance should lead to adjustment in the location of the quoted price. If ΔQ is related to order imbalance as theory suggests, this will provide evidence that ΔQ reflects the transaction cost.

ΔQ = f (HOUR DAY NETT NT RET RETV ΔS)

where HOUR = a set of dummy variables representing the hour of the day

DAY = a set of dummy variables representing the day of the week

NETT = net number of trades (investor purchase minus investor sale)

NT = total number of trades

RET = hourly return

RETV = trade-by-trade return variance

[37]Their analysis is conducted on the level of individual stocks while ours focuses more on aggregate statistics.

$\triangle$S = percentage changes in proportional spread

NETT gives the magnitude of the order imbalance. NT proxies for the depth of the market. RET and RETV are proxies for price trend and price risk, respectively. Since the magnitude of the spread is likely not a constant over the estimation period, $\triangle$S is included to control for this factor.

The frequency of the quote adjustment (FREQ) is another potentially useful piece of information. A higher FREQ suggests more information is coming into the market. On the other hand, changes in the transaction cost will likely reflect more of the degree of informational asymmetry between informed traders and the dealer. Therefore, the inclusion of FREQ may help in distinguishing the rate of information flow hypothesis and the AP hypothesis. FREQ is measured by the average number of trades between quotes. However, incidence of consecutive quotes without trades is often observed. For our purpose, in the case of consecutive quotes without trades, only the first quote is kept for our analysis.

Without prior studies of FREQ, the empirical model of FREQ will be specified less precisely. A model similar to that of the $\triangle$Q is adopted for FREQ, with the modification that NETT, RET, and $\triangle$S are now in their absolute values.

FREQ = f (HOUR DAY NETT NT RET RETV $\triangle$S)

Ideally, we want to employ similar analysis of $\triangle Q$ to other variables that are of interest. Unfortunately, factors influencing these variables are not known. We include only the hour effect and the day effect for the analysis. The two-factor analysis of variance (ANOVA) approach is adopted to test whether the means of NT, VOL, RETV, RET, Z, C, AVGS, and BLOCK are different across hours. The main focus is on the hour of the day effect and the day of the week effect is included as an additional factor. No interaction is assumed to exist between the day and the hour effects.[38]

The same data set used in chapter III is employed here. To study the hour of the day effect, a day's trading is split into six trading periods. During the data period - September and October 1987, two months or 44 trading days - the exchange normally is open six and a half hours[39]. Therefore, a day's trading is partitioned into six sixty-five minute periods. (For convenience, hours and periods will be used as interchangeable terms hereinafter.)

---

[38]The methodology is equivalent to the regression approach using purely dummy variables. If interaction effect is present, it is less meaningful to compare factor level means. Therefore, the exclusion of the interaction effect simplifies the analysis to a great extent. Furthermore, our general results are the same even if the interaction effect is incorporated.

[39]This is verified by the Wall Street Journal.

The nature of this study (i.e., exploratory) requires a large sample. We arbitrarily choose the first 500 stocks in the data base that meet the following criteria : 1) NYSE stocks, 2) no stock dividends or special dividends issued over the data period, and 3) minimum price increment is one-eighth over the entire data period. Some 253 stocks are eliminated by 1), 5 are eliminated by 2) and 87 are excluded due to 3), when reaching the 500 mark[40].

A large sample also requires all 44 days be included. As argued before, however, we present separate analysis for the pre- and post-crash periods. Pre-crash period covers August 31 through October 15. Post-crash period begins with October 16 (the day before the crash, which also exhibits substantial price variability and high volume) and ends with October 30. For the pre-crash period, there are 6 Mondays, 6 Fridays, and 7 each for the other weekdays. For the post-crash period, there are two each for every weekday with the exception of Friday (three).

For each stock, each hour and each day, NT, VOL, RETV, RET, Z1, Z2, C1, C2, AVGS, and BLOCK are computed. RET is the absolute percentage changes between the first and the last transaction prices over an hour. RETV is the trade-by-

---

[40]Additionally, four stocks do not yield convergence when using Stoll's approach (visual inspection indicates this is a result of some doubtful prices) and are excluded. Also, some transactions are excluded, as detailed in chapter III, section B.

trade return variance over an hour. Since at least three transactions are needed to compute covariance of price changes (and Z1, C1), this essentially limits our ability to analyze low-activity stocks. Nevertheless, to be consistent in computing NT, VOL, RETV, RET, Z, C, AVGS, and BLOCK,[41] low trades observations (i.e., less than three trades and three quotes in an hour) are excluded. Therefore, not all stocks have an observation for every one-hour period. This is especially true for a sample of low trading stocks.

Three samples are created based on cutoff points of 600 (approximately 25% quantile of the data set) and 3,800 (approximately 75% quantile of the data set) trades over the data period.[42] The exchange closes two hours early for the last six days of the data period. Therefore, there are a maximum of 252 observations for each stock. High-activity sample contains 135 stocks, 30249 observations (pre-crash : 23786, post-crash : 6463), with the observations for individual stocks ranging from 101 to 252 (average = 224). Medium-activity sample consists of 252 stocks, 31252 observations (pre-crash : 21705, post-crash : 9547), and

[41]In computing RETV, RET, Z, C, and AVGS, prices are adjusted for cash dividends. Namely, the amount of dividend is added to the prices subsequent to and including the ex-dividend date.

[42]The medium sample is then artificially created to be larger than the high and the low samples. This is to create a large enough medium sample so that there is confidence for the medium sample at least.

individual observations range from 28 to 238 (average = 124). Low-activity sample contains 92 stocks, 2115 observations (pre-crash : 943, post-crash : 1172), with individual observations ranging from 1 to 60 (average = 23). Some 21 stocks are excluded from the low activity sample due to insufficient trading over all hours. We know, therefore, the inference based on the low activity sample will be less powerful. The average volume across stocks and hours are 7,080 shares, 15,541 shares, and 63,158 shares for the low-activity sample, medium-activity sample, and high-activity sample, respectively.

The generation of these samples should be considered as a conservative approach. Specifically, by systematically discarding low trading observations, mean statistics for NT and VOL should show less contrast across hours than if all observations are included. If significant differences were to be found across hours, such differences should not be due to sampling biases.

The data set is quite large. Since little effort is spent in detecting potential reporting errors, some more robust procedure that is insensitive to data errors is desired. To this end, observations are ranked and the same ANOVA approach is used on the rank transformed data. Conover (1980, p. 337) comments on the rank transformed data : "The result is a robust regression method that is not sensitive to outliers or nonnormal distributions to the extent that regular regression methods on the data are affected." If the two methods give different results, then : " ... the analysis on ranks is probably more accurate than the analysis on the data and should be preferred."

The AP model predicts concentration of trading activity and price variability for some periods. The argument for the U-shaped pattern is less convincing. To gain adequate inferences of differences in hourly means, multiple mean comparisons procedure is required. Multiple comparisons procedure is necessary because a family of comparisons are of interest here. There are a number of methods for performing multiple comparisons. The Tukey method[43] is chosen here because it gives narrower confidence limits if all or a large number of pairwise comparisons are of interest. Moreover, "The Tukey method is exact if all factor level sample sizes are the same, is a conservative method when the sample sizes are unequal." (Neter, Wasserman, and Kutner (1985).)[44]

---

[43]The Tukey method utilizes the studentized range distribution. The confidence interval for difference in means is the difference in hourly means plus and minus $Ts^2(D)$, where $s^2(D) = (MSE/b^2)\Sigma_j(1/n_{ij}+1/n_{i'j})$, $T = (1/\surd 2)q[1-\alpha;\ a,\ n_T-ab]$, a (b) is number of levels for the hour (day) factor, i (j) is the ith (jth) level for hour (day), n_T is the total number of observations, and q represents the studentized range distribution.

[44]See also Stoline (1981) for a survey of multiple pairwise comparisons methods.

IV. E. Empirical Results

For clarity of the presentation, results for various samples are given separately, in the order of the results for the high-activity sample first (Table 8 for the pre-crash period, Table 9 for the post-crash period), medium-activity sample second (Table 10 for pre-crash, Table 11 for post-crash), and low-activity sample last (Table 12 for pre-crash, Table 13 for post-crash). Since the analysis is the same for all samples, more details will be given in discussing results for the high-activity sample and fewer details will be given for the analyses of other samples. No statistical test is used to test differences among samples, although some discussion will be presented at the latter part of this section.

Table 8 panel A gives the means and standard deviations (they are pooled time series and cross-sectional statistics) for NT, VOL, RETV, RET, Z, C, AVGS and BLOCK. Mean NT, VOL, RETV, RET, Z1, AVGS, BLOCK are the highest, while mean C1 and C2 are the lowest in the first hour. To see the significance of the hour effect as well as the marginal significance of the day effect, ANOVA result is given in panel B. F gives the overall significance of the ANOVA result. F_h represents the F value if only the hour variable is included. $F_{h/w}$ $(F_{w/h})$[45] measures the marginal significance of the hour (day) variable when the day (hour) variable is also included. For all variables, the hour effect is statistically significant at 1% level[46]. The day effect shows some inconsistency for RETV, Z1, Z2, and C2 using the original data and rank transformed data, which suggests that measurement problem is present for these variables. Overall, the day effect is significant.

Panel C gives the signs of hourly mean differences (the differences in means can be inferred from panel A). Except for difference in RETV between hours 1 and 6, the first hour patterns are significant for all variables. NT, VOL, RET, RETV, Z1, Z2, AVGS, and BLOCK are the highest, and C1, C2 are the lowest in the first hour (although Z2 and C2 do not give same conclusion using original data and transformed data). In contrast, remaining hours show mixed results. NT, VOL and RET for hour 2 are higher than those of hours 3-5 but no significant differences are observed for components of spread or AVGS. In general, hour 6's NT, VOL, RETV and RET are higher than those of hours 3-5. With the exception of hour 4, no significant differences between hour 6 and

---

[45]For example, $F_{h/w}$ is $(SSE(R)-SSE(F))/(df_R-df_F)/(SSE(F)/df_F)$, where F represents the full model, R represents the reduced model (i. e., hour effect is excluded). Test for the interaction effect is similar.

[46]For convenience, the term "significant" refers to "significant at 1% level" hereinafter.

hours 2-5 in components of the spread are observed. The results suggest the relationship between trading activity and transaction cost differs across hours. First hour results do not contradict hypotheses (2), (3) and (4). The results for the remaining hours contradict hypotheses (1) through (3). To better distinguish hypotheses (2) and (4), the analysis for $\triangle Q$ is needed.

Since patterns among hours 2-5 are similar, we simplify the analysis of $\triangle Q$ by ignoring differences among hours 2-5. For the day effect, the significant patterns using both the original data and the rank data are that mean NT (Monday : 20.25, Friday : 18.74, others : 22.61), VOL (Monday : 410, Friday : 438, others : 508) and RET (Monday : 0.0052, Friday : 0.0047, others : 0.0061) are lower on Monday and Friday. In addition, mean BLOCK is higher on Friday (0.0477 vs. 0.0437 for other days)[47]. We further simplify the analysis to the following :

$\triangle Q$ = f(H1, H6, W1, W5, NETT, NT, RET, RETV, $\triangle S$)

where H1 is a dummy variable = 1 if hour 1
= 0 otherwise

H6 is a dummy variable = 1 if hour 6
= 0 otherwise

[47]The interaction effect is present for NT (F=6.49), VOL (F=3.09) and RET (F=14.70). Nevertheless, results for hourly mean comparisons are identical in terms of panel C whether the interaction effect is included. This is true for all samples.

W1 is a dummy variable = 1 if Monday

= 0 otherwise

W5 is a dummy variable = 1 if Friday

= 0 otherwise

Similar analysis is conducted for FREQ, except that NETT, RET, and $\triangle$S are now in their absolute values.

FREQ = f(H1, H6, W1, W5, NETT, NT, RET, RETV, $\triangle$S)

The OLS regression approach is adopted for the analysis of $\triangle$Q and FREQ, and the result is shown in Table 8 panel D. The estimated coefficient of H1 measures how much higher the mean $\triangle$Q for hour 1 is than the one for hours 2-5. Similarly, the coefficient of H6 measures how much higher the mean $\triangle$Q for hour 6 is than the one for hours 2-5. The difference between the coefficient of H1 and the coefficient of H6 gives how much higher $\triangle$Q for hour 1 is than that for hour 6. Similar interpretation is applied to W1 and W5.

H1 is significantly negatively related to $\triangle$Q, indicating that the sensitivity of dealer's price adjustment is lower at hour 1. H1 (H6) is positively (negatively) related to FREQ, suggesting specialists adjust quotes more (less) frequently at hour 1 (hour 6). Since $\triangle$Q is lower at hour 1, the depth of the market is inferred to be greater at hour 1. Combining with previous results (higher NT, VOL, RETV, RET, Z1, Z2, AVGS, BLOCK and lower C1, C2 at hour 1), these results support hypotheses (2) and (3). The last hour increase in NT, VOL and RET without significant changes in

Z, C and $\triangle Q$ is more consistent with hypothesis (4). It suggests that the rate of information flow is faster at the last hour. A potential explanation is that because the information is uncertain, investors unwind their positions before the close in order to avoid risk that may incur overnight[48].

Regression result in panel D confirms that $\triangle Q$ is positively related to NETT, which is the underlying behavioral assumption that is commonly employed in the studies of the component of the spread. That is, net investor purchase (sale) will lead to an upward (downward) adjustment in quoted price. The FREQ regression is confusing; in particular, the negative relationship between FREQ and the absolute value of NETT contradicts competitive behavior among specialists. It is quite possible that FREQ model is not well specified.

In contrast, the post-crash results in Table 9 demonstrate inconsistent patterns (inconsistency here refers to contradicting patterns for similar variables such as NT and VOL). Mean NT, RET are the highest and C2 is the lowest at hour 6, but mean VOL, RETV, Z1, Z2, AVGS, BLOCK are the highest and C1 is the lowest at hour 1. Panel B indicates that the hour effect is insignificant for RETV, AVGS, and for original Z2 and C2. The day effect is particularly

[48]Jain and Joh (1988) make similar argument.

strong for NT, VOL, transformed RETV, RET, AVGS and BLOCK[49]. Multiple comparisons results (panel C) are generally inconsistent. The only significant and consistent pattern in trading activity and transaction cost occurs between hour 1 and hours 3-4, between hour 2 and hours 3-4 (hours 1-2 have higher NT, VOL, lower C1 and C2). A possible explanation of the weaker concentration in trading activity and transaction cost is that informed trading dominates uninformed trading for the post-crash period. This explanation predicts that the concentration of volume, price variability and transaction cost is less likely to be present for the post-crash period, according to the transaction cost and the overnight demands hypotheses. The insignificance of H1 and H6 in $\triangle Q$ regression (panel D) leads to similar conclusion.

The pre-crash results for the medium-activity sample are shown in Table 10. Mean statistics in panel A demonstrate some inconsistency in NT and VOL, and in Z1 and Z2. Nevertheless, it seems clear that more obvious patterns occur at hours 1 and 6. Mean NT is the highest at hour 6. Mean VOL, RET, Z1, AVGS are the highest, and C1, C2 are the lowest at hour 1. Mean BLOCK is the highest at hour 3.

---

[49]Mean NT and VOL are lower on Thursday (33.96, 1086) and Friday (35.69, 934) than on other days (42.74, 1319). Mean RET, AVGS and BLOCK are the lowest on Friday (0.0131, 0.0115, 0.0573) than other days (0.0280, 0.0163, 0.0682). The interaction effect is present for RET and AVGS (F=26.81, 2.25).

Additionally, mean Z2 and C2 are quite large, compared to the high-activity sample. The average cost for 100 shares round-trip transactions according Glosten and Harris's approach is approximately 14 (=Z2+C2) times the quoted spread (about more than one-eighth of a dollar), which seems extremely high. Panel B ANOVA results show that, for RETV, Z2, C2, and BLOCK, the hour effect is insignificant using original data while significant using rank data. Hour effects for NT, VOL, RET, Z1, and AVGS are significant and more robust. The day effect is significant only for NT and VOL[50].

Multiple comparisons results (panel C) show some contradicting signs for difference in NT and VOL between hours 1 and 6. The consistent and significant patterns occur between hour 1 and hours 2-5. Like the results for high-activity sample, NT, VOL, RETV, RET, Z1, Z2, and BLOCK are generally higher, and C1 is lower at hour 1 than those at hours 2-5. This is consistent with hypotheses (2) and (3). NT, VOL, RETV, and RET for hours 4-5 are lower than those of hour 6, which is more consistent with hypothesis (4). In panel D, the analysis of $\triangle Q$ indicates $\triangle Q$ is lower at hour 1, which is again more consistent with hypotheses (2) and (3). W5 is significantly positively related to $\triangle Q$.

---

[50]Mean NT is 7.07 on Monday, 7.00 on Friday versus 7.51 for other days. The interaction effect is present for NT and RET (F=4.60, 1.89).

This is inconsistent with Foster and Viswanathan (1989), since their model predicts only Monday is different from other weekdays.

For the post-crash period, mean NT is the highest at hour 6, VOL, RET, Z2, AVGS, and BLOCK are the highest at hour 1. On the other hand, RETV is the highest at hour 3, Z1 is the highest at hour 5, and C1 (C2) is the lowest at hour 5 (hour 2). The hour effects for all transformed variables are significant. Like the results for the post-crash high-activity sample, the day effect is particularly strong for NT, VOL, transformed RETV, RET, and AVGS (panel B)[51]. Investigation of panel C indicates that the only consistent and significant pattern in trading activity, price variability and transaction cost occurs between hour 1 and hour 4. NT, VOL, RETV, RET, Z1, Z2, BLOCK are higher, and C1 is lower at hour 1 than at hour 4. Other patterns are generally either insignificant or inconsistent, this reinforces our previous conjecture that the relative importance of uninformed trading decreases in the post-crash period.

Table 12 presents results for the low-activity sample before the crash. The results are relatively straightforward, compared to those of medium and high

---

[51]Mean NT, VOL, RET and AVGS are lower on Friday (9.67, 173, 0.0151, 0.0190) than on other days (12.17, 231, 0.0275, 0.0266). The interaction effect is present for NT, RET, Z1 and AVGS (F=3.82, 15.94, 2.09, 4.62).

activity samples. Simply put, there are no significant intraday patterns in NT, VOL, RETV, Z, C, AVGS, and BLOCK (Panels B and C). Panel A shows that mean NT and RETV are the highest at hour 1. VOL is the highest at hour 5. Z1, Z2 AVGS, BLOCK are the highest, and C1 is the lowest at hour 4. The analysis of $\triangle$Q shows no significant difference in the price sensitivity between hour 1, hour 6 and the rest of the day[52]. The non-pattern in the low-activity sample can be explained if the relative importance of the discretionary traders is less for low-activity sample. If institutional traders are discretionary traders and if they are less interested in low activity stocks, AP model will predict little or no concentration in volume and price variability.

Post-crash results for the low-activity sample are again insignificant (Table 13). Panel A shows that mean NT, VOL, RET, C2 and BLOCK are the highest at hour 1. RETV and Z2 are higher at hour 2. Panel B indicates that the hour effect is significant only for VOL and for transformed

---

[52]For all regressions in this chapter, the analysis of eigenvalues of the X'X matrix shows that the multicollinearity problem is not severe.

BLOCK. Panel C indicates the only significant pattern occurs for VOL between hour 1 and hours 2-4.[53]

One noteworthy observation is the variations in BLOCK before and after the crash. For high-activity sample, BLOCK generally increases in the post-crash period. The opposite pattern occurs for the low-activity sample. As we have argued before, BLOCK may represent either or both of the institutional trading and informed trading. The observation here contradicts BLOCK is solely a measure of informed trading, since it is difficult to explain why the relative importance of informed trading decreases for some stocks while increases for some other stocks, unless the crash is a result of company-specific information, rather than market-wide information. If this conjecture is true, it reinforces hypotheses (2) and (3), since we find that BLOCK is higher in periods of higher trading activity, price variability, and lower $\triangle Q$.

For the pre-crash high-activity sample, VOL for the first hour is relatively high. Mean VOL for the first hour (=761.24) is about twice the average of those of other hours, although standard deviation of VOL is also the

---

[53]Mean NT, RETV, RET, and AVGS are lower on Monday and Friday. The interaction is present for RET (F=2.02). If 5% significance level is used instead of 1%, there are some small number of changes. For the post-crash low-activity sample, the only change is that VOL is significantly higher at hour 1 than at hours 2-5. For all samples, the general results are unaffected even if 5% is used instead.

highest at hour 1 (=1239, which is about twice the average of those of other hours). In contrast, Medium-activity sample shows less concentration in VOL. Mean VOL is the highest at hour 1 (=144.66) but the average VOL for other hours is only slightly below that of hour 1 (average is about 120). The low-activity sample is in even sharper contrast to the high- and medium-activity samples. No significant patterns are observed for the low-activity sample. This is consistent with the level of trading activity being a proxy for the degree of competition among informed traders (hypothesis (1)). It is also consistent with hypothesis (2) or (3) if the relative importance of discretionary uninformed trading (hypothesis (2)) or liquidity trading (hypothesis (3)) is higher for higher-activity stocks.

In all $\triangle Q$ regressions, $\triangle Q$ is significantly positively related to NETT and RET. Only in the post-crash high-activity sample is $\triangle Q$ significantly negatively related to NT. The negative relationship between $\triangle Q$ and NT is consistent with the notion that dealers response less sensitively to trades when the market depth is greater. Also, there is a significantly negative relationship between $\triangle Q$ and $\triangle S$ for the medium sample, which suggests some tradeoff between $\triangle Q$ and $\triangle S$. These results seem worthwhile for future research.

The caveat in Chapter III applies here. That is, there are some measurement problems involved, especially for the components of the spread. In this chapter, we find that Glosten and Harris's approach may be sensitive to sample size. Average round-trip Z2 or C2 is within the range of 0.1 to 2 for the high-activity sample. For the low-activity sample, the average number of Z2 or C2 is above 10 in most cases. In contrast, Stoll's approach (C1 and Z1) does not appear to be sensitive to sample size.

As in Chapter III where we provide aggregate statistics for various components of the spread across stocks, the independence among stocks is implicitly assumed. Nevertheless, we find the high- and the medium-activity samples do not appear to behave much differently from each other. The difference between the low sample and the high (medium) sample may be due to the degree of measurement problem, or may be due to some differences in stocks characteristics. Either explanation suggests more caution should be taken with respect to results for low-trading stocks.

IV. F. Summary and Conclusions

Intraday variations in means of number of trades, volume, return variance, absolute value of return, the information component, the transitory component, the sensitivity of dealer's price adjustment, and the proportion of block trades are investigated for the period of September and October, 1987. Three samples, separated by the number of trades, show different patterns. High-activity sample shows that trading activity, price variability, the information component, and proportion of block trades are the highest, and the transitory component, the sensitivity of price adjustment are the lowest at the first hour. Medium activity sample shows similar results. In both samples, trading activity increases at the last hour without significant changes in the information component or the price sensitivity. In addition, the concentration of trading activity, price variability and transaction cost weakens in the post-crash period. In contrast, low-activity sample demonstrates no significant intraday patterns, for both the pre- and post-crash periods.

The first hour results are consistent with the notion that liquidity traders time their trades to minimize transaction costs, inconsistent with Admati and Pfleiderer's model that predicts a negative relationship between information component and trading activity. The results are also consistent with overnight liquidity demands hypothesis. Our test unfortunately cannot fully distinguish the transaction cost hypothesis and the overnight demands hypothesis. Nevertheless, according to the overnight liquidity demands hypothesis, the differences among samples suggest liquidity traders are more important for higher-activity stocks, which lacks both theoretical and empirical support. In contrast, the transaction cost hypothesis says that the relative importance of discretionary uninformed trading is more for higher-activity stocks. This explanation seems more appealing. The two hypotheses nevertheless point to the importance of trading not motivated by information.

The patterns of the remaining hours are more consistent with the rate of information flow hypothesis. Of course, the rate of information flow hypothesis does not predict a particular intraday pattern. The significant pattern at the last hour can be explained by investors' avoiding overnight risk exposure by unwinding their positions before the close. Since these transactions involve little precise information, no substantial changes in the information component is predicted. The difference between the pre- and post-crash results is consistent with an increase in the relative importance of informed trading for the post-crash period.

There are two other observations that should be considered when interpreting the results. One is the difference between trading procedures for the opening and the remaining markets. The opening operates like a call market while the remaining is market in which dealers perform a more critical role. It is unclear that whether difference in trading procedures may affect our results. To this, we point out a finding by Garbade and Sekaran (1981) that opening prices are unbiased estimators of subsequent (up to an hour) intraday prices. Another observation is the Harris's (1989) finding of a tendency for the last trade to occur at the ask price. Nevertheless, to the extent that opening and closing minutes are not representative of remaining minutes, the effect should be minimal for our hourly statistics. It is imaginable that the effect may be more severe for low-trading stocks, for which we do not find significant patterns.

Additional results also point to areas for future research. In particular, changes in quoted prices relate more to order imbalance than to the total number of trades. This result provides empirical support for the framework adopted in the model in the third essay.

Chapter V

The Components of Trading Volume and The Components of the Bid-Ask Spread : Model and Application

V. A. Introduction

A simple model is structured to present the simple intuition that the composition of trading volume (informed and uninformed components) conveys information to the dealer about the degree of the information asymmetry. Thus, there is a link between the components of trading volume and the components of the spread. We assume a one-period model and structure the model generally enough to avoid detailing the trading process and the process resolving information differences among market participants. An application of the model then follows.

V. B. Related Issues

Volume is empirically easy to measure. However, it is not so easy to model volume theoretically. Volume is the combined absolute purchase and absolute sale. In itself, volume reflects neither the demand nor the supply of the security. This may partly explain the largely ignorance of volume by financial economists[54]. Some theoretical models construct complex measure of volume, but often do so in an ad-hoc fashion. Since transaction-by-transaction data become increasingly available to researchers, net volume can be empirically estimated. Furthermore, net volume has a clearer economic interpretation - the order imbalances.

Since at a given point in time, only a small proportion of investors can participate in trades, prices likely will not reflect the equilibrium value of the security. The disequilibrium model such as that developed by Goldman and Beja (1979) makes this point clearly. However, this approach is not taken here because it is extremely complex to model interaction among various trading parties using this approach. Rather, the market is structured to resemble a combination of the call market and continuous market to allow the equilibrium type of approach.

V. C. The Model

Consider that we are in a market with three types of participants : a risk-neutral dealer, informed traders and uninformed traders. The demand and supply functions of the

[54]See Karpoff (1987) for a summary of literature of volume and price variability.

informed and uninformed traders can be specified as follows.

$$P_t = X_{it} - E_i V_{it} \quad \text{(C1)}$$

$$P_t = X_{mt} - E_m V_{mt} + e_t \quad \text{(C2)}$$

where

P_t = transaction price at time t

X_{it} = true price known only to the informed trader

X_{mt} = price perceived by the uninformed and dealer

V_t = trading volume at time t (can be either positive or negative)

e_t = an error term representing noise in the trade

E_i = sensitivity of informed demand/supply to prices

E_m = sensitivity of uninformed demand/supply to prices

Here the cause(s) for the noise trading is(are) not specified. It may be due to, for example, the random liquidity need of uninformed traders. On the other hand, informed traders trade only when their perceived true price is different from that of the rest of the market. The reason for informed traders knowing the true price need not be specified. It may be due to the better analytic ability or insider information.

A further assumption is made that e_t and X_{it} are normally distributed with mean and variance (0, var(e)) and (X_{mt}, var(i)), respectively. In addition, noise in the trade and the true price are assumed to be uncorrelated;

that is, cov (e_t, X_{it}) = 0. E_i, E_m, var(e), and var(i) are assumed to be exogenous and are of fixed positive values over the trading period.

Following Glosten and Harris (1988), the dealer will set prices as follows.

$$P_t = X_{mt} + (Z_t + C)\ V_t \quad \text{(C3)}$$

where

Z_t = information component

C = ordering/inventory component

Market clearing requires that the trading volume is the combined quantity desired by informed and uninformed traders.

$$V_t = V_{it} + V_{mt} \quad \text{(C4)}$$

Since we specify that the volume can be either positive or negative, V_t is better viewed as the net quantity. The market setting is similar to that of a call market. In a call market, orders are accumulated and executed at some prespecified time.[55] It is similar to a call market because

---

[55]Some models of the bid-ask spread do assume a call market environment. For example, Mendelson (1982), Ho, Schwartz and Whitcomb (1985) and O'Hara and Oldfield (1986) examine the trading process in a call market. See also Amihud and Mendelson (1987) for an empirical comparison of the opening and the closing data.

the dealer can deal with more than one trade at a time. Additionally, informed traders and uninformed traders can make trades between themselves without going through the dealer. The dealer takes up only the net quantity. However, it is implicitly assumed that traders know the bid-ask quote when they submit the orders. In this sense, it is also similar to a continuous market where the dealer maintains continuous trading by standing ready to trade with the incoming orders based on his/her quotes.

Let $d_t = X_{it} - X_{mt}$, then $E(d_t) = 0$, $var(d_t) = var(i)$. Further, substitute Equations (C1), (C2) and (C3) into Equation (C4), then :

$$V_t = (d_t-(Z_t+C)V_t)/E_i + (e_t-(Z_t+C)V_t)/E_m \quad \text{............. (C5)}$$

A risk-neutral dealer will set the quote so that he/she will break even on the trades. That is, the dealer will set the quoted price equal to the inferred true price plus the ordering/inventory cost.

$$E(X_{it}|V_t) + CV_t = X_{mt} + (Z_t + C)V_t$$

Or,

$$E(d_t|V_t) = Z_tV_t \quad \text{.................................... (C6)}$$

The inference problem faced by the dealer is to solve for the expected difference between the true price and the price perceived by the dealer (and uninformed traders). He can do that only through inferring the net quantity of trades. For example, a higher net quantity implies a greater degree of informational asymmetry. The net quantity is not known to the dealer before the actual transactions take place. However, given V_t, the expected value of d_t can be solved.

$$E(d_t | V_t) = V_t(\mathrm{var}(V_t))^{-1}(\mathrm{cov}(V_t, d_t)) \quad \text{(C7)}$$

Let $a = (1 + (Z_t + C)/E_i + (Z_t + C)/E_m)$ then from Equation (C5), the following obtain.

$$\mathrm{var}(V_t) = \mathrm{var}(i)/E_i^2a^2 + \mathrm{var}(e)/E_m^2a^2 \quad \text{(C8)}$$

$$\mathrm{cov}(V_t, d_t) = \mathrm{var}(i)/E_ia \quad \text{(C9)}$$

Substitute (C7), (C8) and (C9) into (C6), an equilibrium schedule for setting the Z in relation to C is specified as follows.

$$Z_t = \frac{\mathrm{var}(i)E_m(CE_i + CE_m + E_iE_m)}{\mathrm{var}(e)E_i^2 - \mathrm{var}(i)E_iE_m} \quad \text{(C10)}$$

Some comparative statics can be observed from Equation (C10). The information component is positively related to the ordering/inventory component. Note that the positive relationship will be even stronger when the volatility of true price is greater. This is consistent with the notion that, when information is more uncertain, the dealer is more concerned with the information problem, which also aggravates the ordering/inventory cost problem.

The information cost is negatively related to noise trading and the sensitivity of uninformed traders. A possible explanation for the negative relationship between noise trading and the information cost is that when the noise is greater, the dealer cares less about the possible loss to the informed traders. When the demand/supply of uninformed traders is insensitive, it is easier to infer the information conveyed through net quantity. Therefore, the dealer will penalize a large net quantity of trade with a higher information component. This differs from Glosten and Milgrom (1985) in which the more inelastic the demand of the uninformed, the easier can the dealer make back the losses to the informed. Thus, the dealer will set a smaller spread. The two models may not necessarily be inconsistent with each other, however. If the uninformed traders are insensitive to prices, then the dealer will have a 'normal' volume in mind. Excess of normal volume will be interpreted as originating from informed traders and thus the dealer will set a higher information component. On the other hand, a large normal volume also implies a smaller total spread.

From Equations (C8) and (C9), the following conclusions obtain. Variance of the net quantity does not necessarily increase (decrease) with the increase (decrease) in the variance of the true price. To break even on the trades, the dealer will set a higher (lower) spread when the variance of the true price is greater (lower). The mechanism will discourage (encourage) trades which in turn makes the relationship between the variance of true price and variance of net quantity ambiguous.

V. D. Application of the Model

When setting the spread, the dealer can infer the degree of the information problem by observing the order imbalance in trades and the sequence of trades. While previous studies of the components of the spread primarily utilize the latter information, our model focuses on the former.

Trading volume may result either from informed traders or from uninformed traders. The composition of trading volume depends upon the elasticity of traders, the variance of noise, and the variance of the true price, which are all unobservable. However, by combining the information about the sequence of trades (to infer the components of the spread) and the information about the variance of trading volume, a new insight concerning the relative importance of noise and private information in price variability can be gained.

An application of the model follows. Specifically, given E_i, E_m, Equations (C8) and (C10) can be used to infer var(i), and var(e) for the period of interest.

The actual implementation of our experiment involves some difficulties and ambiguities, mainly because no specific functional relationship between price changes and order imbalance is specified. To apply the model, some assumption regarding E_i and E_m has to be made. There is a reason to suspect that E_i and E_m are not equal. That is, there is a possible trade size clientele suggested by Easley and O'Hara (1987). Informed traders will prefer to trade larger quantities to maximize their profit. In our empirical experiment, we employ a range of values for E_i and E_m with the assumption that E_i is one half of E_m.

We choose the ten most actively traded stocks in 1987, with the reasoning that measurement problems for these stocks will be less and that the investors' characteristics may be similar. If employing our methodology, the more reasonable values of E's fall into the same range, it will increase our confidence of the experiment. These stocks are : American Express, Eastman Kodak, General Electric, General

Motors, IBM, Navistar International, AT&T, Texaco, USX, and Exxon[56].

Our model indicates that net volume is the critical information in inferring the relative importance of private information and noises. Series of hourly net volume and hourly ending prices are extracted from the same data set used in Chapters III and IV. However, since no specific functional relationship between net volume and price changes[57] is specified, we transform the net volume into values of +1 (if positive net volume), -1 (if negative net volume) and 0. This is quite an ad-hoc approach indeed. Nevertheless, it readily fits into Glosten and Harris's (1988) estimation approach. Recall the Glosten and Harris's basic equation assuming all transactions are unit trades and the inventory component is zero :

$$P_t - P_{t-1} = C(Q_t - Q_{t-1}) + ZQ_t + e_t$$

where P_t = transaction price at time t

Q_t = 1 if investor purchase, = -1 if sale

e_t = an error term describing the arrival of public information

---

[56]Source : 1988 NYSE Fact Book.

[57]Prices are adjusted for cash dividends, stock dividends, and stock splits.

That is, we are applying Glosten and Harris's approach using transformed data. Q_t is now hourly net volume with values of +1, 0, or -1, and P_t is now the hourly ending price. We can argue that this transformed data is less noisy than individual transactions, because dealers sometimes may not adjust quotes immediately and because there is evidence of a strong relationship between net trades and quote adjustment (see Chapter IV Section E). To be consistent, variance of net volume is also computed using the transformed data. Therefore, some information regarding var(V) will be lost than if the original data is used.

According to Glosten and Harris, variance of e_t reflects price variability induced by public information. This is fairly important for our test. The variance of e_t is served as the lower bound for var(i), since var(i) reflects the combination of public information and private information. If applying different values of E, the resulting var(i) and variance of e_t are close, more reasonable values of E_i, E_m, var(e) and var(i) can be simultaneously inferred. This is of course a rather imprecise approach. Except in presenting the estimation results for Z and C, we will not give statistical statements concerning the precision of the estimates of E, var(e), and var(i).

The period investigated is arbitrarily chosen to be the 18 days August 31 through September 24. The number of hourly price changes is then 107. The estimations of Z and C for each of the ten stocks are shown in Table 14. The t values for the estimated Z are significant at 1% for all stocks with the exception of Navistar International that is a relatively lower-priced stock (about $5, compared to other stocks, $35 - $160) and may involve a greater degree of measurement problem. This indicates that the information problem is important to the specialist. In contrast, C is significant at 1% only for Eastman Kodak and General Electric, significant at 10% for five other stocks. If Q_t is an unit trade, estimated Z and C are clearly related to the magnitude of spread for an unit trade. In our case, the interpretation of the estimated Z and C is unfortunately not so clear. The range of values for R-square is 0.09 to 0.34, and is 10.9 to 54.6 for F values.

Table 15 gives the values of var(e) and var(i) assuming different values of E_i and E_m. The variance of the error term (denoted as redv) for our previous regression is to be used as the lower bound of var(i). For EXXON, the variance of the error term seems relatively high compared to its price level, this may be due to a stock split (two for one) that goes into effect during the estimation period. In general, the values of var(i) and var(e) increase with increases in the values of E_i and E_m. The intuition is that when investors are less sensitive to prices, larger values of var(i) and var(e) are implied for a given value of var(V). The patterns in var(i) and var(e) are similar across stocks, which may be due to lack of variations in var(V) and redv across stocks.

The var(i) that are below redv can be ruled out as reasonable values. Using this rule of thumb, more reasonable values of var(i) and var(e) usually fall in the range of $E_i = 0.5$ ($E_m = 1.0$) to $E_i = 1.0$ ($E_m = 2.0$). The results suggest that, for our sample period and sample stocks and with our added assumptions, variance of true prices is lower than variance of noises. Although not based on information, uninformed trading still induces price changes because dealers are uncertain whether trades contain information or not.

V. E. Conclusions

The theory in the area of the components of the spread is not fully integrated. This essay presents a model that incorporates one important piece of information in trading : order imbalances (net volume). Although some simplifying assumptions have to be made, the concept this model introduces may be considered as a small contribution to our understanding regarding prices, volume, spread, and information.

Chapter VI

Conclusions

This dissertation is composed of three essays. Although each essay has a different focus, the central theme of the three essays lies in the relationship between the components of the spread and transaction prices. The first essay (Chapter III) and the second essay (Chapter IV) apply existing approaches to decompose the spread into ordering component, the inventory component, and the information component. On the other hand, the third essay (Chapter V) points to a direction for improving existing approaches by bringing in additional information about order imbalances.

The first essay employs a slight variant of Stoll's (1989) approach to decompose the spread. Then each component of the spread is related to factors that are important to each component. The results should provide a clearer understanding of the nature of costs faced by dealers, thereby extend the literature of the determinants of the spread. In addition, some considerations about the appropriate data are given. These considerations include the use of transaction data versus daily data, the appropriate estimation period, etc.

Problems remain in the decomposition of the spread though. In particular, the decomposition is sensitive to obtained values of the covariance of price changes. Currently, there is no rigorous theory for positive values of covariance of price changes, although we sometimes encounter such cases. Also, there are some substantial differences between covariances of price changes based on daily data and transaction data. We have largely ignored this issue since transaction prices are more compatible with the study of market microstructure. Nevertheless, the cause of this difference is potentially important. It may be related to, for example, noises in trades. In all, the fact is that our understanding of the nature of covariance of price changes is far from complete. Stoll's (1989) approach relies heavily on the covariance of price changes. Therefore, care must be taken in interpreting results obtained from using Stoll's approach. Glosten and Harris's (1988) approach utilizes similar information. Furthermore, their approach cannot fully decompose the spread, which limits its uses.

Another problem in the decomposition of the spread is the ignorance of monopoly rent, which suggests another component of the spread. A complete investigation of this issue would likely require more information regarding costs and revenues of specialists. The potential diversification effect among specialists within a specialist unit likely requires similar information.

Even without accounting for theoretical problems, there are some institutional factors that may affect the results, such as the simultaneous existence of the limit orders and market orders, the difference between trading procedures for the opening and the remaining of the day.

Nevertheless, the cross-sectional results should be valid unless measurement problems also vary across stocks. The first essay uses the price level to proxy and control for the discreteness problem. If future research indicates that other measurement problems affect stocks differently, then our empirical test should adjust accordingly.

The cross-sectional test in the first essay is more extensive than previous studies. The extension of the cross-sectional test requires better understanding of factors influencing the components of the spread and appropriate measures of these factors.

The second essay also offers partial relief to the measurement problem. Various measures of trading activity, price variability and transaction cost are given and the results are robust with respect to uses of different measures. Our main results say that trading activity, price variability and the information component are the highest in the first trading hour. On the other hand, the transitory component and the percentage changes in quoted prices are the lowest in the first hour. These results are consistent with the notion that liquidity traders time their trades to minimize transaction costs. This notion is based on Admati and Pfleiderer (1988), which presents a more restrictive model based on the minimization of information cost. The last hour increases in trading activity and price variability without significant changes in information component and the percentage changes in quoted changes are consistent with investors' unwinding position to avoid risk incurred overnight. Timing of trades is a relatively unexploratory area in finance. To the extent that timing of trades affects transaction prices, volume and transaction cost, it adds another dimension to the traditional analysis.

The test employed in the second essay gives a simple analysis of hourly means for variables that are of interest. It is likely that trading activity, price variability and the transaction cost are determined simultaneously. Therefore, possible interactions need to be considered in future research.

The differences in results between the pre-crash and the post-crash periods are interesting to say the least. If specialists' quoting behavior varies over time, the return generating process might well be different. This points to another potential explanation for nonstationarity in the return generating process.

The quote adjustment can take place either in the form of adjustment in the location of quoted prices or in the form of adjustment in the magnitude of the spread.

Essentially, the dealer's decision is multi-dimensional. Evidence provided in the second essay (negative relationship between changes in the magnitude of the spread and changes in quoted prices) suggests this may be the case. Implicitly, a constant magnitude of the spread suggests the degree of competition is so great and/or the monitoring function of the exchange is in great constraint to the dealer that the dealer has forgiven manipulating the magnitude of the spread. This is perhaps one of the most difficult problems in this area.

The problems encountered in the decomposition of the spread may be solved by bringing in additional information. This is the primary motivation of the third essay. The components of the spread reflects the degree of informational asymmetry between the dealer and informed traders, and the components of the net volume reflects the degree of informational asymmetry between the informed and the uninformed. Information about the components of the spread and information about the components of net volume thus should compliment each other.

Information about order imbalance is less ambiguous relative to volume or trade size. Volume reflects the combination of absolute purchase and absolute sale. Volume is therefore not compatible with traditional economic analysis. Trade size is another potentially useful piece of information, but different models give different interpretations of the role of trade size (Easely and O'Hara (1987) and Kyle (1985)). On the other hand, order imbalance is more readily interpretable (excess supply or demand). The intuition behind the third essay is that the variability of excess supply/demand gives information concerning the degree of informational asymmetry existed in financial markets.

There will be more to come in the future.

Table 1

Summary Statistics Describing the Sample Stocks
NYSE Transactions
(100 stocks)

		870918 Friday	870921 Monday	870922 Tuesday	870923 Wednesday	870924 Thursday
NT	mean	76.03	92.51	107.30	116.75	78.05
	median	52.50	71.50	92.00	89.50	60.00
	s. d.	58.52	67.47	78.50	87.53	62.38
	skewness	1.84	1.86	1.57	1.41	1.96
NQ	mean	54.82	64.78	75.91	78.53	57.04
	median	42.50	52.00	60.00	64.50	45.50
	s. d.	38.95	41.91	52.51	53.75	40.20
	skewness	2.15	1.54	1.46	1.69	1.79
S	mean	0.2549	0.2643	0.2779	0.2760	0.2789
	median	0.2381	0.2376	0.2589	0.2631	0.2582
	s. d.	0.1181	0.1345	0.1537	0.1391	0.1508
	skewness	4.75	5.19	6.21	4.86	5.41
vol	mean	200111	192778	225582	250059	188473
	median	130800	142800	163950	166800	111800
	s. d.	212811	171286	215885	233103	291065
	skewness	2.60	1.45	1.82	1.49	5.52
price	mean	45.40	45.16	44.78	45.62	45.73
	median	35.58	35.55	35.73	36.29	35.94
	s. d.	(47.15)	(46.81)	(46.30)	(46.85)	(46.86)
	skewness	5.88	5.84	5.85	5.82	5.79

NT = number of trades
NQ = number of quotes
S = average quoted spread (in dollars)
vol = volume of trading (in number of shares)
price = average stock price (in dollars)

Table 2

Components of the Bid-Ask Spread
NYSE Transactions
(100 stocks)

		870918 Friday	870921 Monday	870922 Tuesday	870923 Wednesday	870924 Thursday
δ	mean	0.4257	0.4248	0.4229	0.4177	0.4376
	s. d.	(0.0902)	(0.1199)	(0.0789)	(0.0755)	(0.1067)
π	mean	0.5096	0.5036	0.4914	0.5084	0.5217
	s. d.	(0.0514)	(0.0960)	(0.0748)	(0.0462)	(0.0840)
C	mean	0.1486	0.1505	0.1541	0.1646	0.1248
	s. d.	(0.1804)	(0.2397)	(0.1578)	(0.1510)	(0.2135)
H	mean	0.0192*	0.0071#	-0.0171#	0.0168*	0.0435
	s. d.	(0.1027)	(0.1920)	(0.1496)	(0.0925)	(0.1680)
Z	mean	0.8322	0.8424	0.8630	0.8186	0.8317
	s. d.	(0.1729)	(0.1427)	(0.1752)	(0.1647)	(0.1742)
S	mean	0.0080	0.0083	0.0086	0.0085	0.0087
	s. d.	(0.0059)	(0.0058)	(0.0061)	(0.0066)	(0.0074)
covP	mean	-0.0020	-0.0022	-0.0012*	-0.0026	-0.0025
	s. d.	(0.0025)	(0.0019)	(0.0067)	(0.0029)	(0.0048)
covQ	mean	-0.0003	-0.0013#	0.0009#	+0.0000#	-0.0044*
	s. d.	(0.0014)	(0.0168)	(0.0088)	(0.0026)	(0.0319)

: insignificantly different from zero by both the t test and the sign rank test (at 5% level)
* : insignificantly different from zero by the t test
$1-\delta$ = magnitude of price reversal (proportion of spread)
π = probability of reversal
C = order cost (proportion of spread)
H = inventory cost (proportion of spread)
Z = information cost (proportion of spread)
S = quoted spread (proportion of price)
covP = covariance of transaction price changes
covQ = covariance of quoted price changes

Table 3

Components of the Bid-Ask Spread
All Transactions
(100 stocks)

		870918 Friday	870921 Monday	870922 Tuesday	870923 Wednesday	870924 Thursday
δ	mean	0.5275	0.5275	0.5629	0.5130	0.5292
	s. d.	(0.0919)	(0.1067)	(0.3338)	(0.1363)	(0.1013)
π	mean	0.6085	0.6129	0.6448	0.6143	0.6180
	s. d.	(0.0742)	(0.0836)	(0.3205)	(0.0767)	(0.0809)
C	mean	-0.0550	-0.0549	-0.1257	-0.0259	-0.0583
	s. d.	(0.1837)	(0.2135)	(0.6675)	(0.2725)	(0.2026)
H	mean	0.2169	0.2258	0.2897	0.2285	0.2359
	s. d.	(0.1483)	(0.1672)	(0.6411)	(0.1533)	(0.1618)
Z	mean	0.8381	0.8291	0.8361	0.7974	0.8224
	s. d.	(0.1293)	(0.1176)	(0.1413)	(0.2353)	(0.1397)
S	mean	0.0098	0.0103	0.0108	0.0104	0.0106
	s. d.	(0.0062)	(0.0063)	(0.0069)	(0.0072)	(0.0079)
covP	mean	-0.0032	-0.0049	-0.0035	-0.0051	-0.0041
	s. d.	(0.0026)	(0.0059)	(0.0056)	(0.0095)	(0.0065)
covQ	mean	-0.0093	-0.0120	-0.2440	-0.0103	-0.0145
	s. d.	(0.0133)	(0.0251)	(2.1397)	(0.0132)	(0.0289)

$1-\delta$ = magnitude of price reversal (proportion of spread)
π = probability of reversal
C = order cost (proportion of spread)
H = inventory cost (proportion of spread)
Z = information cost (proportion of spread)
S = quoted spread (as proportion of price)
covP = covariance of transaction price changes
covQ = covariance of quoted price changes

Table 4
Panel A
Comparison of Option versus Non-Option Stocks on September 23(NYSE Transactions)

		Option (66 stocks)	Non-Option (34 stocks)
δ	mean	0.4211	0.4110
	s. d.	(0.0567)	(0.1034)
π	mean	0.5007	0.5234*
	s. d.	(0.0461)	(0.0433)
C	mean	0.1578	0.1780
	s. d.	(0.1134)	(0.2069)
H	mean	0.0013	0.0467*
	s. d.	(0.0922)	(0.0866)
Z	mean	0.8409	0.7753
	s. d.	(0.1318)	(0.2103)
S	mean	0.2734	0.2808
	s. d.	(0.1452)	(0.1283)
vol	mean	332268	90476*
	s. d.	(243758)	(78871)
size	mean	2294	1654*
	s. d.	(1105)	(1180)
price	mean	52.20	32.86*
	s. d.	(54.80)	(20.41)
covP	mean	-0.0024	-0.0029
	s. d.	(0.0021)	(0.0040)
covQ	mean	0.0003	-0.0006*
	s. d.	(0.0027)	(0.0021)

* indicates different between two groups at 10% level
$1-\delta$ = magnitude of price reversal (proportion of spread)
π = probability of reversal
C = order cost (proportion of spread)
H = inventory cost (proportion of spread)
Z = information cost (proportion of spread)
S = average quoted spread (in dollars)
vol = volume of trading (in number of shares)
size = average size of trade (in number of shares)
price = average stock price (in dollars)
covP = covariance of transaction price changes
covQ = covariance of quoted price changes

Table 4 (continued)
Panel B
Comparison of Option versus Non-Option Stocks
on September 24(NYSE Transactions)

		Option (66 stocks)	Non-Option (34 stocks)
δ	mean	0.4466	0.4201
	s. d.	(0.1097)	(0.0999)
π	mean	0.5227	0.5199
	s. d.	(0.0890)	(0.0747)
C	mean	0.1068	0.1597
	s. d.	(0.2195)	(0.1999)
H	mean	0.0453	0.0398
	s. d.	(0.1779)	(0.1495)
Z	mean	0.8479	0.8004
	s. d.	(0.1401)	(0.2255)
S	mean	0.2796	0.2775
	s. d.	(0.1714)	(0.1018)
vol	mean	274153	74565*
	s. d.	(341435)	(65416)
size	mean	2524	1847
	s. d.	(2320)	(1776)
price	mean	52.35	32.89*
	s. d.	(54.84)	(20.16)
covP	mean	-0.0018	-0.0040*
	s. d.	(0.0035)	(0.0063)
covQ	mean	-0.0059	-0.0016
	s. d.	(0.0390)	(0.0071)

* indicates different between two groups at 10%
$1-\delta$ = magnitude of price reversal (proportion of spread)
π = probability of reversal
C = order cost (proportion of spread)
H = inventory cost (proportion of spread)
Z = information cost (proportion of spread)
S = average quoted spread (in dollars)
vol = volume of trading (in number of shares)
size = average size of trade (in number of shares)
price = average stock price (in dollars)
covP = covariance of transaction price changes
covQ = covariance of quoted price changes

Table 5
Panel A
Comparison of Different Volume Subgroups
on September 23 (NYSE Transactions)

		Low (21 stocks)	Median (47 stocks)	High (32 stocks)
δ	mean	0.3945	0.4281	0.4176
	s. d.	(0.1171)	(0.0634)	(0.0530)
π	mean	0.5049	0.5082	0.5111
	s. d.	(0.0673)	(0.0427)	(0.0341)
C	mean	0.2111	0.1437	0.1648
	s. d.	(0.2342)	(0.1268)	(0.1061)
H	mean	0.0097	0.0163	0.0222
	s. d.	(0.1346)	(0.0854)	(0.0683)
Z	mean	0.7792	0.8400	0.8130
	s. d.	(0.2452)	(0.1395)	(0.1310)
S	mean	0.3454	0.2720	0.2362
	s. d.	(0.2659)	(0.0765)	(0.0506)
vol	mean	39857	153911	529222
	s. d.	(17900)	(62200)	(207815)
size	mean	1041	1961	2925
	s. d.	(558)	(866)	(1239)
price	mean	54.56	42.07	44.97
	s. d.	(86.79)	(32.54)	(22.48)
covP	mean	-0.0032	-0.0024	-0.0023
	s. d.	(0.0050)	(0.0022)	(0.0016)
covQ	mean	0.0006	-0.0002	-0.0001
	s. d.	(0.0051)	(0.0015)	(0.0009)

$1-\delta$ = magnitude of price reversal (proportion of spread)
π = probability of reversal
C = order cost (proportion of spread)
H = inventory cost (proportion of spread)
Z = information cost (proportion of spread)
S = average quoted spread(in dollars)
vol = volume of trading (in number of shares)
size = average size of trade (in number of shares)
price = average stock price (in dollars)
covP = covariance of transaction price changes
covQ = covariance of quoted price changes

Table 5(continued)
Panel B
Comparison of Different Volume Subgroups
on September 24 (NYSE Transactions)

		Low (37 stocks)	Median (51 stocks)	High (12 stocks)
δ	mean	0.4262	0.4494	0.4228
	s. d.	(0.0957)	(0.1225)	(0.0552)
π	mean	0.5171	0.5287	0.5062
	s. d.	(0.0667)	(0.0982)	(0.0674)
C	mean	0.1476	0.1013	0.1544
	s. d.	(0.1914)	(0.2450)	(0.1104)
H	mean	0.0342	0.0574	0.0124
	s. d.	(0.1334)	(0.1965)	(0.1348)
Z	mean	0.8182	0.8413	0.8332
	s. d.	(0.2266)	(0.1403)	(0.1200)
S	mean	0.3085	0.2693	0.2248
	s. d.	(0.2237)	(0.0857)	(0.0377)
vol	mean	43116	165578	733958
	s. d.	(22655)	(58968)	(588962)
size	mean	1159	2513	4859
	s. d.	(631)	(1441)	(4444)
price	mean	48.52	44.49	42.40
	s. d.	(72.05)	(22.89)	(18.67)
covP	mean	-0.0023	-0.0027	-0.0022
	s. d.	(0.0067)	(0.0035)	(0.0019)
covQ	mean	-0.0017	-0.0074	-0.0002
	s. d.	(0.0074)	(0.0442)	(0.0012)

$1-\delta$ = magnitude of price reversal (proportion of spread)
π = probability of reversal
C = order cost (proportion of spread)
H = inventory cost (proportion of spread)
Z = information cost (proportion of spread)
S = average quoted spread (in dollars)
vol = volume of trading (in number of shares)
size = average size of trade (in number of shares)
price = average stock price (in dollars)
covP = covariance of transaction price changes
covQ = covariance of quoted price changes

Table 6

Correlation Coefficients for Independent Variables
in Regressions (A1)-(A3)

	AZ	TR	UR	OP	IN	IT	FZ	P	VOL
NT	0.085 (0.413)	-0.073 (0.490)	-0.075 (0.479)	0.458 (0.001)	0.245 (0.017)	-0.203 (0.049)	0.624 (0.000)	0.061 (0.560)	0.854 (0.000)
AZ		-0.063 (0.553)	-0.071 (0.501)	0.299 (0.004)	0.229 (0.026)	-0.221 (0.032)	0.140 (0.177)	-0.210 (0.043)	0.495 (0.000)
TR			0.997 (0.000)	-0.254 (0.015)	-0.251 (0.016)	0.196 (0.063)	-0.236 (0.025)	-0.212 (0.044)	-0.094 (0.375)
UR				-0.260 (0.013)	-0.267 (0.010)	0.200 (0.058)	-0.235 (0.025)	-0.210 (0.046)	-0.098 (0.354)
OP					0.363 (0.001)	-0.393 (0.001)	0.405 (0.001)	0.173 (0.095)	0.465 (0.001)
IN						-0.492 (0.000)	0.190 (0.066)	0.324 (0.002)	0.288 (0.005)
IT							-0.230 (0.026)	-0.052 (0.617)	-0.247 (0.016)
FZ								0.317 (0.002)	0.589 (0.000)
P									-0.041 (0.694)

**Numbers in parentheses are the p values for H_o: coefficient =0

*Note: sample size = 91

NT=number of trades
AZ=transaction size (in hundreds of shares)
TR=return variance
UR=unsystematic risk
OP=1, if options traded, =0 otherwise
IN=institutional holding (%)
IT=insider holding (%)
FZ=firm size (in millions)
P=average transaction price (in dollars)
VOL=volume (in hundreds of shares)

Table 7

Panel A

OLS Estimation of Equation (A1)

$$C = \alpha_1 + \alpha_2 NT + \alpha_3 AZ + \alpha_4 IN + \alpha_5 P + \alpha_6 VOL$$

	(A1.1)	(A1.2)	(A1.3)
α_1	0.1576 (0.0001)	0.1709 (0.0001)	0.2169 (0.0001)
α_2	0.0003 (0.0509)	0.0003 (0.0507)	-0.0003 (0.4893)
α_3	0.0008 (0.4396)	0.0002 (0.8541)	-0.0018 (0.2825)
α_4	-0.1266 (0.0201)	-0.0880 (0.1274)	-0.0845 (0.1410)
α_5		-0.0004 (0.0823)	-0.0004 (0.0751)
α_6			$2.6*10^{-5}$ (0.1394)
F	2.691	2.840	2.750
R^2	0.0849	0.1167	0.1392
adj. R^2	0.0534	0.0756	0.0886
cn	50.5	58.3	344.6
N	91	91	91

Numbers in parentheses are p values for H_o : coefficient=0
NT=number of trades
AZ=transaction size (in hundreds of shares)
IN=institutional holding (%)
P=average transaction price (in dollars)
VOL=volume (in hundreds of shares)
cn=maximum eigenvalue/minimum eigenvalue
N=sample size

Table 7 (continued)

Panel B

OLS Estimation of Equation (A2)

$$H = \beta_1 + \beta_2 \, TR + \beta_3 \, OP + \beta_4 \, IN + \beta_5 \, P + \beta_6 \, VOL$$

	(A2.1)	(A2.2)	(A2.3)
β_1	0.0549 (0.0183)	0.0582 (0.0131)	0.0589 (0.0126)
β_2	8.6226 (0.4010)	6.9927 (0.4985)	7.1101 (0.4935)
β_3	0.0081 (0.6337)	0.0088 (0.6024)	0.0123 (0.5105)
β_4	-0.0976 (0.0115)	-0.0858 (0.0307)	-0.0824 (0.0419)
β_5		-0.0002 (0.2436)	-0.0002 (0.2212)
β_6			$-1.9*10^{-6}$ (0.6530)
F	2.963	2.576	2.083
R^2	0.0927	0.1070	0.1091
adj. R^2	0.0614	0.0655	0.0567
cn	50.4	57.8	68.8
N	91	91	91

Numbers in parentheses are p values for H_0 : coefficient=0
TR=return variance
OP=1, if options traded, =0 otherwise
IN=institutional holding(%)
P=average transaction price (in dollars)
VOL=volume (in hundreds of shares)
cn=maximum eigenvalue/minimum eigenvalue
N=sample size

Table 7 (continued)

Panel C

OLS Estimation of Equation (A3)

$$Z = \theta_1 + \theta_2 RZ + \theta_3 UR + \theta_4 IN + \theta_5 FZ + \theta_6 IT + \theta_7 P + \theta_8 VOL$$

	(A3.1)	(A3.2)	(A3.3)
θ_1	0.6085 (0.0001)	0.6182 (0.0001)	0.6293 (0.0001)
θ_2	4.0641 (0.0044)	3.8165 (0.0060)	3.4342 (0.0284)
θ_3	-1.3473 (0.9303)	1.9358 (0.8977)	2.0314 (0.8932)
θ_4	0.2787 (0.0001)	0.2309 (0.0005)	0.2386 (0.0005)
θ_5	$5.0*10^{-6}$ (0.1470)	$2.7*10^{-6}$ (0.4343)	$3.7*10^{-6}$ (0.3499)
θ_6	0.1825 (0.0218)	0.1454 (0.0643)	0.1438 (0.0685)
θ_7		0.0006 (0.0186)	0.0005 (0.0499)
θ_8			$-4.4*10^{-6}$ (0.5974)
F	5.489	5.790	4.960
R^2	0.2441	0.2926	0.2950
adj. R^2	0.1996	0.2420	0.2355
cn	116.4	138.8	171.9
N	91	91	91

Numbers in parentheses are p values for H_o : coefficient=0
RZ=relative transaction size (=AZ/VOL)
UR=unsystematic risk
IN=institutional holding(%)
FZ=firm size (in millions)
IT=insider holding(%)
P=average transaction price (in dollars)
VOL=volume (in hundreds of shares)
cn=maximum eigenvalue/minimum eigenvalue
N=sample size

Table 8

Hourly Statistics of Trading Activity, Price Changes and Transaction Cost for the Pre-Crash High-Activity Sample

Panel A

High-Activity Sample Hourly Summary Statistics

Pre-Crash Period

Numbers are Means and Standard Deviation (in parentheses) across days and stocks

Hour	1	2	3	4	5	6
NT	27.11	22.31	18.70	16.34	18.55	24.71
	(22.77)	(19.11)	(15.57)	(13.87)	(14.86)	(19.41)
VOL	761.24	492.51	402.63	336.46	344.23	485.85
	(1239.7)	(784.15)	(567.40)	(505.39)	(452.89)	(808.31)
RETV	1.3E-5	1.1E-5	1.1E-5	1.1E-5	1.1E-5	1.2E-5
	(3.7E-5)	(2.6E-5)	(2.4E-5)	(4.5E-5)	(2.4E-5)	(2.7E-5)
RET	0.0072	0.0054	0.0047	0.0044	0.0048	0.0067
	(0.0072)	(0.0062)	(0.0052)	(0.0049)	(0.0053)	(0.0062)
Z1	0.8475	0.8001	0.7986	0.8079	0.7935	0.8015
	(0.2291)	(0.3003)	(0.2804)	(0.2836)	(0.2898)	(0.2590)
Z2	0.1144	0.1006	1.1495	0.5707	1.2889	0.4368
	(5.4458)	(2.3930)	(25.690)	(16.388)	(27.630)	(14.147)
C1	0.1525	0.1999	0.2014	0.1921	0.2065	0.1985
	(0.2291)	(0.3003)	(0.2804)	(0.2836)	(0.2898)	(0.2590)
C2	0.5810	1.0120	1.8506	2.3379	1.2027	1.0163
	(3.2363)	(15.700)	(24.189)	(27.224)	(14.145)	(12.247)
AVGS	0.0077	0.0070	0.0069	0.0069	0.0068	0.0073
	(0.0055)	(0.0042)	(0.0049)	(0.0040)	(0.0040)	(0.0043)
BLOCK	0.0580	0.0442	0.0459	0.0433	0.0360	0.0378
	(0.0721)	(0.0702)	(0.0763)	(0.0746)	(0.0642)	(0.0650)
sample size	4363	4071	3821	3646	3824	4061

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (as proportion of quoted spread)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 8 (continued)
Panel B

High-Activity Sample ANOVA Results

Pre-Crash Period

Dependent Variables	F	F_h	$F_{h/w}$	$F_{w/h}$	R^2	n
NT	139.99 (224.58)	210.66 (335.37)	212.36 (338.08)	51.67 (86.09)	0.0503 (0.0784)	23786
VOL	103.85 (185.86)	165.94 (303.03)	166.91 (305.02)	26.22 (39.40)	0.0378 (0.0657)	23786
RETV	4.49 (29.42)	6.57 (47.02)	6.59 (47.36)	1.90+ (7.42)	0.0017 (0.0110)	23786
RET	95.38 (86.11)	140.00 (126.48)	140.93 (127.50)	39.60 (35.65)	0.0348 (0.0316)	23786
Z1	15.06 (12.52)	22.32 (21.13)	22.48 (21.19)	5.98 (1.76)+	0.0057 (0.0047)	23786
Z2	1.94+ (15.05)	3.27 (18.82)	3.27 (18.97)	0.29+ (10.32)	0.0007 (0.0057)	23786
C2	3.86 (9.51)	5.17 (16.22)	5.20 (16.25)	2.21 (1.13)+	0.0015 (0.0036)	23786
AVGS	15.93 (22.10)	25.81 (32.09)	25.92 (32.26)	3.57 (9.62)	0.0060 (0.0083)	23786
BLOCK	31.08 (63.59)	51.09 (110.00)	50.90 (110.04)	6.07 (5.57)	0.0116 (0.0235)	23786

Independent variables = hour of the day and day of the week

F = overall F value of the ANOVA
F_h = F value when only the hour effect is included
$F_{h/w}$ = partial F value of the hour effect
$F_{w/h}$ = partial F value of the day effect

+ : insignificant at the 1% level

Numbers in parentheses are the corresponding statistics using rank transformed data

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (not included as the result is the same as Z1)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 8 (continued)
Panel C

High-Activity Sample Hourly Mean Comparisons
Pre-Crash Period

Signs of row minus column means : +++ significant at 1% using both original and rank data, ++ significant only with ranks, + insignificant or significant only when using original data (where contradiction occurs, sign for ranks is used); same signing convention is used for negative values.

HOUR		2	3	4	5	6	
1	NT	+++	+++	+++	+++	+++	NT
	VOL	+++	+++	+++	+++	+++	VOL
	RETV	+++	+++	+++	+++	+	RETV
	RET	+++	+++	+++	+++	+++	RET
	Z1	+++	+++	+++	+++	+++	Z1
	Z2	++	++	++	++	++	Z2
	C1	---	---	---	---	---	C1
	C2	--	--	--	--	--	C2
	AVGS	+++	+++	+++	+++	+++	AVGS
	BLOCK	+++	+++	+++	+++	+++	BLOCK
2	NT		+++	+++	+++	---	NT
	VOL		+++	+++	+++	+	VOL
	RETV		+	++	+	--	RETV
	RET		+++	+++	+++	---	RET
	Z1		+	-	+	+	Z1
	Z2		+	+	-	-	Z2
	C1		-	+	-	-	C1
	C2		-	+	-	-	C2
	AVGS		+	+	+	-	AVGS
	BLOCK		+	++	+++	+	BLOCK
3	NT			+++	-	---	NT
	VOL			+++	++	---	VOL
	RETV			+	-	--	RETV
	RET			+	+	---	RET
	Z1			-	+	+	Z1
	Z2			+	--	--	Z2
	C1			+	-	-	C1
	C2			+	-	-	C2
	AVGS			+	+	--	AVGS
	BLOCK			+	+++	+	BLOCK
4	NT				---	---	NT
	VOL				-	---	VOL
	RETV				--	--	RETV
	RET				-	---	RET
	Z1				+	++	Z1
	Z2				--	--	Z2
	C1				-	--	C1
	C2				-	--	C2
	AVGS				+	---	AVGS
	BLOCK				+	-	BLOCK
5	NT					---	NT
	VOL					---	VOL
	RETV					--	RETV
	RET					---	RET
	Z1					+	Z1
	Z2					+	Z2
	C1					-	C1
	C2					+	C2
	AVGS					---	AVGS
	BLOCK					--	BLOCK

Table 8 (continued)
Panel D

High-Activity Sample Regression Result
Pre-Crash Period

	Dependent Variable	
	△Q	FREQ
Independent Variables		
intercept	-2.6E-6	0.9073
	(-0.045)	(131.213)*
H1	-0.00021	0.1288
	(-2.544)*	(14.698)*
H6	8.8E-5	-0.0342
	(1.040)	(-3.842)*
W1	-6.4E-5	0.0066
	(-0.774)	(0.764)
W5	2.9E-5	0.0023
	(0.341)	(0.265)
NETT	9.4E-5	-0.0142
	(18.417)*	(-17.244)*
NT	-2.8E-6	-0.0018
	(-1.609)	(-8.109)*
△S	2.6E-5	0.0358
	(0.500)	(5.272)*
RET	0.7863	2.1482
	(201.262)*	(3.945)*
RETV	0.0478	-621.799
	(0.049)	(-5.932)*
F	5376.93	114.878
R^2	0.6705	0.0417
adj. R^2	0.6704	0.0413
n	23786	23786

Number in parentheses are the t values for the null hypothesis that corresponding coefficient = 0,
* indicates significant at the 1% level

Dependent variables :
△Q = percentage changes in quoted prices over an hourly interval
FREQ = number of quotes / number of trades

Independent variables : (Note : NETT, △S and RET are in their absolute values in FREQ regression)
H1 : an indicator = 1 if hour 1, = 0 otherwise
H6 : an indicator = 1 if hour 6, = 0 otherwise
W1 : an indicator = 1 if Monday, = 0 otherwise
W5 : an indicator = 1 if Friday, = 0 otherwise
NETT : number of purchase minus number of sale over an hourly interval
NT : hourly number of trades
△S : percentage changes in the magnitude of the spread
RET : hourly return
RETV : (trade-by-trade) return variance

Table 9

Hourly Statistics of Trading Activity, Price Changes and Transaction Cost for the Post-Crash High-Activity Sample

Panel A

High-Activity Sample Hourly Summary Statistics

Post-Crash Period

Numbers are Means and Standard Deviation (in parentheses) across days and stocks

Hour	1	2	3	4	5	6
NT	39.80 (33.05)	43.14 (33.16)	35.24 (29.49)	35.86 (29.36)	41.00 (32.34)	46.49 (35.96)
VOL	1638.50 (1880.9)	1313.10 (1629.0)	929.69 (1294.0)	972.59 (1461.5)	1056.50 (1214.8)	1198.00 (1503.4)
RETV	0.0601 (2.1395)	5.9E-5 (1.4E-4)	6.1E-5 (1.5E-4)	5.8E-5 (1.5E-4)	5.0E-5 (9.7E-5)	5.9E-5 (1.4E-4)
RET	0.0279 (0.0319)	0.0322 (0.0398)	0.0178 (0.0238)	0.0179 (0.0207)	0.0232 (0.0272)	0.0348 (0.0407)
Z1	0.8997 (0.2382)	0.8879 (0.2815)	0.8487 (0.2548)	0.8564 (0.2971)	0.8903 (0.2625)	0.8837 (0.1969)
Z2	0.5249 (18.381)	0.0064 (0.0112)	0.0055 (0.0163)	0.0072 (0.0234)	0.1571 (3.8006)	0.0076 (0.0132)
C1	0.1003 (0.2382)	0.1121 (0.2815)	0.1513 (0.2548)	0.1436 (0.2971)	0.1097 (0.2625)	0.1163 (0.1969)
C2	0.6171 (5.3939)	0.5881 (4.4438)	0.4986 (0.2138)	1.2481 (27.322)	0.4855 (0.2440)	0.4650 (0.1862)
AVGS	0.0157 (0.0111)	0.0155 (0.0107)	0.0153 (0.0115)	0.0151 (0.0111)	0.0149 (0.0104)	0.0154 (0.0105)
BLOCK	0.0898 (0.0778)	0.0671 (0.0728)	0.0600 (0.0745)	0.0590 (0.0729)	0.0544 (0.0635)	0.0538 (0.0579)
sample size	1271	1313	1292	1295	647	645

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (as proportion of quoted spread)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 9 (continued)
Panel B

High-Activity Sample ANOVA Results

Post-Crash Period

Dependent Variables	F	F_h	$F_{h/w}$	$F_{w/h}$	R^2	n
NT	21.86 (41.61)	18.18 (35.82)	18.14 (35.77)	26.46 (48.84)	0.0296 (0.0548)	6463
VOL	30.27 (55.28)	37.23 (67.63)	37.59 (68.41)	21.58 (39.83)	0.0405 (0.0716)	6463
RETV	0.88+ (49.66)	0.82+ (1.75)+	0.82+ (1.84)+	0.97+ (109.55)	0.0012 (0.0648)	6463
RET	109.92 (90.76)	63.02 (62.32)	63.03 (62.67)	168.55 (126.30)	0.1329 (0.1124)	6463
Z1	6.26 (7.89)	7.39 (9.43)	7.43 (9.50)	4.86 (5.98)	0.0087 (0.0109)	6463
Z2	0.87+ (5.62)	0.79+ (9.80)	0.79+ (9.79)	0.98+ (0.40)+	0.0012 (0.0078)	6463
C2	0.91+ (9.79)	0.67+ (9.66)	0.67+ (9.76)	1.21+ (9.95)	0.0013 (0.0135)	6463
AVGS	35.83 (56.81)	0.75+ (2.22)+	0.80+ (2.35)+	79.68 (125.04)	0.0476 (0.0734)	6463
BLOCK	25.74 (35.53)	39.40 (57.41)	39.06 (57.81)	8.66 (8.18)	0.0347 (0.0472)	6463

Independent variables = hour of the day and day of the week

F = overall F value of the ANOVA
F_h = F value when only the hour effect is included
$F_{h/w}$ = partial F value of the hour effect
$F_{w/h}$ = partial F value of the day effect

+ : insignificant at the 1% level

Numbers in parentheses are the corresponding statistics using rank transformed data

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (not included as the result is the same as Z1)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 9 (continued)
Panel C

High-Activity Sample Hourly Mean Comparisons
Post-Crash Period

Signs of row minus column means : +++ significant at 1% using both original and rank data, ++ significant only with ranks, + insignificant or significant only when using original data (where contradiction occurs, sign for ranks is used); same signing convention is used for negative values.

HOUR		2	3	4	5	6	
1	NT	---	+++	++	-	---	NT
	VOL	+++	+++	+++	+++	+++	VOL
	RETV	+	+	+	+	+	RETV
	RET	-	+++	+++	+	---	RET
	Z1	+	+++	+++	+	+	Z1
	Z2	+	++	+	+	-	Z2
	C1	-	---	---	-	-	C1
	C2	-	--	--	-	-	C2
	AVGS	+	+	+	+	+	AVGS
	BLOCK	+++	+++	+++	+++	+++	BLOCK
2	NT		+++	+++	-	-	NT
	VOL		+++	+++	+++	+	VOL
	RETV		+	+	-	-	RETV
	RET		+++	+++	+++	-	RET
	Z1		+++	+	-	+	Z1
	Z2		++	+	-	-	Z2
	C1		---	-	+	-	C1
	C2		--	--	-	-	C2
	AVGS		+	+	+	-	AVGS
	BLOCK		++	++	+++	+++	BLOCK
3	NT			-	---	---	NT
	VOL			-	--	---	VOL
	RETV			+	-	-	RETV
	RET			-	---	---	RET
	Z1			-	--	-	Z1
	Z2			--	--	--	Z2
	C1			+	++	+	C1
	C2			+	+	+	C2
	AVGS			+	-	-	AVGS
	BLOCK			+	+	+	BLOCK
4	NT				---	---	NT
	VOL				-	--	VOL
	RETV				-	-	RETV
	RET				---	---	RET
	Z1				-	-	Z1
	Z2				-	-	Z2
	C1				+	+	C1
	C2				+	+	C2
	AVGS				-	-	AVGS
	BLOCK				+	+	BLOCK
5	NT					--	NT
	VOL					-	VOL
	RETV					+	RETV
	RET					---	RET
	Z1					+	Z1
	Z2					-	Z2
	C1					-	C1
	C2					+	C2
	AVGS					-	AVGS
	BLOCK					-	BLOCK

Table 9 (continued)
Panel D

High-Activity Sample Regression Result
Post-Crash Period

Independent Variables	Dependent Variable: △Q	Dependent Variable: FREQ
intercept	0.00015	0.6207
	(0.541)	(98.086)*
H1	8.1E-5	0.0514
	(0.218)	(6.567)*
H6	-2.1E-5	-0.0048
	(-0.043)	(-0.455)
W1	0.00080	-0.0009
	(2.022)	(-0.115)
W5	0.00098	0.0686
	(2.585)*	(8.604)*
NETT	4.9E-5	-0.0051
	(3.582)*	(-11.914)*
NT	-1.7E-5	-0.0013
	(-3.618)*	(-10.272)*
△S	-0.00027	0.0122
	(-1.371)	(2.492)*
RET	0.9620	0.8634
	(247.627)*	(8.564)*
RETV	-1.1E-6	0.0017
	(-0.007)	(0.537)
F	8022.94	88.249
R^2	0.9180	0.1096
adj. R^2	0.9179	0.1084
n	6463	6463

Number in parentheses are the t values for the null hypothesis that corresponding coefficient = 0,
* indicates significant at the 1% level

Dependent variables :
△Q = percentage changes in quoted prices over an hourly interval
FREQ = number of quotes / number of trades

Independent variables : (Note : NETT, △S and RET are in their absolute values in FREQ regression)
H1 : an indicator = 1 if hour 1, = 0 otherwise
H6 : an indicator = 1 if hour 6, = 0 otherwise
W1 : an indicator = 1 if Monday, = 0 otherwise
W5 : an indicator = 1 if Friday, = 0 otherwise
NETT : number of purchase minus number of sale over an hourly interval
NT : hourly number of trades
△S : percentage changes in the magnitude of the spread
RET : hourly return
RETV : (trade-by-trade) return variance

Table 10

Hourly Statistics of Trading Activity, Price Changes and Transaction Cost for the Pre-Crash Medium-Activity Sample

Panel A

Medium-Activity Sample Hourly Summary Statistics

Pre-Crash Period

Numbers are Means and Standard Deviation (in parentheses) across days and stocks

Hour	1	2	3	4	5	6
NT	8.02	7.38	6.79	6.35	6.66	8.06
	(5.48)	(4.71)	(4.07)	(3.95)	(3.88)	(5.07)
VOL	144.66	130.18	118.22	113.57	108.48	126.22
	(249.20)	(260.90)	(229.17)	(242.02)	(194.86)	(219.75)
RETV	4.4E-5	3.9E-5	4.0E-5	4.4E-5	3.4E-5	3.7E-5
	(1.6E-4)	(1.4E-4)	(1.4E-4)	(2.2E-4)	(1.1E-4)	(1.2E-4)
RET	0.0091	0.0075	0.0074	0.0070	0.0068	0.0077
	(0.0120)	(0.0096)	(0.0101)	(0.0103)	(0.0095)	(0.0098)
Z1	0.8246	0.7956	0.7945	0.7968	0.7861	0.7918
	(0.3589)	(0.3676)	(0.3526)	(0.3514)	(0.3571)	(0.3370)
Z2	8.0479	8.3509	7.3269	9.6486	9.3004	8.5494
	(59.546)	(74.390)	(55.010)	(63.916)	(64.423)	(70.593)
C1	0.1754	0.2044	0.2055	0.2032	0.2139	0.2082
	(0.3589)	(0.3676)	(0.3526)	(0.3514)	(0.3571)	(0.3370)
C2	5.4236	6.6459	7.3644	7.3385	8.1850	6.9190
	(46.293)	(51.889)	(62.114)	(53.200)	(61.088)	(53.129)
AVGS	0.0138	0.0129	0.0130	0.0129	0.0126	0.0127
	(0.0091)	(0.0087)	(0.0089)	(0.0098)	(0.0084)	(0.0080)
BLOCK	0.0350	0.0346	0.0366	0.0333	0.0330	0.0298
	(0.0864)	(0.0883)	(0.0942)	(0.0904)	(0.0894)	(0.0808)
sample size	5495	3913	3162	2527	2954	3654

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (as proportion of quoted spread)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 10 (continued)
Panel B

Medium-Activity Sample ANOVA Results

Pre-Crash Period

Dependent Variables	F	F_h	$F_{h/w}$	$F_{w/h}$	R^2	n
NT	53.51 (56.21)	82.86 (87.61)	82.59 (87.17)	16.83 (16.97)	0.0217 (0.0228)	21705
VOL	8.71 (25.28)	12.43 (41.48)	12.54 (41.70)	4.05 (5.04)	0.0036 (0.0104)	21705
RETV	1.91+ (13.07)	2.38+ (22.96)	2.34+ (22.99)	1.31+ (0.71)+	0.0008 (0.0054)	21705
RET	15.48 (20.35)	25.74 (33.22)	25.69 (33.09)	2.65+ (4.27)	0.0064 (0.0084)	21705
Z1	4.74 (4.95)	6.85 (6.77)	6.89 (6.83)	2.10+ (2.67)+	0.0020 (0.0020)	21705
Z2	0.53+ (15.46)	0.47+ (26.29)	0.47+ (26.27)	0.60+ (1.93)+	0.0002 (0.0068)	21705
C2	0.80+ (3.72)	1.22+ (6.31)	1.21+ (6.31)	0.29+ (0.48)+	0.0003 (0.0016)	21705
AVGS	6.43 (11.51)	10.85 (19.95)	10.81 (19.92)	0.91+ (0.97)+	0.0027 (0.0048)	21705
BLOCK	3.24 (4.01)	2.49+ (5.26)	2.43+ (5.29)	4.18 (2.43)+	0.0013 (0.0017)	21705

Independent variables = hour of the day and day of the week

F = overall F value of the ANOVA
F_h = F value when only the hour effect is included
$F_{h/w}$ = partial F value of the hour effect
$F_{w/h}$ = partial F value of the day effect

+ : insignificant at the 1% level

Numbers in parentheses are the corresponding statistics using rank transformed data

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (not included as the result is the same as Z1)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 10 (continued)
Panel C

Medium-Activity Sample Hourly Mean Comparisons
Pre-Crash Period

Signs of row minus column means : +++ significant at 1% using both original and rank data, ++ significant only with ranks, + insignificant or significant only when using original data (where contradiction occurs, sign for ranks is used); same signing convention is used for negative values.

HOUR		2	3	4	5	6	
1	NT	+++	+++	+++	+++	-	NT
	VOL	++	+++	+++	+++	+++	VOL
	RETV	++	++	++	++	+	RETV
	RET	+++	+++	+++	+++	+++	RET
	Z1	+++	+++	+	+++	+++	Z1
	Z2	++	++	++	++	++	Z2
	C1	---	---	-	---	---	C1
	C2	-	-	-	-	--	C2
	AVGS	+++	+++	+++	+++	+++	AVGS
	BLOCK	+	+	++	++	++	BLOCK
2	NT		+++	+++	+++	---	NT
	VOL		+	++	+++	-	VOL
	RETV		+	++	+	-	RETV
	RET		+	+	++	-	RET
	Z1		+	-	+	+	Z1
	Z2		+	+	+	-	Z2
	C1		-	+	-	-	C1
	C2		-	+	-	-	C2
	AVGS		+	+	+	+	AVGS
	BLOCK		+	+	+	+	BLOCK
3	NT			+++	+	---	NT
	VOL			+	+	--	VOL
	RETV			+	+	-	RETV
	RET			+	+	-	RET
	Z1			-	-	+	Z1
	Z2			+	-	-	Z2
	C1			+	+	-	C1
	C2			+	-	-	C2
	AVGS			+	+	+	AVGS
	BLOCK			+	+	+	BLOCK
4	NT				--	---	NT
	VOL				-	--	VOL
	RETV				-	--	RETV
	RET				+	--	RET
	Z1				+	+	Z1
	Z2				-	--	Z2
	C1				-	-	C1
	C2				-	-	C2
	AVGS				+	-	AVGS
	BLOCK				+	+	BLOCK
5	NT					---	NT
	VOL					--	VOL
	RETV					--	RETV
	RET					---	RET
	Z1					+	Z1
	Z2					-	Z2
	C1					-	C1
	C2					-	C2
	AVGS					-	AVGS
	BLOCK					-	BLOCK

Table 10 (continued)
Panel D

Medium-Activity Sample Regression Result
Pre-Crash Period

Independent Variables	Dependent Variable ΔQ	FREQ
intercept	0.00016	1.1589
	(1.182)	(124.739)*
H1	-0.00084	0.1821
	(-5.608)*	(18.188)*
H6	9.0E-5	-0.0273
	(0.524)	(-2.352)
W1	7.2E-5	0.0374
	(0.431)	(3.346)*
W5	0.00048	0.0563
	(2.789)*	(4.841)*
NETT	0.00039	-0.0606
	(15.333)*	(-23.806)*
NT	2.7E-5	-0.0149
	(2.036)	(-13.526)*
ΔS	-0.00239	0.0273
	(-25.220)*	(3.699)*
RET	0.7045	4.6699
	(143.122)*	(10.498)*
RETV	0.6126	-92.984
	(1.486)	(-3.099)*
F	2659.38	226.89
R^2	0.5245	0.0860
adj. R^2	0.5243	0.0856
n	21705	21705

Number in parentheses are the t values for the null hypothesis that corresponding coefficient = 0,
* indicates significant at the 1% level

Dependent variables :
ΔQ = percentage changes in quoted prices over an hourly interval
FREQ = number of quotes / number of trades

Independent variables : (Note : NETT, ΔS and RET are in their absolute values in FREQ regression)
H1 : an indicator = 1 if hour 1, = 0 otherwise
H6 : an indicator = 1 if hour 6, = 0 otherwise
W1 : an indicator = 1 if Monday, = 0 otherwise
W5 : an indicator = 1 if Friday, = 0 otherwise
NETT : number of purchase minus number of sale over an hourly interval
NT : hourly number of trades
ΔS : percentage changes in the magnitude of the spread
RET : hourly return
RETV : (trade-by-trade) return variance

Table 11

Hourly Statistics of Trading Activity, Price Changes and Transaction Cost for the Post-Crash Medium-Activity Sample

Panel A

Medium-Activity Sample Hourly Summary Statistics

Post-Crash Period

Numbers are Means and Standard Deviation (in parentheses) across days and stocks

Hour	1	2	3	4	5	6
NT	12.46	12.63	10.93	10.52	11.18	12.64
	(7.26)	(7.52)	(6.73)	(6.10)	(6.24)	(7.28)
VOL	275.57	228.66	193.25	193.41	178.38	225.31
	(352.43)	(395.83)	(572.99)	(615.68)	(261.12)	(935.93)
RETV	1.6E-4	1.5E-4	1.0E-3	1.5E-4	1.5E-4	1.4E-4
	(2.7E-4)	(2.7E-4)	(3.6E-2)	(2.7E-4)	(4.8E-4)	(2.9E-4)
RET	0.0320	0.0266	0.0229	0.0198	0.0202	0.0267
	(0.0340)	(0.0292)	(0.0303)	(0.0226)	(0.0235)	(0.0325)
Z1	0.8664	0.8577	0.8214	0.8108	0.8672	0.8340
	(0.3016)	(0.2843)	(0.3146)	(0.3378)	(0.2894)	(0.3814)
Z2	2.7235	1.6986	2.2479	2.4130	2.5390	1.4660
	(34.784)	(28.132)	(31.269)	(25.961)	(30.849)	(20.161)
C1	0.1336	0.1423	0.1786	0.1892	0.1328	0.1660
	(0.3016)	(0.2843)	(0.3146)	(0.3378)	(0.2894)	(0.3814)
C2	2.5730	1.4041	1.5399	2.4130	3.4368	0.8778
	(28.447)	(17.391)	(15.333)	(25.961)	(33.509)	(7.4512)
AVGS	0.0262	0.0252	0.0253	0.0252	0.0238	0.0243
	(0.0158)	(0.0148)	(0.0165)	(0.0166)	(0.0158)	(0.0154)
BLOCK	0.0469	0.0300	0.0296	0.0315	0.0299	0.0316
	(0.0807)	(0.0675)	(0.0707)	(0.0740)	(0.0690)	(0.0679)
sample size	2149	2015	1809	1776	874	924

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (as proportion of quoted spread)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 11 (continued)
Panel B

Medium-Activity Sample ANOVA Results

Post-Crash Period

Dependent Variables	F	F_h	$F_{h/w}$	$F_{w/h}$	R^2	n
NT	60.85 (65.68)	32.84 (33.72)	34.13 (35.19)	95.87 (105.63)	0.0543 (0.0584)	9547
VOL	8.96 (58.79)	7.66 (65.76)	7.97 (67.56)	10.58 (50.07)	0.0084 (0.0526)	9547
RETV	0.91+ (63.39)	0.87+ (7.69)	0.85+ (7.60)	0.96+ (133.02)	0.0009 (0.0564)	9547
RET	74.40 (74.74)	46.03 (44.75)	47.51 (46.56)	109.87 (112.21)	0.0656 (0.0659)	9547
Z1	8.76 (6.28)	9.88 (7.55)	10.02 (7.65)	7.36 (4.69)	0.0082 (0.0059)	9547
Z2	1.70+ (9.14)	0.50+ (13.66)	0.49+ (13.82)	3.19+ (3.50)	0.0017 (0.0089)	9547
C2	2.71 (5.61)	1.92+ (6.45)	1.89+ (6.50)	3.71 (4.56)	0.0026 (0.0053)	9547
AVGS	53.64 (78.18)	4.09 (7.43)	4.10 (7.43)	115.59 (166.63)	0.0482 (0.0687)	9547
BLOCK	11.25 (23.71)	17.19 (33.80)	17.23 (34.15)	3.83 (11.10)	0.0105 (0.0219)	9547

Independent variables = hour of the day and day of the week

F = overall F value of the ANOVA
F_h = F value when only the hour effect is included
$F_{h/w}$ = partial F value of the hour effect
$F_{w/h}$ = partial F value of the day effect

+ : insignificant at the 1% level

Numbers in parentheses are the corresponding statistics using rank transformed data

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (not included as the result is the same as Z1)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 11 (continued)
Panel C

Medium-Activity Sample Hourly Mean Comparisons
Post-Crash Period

Signs of row minus column means : +++ significant at 1% using both original and rank data, ++ significant only with ranks, + insignificant or significant only when using original data (where contradiction occurs, sign for ranks is used); same signing convention is used for negative values.

HOUR		2	3	4	5	6	
1	NT	-	+++	+++	+++	-	NT
	VOL	++	+++	+++	+++	++	VOL
	RETV	+	+	++	++	++	RETV
	RET	+++	+++	+++	+++	+++	RET
	Z1	+	+++	+++	+	+	Z1
	Z2	++	++	++	++	+	Z2
	C1	-	---	---	-	-	C1
	C2	-	-	--	-	-	C2
	AVGS	+	+	+	+++	++	AVGS
	BLOCK	+++	+++	+++	+++	+++	BLOCK
2	NT		+++	+++	+++	-	NT
	VOL		++	++	++	+	VOL
	RETV		+	+	++	+	RETV
	RET		+++	+++	+++	+	RET
	Z1		+	+	-	+	Z1
	Z2		+	+	+	-	Z2
	C1		---	-	+	-	C1
	C2		-	--	-	-	C2
	AVGS		+	+	++	+	AVGS
	BLOCK		+	+	+	-	BLOCK
3	NT			-	-	---	NT
	VOL			+	-	--	VOL
	RETV			+	++	+	RETV
	RET			+	+	--	RET
	Z1			+	-	-	Z1
	Z2			+	-	--	Z2
	C1			+	++	+	C1
	C2			-	-	+	C2
	AVGS			+	+	+	AVGS
	BLOCK			-	-	-	BLOCK
4	NT				-	---	NT
	VOL				-	--	VOL
	RETV				+	+	RETV
	RET				+	---	RET
	Z1				---	-	Z1
	Z2				-	--	Z2
	C1				+	+	C1
	C2				+	+	C2
	AVGS				+	+	AVGS
	BLOCK				-	-	BLOCK
5	NT					---	NT
	VOL					-	VOL
	RETV					-	RETV
	RET					---	RET
	Z1					+	Z1
	Z2					-	Z2
	C1					-	C1
	C2					+	C2
	AVGS					-	AVGS
	BLOCK					-	BLOCK

Table 11(continued)
Panel D

Medium-Activity Sample Regression Result
Post-Crash Period

Independent Variables	Dependent Variable ΔQ	FREQ
intercept	-4.1E-6	0.8802
	(-0.013)	(116.957)*
H1	-0.00024	0.0582
	(-0.702)	(7.405)*
H6	0.00011	-0.0025
	(0.237)	(-0.226)
W1	-0.00076	-0.0171
	(-2.057)	(-2.098)
W5	0.00017	0.0453
	(0.445)	(5.232)*
NETT	0.00041	-0.0239
	(9.991)*	(-17.577)*
NT	-2.8E-6	-0.0114
	(-0.140)	(-19.952)*
ΔS	-0.00072	0.0265
	(-3.139)*	(4.233)*
RET	0.8889	1.2186
	(224.805)*	(10.454)*
RETV	0.0029	-0.0336
	(0.329)	(-0.165)
F	6977.24	180.96
R^2	0.8681	0.1459
adj. R^2	0.8680	0.1451
n	9547	9547

Number in parentheses are the t values for the null hypothesis that corresponding coefficient = 0,
* indicates significant at the 1% level

Dependent variables :
ΔQ = percentage changes in quoted prices over an hourly interval
FREQ = number of quotes / number of trades

Independent variables : (Note : NETT, ΔS and RET are in their absolute values in FREQ regression)
H1 : an indicator = 1 if hour 1, = 0 otherwise
H6 : an indicator = 1 if hour 6, = 0 otherwise
W1 : an indicator = 1 if Monday, = 0 otherwise
W5 : an indicator = 1 if Friday, = 0 otherwise
NETT : number of purchase minus number of sale over an hourly interval
NT : hourly number of trades
ΔS : percentage changes in the magnitude of the spread
RET : hourly return
RETV : (trade-by-trade) return variance

Table 12

Hourly Statistics of Trading Activity, Price Changes and Transaction Cost for the Pre-Crash Low-Activity Sample

Panel A

Low-Activity Sample Hourly Summary Statistics

Pre-Crash Period

Numbers are Means and Standard Deviation (in parentheses) across days and stocks

Hour	1	2	3	4	5	6
NT	4.47	4.35	4.27	4.13	4.13	4.51
	(1.94)	(1.86)	(1.76)	(2.00)	(1.41)	(1.70)
VOL	84.84	73.82	76.92	60.01	127.31	47.95
	(207.90)	(192.96)	(270.69)	(99.06)	(516.54)	(56.60)
RETV	7.1E-5	3.8E-5	3.9E-5	4.3E-5	2.3E-5	3.8E-5
	(2.5E-4)	(6.7E-5)	(7.7E-5)	(8.8E-5)	(3.9E-5)	(8.1E-5)
RET	0.0105	0.0104	0.0082	0.0104	0.0079	0.0105
	(0.0139)	(0.0133)	(0.0078)	(0.0126)	(0.0088)	(0.0112)
Z1	0.8108	0.8038	0.8141	0.8823	0.8031	0.8593
	(0.3301)	(0.3247)	(0.3009)	(0.1821)	(0.2872)	(0.1751)
Z2	15.868	17.347	11.527	17.612	12.643	13.205
	(73.027)	(82.535)	(53.401)	(63.300)	(73.907)	(56.177)
C1	0.1892	0.1962	0.1859	0.1177	0.1969	0.1407
	(0.3301)	(0.3274)	(0.3009)	(0.1821)	(0.2872)	(0.1751)
C2	15.836	8.6443	15.064	18.747	12.701	7.1014
	(75.770)	(42.132)	(73.787)	(71.044)	(59.344)	(36.727)
AVGS	0.0182	0.0165	0.0167	0.0190	0.0141	0.0157
	(0.0119)	(0.0090)	(0.0092)	(0.0129)	(0.0076)	(0.0076)
BLOCK	0.0297	0.0275	0.0215	0.0373	0.0320	0.0191
	(0.0948)	(0.0972)	(0.0873)	(0.1203)	(0.1035)	(0.0704)
sample size	384	217	123	72	64	83

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (as proportion of quoted spread)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 12 (continued)
Panel B

Low-Activity Sample ANOVA Results

Pre-Crash Period

Dependent Variables	F	F_h	$F_{h/w}$	$F_{w/h}$	R^2	n
NT	1.24+ (1.43)+	0.88+ (1.30)+	0.91+ (1.36)+	1.68+ (1.58)+	0.0118 (0.0134)	943
VOL	0.78+ (0.91)+	1.00+ (0.37)+	1.10+ (0.40)+	0.51+ (1.58)+	0.0075 (0.0087)	943
RETV	1.24+ (1.26)+	1.98+ (2.21)+	2.01+ (2.19)+	0.32+ (0.07)+	0.0119 (0.0120)	943
RET	1.13+ (0.60)+	1.10+ (0.38)+	1.22+ (0.43)+	1.17+ (0.88)+	0.0108 (0.0058)	943
Z1	0.76+ (0.60)+	1.13+ (0.66)+	1.17+ (0.74)+	0.30+ (0.52)+	0.0073 (0.0057)	943
Z2	1.60+ (1.86)+	0.14+ (1.45)+	0.19+ (1.46)+	3.42 (2.38)+	0.0171 (0.0200)	943
C2	0.76+ (0.65)+	0.60+ (0.51)+	0.63+ (0.47)+	0.97+ (0.82)+	0.0075 (0.0063)	943
AVGS	2.41+ (1.64)+	2.89+ (2.04)+	3.18 (2.21)+	1.82+ (1.15)+	0.0228 (0.0156)	943
BLOCK	0.86+ (0.80)+	0.45+ (0.34)+	0.39+ (0.32)+	1.39+ (1.37)+	0.0083 (0.0076)	943

Independent variables = hour of the day and day of the week

F = overall F value of the ANOVA
F_h = F value when only the hour effect is included
$F_{h/w}$ = partial F value of the hour effect
$F_{w/h}$ = partial F value of the day effect

+ : insignificant at the 1% level

Numbers in parentheses are the corresponding statistics using rank transformed data

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (not included as the result is the same as Z1)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 12 (continued)
Panel C

Low-Activity Sample Hourly Mean Comparisons
Pre-Crash Period

Signs of row minus column means : +++ significant at 1% using both original and rank data, ++ significant only with ranks, + insignificant or significant only when using original data (where contradiction occurs, sign for ranks is used); same signing convention is used for negative values.

HOUR		2	3	4	5	6	
1	NT	+	+	+	+	-	NT
	VOL	+	+	+	+	+	VOL
	RETV	+	+	+	+	+	RETV
	RET	-	+	+	+	-	RET
	Z1	+	-	-	+	-	Z1
	Z2	+	+	+	+	+	Z2
	C1	-	+	+	-	+	C1
	C2	+	-	-	-	-	C2
	AVGS	+	+	-	+	+	AVGS
	BLOCK	+	+	-	-	+	BLOCK
2	NT		+	+	+	-	NT
	VOL		-	+	-	+	VOL
	RETV		+	+	+	+	RETV
	RET		+	+	+	-	RET
	Z1		-	-	-	-	Z1
	Z2		+	+	+	+	Z2
	C1		+	+	+	+	C1
	C2		-	-	-	-	C2
	AVGS		-	-	+	+	AVGS
	BLOCK		+	-	-	+	BLOCK
3	NT			+	+	-	NT
	VOL			+	-	+	VOL
	RETV			+	+	+	RETV
	RET			-	+	-	RET
	Z1			-	+	+	Z1
	Z2			+	+	-	Z2
	C1			+	-	-	C1
	C2			+	-	-	C2
	AVGS			-	+	+	AVGS
	BLOCK			-	-	+	BLOCK
4	NT				-	-	NT
	VOL				-	+	VOL
	RETV				+	-	RETV
	RET				+	-	RET
	Z1				+	+	Z1
	Z2				-	-	Z2
	C1				-	-	C1
	C2				-	-	C2
	AVGS				+	+	AVGS
	BLOCK				+	+	BLOCK
5	NT					-	NT
	VOL					+	VOL
	RETV					-	RETV
	RET					-	RET
	Z1					-	Z1
	Z2					-	Z2
	C1					+	C1
	C2					-	C2
	AVGS					-	AVGS
	BLOCK					+	BLOCK

Table 12 (continued)
Panel D

Low-Activity Sample Regression Result
Pre-Crash Period

Independent Variables	Dependent Variable $\triangle Q$	FREQ
intercept	-0.00030	1.3524
	(-0.399)	(20.913)*
H1	-0.00061	0.1241
	(-1.075)	(2.712)*
H6	0.00024	-0.0031
	(0.245)	(-0.039)
W1	3.7E-5	0.1206
	(0.051)	(2.048)
W5	0.00070	0.0728
	(0.880)	(1.139)
NETT	0.00079	-0.1729
	(4.323)*	(-8.793)*
NT	-3.0E-5	-0.0163
	(-0.203)	(-1.170)*
$\triangle S$	0.00043	0.0300
	(0.901)	(0.621)
RET	0.7564	1.6360
	(43.897)*	(0.868)
RETV	10.5123	-217.743
	(6.521)*	(-1.573)
F	243.75	14.97
R^2	0.7016	0.1262
adj. R^2	0.6987	0.1178
n	943	943

Number in parentheses are the t values for the null hypothesis that corresponding coefficient = 0,
* indicates significant at the 1% level

Dependent variables :
$\triangle Q$ = percentage changes in quoted prices over an hourly interval
FREQ = number of quotes / number of trades

Independent variables : (Note : NETT, $\triangle S$ and RET are in their absolute values in FREQ regression)
H1 : an indicator = 1 if hour 1, = 0 otherwise
H6 : an indicator = 1 if hour 6, = 0 otherwise
W1 : an indicator = 1 if Monday, = 0 otherwise
W5 : an indicator = 1 if Friday, = 0 otherwise
NETT : number of purchase minus number of sale over an hourly interval
NT : hourly number of trades
$\triangle S$: percentage changes in the magnitude of the spread
RET : hourly return
RETV : (trade-by-trade) return variance

Table 13

Hourly Statistics of Trading Activity, Price Changes and Transaction Cost for the Post-Crash Low-Activity Sample

Panel A

Low-Activity Sample Hourly Summary Statistics

Post-Crash Period

Numbers are Means and Standard Deviation (in parentheses) across days and stocks

Hour	1	2	3	4	5	6
NT	5.53	5.26	5.16	5.14	4.92	5.31
	(2.46)	(2.44)	(2.13)	(2.10)	(2.03)	(2.11)
VOL	80.68	57.25	55.03	54.08	47.87	60.88
	(113.60)	(81.83)	(88.73)	(62.64)	(51.65)	(74.34)
RETV	2.6E-4	3.0E-4	2.4E-4	2.9E-4	2.4E-4	2.6E-4
	(7.0E-4)	(1.2E-3)	(6.2E-4)	(5.9E-4)	(5.8E-4)	(7.2E-4)
RET	0.0264	0.0237	0.0233	0.0234	0.0230	0.0237
	(0.0267)	(0.0252)	(0.0230)	(0.0238)	(0.0206)	(0.0250)
Z1	0.8153	0.8306	0.8174	0.8026	0.8318	0.8219
	(0.3741)	(0.3816)	(0.2522)	(0.2830)	(0.2473)	(0.2120)
Z2	12.266	20.553	7.2647	11.663	8.4469	2.0552
	(57.552)	(86.319)	(44.873)	(59.712)	(41.441)	(17.658)
C1	0.1847	0.1694	0.1826	0.1974	0.1682	0.1780
	(0.3741)	(0.3816)	(0.2522)	(0.2830)	(0.2473)	(0.2120)
C2	11.553	4.0376	7.5911	9.7731	8.7095	3.2560
	(54.557)	(22.504)	(38.024)	(61.609)	(53.846)	(26.183)
AVGS	0.0332	0.0342	0.0329	0.0351	0.0334	0.0327
	(0.0251)	(0.0264)	(0.0229)	(0.0250)	(0.0250)	(0.0215)
BLOCK	0.0239	0.0135	0.0079	0.0146	0.0157	0.0157
	(0.0754)	(0.0739)	(0.0400)	(0.0577)	(0.0738)	(0.0621)
sample size	405	253	183	172	75	84

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (as proportion of quoted spread)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 13 (continued)
Panel B

Low-Activity Sample ANOVA Results

Post-Crash Period

Dependent Variables	F	F_h	$F_{h/w}$	$F_{w/h}$	R^2	n
NT	3.31 (3.22)	1.54+ (1.56)+	1.70+ (1.67)+	5.53 (5.30)	0.0250 (0.0243)	1172
VOL	3.08 (4.69)	4.26 (6.19)	4.35 (6.36)	1.61+ (2.82)+	0.0233 (0.0351)	1172
RETV	1.89+ (3.99)	0.18+ (0.21)+	0.22+ (0.30)+	4.03 (8.71)	0.0145 (0.0300)	1172
RET	2.78 (3.24)	0.74+ (0.62)+	0.85+ (0.77)+	5.34 (6.50)	0.0211 (0.0244)	1172
Z1	0.40+ (0.98)+	0.18+ (1.04)+	0.19+ (1.06)+	0.68+ (0.91)+	0.0031 (0.0075)	1172
Z2	1.05+ (1.21)+	1.56+ (2.04)+	1.54+ (2.07)+	0.42+ (0.18)+	0.0089 (0.0102)	1172
C2	1.25+ (0.78)+	1.05+ (0.45)+	0.99+ (0.47)+	1.50+ (1.20)+	0.0097 (0.0061)	1172
AVGS	2.59 (4.51)	0.23+ (0.32)+	0.29+ (0.44)+	5.54 (9.74)	0.0197 (0.0337)	1172
BLOCK	1.77+ (2.51)	1.72+ (3.12)	1.66+ (3.07)	1.84+ (1.75)+	0.0135 (0.0191)	1172

Independent variables = hour of the day and day of the week

F = overall F value of the ANOVA
F_h = F value when only the hour effect is included
$F_{h/w}$ = partial F value of the hour effect
$F_{w/h}$ = partial F value of the day effect

+ : insignificant at the 1% level

Numbers in parentheses are the corresponding statistics using rank transformed data

NT : hourly number of trades
VOL : hourly volume (in hundreds of shares)
RETV : (trade-by-trade) return variance
RET : absolute value of hourly return
Z1 : information component using Stoll's approach (as proportion of quoted spread)
Z2 : information component using Glosten and Harris's approach (as proportion of quoted spread)
C1 : transitory component using Stoll's approach (not included as the result is the same as Z1)
C2 : transitory component using Glosten and Harris's approach (as proportion of quoted spread)
AVGS : average quoted spread (as proportion of price)
BLOCK : proportion of block trades (#block trades/#total trades)

Table 13 (continued)
Panel C

Low-Activity Sample Hourly Mean Comparisons
Post-Crash Period

Signs of row minus column means : +++ significant at 1% using both original and rank data, ++ significant only with ranks, + insignificant or significant only when using original data (where contradiction occurs, sign for ranks is used); same signing convention is used for negative values.

HOUR		2	3	4	5	6	
1	NT	+	+	+	+	+	NT
	VOL	++	++	++	+	+	VOL
	RETV	-	-	-	-	-	RETV
	RET	+	+	+	+	+	RET
	Z1	-	+	+	+	+	Z1
	Z2	+	+	+	+	+	Z2
	C1	+	-	-	-	-	C1
	C2	+	+	-	+	+	C2
	AVGS	-	-	-	-	-	AVGS
	BLOCK	+	+	+	+	+	BLOCK
2	NT		-	-	+	-	NT
	VOL		+	+	+	-	VOL
	RETV		+	-	+	-	RETV
	RET		-	-	-	-	RET
	Z1		+	+	+	+	Z1
	Z2		+	+	+	+	Z2
	C1		-	-	-	-	C1
	C2		-	-	-	+	C2
	AVGS		+	-	+	+	AVGS
	BLOCK		+	-	-	-	BLOCK
3	NT			+	+	-	NT
	VOL			-	+	-	VOL
	RETV			-	+	-	RETV
	RET			-	-	-	RET
	Z1			+	-	+	Z1
	Z2			+	-	+	Z2
	C1			-	+	-	C1
	C2			-	-	+	C2
	AVGS			-	-	-	AVGS
	BLOCK			-	-	-	BLOCK
4	NT				+	-	NT
	VOL				+	-	VOL
	RETV				+	-	RETV
	RET				-	+	RET
	Z1				-	+	Z1
	Z2				-	-	Z2
	C1				+	-	C1
	C2				+	+	C2
	AVGS				+	+	AVGS
	BLOCK				+	-	BLOCK
5	NT					-	NT
	VOL					-	VOL
	RETV					-	RETV
	RET					+	RET
	Z1					+	Z1
	Z2					+	Z2
	C1					-	C1
	C2					+	C2
	AVGS					+	AVGS
	BLOCK					-	BLOCK

Table 13(continued)
Panel D

Low-Activity Sample Regression Result
Post-Crash Period

Independent Variables	Dependent Variable: ΔQ	Dependent Variable: FREQ
intercept	-0.00052	1.0911
	(-0.417)	(36.054)*
H1	-0.00076	0.0093
	(-0.777)	(0.410)
H6	0.00179	0.0152
	(0.987)	(0.365)
W1	-0.00253	-0.0139
	(-2.304)	(-0.560)
W5	-0.00019	0.0233
	(-0.811)	(0.688)
NETT	0.00117	-0.0835
	(4.276)*	(-9.617)*
NT	8.3E-5	-0.0253
	(0.412)	(-4.576)*
ΔS	-0.00060	0.0403
	(-0.792)	(1.883)
RET	0.8240	1.5656
	(57.189)*	(3.292)*
RETV	-0.2135	-7.0052
	(-0.364)	(-0.490)
F	457.58	23.19
R^2	0.7799	0.1523
adj. R^2	0.7782	0.1457
n	1172	1172

Number in parentheses are the t values for the null hypothesis that corresponding coefficient = 0,
* indicates significant at the 1% level

Dependent variables :
ΔQ = percentage changes in quoted prices over an hourly interval
FREQ = number of quotes / number of trades

Independent variables : (Note : NETT, ΔS and RET are in their absolute values in FREQ regression)
H1 : an indicator = 1 if hour 1, = 0 otherwise
H6 : an indicator = 1 if hour 6, = 0 otherwise
W1 : an indicator = 1 if Monday, = 0 otherwise
W5 : an indicator = 1 if Friday, = 0 otherwise
NETT : number of purchase minus number of sale over an hourly interval
NT : hourly number of trades
ΔS : percentage changes in the magnitude of the spread
RET : hourly return
RETV : (trade-by-trade) return variance

Table 14

Estimation of the Information and transitory components
for the ten most active stocks in 1987
estimation period : August 31 - September 24

Company Name	Symbol	Z	C	R^2	F	n	P
American Express	AXY	0.096 (2.952)*	0.054 (1.648)+	0.153	18.988	107	35
Eastman Kodak	EK	0.176 (3.327)*	0.177 (3.367)*	0.287	42.194	107	70
General Electric	GE	0.208 (4.597)*	0.124 (2.752)*	0.342	54.582	107	60
General Motors	GM	0.248 (5.459)*	0.034 (0.741)	0.289	42.753	107	90
IBM	IBM	0.408 (4.309)*	0.124 (1.309)+	0.286	42.047	107	160
Navistar International	NAV	0.019 (1.622)+	0.016 (1.470)+	0.094	10.889	107	5
AT&T	T	0.095 (4.468)*	0.035 (1.645)+	0.274	39.531	107	35
Texaco	TX	0.096 (4.003)*	0.053 (2.238)+	0.248	34.699	107	40
USX Corp	X	0.143 (5.864)*	0.002 (0.074)	0.287	42.208	107	40
Exxon Corp	XON	0.265 (4.247)*	0.060 (0.971)	0.222	29.898	107	50

Numbers in parentheses are the t values for the null hypothesis that the corresponding coefficient = 0, and * (+) indicates that the regression coefficient is significantly different from zero at the 1% (10%) level

Z = information component (in dollars)
C = transitory component (in dollars)
n = sample size = number of days * hours per day - 1 = 107
P = approximate stock price (in dollars)

The Z and C are estimated from the OLS regression (Glosten and Harris (1988))

$$P_t - P_{t-1} = C(Q_t - Q_{t-1}) + ZQ_t + e_t$$

The P_t and Q_t are different from Glosten and Harris in that :

P_t = ending price at hour t
Q_t = net volume (investor purchase - investor sale) in hour t
= 1 if positive net volume
= 0 if zero net volume
=-1 if negative net volume

Table 15

The Implied Values of the Variances of Noises and Variances of True Prices
Using Equations (C8) and (C10)
assuming different values of sensitivities of informed and uninformed traders

for the ten most active stocks in 1987
estimation period : August 31 - September 24

	E_i=0.05 E_m=1.0 var(i)	var(e)	E_i=0.25 E_m=0.5 var(i)	var(e)	E_i=0.5 E_m=1.0 var(i)	var(e)	E_i=1.0 E_m=2.0 var(i)	var(e)	redv	P	var(V)
AXY	0.025	0.019	0.044	0.693	0.067	1.755	0.113	5.324	0.091	35	0.963
EK	0.099	0.909	0.133	1.829	0.176	3.415	0.261	8.045	0.229	70	0.971
GE	0.115	0.751	0.156	1.627	0.209	3.176	0.313	7.785	0.163	60	1.007
GM	0.118	0.428	0.168	1.512	0.231	2.510	0.356	6.741	0.169	90	1.009
IBM	0.349	1.509	0.432	2.711	0.535	4.667	0.740	10.093	0.585	160	1.009
NAV	0.001	0.025	0.004	0.254	0.008	0.871	0.015	3.214	0.008	5	0.738
T	0.023	0.147	0.042	0.621	0.066	1.661	0.113	5.232	0.036	35	0.995
TX	0.025	0.187	0.044	0.686	0.066	1.741	0.112	5.287	0.049	40	0.957
X	0.038	0.132	0.066	0.602	0.102	1.638	0.173	5.201	0.052	40	0.995
XON	0.143	0.587	0.196	1.396	0.262	2.845	0.394	7.272	0.312	50	0.999

E_i (E_m) = sensitivity of informed (uninformed) traders.

var(e) = variance of noises

var(i) = variance of true prices

redv = variance of error terms for the regression in Table 14

P = approximate stock price

var(V) = variance of net volume

BIBLIOGRAPHY

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

Akerloff, George, " The Market for "Lemons" : Quality Uncertainty and the Market Mechanism, " Quarterly Journal of Economics 84, (1970), 488-500.

Amihud, Y., and H. Mendelson, " Dealership Market : Market Making with Uncertainty, " Journal of Financial Economics 8, (1980), 31-54.

______________________________, " Asset Pricing and the Bid-Ask Spread, " Journal of Financial Economics 17, (1986), 223-250.

______________________________, " Trading Mechanisms and Stock Returns : An Empirical Investigation, " Journal of Finance 42, (1987), 533-553.

Bagehot, Walter, [pseud.] " The Only Game in Town," Financial Analysts Journal 22, (1971), 12-14.

Ball, Clifford, " Estimation Bias Induced by Discrete Security Prices, " Journal of Finance 43, (1988), 841-865.

Banz, Rolf, " The Relationship between Return and Market Value of Common Stocks," Journal of Financial Economics 9, (1981), 3-18.

Barnea, Amir, " Reply : Specialists' Performance and Serial Dependence of Stock Price Changes, " Journal of Financial and Quantitative Analysis 11, (1976), 909-911.

Barry, C., and S. Brown, " Differential Information and the Small Firm Effect, " Journal of Financial Economics 13, (1984), 283-294.

Benston, G., and R. Hagerman, " Determinants of Bid-Ask Spreads in the Over-The-Counter Market, " Journal of Financial Economics 1, (1974), 353-364.

Barclay, M., and C. Smith, " Corporate Payout Policy : Cash Dividends versus Open Market Share Repurchases, " Journal of Financial Economics 22, (1988), 61-82.

Bhattacharya, Mihir, " Direct Tests of Bid/Ask Spread Models, " working paper, University of Michigan, (1985).

Black, Fischer, " Noise, " Journal of Finance 41, (1986), 529-543.

Black, F., and M. Scholes, " The Pricing of Options and Corporate Liabilities, " Journal of Political Economy 81, (1973), 637-654.

Bradfield, James, " A Formal Dynamic Model of Market Making, " Journal of Financial and Quantitative Analysis 2, (1979), 275-291.

Branch, B., and W. Freed, " Bid-Ask Spreads on the AMEX and the Big Board, " Journal of Finance 32, (1977), 159-163.

Breusch, T. S., and A. R. Pagan, " A Simple Test for Heteroscedasticity and Random Coefficient Variation, " Econometrica 47, (1979), 1287-1294.

Brock, W., and A. Kleidon, " Exogenous Demand Shocks and Trading Volume : A Model of Intraday Bids and Asks, " working paper, Stanford University, (1989).

Cho, D., and E. Frees, " Estimating the volatility of Discrete Stock Prices, " Journal of Finance 43, (1988), 451-466.

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

Cohen, K. J., S. F. Maier, R. A. Schwartz and D. K. Whitcomb, " Limit Orders, Market Structure and the Return Generating Process, " Journal of Finance 33, (1978), 723-736.

________, " Transaction Costs, Order Placement Strategy, and the Existence of the Bid-Ask Spread, " Journal of Political Economy 89, (1981), 287-305.

Copeland, T., and D. Galai, " Information effects on the Bid-Ask Spread, " Journal of Finance 38, (1983), 383-417.

Conover, W. J. Practical Nonparametric Statistics. John Wiley. 1980.

Conroy, R., and R. Winkler, " Informational Differences between Limit and Market Orders for a Market Maker, " Journal of Financial and Quantitative Analysis 16, (1981), 703-724.

Demsetz, Harold, " The Cost of Transacting, " Quarterly Journal of Economics 82, (1968), 33-53.

Easley, D., and M. O'Hara, " Volume and Market Efficiency, " working paper, UCLA/Cornell University,(1985).

_________________________, " Price, Trade Size and Information in Securities Markets, " Journal of Financial Economics 19, (1987), 69-90.

Foster, F. D., and S. Viswanathan, " Variations in Volumes, Spreads and Variances, " working paper, Duke University, (1988).

_________________________________, " A Theory of the Interday Variations in Volumes, Variances and Trading Costs in Securities Markets," working paper, Duke University, (1989).

French, Kenneth, " Stock Returns and the Weekend Effect," Journal of Financial Economics 8, (1980),55-70.

French, K., and R. Roll, " Stock Return Variances : The Arrival of Information and the Reaction of Traders, " Journal of Financial Economics 17, (1986), 5-26.

Fuller, Wayne. Introduction to Statistical Time Series. John Wiley. 1976.

Garbade, K., and Z. Lieber, " On the Independence of Transactions on the New York Stock Exchange, " Journal of Banking and Finance 1, (1977), 151-172.

Garbade, K., and C. Sekaran, " Opening Prices on the New York Stock Exchange, " Journal of Banking and Finance 5, (1981), 345-355.

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

Glosten, Lawrence, " Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices, ' Journal of Finance 42, (1987), 1293-1308.

_________________, " Insider Trading, Liquidity, and the Role of the Monopolist Specialist, " Journal of Business 62, (1989), 211-235.

Glosten, L., and L. Harris, " Estimating the Components of the Bid/Ask Spread," Journal of Financial Economics 21, (1988), 123-142.

Glosten, L., and P. Milgrom, " Bid, Ask and Transaction Prices in A Specialist Market with Heterogeneously Informed Traders, " Journal of Financial Economics 14, (1985), 71-100.

Goldman, M. B., and A. Beja, " Market Prices vs. Equilibrium Prices : Return Variance, Serial Correlation, and the Role of the Specialist, " Journal of Finance 34, (1979), 595-607.

Gottlieb, G., and A. Kalay, " Implications of the Discreteness of Observed Stock Prices, " Journal of Finance 40, (1985),135-154.

Grossman, S., and M. H. Miller, " Liquidity and Market Structure, " Journal of Finance 43, (1988), 617-633.

Grossman, S., and J. Stiglitz, " On the Impossibility of Informationally Efficient Markets, " American Economic Review 70, (1980), 393-407.

Hamilton, J., " Competition, Scale Economics, and Transaction Cost in the Stock Market, " Journal of Financial and Quantitative Analysis 11, (1976), 779-802.

____________, " The Dealer and Market Concept of Bid-Ask Spread : A Comparison for NASDAQ Stocks, " working paper, Wayne State University, (1988).

Harris, Lawrence, " A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns, " Journal of Financial Economics 16, (1986), 99-118.

________________, " A Day-End Transaction Price Anomaly," Journal of Financial and Quantitative Analysis 24, (1989), 29-45.

Hasbrouck, Joel, " Trade, Quote, Inventories and Information," Journal of Financial Economics 22, (1988), 229-252.

Hasbrouck, J., and T. Ho, " Order Arrival, Quote Behavior, and the Return Generating Process, " Journal of Finance 42, (1987), 1035-1048.

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

Ho, T., R. Schwartz and D. Whitcomb, " The Trading Decision and Market Clearing under Transaction Price Uncertainty, " Journal of Finance 40, (1985), 21-42.

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

Jain, P., and G. Joh, " The Dependence between Hourly Prices and Trading Volume, " Journal of Financial and Quantitative Analysis 23, (1988), 270-283.

Jameson, Mel, " Notes on Estimating the Components of Bid-Ask Spread, " Louisiana State University,(June 1988).

Judge, G., C. Hill, W. Griffiths, H. Lütkepohl, and T. C. Lee. Introduction to the Theory and Practice of Econometrics. John Wiley. 1982.

Karpoff, Jonathan, " The Relation between Price Changes and Trading Volume : A Survey, " Journal of Financial and Quantitative Analysis 22, (1987), 109-126.

Keim, D., " Size-related Anomalies and Stock Return Seasonality : Further Empirical Evidence, " Journal of Financial Economics 12, (1983), 13-32.

Kross, W., " The Size Effect is Primarily a Price Effect, " Journal of Financial Research 8, (1985), 169-180.

Kyle, Albert, " Continuous Auctions and Insider Trading," Econometrica 53, (1985), 1315-1335.

Lamoureux, G., and G. Sanger, " Firm Size and Turn-of-the-year Effects in the OTC/NASDAQ Market, " Journal of Finance, forthcoming.

Litner, John, " Security Prices, Risk, and Maximal Gains from Diversification, " Journal of Finance 20, (1965), 587-615.

Marsh, T., and K. Rock, " Exchange Listing and Liquidity : A Comparison of the American Stock Exchange with the NASDAQ National Market System, " American Stock Exchange Transactions Data Research Project, (1986).

Mendelson, Haim, " Market Behavior in A Clearing House," Econometrica 50, (1982), 1505-1524.

Mildenstein, E., and H. Schleef, " The Optimal Pricing Policy of a Monopolistic Marketmaker in the Equity Market, " Journal of Finance 38, (1983), 218-231.

Mossin, Jan, " Equilibrium in a Capital Asset Market, " Econometrica 34, (1966), 768-783.

Mulherin, J., M. Gerety, " Information and Trading Volume in Securities Markets : An Analysis of Hourly Data for 1933-1988," working paper, Clemson University, (1989).

Neter, J., W. Wasserman, and M. H. Kutner. Applied Linear Statistical Models. Richard D. Irwin. 1985.

Niederhoffer, V., and M. F. M. Osborne, " Market Making and Reversals on the Stock Exchange, " Journal of the American Statistical Association 61, (1966), 897-916.

NYSE Fact Book. New York : New York Stock Exchange. 1988.

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

Reinganum, Marc, " Misspecification of Capital Asset Pricing : Empirical Anomalies based on Earnings Yields and Market Values, " Journal of Financial Economics 9, (1981), 19-46.

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

Schultz, J., S. Gustavson, and F. Reily, " Factors Influencing the New York Stock Exchange Specialists' Price-Setting Behavior : An Experiment, " Journal of Financial Research 8, (1985), 137-144.

Sharpe, William, " Capital Asset Pricing : A Theory of Market Equilibrium under Conditions of Risk, " Journal of Finance 19, (1964), 425-442.

Smidt, Seymour, " Which Road to an Efficient Stock Market, " Financial Analysts Journal, September-October (1971).

_______________, " The Bid Ask Spread and the Cost of Transacting, " working paper, Cornell University, (1988).

Stoline, Michael, " The Status of Multiple Comparisons : Simultaneous Estimation of All Pairwise Comparisons in One-

Way ANOVA Design, " American Statistician 35, (1981), 134-141.

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

___________, " The Pricing of Security Dealer Services : An Empirical Study of NASDAQ Stocks, " Journal of Finance 33, (1978a), 1153-1172.

___________, " The Stock Exchange Specialist System : An Economic Analysis, " Monograph Series in Finance and Economics, New York University, Salomon Brothers Center. (1985).

___________, " Inferring the Components of the Bid-Ask Spread : Theory and Empirical Tests, " Journal of Finance 44, (1989), 115-134.

Tinic, S.," The Economics of Liquidity Services, " Quarterly Journal of Economics 86, (1972), 70-93.

Tinic, S., and R. West, " Competition and the Pricing of Dealer Service in the Over-The-Counter Stock Market, " Journal of Financial and Quantitative Analysis 7, (1972), 1707-1727.

Wood, R., T. McInish and K. Ord, " An Investigation of Transactions Data for NYSE Stocks, " Journal of Finance 40, (1985), 723-741.

APPENDIX A

List of Company Names for the Sample used in Chapter III

Aluminum Co. Amer.
Anacomp Inc.
Alexander & Alexander Services
Abbott Labs
American Barrick Res. Corp.
ACM Government Income Fund Inc.
Armstrong World Industry Inc.
American Cyanamid Co.
Ames Dept. Stores Inc.
Analog Devices Inc.
Archer Daniels Midland Co.
Allegis Corp.
American Elec. Power Inc.
Aetna Life & Gas Co.
American Family Corp.
Affiliated Publications Inc.
Amfesco Inds Inc.
American General Corp.
Edward A. G. Inc.
AGS Computers Inc.
Amerada Hess Corp.
Ahmanson H. F. & Co.
American Home Products Corp.
American International Group Inc.
Allied Signal Inc.
Arkla Inc.
Amfac Inc.
American Brands Inc.
Advanced Micro Devices Inc.
American Medical International Inc.
AMP inc.
AMR Corp.
Amoco Corp.
Anchor Glass Container Corp.
Anadarko Petroleum Corp.
Air Products & Chemicals Inc.
Applied Magnetics Corp.
Apache Petroleum Co.
QMS inc.
Asarco Inc.
Atlantic Richfield Co.
Aristech Chem. Corp.
Arvin Industries Inc.
Armco Inc.
ASA Ltd.
American Stores Co.
American Standard Inc.
Alltel Corp.

Automatic Data Processing Inc.
Augat Inc.
Avon Products Inc.
Avnet Inc.
Avery International Corp.
American Express Co.
Boeing Co.
Baxter Travenol Labs Inc.
Barnett Banks Inc.
Best Buy Inc.
Brunswick Corp.
Boise Cascade Corp.
BCE Inc.
Biocraft Labs Inc.
Bard (CR) Inc.
Becton Dickinson & Co.
BET Pub. Ltd. Co.
Browning Ferris Industries Inc.
Baltimore Gas & Electrics Co.
Baker Hughes Inc.
Bell & Howell Co.
Bank New York Inc.
Ball Corp.
Battle Mountain Gold Co.
Borden Inc.
Burlington Northern Inc.
Beneficial Corp.
Bausch & Lomb Inc.
Bowater Inc.
British Petroleum Ltd.
Brockway Inc.
Bethelehem Steel Corp.
Bear Sterns Cos Inc.
Brooklyn Union Gas Co.
Anheuser Busch Cos Inc.
Chrysler Corp.
Computer Association International Inc.
Furs Bishops Cafeterias
ConAgra Inc.
CalFed Inc.
Cannon Group Inc.
Carter Wallace Inc.
Caterpillar Inc.
Caesars World Inc.
Cooper Industries Inc.
CBS Inc.
Cabot Corp.
Commodore International Ltd.
Capital Cities ABC inc.
Commercial Cr. Co.
Coca Cola Enterprises Inc.
Citicorp

APPENDIX B

Derivation of Serial Covariance Allowing Changes in the Spread

Based on Stoll (1989), consider that the spread changes at time t, and let $dS = S_t - S_{t-1}$, $dP_t = P_t - P_{t-1}$ then the joint distribution of successive transaction price changes may be shown as follows. The notation BB here refers to that a bid price is followed by a bid price.

dP_t

		Initial Trade at Bid		Initial Trade at Ask	
		BB $\delta*$ $(-S_{t-1}-dS/2)$	BA $(1-\delta)*$ $(S_{t-1}+dS/2)$	AA $\delta*$ $(S_{t-1}+dS/2)$	AB $(1-\delta)*$ $(-S_{t-1}-dS/2)$
dP_{t+1} BB	$-\delta S$	$(1-\pi)^2$	0	0	$\pi(1-\pi)$
BA	$(1-\delta)S$	$\pi(1-\pi)$	0	0	π^2
AA	δS	0	$\pi(1-\pi)$	$(1-\pi)^2$	0
AB	$-(1-\delta)S$	0	π^2	$\pi(1-\pi)$	0

Then

$$\begin{aligned} covP &= (1-\pi)^2\delta^2(S_tS_{t-1}+S_tdS/2)-\pi^2(1-\delta)^2(S_tS_{t-1}+S_tdS/2) \\ &= (S_tS_t+S_tS_{t-1})/2*(\delta^2(1-2\pi)-\pi^2(1-2\delta)) \end{aligned}$$

If dS is random and small, then the equation reduces to Equation (III).

Vita

Pei-Hwang Philip Wei

Areas of Interest

Teaching : Business Finance, Financial Institutions, Investments

Research : Market Microstructure, Information and Prices

Personal Data

Birth Date : December 6, 1957

Marital Status : Single

Education

Ph. D.	Louisiana State University	Major : Finance Minor : Economics
M. S.	Louisiana State University Completed December, 1984	Major : Finance
B. B. A.	National Taiwan University Completed July, 1980	Major : Business

Dissertation

Essays on the Components of the Bid-Ask Spread

Experience

August 1984- May 1989	Department of Finance, Louisiana State University Graduate Research Assistant
January 1988- August 1988	Department of Finance, Louisiana State University Graduate Teaching Assistant
May 1982- June 1982	Shinkong Life Insurance Company, Taiwan Staff in Securities Department

August 1980- May 1982	Army, Taiwan Lieutenant (quarter Master)

Professional Associations

FMA, ASSA

Dissertation Committee

Gary C. Sanger (Chairman)
Department of Finance
Louisiana State University
Baton Rouge, LA 70803
(504) 388-6353

G. Geoffrey Booth
Department of Finance
Louisiana State University
Baton Rouge, LA 70803
(504) 388-6291

John S. Howe
Department of Finance
Louisiana State University
Baton Rouge, LA 70803
(504) 388-6254

William R. Lane
Department of Finance
Louisiana State University
Baton Rouge, LA 70803
(504) 388-6367

R. Carter Hill
Department of Economics
Louisiana State University
Baton Rouge, LA 70803

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: Pei-Hwang Wei

Major Field: Business Administration (Finance)

Title of Dissertation: Essays on the Components of the Bid-Ask Spread

Approved:

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

Date of Examination: July 19, 1989