19.

On the Trading Volume and Time Effects on the Bid-Ask Spread Components of NYSE and NASDAQ Common Stocks

by

Spiridon Spirakos-Papastavridis

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Finance

Warwick Business School
University of Warwick

March 2004

Acknowledgements

I wish to express my sincere thanks to my supervisor Dr. Nick Webber for his invaluable assistance, his eagerness to respond as well as for his constructive comments during the preparation of this thesis. His professionalism and excellent sense of timing have helped me a lot. I also wish to express my gratitude to Dr. Jack Broyles for his encouragement, for enriching my understanding of the mechanisms of the stock exchange and for instilling into me the charm for its 'mysteries' during our long discussions in the early stages of my research. Finally, I wish to express my appreciation for the support of Ms. Jan Woodley, administrative director of the postgraduate programmes, for guiding me through the formal procedures of this undertaking.

To my father Emmanuel and
my mother Despina for their
unconditional love

TABLE OF CONTENTS

Appendix F

LIMIT ORDERS

Appendix G

TABLES OF THE RESULTS AND THE SPECIFICATION TESTS OF THE MODELS

LIST OF FIGURES

Figure	*Title*	*Page*

Chapter Five

Chapter Six

Appendix D

Appendix G

ABBREVIATIONS

AMEX	:	American Stock Exchange
BBO	:	Best Bid Offered
BIC	:	Bayesian Information Criterion
CBOE	:	Chicago Board Options Exchange
CUSIP	:	Committee on Uniform Security Identification Procedure
DOT	:	Designated Order Turnaround System (NYSE)
ECN	:	Electronic Communications Network
GMM	:	Generalized Method of Moments
IPO	:	Initial Public Offering
ISE	:	International Stock Exchange (London)
ISSM	:	Institute for the Study of Security Markets
LM	:	Lagrange Multiplier
LR	:	Likelihood Ratio
LR rule	:	Lee and Ready (1991) Rule
LSE	:	London Stock Exchange
MQP	:	Mandatory Quote Period
MSC	:	Moment Selection Criterion
NASD	:	National Association of Securities Dealers
NASDAQ	:	National Association of Securities Dealers Automated Quotations System
NMS	:	National Market System
NYSE	:	New York Stock Exchange
OARS	:	Opening Automated Report Service (NYSE)
OLS	:	Ordinary Least Squares
OTC	:	Over the Counter (Market)

SEC	:	Securities and Exchange Commission
SOD	:	System Order Database
TAQ	:	Trades and Quotes (Database)
TORQ	:	Trades, Orders, Reports and Quotes (Database)
US	:	United States
USD	:	United States Dollar
VAR	:	Vector Auto-Regression

SUMMARY

The study of the bid-ask spreads of stocks is important since they constitute the mechanism through which trading costs are incorporated into prices and recovered by market-makers. Realized spreads are a measure of the trading costs which private and institutional investors have to cover whereas quoted spreads are important in revealing the price-generating mechanisms involved. A comprehensive trade-indicator model of the bid-ask spread of common stocks has been developed which, unlike previous research, has incorporated trading volume, the depth at the quoted prices, the waiting-time between trades, as well as the fixed-cost-of-trade into equations for the changes in the quoted-spread, the ask, the bid and the transaction prices. The parameters of the models have been estimated using intraday data of NYSE and NASDAQ stocks, split into deciles on the basis of their trading activity as measured by the number of shares traded.

For both exchange mechanisms I find that a large adverse-selection cost, which depends on the trading volume and a smaller inventory-holding cost are present in the quoted spread and that the parameters of both of these vary with trading activity. These costs are asymmetric in the bid and ask sides of the quoted spread, a feature not analyzed in previous empirical work. Only a small part of these costs is recovered in the realized spread through trading. My estimates of the adverse-selection cost present in the realized spread are close to the values given by other researchers but the size of the inventory-holding cost is found to be much lower probably owing to the shorter time-horizon of the data employed. The depth at the quotes is found to be symmetric, to affect the size of the spread of both the NYSE and NASDAQ stocks and also to be present in the realized spread. The parameters of the above components, as well as the fixed-cost-of-trade, are estimated and their patterns for the two trading mechanisms examined are compared and contrasted. Weak evidence is found for the waiting-time between trades both in the quoted and the realized spread.

The results of this thesis, apart from offering support for recent empirical evidence which indicates that information first enters the price-process through the depths of the quotes and not the spread, also contribute to the formation of a more theoretically sound explanation. Moreover, the finding in this thesis that the adverse-selection cost for NASDAQ is larger compared to that of NYSE stocks is in line with other recent empirical research.

Chapter 1

Introduction

1.1 Introduction

Trading in real exchanges involves transaction costs which are necessary in order that market-makers, who are responsible for providing liquidity in the market, could cover certain expenses incurred as well as be compensated for their risks. Such risks are normally associated with the inventory they are obliged to hold so that trading can take place, the risks of the assets traded as well as the danger that the counter-party may be better informed in which case they are likely to enter into an unprofitable transaction. The means by which transaction costs, other than commissions, are recovered in the financial markets is through the bid-ask spread the mechanics of the formation of which have occupied a large amount of empirical as well as theoretical research. Such research has also attempted to investigate the factors which contribute to its formation through the price-generating process and the mechanisms of the exchanges and have referred to components of the spread as building blocks of the total difference between the bid and the ask price. The study of the bid-ask spread and its components is thus necessary for accurately estimating transaction costs which in turn is important in studies of market efficiency, asset pricing as well as market microstructure in general.

Existing research has produced a number of models of the spread and its components using two approaches in modeling it. The first uses the covariance of transaction and/or quote returns whereas the second examines the data to classify trades as buyer- or seller- initiated and thus assign a trade-indicator variable to each trade when there is lack of more sophisticated data which could indicate the origin and nature of each trade. All of the above models are essentially models of the return-generating process in that they make assumptions about the mechanism by which transaction prices arise through the interaction of the actions of market participants, their information and other characteristics which lead to these actions as well as the mechanisms imposed by the particular trading locale.

Models of intraday price formation describe the change in price in terms of a commonly perceived 'True Price' of the stock as well as innovations in information from public shocks and through the innovation which is implicit in the trade. These models capture many important characteristics of the trading process but in order that they are tractable they need to impose a number of simplifying assumptions and ignore effects associated with strategic trading, irregular time between trades, price discreetness and other microstructure imperfections.

Some of the issues which have not been tackled in the existing literature are:

- Even though the price of each quote has been modeled as a factor affecting price formation little emphasis has been given to the size of the bid and ask quotes which are likely to affect the behavior of traders and thus the return-generating mechanism.
- Similarly, little attention has been paid to the possible imbalances between supply and demand, as evidenced by the schedules of orders over a period preceding each trade. Even though theoretical models have tackled this possibility empirical models have not.
- The volume of each trade has an effect on the perception by traders of the true price affecting the adverse selection component of the spread, which is that part of the spread charged by the market-maker in his attempt to minimize his losses when trading with potentially informed traders. Some models have used ranges of sizes within which they estimate their

parameters but no detailed study has incorporated volume directly into the spread equation.

- The time between trades is important and should be accounted for in the trading models since long periods lacking trades may signal the absence of private information whereas frequent trading may often carry the opposite information.
- The speed of assimilation of information could also be modeled by means of a time-adjustment factor which could in turn depend on volume.
- The behavior of the components of the spread could vary with the liquidity of the particular stock and therefore a thorough analysis should estimate these separately for varying levels of trading activity.
- Apart from the quote-setting behaviour of the dealer a certain amount of information could also probably be revealed by the quote-revision process and if so would affect the order-placement decisions of market participants.
- The pricing strategy of dealers is likely to be affected by the size of the bid-ask spread and the probability of a trade inside the spread. However explicitly incorporating pricing strategies in empirical analysis is difficult.
- Apart from the exogenous factors which affect the order-arrival process, such as macroeconomic factors and traders' characteristics, certain endogenous factors may also affect this process since characteristics of the market may invoke a certain volume of trade irrespective of the usual demand / supply-initiated trading.
- Most models of the spread assume that the market is efficient which implies that new public information is accurately and promptly incorporated into prices. A consequence of this assumption is that there is a 'true price' of the

stock which is the same for all participants. However, this assumption may be unrealistic and one may not be always able to arrive at a full-information value. Therefore different results might be obtained if mathematical models of the spread adopted more than one 'true price'.

- The variance of return per unit volume of trade has not been investigated (this variable can be a proxy for order-arrival rates).

1.2 Contribution of the Thesis

In this thesis an attempt has been made to introduce some of the most important of the above features into existing models of the bid-ask formation. In particular, we have modeled explicitly the volume of each trade by incorporating its actual size directly into the bid-ask equation together with the sizes of the bid and ask quotes, as well as the waiting time between trades. While most research work has centered around one particular exchange we have extended our analysis on both the New York Stock Exchange (NYSE), which operates under a double-auction specialist trading mechanism and the National Association of Securities Dealers Automated Quotations (NASDAQ) System which is a quote driven multiple dealer system. These two systems are representative of the majority of the trading mechanisms under which shares around the world are traded and the parallel analysis contributes to highlighting the differences in our results which arise between the two types of exchange. Moreover our study breaks down our dataset into deciles of trading activity so as to gain an understanding of the differences in our results among firms which have varying degrees of liquidity.

Two models have been developed in this thesis. The first describes the change in the quoted bid-ask spread as a function of a number of variables related to its components while the second attempts to model the change in price after the trade has occurred as a function of the same components. The aim has been to examine the differences in the components of the spread prior to and following each trade.

Finally we have used the Generalized Method of Moments for estimating our models and have carried out specification tests, which scarcely appear in other work in finance which makes use of this method, in order to determine the significance of our results.

The contribution of the research work in this thesis has been to show that:

1. There is an adverse-selection component implicit in the quoted-spread which depends directly on the number of shares traded, unlike previous studies which only use dummy variables for ranges of size. It has also been shown that this effect is asymmetric for the bid and ask sides of the spread. Evidence has indicated that trading volume influences the adverse-selection component of the realized spread but the cost recovered through the trade is small. The results in this thesis show that adverse-selection costs in the realized spread are larger for NASDAQ stocks than for NYSE which is in line with recent evidence by Van Ness, Van Ness and Warr (2002).

2. There is an inventory-holding component implicit in the quoted spread which varies with the trading activity of both NYSE and NASDAQ common stocks. It has been shown that this component, like the adverse-selection one, is not symmetric for the ask and bid sides of the spread. Evidence has also been found that this cost is recovered through the realized spread and thus the price but it is doubtful that such could be completely uncovered in the short horizon which our datasets span.

3. Even though the adverse-selection cost constitutes a large part of the quoted spread it appears to be a very small part of the realized spread. The results of this thesis for both the adverse-selection and the inventory-holding cost show that they are smaller compared to those estimated in other papers.
4. The volume of shares quoted at the bid and ask (depth) plays a significant role in the formation of the quoted spread for both NYSE and NASDAQ common stocks, a feature not considered in earlier work. This effect is found to be symmetric for the ask and bid sides of the spread. Depth also affects the price-formation process of NYSE stocks its effect decreasing with increasing trading activity in the stock. The evidence in this thesis is in line with recent findings which indicate that new information first enters the price-process through the depths of the quotes and not the size of the spread, however a thorough analysis of this effect is provided in this thesis for two separate market mechanisms and also for stocks of varying trading activity.
5. There is a significant fixed-cost-of-trade component in both NYSE and NASDAQ stocks which decreases as trading activity increases.
6. The waiting-time between trades may account for a very small part of the quoted spread but may also play a significant role in the price-formation process of those NYSE stocks which have very high liquidity.
7. The above analysis has been carried out for two different trading mechanisms, at both the pre-trade and post-trade stages, using deciles of trading activity which help determine the patterns of the parameters of the above components.

8. A thorough analysis of the components of the spread cannot be carried out without explicitly incorporating the effect of volume whether it is in the form of shares traded or shares quoted prior to a particular trade.
9. The results in this thesis have demonstrated that, in spite of the competitive nature of the multiple-dealer market of NASDAQ, its dealers overreact to adverse information extracting higher rents from traders compared to what a specialist in NYSE would charge. This behaviour could lead to a liquidity crisis easier than under the specialist system of NYSE where the monopoly power of the specialist and his privileged position, owing to the consolidation of the order flow for a stock through him, could help him moderate his losses. Even though quotes in NASDAQ are competitive, the lack of a limit order book, as was the case in the period where our data originate from, leads to higher effective spreads. However, the ability of NASDAQ dealers to share the risks of inventory imbalances allows them to charge lower inventory costs as has been found by our analysis.

The most relevant papers to our research are those of Huang and Stoll (1997) and Madhavan, Richardson and Roomans (1997). Neither of these papers considers the effect of the quoted volume of shares on the spread. Moreover, the latter uses quotes only in order to determine the trade indicator variable and does not consider the inventory cost effect. The former model assumes that the spread is constant and estimates it using all the data. None of these studies incorporates the volume of shares traded explicitly like we have done and none attempts to study the behaviour of these components based on activity deciles. They also do not contrast the behaviour of the two exchanges in the two separate trading mechanisms. Therefore our results, especially those pertaining to the quoted spread, are not directly comparable to the findings of the above papers since in our models a larger number of parameters is involved with fewer assumptions.

1.3 Structure of the Thesis

Apart from the introduction this thesis consists of six more chapters a description of which is provided below. Chapters two and three review the existing literature on market microstructure, the return generating process and the bid-ask spread. In chapter two a presentation is made of the theoretical underpins and the research carried out regarding certain aspects of the return-generating process such as inventory and dealer models, specialist, single and multiple dealer models as well as pertinent empirical work. Reference is also made to differences between exchange mechanisms, types of market participants, types of orders as well as of the importance of volume in the return-generating process.

In chapter three research on the intraday behavior of the bid-ask spread is presented together with the theories which have been put forward to explain it which are also those explaining the components of the spread to a certain extent. Existing models of the spread and its components are also discussed since they form the foundations upon which our models have been constructed.

Chapter four discusses the design of our empirical approach, analyses the database employed and explains the procedure followed for selecting the data used for the estimation of the models together with the methodology for the classification of trades into buyer- and seller-initiated. Summary statistics, which are presented in Appendix D, are also discussed.

Chapters five and six constitute our two empirical chapters. In chapter five the Quoted Spread Model is developed from first principles, based on the theories discussed in chapters two and three as well as on our own assumptions. To refine our analysis and diagnose differences between the components pertaining to the quoted bid and the quoted ask two more equations are developed and estimated, the Offer- and the Bid-Change equations. Summary statistics of the variables used are discussed together with the results from the estimation of the parameters of the models. Some discussion of the estimation method is made but more details are provided in the appendices. Results of specification tests of the models are also discussed and conclusions are drawn. Tables of the summary statistics, the estimation as well as the specification test results are presented in Appendix G.

A similar procedure has been followed for the presentation of the Price-Change model in chapter six. Based on principles discussed in chapter five the model is developed, summary statistics of its variables are discussed, estimation is carried out together with pertinent specification tests and finally results are discussed and conclusions are drawn. Tables of the summary statistics, the estimation as well as the specification test results of this model are also presented in Appendix G.

Finally in chapter seven the conclusions of this thesis are presented, the limitations of our models and the database used are criticized, the policy implications of the results are discussed and ideas for further empirical research work in the future are proposed.

A number of appendices have been included to clarify our methodology as well as explain the procedures used for extracting the data from our database.

Appendix A presents an outline of the basic principles of the GMM methodology which has been used for estimating the parameters of our models.

Appendix B discusses the principles behind the tests employed for the specification of the models in chapters five and six.

Appendix C outlines the computational procedure which has been followed in order to extract the data for our estimations from the Trade and Quote (TAQ) database of NYSE used in this thesis.

Appendix D presents tables and plots of the summary statistics of the variables in the models for both the NYSE and NASDAQ datasets, split into deciles of trading activity so that their variation can be seen together with their differences from one exchange mechanism to the other.

Appendix E presents some condition and correction codes from the TAQ manual which may be necessary so that the discussion presented in chapter four can be easily followed.

Appendix F provides a literature survey on the various aspects of the mechanisms of limit-orders which are essential for a proper understanding of the price-formation process.

Finally, Appendix G presents all the tables for the summary statistics, the results and the specification tests from the models in chapters five and six.

Chapter 2

Literature Survey

2.1 Market Microstructure and the Bid Ask Spread

Early research on the operation of financial markets has assumed that the latter operate in a frictionless environment, where no transaction costs[1] exist and where all participants share the same information. In such an environment orders transmitted to the market were assumed to be carried out at whatever price was established and there was no explanation as to how that equilibrium price had been arrived at. The availability of transaction data together with advances in theoretical concepts such as the capital asset pricing and the options pricing theories, as well as progress in the study of information economics have stimulated research into a number of aspects of market microstructure which has dealt with the price-formation, or return-generation process under the impact of transaction costs, asymmetric information and traders' risks.

To construct models for the study of market microstructure issues, such as the price-formation process, researchers have employed the characteristics of actual trading mechanisms and in particular the rules which determine how trading takes place in conventional exchanges.

Such exchanges are real markets where traders demand transactions, or liquidity and dealers supply it to them at a cost which compensates them for assuming unnecessary risks when adopting non-optimal inventory positions, for processing traders' orders and for the risk they face of losing to informed traders, that is those traders who have superior information. The way these costs are recovered by the

---

[1] Transaction cost literature has been surveyed by Beebower (1989) and Harris (1990).
Most of the pertinent studies focus on realized trades and as Perold (1988) notes, this ignores the cost of foregone trades. Harris and Hasbrouck (1996) focused on orders and imposed a cost on execution failure, in their attempt to remedy this effect. DeJong et al (1995) investigated the composition of trading costs in a limit order market and a dealer market. Kumar and Seppi (1994) modeled both market and limit orders based on an order-driven mechanism and showed that brokerage costs have important influence on the depth and composition of the limit order book.

dealers is by posting two prices instead of one, a price at which they are willing to buy, the bid and one at which they are willing to sell, the ask.

This chapter together with chapter 3 present the theoretical foundations and pertinent literature on which a study of the bid-ask spread and its components should be based. Whereas chapter 3 presents the literature survey and technical details on some of the models of the bid-ask spread and its components this chapter discusses more general theoretical issues related to the elements of the return generating process.

To facilitate the presentation first the factors affecting the process by which prices are formed are discussed followed by the types of market participants, the most common exchange mechanisms by which trading of stocks takes places and the types of orders available to traders. Particular emphasis then is given to the definition as well as the issues related to the liquidity in the market as well as to price volatility. The main directions which research, both theoretical and empirical, has followed are also presented and particularly those pertaining to the models of the trading behaviour of special types of participants such as dealers and specialists. Finally discussion concludes with the presentation of the issues related to trading volume which is one of the most important variables of the trading process since it is associated with information.

2.2 The Return-Generation Process

A great amount of research in finance in the past thirty five years has concentrated on the efficiency of the market and the predictability of stock returns with the aim of exploiting profit opportunities. This research has mainly focused on price patterns and the statistical analysis of the time-series of stock returns but has dealt only partially with the factors which drive the trading process and lead to the formation of prices, the most important of which are the following:

- The needs and characteristics of traders and dealers like profit, liquidity levels, portfolio re-balancing, wealth constraints, etc.
- The amount and quality of information, and the way it is shared among participants.
- The nature and frequency of demand as expressed by the type and the arrival rate of orders.
- The trading mechanism employed like for example the continuous-auction market with, or without a specialist, the dealer market and the limit order market.
- The transaction costs which can be explicit (commissions, clearing and processing costs) or implicit, through the bid-ask spread.
- The trading strategies of market participants which depend on the timing, type and number of the orders given, the pricing rules traders expect the dealer to use and the varying horizons of players in the market.
- The risk and its sources (transactions rate, future value of the asset, information trading) and the participants' attitudes towards it.

2.3 Types of Market Participants

Market participants in real markets consist of private investors, institutional investors, dealers and specialists. Not all types of participants exist in every market however. In physical exchanges, such as the NYSE, AMEX and the Toronto Stock Exchange, trading is organized around specialists, that is market professionals who function both as principals (dealers) and agents (broker's broker) who are assigned a particular stock. In decentralized, automated exchanges like NASDAQ and the International Stock Exchange (ISE) in London, where the physical presence of members is not necessary, dealers make a market in certain stocks they are assigned by posting bid and ask quotes at regular intervals. Since more than one dealer are allowed to make a market in a stock such exchanges are also called competitive-dealer trading systems.[2]

2.3.1 The Specialist

The specialist in the NYSE has both responsibilities and restrictions set by the exchange which ensure that he does not take advantage of his dual capacity role, thus protecting the interests of the public. As an auctioneer he has to adhere to the order-execution regulations. He is also obliged to make a 'fair and orderly market' which implies that he helps to avoid excessively large and erratic price changes by buying and selling for his own account in his capacity as dealer. This has both the effect of reducing price volatility and making price transitions orderly whenever there is a trend, a process which is referred to as price-stabilization and gives the specialist the opportunity to exploit his information advantage. Moreover, by posting bid and ask prices before seeing the public order flow leads to continuity in prices since public traders can trade at any time with the dealer. As a broker, the specialist executes limit orders left to him by other brokers.[3]

---

[2] In contrast to NASDAQ the London International Stock Exchange is characterized by the domination of its order flow by orders of large institutional investors (large, infrequent trades) since, unlike NASDAQ, the latter are allowed to bring their orders for the largest, most liquid (alpha) stocks directly to the market-makers for immediate filling, instead of having them negotiated in an 'upstairs' market as in the U.S.

[3] Stoll (1985) has pointed out that during block trading in the NYSE the specialist remains a bystander even though he can trade some of the shares in a block either for his own account or for the account of limit orders in the book.

2.3.2 Dealers

The NASDAQ market consists of a large number of dealers and member firms which are linked electronically. This market has simpler order-handling and trade-execution rules and relies more on competitive forces and less on explicit regulations in order to attain a fair and orderly market. Market in a stock is created by a large number of dealers (minimum is two) which helps keep spreads tight but makes the price-discovery process inaccurate. Dealers registered as market makers in a stock are obliged to continuously post firm quotes for at least one round lot at both quotes and honour them.

2.3.3 Traders

In real markets traders can be floor brokers, brokers and the public. Traders can be characterized according to the type of orders they place and the type of information they possess. Those requiring immediate execution can place market orders whereas those wishing to transact at a given price can place limit orders. Limit-order traders in the NYSE in a way compete with the specialist by offering other traders a narrower spread than the one a specialist would allow. The type of information may also determine the trading strategies of market participants. Usually uninformed traders are assumed to trade for liquidity reasons whereas informed traders are assumed to attempt to exploit their information advantage by concealing their presence using various techniques such as breaking down their large orders into smaller ones. Moreover liquidity traders can be split into discretionary and non-discretionary, the former being those who prefer to trade at certain times.

2.4 Differences between Exchange Mechanisms

The two main institutional differences between the hybrid (call-auction / continuous double-auction) exchange mechanism of NYSE and the quote-driven mechanism of NASDAQ are the intraday market power of the specialist in NYSE and the different ways in which inventory is managed under the two exchange settings. These differences have a significant impact on the formation of the bid-ask spread and thus on transaction costs.

Specialists have an informational advantage over other traders near the open and close of trading that is not shared by NASDAQ market makers. After the Opening Automated Report Service (OARS) of NYSE automatically matches buy and sell orders, the specialist offsets any remaining order imbalance from inventory, i.e. he uses knowledge of market and limit orders in setting the opening price. Stoll and Whaley (1990) conclude that the ability of the specialist to set a low (high) bid (ask) price when the opening order imbalance requires purchases (sales) from inventory results in monopoly profits and excess volatility.

No formal opening procedure is available on NASDAQ where the search for an equilibrium price may force dealers to post wider intraday spreads near the open as protection against the expected losses from facing informed traders, and/or to trade in relatively small units until equilibrium prices are determined, a procedure often referred to as price-discovery.

The specialist can also take advantage of the 'market-on-close' orders that are routed to him in order to learn about the desire of the market to trade and can thus widen his spreads (investors who manage portfolios may exhibit inelastic demand during the open and close; fund managers may prefer to trade near the close since closing prices are used to calculate net asset value for sales and redemptions). Thus if the increased width of inside spreads at the open and close in a specialist market reflects the exploitation of inelastic demand of investors to trade, replacing the specialist with multiple market-makers would reduce the tendency of bid-ask spreads to widen near non-trading intervals (the opening and closing periods of the market).

Inventory management considerations may also affect the width of spreads near the close. Hasbrouck and Sofianos (1993) attribute this to the desire of traders to hedge positions before the close whereas Amihud and Mendelson (1982) predict that the width of the bid-ask spread increases to discourage additional accumulation of inventory when its position deteriorates.[4] However, in NASDAQ dealers who accumulate excess inventory may post bid and ask quotes to attract orders which would otherwise go to competing dealers, thus narrowing the inside spread.

Hasbrouck and Sofianos (1993) and Madhavan and Smidt (1993) suggest that although specialists' inventories are stationary, inventory imbalances are reversed over a number of trading days. Thus specialists appear to require more than a single trading day to control inventory through prices perhaps due to their obligation to maintain an orderly market, which forces them to accept trades on both sides of the spread. Thus the dealer market may provide market makers with a more effective mechanism to control inventory through prices than specialists enjoy on the organized exchanges.

2.5 Types of Orders

The two main types of orders available in organized exchanges, in particular market and limit orders, as well as their characteristics are discussed below. These are important in that they convey the traders' wishes to the market and constitute the main vehicle through which investors' strategies are materialized.

A more thorough discussion of limit orders, which even though more complicated are important for strategic trading, is given in Appendix F.

2.5.1 Market and Limit Orders

Orders are mainly characterized by the price at which they are executed and the time at which they are to be executed. There is a large number of orders allowed in various exchanges but the most common among the price orders are market and limit orders.[5] Market orders are executed at the best available price (highest bid for seller and lowest offer for buyer). Due to their nature these orders demand immediacy of execution and for this reason they consume liquidity. Their placement guarantees execution (almost always) but not the price and for this reason they are preferred by the brokers since their commission is almost always guaranteed. However, in cases when the bid-ask spread is excessively high the cost of the order to the trader is also high.

2.5.2 The Importance of Limit Orders

Limit orders are one of the most important types of orders available in most exchanges and in particular the NYSE. Harris and Hasbrouck (1996) report that they constitute about 45 % of the orders submitted through SuperDOT.[6] [7] For the Paris Bourse, which operates as a limit order market, Biais, Hillion and Spatt (1995) report that order flow consists of 41.3 % limit orders and 47.2 % market orders.

---

[4] Harris and Hasbrouck (1992) provide an alternative explanation for the wider spreads near the close for NYSE based on the depth of the limit order book.

[5] Other types of price orders include stop, stop-limit, on the opening, on the close, immediate or cancel, fill or kill, all or none and discretionary orders. Time orders include day and good till canceled orders. Other types of orders, characterized by the conditions attached to them, are also used.

[6] SuperDOT is an electronic system which handles the bulk of individual market and limit orders submitted to the NYSE by routing approximately 85% of all orders entered to the trading posts. Its name derives from its predecessor system DOT which is an acronym for Designated Order Turnaround (system).

[7] The rest consist of 50 % market and 5 % stop orders.

The importance of limit orders lies in their use as providers of liquidity in exchanges where market orders are allowed and which consume liquidity in order to satisfy the demand of investors for immediacy. Traders place limit orders in order to obtain prices better than if they had submitted a market order, in which case they would have paid a cost proportional to the spread. Moreover, these orders free the traders from watching the market. Their execution is not certain since for a limit order to execute, prices have to move to the direction of the limit price. Contrary to market, limit orders can guarantee a price but not a quantity since not all of the quantity of the limit order can be satisfied whereas market orders guarantee the quantity, in a sufficiently deep market, but not the price. However, limit orders provide liquidity to the market and help it move to the direction of the market orders. They may then compete with the market-maker who also serves this function. In a market which is not automated however the trader runs the risk of "missing the market".[8]

A more detailed discussion, which tackles the crucial questions concerning market microstructure research on limit orders and presents the most important findings, is given in Appendix F since it is believed that knowing the characteristics of limit orders is crucial for a deep understanding of the mechanisms involved in the price-generating process.

2.6 Liquidity

Liquidity is the ability to trade large size, quickly at low cost and is the most important characteristic of the markets.

It is the result of a bilateral search process where buyers and sellers attempt to arrive at mutually acceptable terms so that a trade can take place. While patient traders can

[8] In contrast to the NYSE limit order book of a specialist who is assigned a particular stock exclusively, in NASDAQ there is no consolidated limit order book for each stock. Under the more recent NASD rules (IM-

supply liquidity by placing limit orders for anxious market-order traders, there exist large traders willing to offer it when asked in which case it is called latent liquidity.

Liquidity comprises of three main dimensions: Immediacy, width and depth. *Immediacy* refers to how quickly trades of a given size can be carried out at a given cost. *Width*, also called market depth, is the cost of effecting a trade of a given size. For small trades it is equivalent to the bid-ask spread plus the brokerage fees. It essentially describes the cost per unit of liquidity. Finally *depth* refers to the size of a trade that can be arranged at a given cost and describes the units which are available at a given price of liquidity. Depth and width are closely related notions. Another dimension of liquidity, often encountered in the pertinent literature, is *resiliency* which refers to how quickly prices revert to former levels after they change in response to large order flow imbalances which are initiated by uninformed traders.

Since liquidity is the ability to trade it can be characterized as a function that gives the probability of trading a given size at a given price, given the time available for search. This probability is influenced by the type of instrument traded, the types and the quality of traders in the market, the quality of information regarding fundamental values, imbalances between buy and sell orders, the actions of other market participants as well as the state of the market that is whether it is open or closed.

The types of traders rather than the types of orders determine whether liquidity is supplied in the market. The main liquidity suppliers are the dealers making a market and the value traders, who both act as passive liquidity suppliers and the pre-committed liquidity suppliers who place limit orders so as to reduce their trading costs but can easily become liquidity demanders if their limit orders are not fulfilled. Another not so evident liquidity supplier is the arbitrageur who can increase the depth

2110-2 and 2320), however, limit orders of customers which improve (reduce) the spread must take precedence to dealer's quotes.

in a market by offering liquidity from another market and in this capacity he competes with dealers.

One of the most important issues in the design of market mechanisms is the way in which various mechanisms differ in their provision of liquidity.

Whereas when the aggregate order flow is considered the bid ask spread is an adequate measure of liquidity, when price varies with trade size large trades may involve larger spreads than small trades in the same market and the bid ask spread cannot be an accurate measure of liquidity. A more accurate measure may be derived from the fact that in liquid markets prices move after trades reflecting the costs of trading. This implies a time series dimension of liquidity which is based on the fact that traders willing to wait can obtain better prices compared to those who demand immediate execution. Grossman and Miller (1988) made use of this time dimension of liquidity to construct a model in which prices are determined by inventory imbalances and there is no private information . Prices change because of liquidity shocks. Market makers are assumed to act as speculators who take positions in the risky asset but they do not quote bid and ask prices. Liquidity in this market is provided by speculators and the price of this liquidity depends on the price movement, the total number of speculators in the market and the common risk-aversion coefficient. The equilibrium number of speculators in turn depends on the cost of being a speculator, the risk-aversion coefficient and the variances of the prices and the liquidity shocks. The greater the number of speculators willing to provide immediacy the greater the liquidity in the market. Since the return to speculation depends on the variance of prices, more volatile markets attract more speculators but require higher returns to compensate them for the risk they undertake. A consequence of this is that if speculator returns can be improved, or the speculators' risk-bearing ability increased, liquidity would be enhanced. Since liquidity depends on the number of traders and the number of traders depends on liquidity there appears to be circularity which complicates this issue.

Pagano (1989) has investigated whether the fact that the scale of trading affects liquidity is against the existence of multiple markets. In a model without asymmetric information he concluded that in the absence of transaction costs the larger market, in terms of which market the trader believes attracts more traders, will dominate trading. If transaction costs exist traders will base their decisions both on the level of transaction costs and the size of the market but the continued existence of the two markets cannot be guaranteed since it is likely that traders will eventually consolidate. In the absence of informed trading liquidity differences between markets can arise and persist as a result of differences in the portfolio of traders and the existence of exogenous transaction costs.

2.7 Volatility

Volatility is the tendency of prices to change unexpectedly. Volatility changes through time and can become very dangerous when large price changes appear in short intervals of time (episodic volatility). When there is excessive volatility markets do not function properly because the prices of the assets in the market do not accurately reflect their value and decisions based on these can be risky. Excessive volatility can also lead to market crashes and for this reason it is of great concern to regulators who attempt to design markets in such a way as to avoid it.

Moreover, volatility is highly related with the risks and the profits in the market and is of great concern to option traders and technical traders. It is often correlated with trading volume but this relation is complicated.

There are two types of volatility. *Fundamental volatility* which is caused by unanticipated changes in the values of instruments and *transitory volatility* which is due to trading activity by uninformed traders. It is important for traders to be able to predict future volatility so that they can assess the profitability of different strategies and estimate their likely transaction costs. Moreover, regulators should be able to remedy transitory volatility using appropriate policies which they can incorporate into the trading mechanisms.

Since prices are used in the economy to allocate resources it is important that they reflect fundamental values. These values change when factors which affect them, called fundamental factors, change. Such factors can be the quality of the management of a company, its resources and level of technology, supply and demand conditions for its resources and products as well as interest rates. Dealers try to infer information about fundamental values from the order flows of traders who they perceive to be informed. In this way they charge a higher price to compensate themselves for the risk of transacting with informed traders increasing the adverse selection component of the spread, which will be presented in the next chapter and contributing to fundamental volatility. When the number of traders having fundamental information is small they attempt to take advantage of their privileged position by trading in large quantities quickly thus increasing volatility.

Fundamental volatility is unpredictable since it is caused by unexpected events. Other factors which lead to fundamental volatility is the uncertainty about fundamental factors, unexpected political action, uncertainties regarding the earnings of high price-earning stocks and the high leverage of firms which makes earnings risky.

Contrary to fundamental, transitory volatility is usually brought about by the actions of impatient uninformed traders who cause prices to diverge from their fundamental values. Prices eventually revert to their fundamental values and their divergence, together with the subsequent price-reversal, contributes to this type of volatility. The simplest form of transitory volatility is due to the bid-ask bounce which arises when trades alternate between the bid and the ask quotes. The bid ask bounce is what causes part of the spread (the transitory cost component which is explained in the next chapter). Large orders and cumulative order imbalances also cause transitory volatility but these are eventually corrected by value traders or arbitrageurs.

Transitory volatility is correlated with the transaction costs of uninformed traders. It is the result of their own impact on prices and is small in liquid markets. When transitory volatility is high markets become illiquid and regulators attempt to remedy this problem by taking appropriate actions. However they have to become certain that it is transitory volatility which causes these problems otherwise their efforts will be futile.

It follows from the above discussion that to measure volatility in a market one has to estimate both transitory as well as fundamental volatility. For the transitory component this is usually accomplished by making use of successive price changes which in the case of uninformed liquidity demanders are negatively correlated since they reverse over time. The horizon during which this occurs may vary from short, in the case of the bid-ask bounce, to periods of minutes or even months for the case of large orders and large order imbalances. However, fundamental price changes may or may not be correlated with the order flows of informed traders owing to the multiplicity of trading strategies which they employ so that they may conceal their information advantage. Since fundamental price changes lead to revisions in beliefs regarding the value of assets they normally do not revert.

2.8 Avenues of Research

There are two general directions along which research in the price-formation process has proceeded. One which has centered around the inventory position of the dealer / specialist and another dealing with the information aspects of trading. This latter approach has produced models which have been called non-asymmetric in contrast with the former approach which has assumed that all market participants share the same information and has led to the formulation of symmetric models. All of the above classes of models depend on the presumed quote-setting behaviour of the dealer.

A) In the inventory-related research there have been three main avenues:

The first is based on the Walrasian[9] auction paradigm of a trading mechanism where demand / supply schedules are submitted to an auctioneer who aggregates them, posts a tentative price in order to tempt traders to revise their orders, repeating the process until an equilibrium price is reached at which trading takes place. Even though this setting may seem unrealistic, it approximates some real markets. What it does not however do is to explain (i.e. it is agnostic) the way by which the factors described above lead to price-formation. Demsetz (1968) was the first to directly examine the factors affecting trading, even though he focused more on the trading costs' aspect of the process. In a study of the 'cost of transacting' he put forward the idea that liquidity has a cost which leads to the formation of the bid-ask spread. By showing how the non-synchronous arrival of buy and sell orders affects market prices he set the foundations of further work on the microstructure of the markets.

[9] Leon Walras (1834 – 1910), a French economist, first thought of buyers and sellers of a resource being brought together by an auctioneer who calls out tentative prices in a sequence of trials until a clearing price is reached at which buyers and sellers trade optimally (tatonnement process).

The second avenue pays less emphasis on trading costs and is founded on Garman's (1976) pioneering paper which examined the nature of the order flow and its effect on security prices. Advancing beyond the Walrassian paradigm he attempted to study the factors behind the process, that is how the collective behaviour of agents in the market affects market behavior. He analyzed the position of a dealer when he is faced with the stochastic arrival of buy and sell orders which demand immediacy of execution while trying to keep his inventory within certain constraints so as to avoid bankruptcy.

A third avenue of research, pioneered by Cohen, Maier, Schwartz and Whitcomb (1981), examined the reasons for the strategic choice of alternative types of orders on the liquidity of the market.

All of the early inventory-control, or inventory models described above have focused on the attempt of agents to keep their inventories within certain constraints, thus minimizing the risk of bankruptcy. When the market-maker faces inventory carrying cost or is risk averse he will actively control his inventory by setting prices to induce movements towards his desired inventory level. Therefore, the effect of his behaviour on prices is temporary and reverts to the true price when order flow is balanced. In this way only the short-term characteristics of the process are given attention in contrast with the impact of information which has a more permanent, longer-term effect. These models focus on the problem that in the simple bid-ask model the dealer's inventories of stock and cash follow a random walk leading to large positive or negative positions.[10] This happens because the dealer is not allowed to have negative inventory and carrying costs (non-negativity constraints). They thus predict that the dealer will set quotes as an inventory-control mechanism, to induce an

[10] Garman (1976), Stoll (1976), Amihud and Mendelson (1980, 1982), Zabel (1981), Ho and Stoll (1980,1981, 1983), O'Hara and Oldfield (1986).

imbalance of buy and sell orders, in order to maintain his inventory around some optimal level.

B) Apart from the inventory concern of the dealer information also affects his behaviour in a different manner leading to a distinct, permanent effect on prices. There have been two approaches in the early non-asymmetric-information microstructure literature :

1) Research which assumes that the return-generation process is consistent with our general understanding of the market structure which, however, is not necessarily based on any principles of individual economic behaviour. This work defines the relationship between the observed distribution of returns and the 'true' underlying distribution [11] and
2) Research that specifies the demand and supply for dealer services whose compensation is reflected in a transaction price different from the true price.[12]

Asymmetric Information Models, emphasize the importance of asymmetric information in analyzing market-maker behaviour. In these models the perceived presence of informed traders with private information regarding fundamental asset values affects price dynamics and the size of the bid – ask spread. Based on an idea by Bagehot (1971), these models drop the assumption of uniform information and address the problem of an uninformed dealer who by offering to buy and sell at the

[11] Clark (1973), Blattberg and Gonedes (1974), Westerfield (1977), Oldfield, Rogalski and Jarrow (1977), Schwartz and Whitcomb (1977), and Goldman and Beja (1979).
[12] A review is provided in Cohen, Maier, Schwartz and Whitcomb (1979); Demsetz (1968), Tinic (1972), Stoll (1978a) dealt with the supply side; Tinic and West (1972), Benston and Hagerman (1974), Stoll (1978b), Smidt (1979) carried out empirical studies; Copeland (1976), Epps (1976), Cohen, Maier, Schwartz and Whitcomb (1978a,b) and Goldman and Sosin (1979) dealt with the demand side and specified why investors trade and how trading affects prices.

quoted prices suffers exposure to informed traders.[13] The dealer views orders as originating from an informed trader with some probability and thus convey information and motivate quote changes. Inventory considerations are usually ignored.

The inventory control and the asymmetric information theories should not be considered as being mutually exclusive since they yield similar predictions about asset returns and volume.

2.9 Inventory and Dealer Models

To facilitate the presentation of the models discussed, a distinction is made below between models which deal directly with the inventory concerns of the dealer, be it a specialist or other type of dealer and those which deal exclusively with the specialist.

2.9.1 Inventory Models

The fundamental idea that immediacy has a price which anxious buyers and sellers should pay was introduced by Demsetz (1968) who first employed a supply / demand framework to show that immediacy has no price if buy and sell orders could arrive simultaneously at the market in which case no waiting is required. In his model the spread measures the price of immediacy in both buying and selling securities.[14] The main force working to reduce the spread is thus the time rate of transactions since the greater the frequency of transacting the lower the cost of waiting and thus the spreads which traders are willing to submit. Potential waiting costs dominate security trading and together with the risk of adverse price changes they are relatively more

---

[13] Bagehot (1971), Copeland and Galai (1983), Glosten and Milgrom (1985), Glosten (1989), Easley and O'Hara (1987). Kyle (1985) presents a model of the strategic problem faced by the informed investor; also Admati and Pfleiderer (1989)

important than commissions[15] as far as exchange members are concerned. Due to competitive forces the spread should remain close to this waiting cost.[16] However he did not explain why anxious buyers can lead to a price increase or why there is not an infinite supply of the asset at the equilibrium price.

In the models of Garman (1976) and Amihud and Mendelson (1980) results depend on the assumptions made regarding the arrival rates of buy and sell orders, the inventory as well as the pricing policy of the dealer. Therefore, since in both models the arrival of orders is stochastic it creates uncertainty due to potential imbalances (transactions uncertainty) and is not informative about future price movements. In both models the dealer attempts to control his inventory by setting bid and ask prices around the equilibrium price but whereas Garman's (1976) dealer runs the risk of becoming bankrupt (in the extreme case when buying and selling prices are equal to the equilibrium price), Amihud and Mendelson (1980) avoid this problem by imposing bounds on the upper and lower levels thus preventing the market-maker from failing. In Garman's (1976) model the inventory per se plays no role in the dealer's decision problem since it was assumed that prices are set at the beginning of trading in order to equate the order-arrival rates. Therefore, where prices are set depends on the profit-maximization objective of the dealer rather than inventory. In the second model, however, where the optimal pricing policy of the market maker depends on the inventory, it was shown that, consistent with the dynamic pricing policies described by Smidt (1971), Barnea and Logue (1975) and Stoll (1978a), the optimal bid and ask prices are monotone decreasing functions of the dealer's

---

[14] As he put it, "the bid-ask spread is the markup that is paid for predictable immediacy of exchange in organized markets".

[15] Commission brokerage fees are not determined by the same procedures.

[16] The main types of competition were identified as: (1) rivalry for the specialist's job, (2) competing markets, (3) outsiders who submit limit rather than market orders, (4) floor traders who may bypass the specialist by crossing buy and sell orders themselves, and (5) other specialists.

inventory position. The existence of positive costs of providing dealership services leads to a positive spread which straddles the market clearing price. Amihud and Mendelson (1980) showed that the preferred bid-ask spread is always greater than the corresponding spread set by Garman's (1976) monopolist (it also straddles the market clearing price at the intersection of demand and supply). It is not optimal for the profit-maximizing market-maker to refrain from making buy or sell transactions, unless he reaches his limiting positions. Amihud and Mendelson (1980) also showed that relaxing Garman's (1976) constraints by expanding the allowed short or long positions strictly increases the market-maker's profits.

Moreover, consistent with later evidence provided by Madhavan and Smidt (1993) and Hasbrouck and Sofianos (1993), which is presented below, the market maker adopts a pricing policy which produces a preferred inventory position, located away from his limiting positions and he reduces the spread as he approaches the preferred inventory position. Any trading rule based on monitoring the behaviour of the market maker (i.e. knowing his current inventory position and his pricing policy) was shown to be useless and certain to produce a loss thus proving that the optimal pricing policy is consistent with the efficient market hypothesis.

Both the Garman (1976) and the Amihud and Mendelson (1980) models dealt with liquidity-motivated traders only since their assumptions precluded the existence of private information.

By treating the collection of market agents as a statistical ensemble Garman (1976) ignored attributes of traders such as motivations, tastes, and endowments and since buy and sell orders were taken as given he did not use economic features which would allow him to model the behaviour of the dealer more realistically. He examined both call and continuous markets[17]and his results on the auction market closely resembled the empirical findings of Niederhoffer and Osborne (1966), which had indicated that the double-auction microstructure might explain these early empirical findings. Garman (1976) did not however compare the outcome of his model to the real characteristics of prices nor did he consider the optimal dynamic strategies which market-makers should pursue to adjust their prices to their changing inventory positions and did not discuss how the price-probability functions he used depend on previous transaction prices in real world exchanges.

Another important point to consider is that the spread in the Amihud and Mendelson (1980) model reflects the market power of the dealer and the transaction costs but plays no role in the viability of the market. The preferred inventory position arises because of the nature of the order-arrival process whereas the underlying asset value is irrelevant. Finally, to gain additional insight into the price-setting problem would require greater emphasis to be placed on the nature of the dealer's decision problem and this would make it necessary to depart from the simple stochastic process approach of Garman (1976) and Amihud and Mendelson (1980).

2.9.2 Empirical Research

Due to the non-availability of databases which distinguish market-maker transactions from those of the other participants very few of the early studies were based on actual trading records of market-makers in any financial markets.[18] Moreover, asymmetric-information was ignored and the entire dynamic relation between trades and prices was attributed to inventory-control. The information effect

[17] He noted that, as exchange transaction volume increases, markets tend to evolve from 'call markets' (trading synchronously at pre-established discrete times) to 'continuous markets' (trading asynchronously during continuous intervals of time) and for this reason he examined two cases: one where the market maker was assumed to be a monopolist and the other where an auction market exists.

was usually modeled as a static component of the bid-ask spread, which is unsatisfactory because in the short-run both effects can have a similar impact on prices. Other studies that recognised both effects[19] found only weak evidence of short-run inventory effects but strong information effects and suggested that inventory behaviour over longer horizons should be examined.

Madhavan and Smidt (1993) analyzed, both theoretically and empirically, the trading behaviour of a market-maker (or specialist) on the NewYork Stock Exchange (NYSE) who, in order to perform his dealership function, must bear unwanted inventories. [20] In a transactional-level study Hasbrouck and Sofianos (1993) analyzed empirically inventory-adjustment, price determination, and trading profits of the NYSE specialists using both inventory and asymmetric information concepts, even though they relied more on the former.[21]

In their model Madhavan and Smidt (1993) showed that as a dealer, the specialist quotes prices that induce mean reversion towards a target inventory level but as an investor, he chooses a desired long-term inventory based on portfolio considerations, and may periodically revise this target. They found that specialist inventories exhibit mean reversion, but the adjustment process is slow if his desired inventories are assumed constant.[22] Shifts in desired inventories could account for the apparent slowness of the inventory-adjustment process and the apparently weak inventory effects found in the literature. Strong support was also found for the hypothesis that

---

[18] Smidt (1971), Stoll (1985), Stoll (1976) and Ho and Macris (1984).

[19] Hasbrouck (1988), Madhavan and Smidt (1991, 1993), Stoll (1989) and Neuberger (1992), Foster and Viswanathan (1993).

[20] Their intertemporal model of market-maker trades and quotes incorporated the effects of both asymmetric information and inventory control .

[21] In the literature, the dealer's inventory control problem becomes a pricing problem because active price control is necessary to keep inventories stationary. Amihud and Mendelson (1980) solve this pricing problem of the monopolistic dealer, Ho and Stoll (1983) solve the problem for multiple dealers and O'Hara and Oldfield (1986) consider the effects of day boundaries on the dealer's pricing problem.

quote revisions are negatively related to specialist trades and positively related to the information conveyed by unanticipated (non-block) order imbalances. Large-block trades appeared to convey little information to the specialist who appeared to possess market information unavailable to most traders. Future order imbalances were also found to affect current price quotations.

The adjustments in inventory found by Hasbrouck and Sofianos (1993) were also slow and inventory autocorrelations were positive and persistent over long lags (in some cases weeks). Since specialists are capable of rapid inventory adjustments, the observed long-term fluctuations in inventory were considered to be the result of changes in long-term investment positions, a conclusion similar to that reached by Madhavan and Smidt (1993). Adjustment was shown to be faster for frequently traded stocks and inventory levels in at least some stocks contained long-term (weekly or longer) components which were attributed to shifts in the desired level of holdings (speculative positions). Adjustment towards a target level was shown to take place at a slow rate and to be faster after exogenous shocks. Specialist profits were found to be almost statistically indistinguishable from zero in the long term. The short- and medium-term components of trading profits, however, were strongly positive and often statistically significant which confirms that specialist trading profits are due to short- and medium-term trading activity. They found evidence that unexpected shocks to the specialist's inventory affect quotes and, since inventory shocks are inversely correlated with the signed trade variables, this finding was consistent with asymmetric information effects.[23]

---

[22] It takes, on average, over 49 trading days for an imbalance in inventory to be reduced by 50% but if a correction is made for periodic, unobserved shifts in desired inventory holdings, half life drops to 7.3 trading days.

[23] No evidence of transient quote overshooting or reversion, as predicted by the classic inventory control mechanism, was found.

Using VAR models[24] Hasbrouck and Sofianos (1993) found that trades in which the specialist participates are associated with larger quote revisions which according to the authors may reflect institutional constraints requiring participation in high-impact trades, or that specialist participation may serve as a proxy variable for the direction of trades that occur at the midpoint of the quotes and cannot thus be signed.

Using data from the Institute for the Study of Security Markets (ISSM) and a NYSE specialist firm Madhavan and Smidt (1991) found strong support for the existence of information asymmetries as perceived by the specialist and evidence that he provides liquidity by absorbing transitory order imbalances. His participation was shown to increase with the degree of information asymmetry, which is a major determinant of this activity, possibly because trades excluding the specialist are easier to arrange when traders agree about asset value. They also found a weak inventory effect, that trades in active stocks have smaller price impact than those in less active ones and that buyer-initiated and seller- initiated trades have an asymmetric price impact.

2.9.3 Dealer Models

The literature on dealer models has developed from the early analysis of the position of a single dealer to the more complicated case of a number of dealers operating in the same market for a stock. In this section specialist models, which are a special case of single-dealer models, are also presented.

[24] Hasbrouck (1991a, 1991b, 1993) discussed microstructure applications of VARs. They analyzed the differential effect of specialist participation in transactions: the asymmetric information effect suggests that a trade will have an impact on the price irrespective of the counterparty's identity whereas an inventory control effect should be evident only in those instances in which the specialist participates.

2.9.3.1 Single-Dealer Models

Single dealer models have been developed by Demsetz (1968), Tinic (1972), Garman (1976), Stoll (1978a), Amihud and Mendelson (1980, 1982), Copeland and Galai (1983), and Ho and Stoll (1981). However, early empirical work which attempted to assess the efficiency of different market organizations and regulatory constraints and to determine the factors underlying dealer costs[25] had not been based on very explicit theoretical foundations.

Stoll (1978a) has dealt with the supply of dealer services[26] whereas Ho and Stoll (1981) with the demand side. In both models the spread does not depend on the dealer's inventory position but its placement affects the pricing policy and its size remains constant. Moreover, no assumptions were made about equilibrium asset prices.

According to Stoll (1978a), the technical details of which are presented in chapter 3, the cost of immediacy to the dealer is the sum of a) holding costs (increasing with decreasing order costs) which arise from holding a suboptimal portfolio position and thus reflect his degree of risk exposure, b) order-processing costs (a constant dollar amount per transaction), which depend on the type of trading mechanism employed and c) information costs arising from asymmetric information (adverse selection, assumed independent of transaction size). Holding costs depend on dealer characteristics such as initial wealth and relative risk aversion, the size of the transaction in the stock, characteristics of the stock like variance of return and covariance of return with the return on the initial trading-account portfolio, and finally on the size of the initial position in the trading account.[27] When the cost function is

[25] Demsetz (1968), Institutional Investor Study Ch. 12 (1971), Tinic (1972), Tinic and West (1972) and Benston and Hagerman (1974).
[26] Copeland (1976) and Epps (1976) also attempted to specify the demand for dealer services.
[27] Easley and O'Hara (1987a) assumed adverse selection costs to increase in trade size.

symmetric, i.e. a sale of a given size costs the dealer the same amount as a purchase of this size, the spread (but not the bid or ask price) is independent of the initial inventory. The cost of dealer services is also affected by the organization of dealers and the equilibrium number of dealers is that number that demand can support if all are operating at minimum average cost. Stoll's (1978a) inventory affects the placement of the spread but not its size, a hypothesis which was later tested by Hasbrouck (1988) and his spread is based on the market-maker's risk aversion, contrary to the model of Amihud and Mendelson (1980) in which it was based on the market power of the specialist and that of Garman (1976) in which it was based on the dealer's effort to avoid bankruptcy.

When transactions uncertainty (stochastic transactions as in Garman (1976)) is added to the above model, as was done in Ho and Stoll (1981) the dealer is faced with two types of risk: uncertainty about when future transactions will occur and uncertainty about the return on his inventory and the rest of his portfolio. Then based on his opinion of the 'true' price, which is exogenously determined by his information set (the public's opinion of the true price influences arrival rates of buy/sell orders together with the pricing strategy of the dealer), the optimal bid and ask prices are set by the dealer in such a way as to maximize the expected utility of his terminal wealth.[28]

The bankruptcy problem of Garman (1976) can be avoided when the true price is assumed to remain constant for a sufficiently short time horizon. Then the spread, as in Stoll (1978a), does not depend on the dealer's inventory position, but price

[28] Unlike the risk neutral models of Garman (1976), Amihud and Mendelson (1980) and Mildenstein and Schleef (1983) the Ho and Stoll (1981) model did not require the dealer to maximize expected profits. The key concern was the risk he faces and how this affects his willingness to supply dealer services. In their model however, the risk attitude of the dealer affects the solution. Garman (1976) was more concerned with the equilibrium price of securities (rather than with the equilibrium price of

adjustment does depend on inventory (the dealer affects the order arrival process by changing the placement of the spread relative to the true price rather than increasing / decreasing the spread itself). The dealer's risk is greater than that implied by Stoll (1978a) because uncertainty of the demand to trade with the dealer is not totally eliminated or offset by the dealer's pricing strategy.

The validity of the above models is limited by the unrealistic assumptions regarding liquidation of the stock at some known point in time (this reduces the underlying inventory risk), the fixed 'true' price of the stock and the exclusion of informed traders (due to the assumption that the order flow follows a specific stochastic process) who would affect the order flow differently.

In Amihud and Mendelson (1982) only liquidity traders are considered and traders who think they are informed but are in fact not informed and the market maker is only a trader, not an investor. His optimal pricing policy consists of all the prices he will quote at all possible inventory positions.[29] The bid-ask spread is positive and the change in prices is a gradual and monotone function of the market maker's inventory.[30] The model's results also supported the suggestion of other researchers like Smidt (1971) , Logue (1975), Barnea and Logue (1975) and Stoll(1978a) of a preferred inventory position which implies negative serial correlation in the transaction by transaction returns at this position and positive when he is away from the preferred level until the preferred position is restored. The model showed that if the market-maker is allowed to take a greater long or short position, the spread will decrease which may indicate that encouragement of market-makers to take larger

---

services) and with the necessary behaviour of transaction prices to meet an exogenous inventory objective of the dealer and avoid bankruptcy.

[29] Explicit expressions for the market-maker optimal policy are given in Amihud and Mendelson (1980).

[30] This analysis assumed that demand and supply are stationary. When demand and supply shift, price adjustment may follow a different pattern.

positions by exchange authorities can reduce spreads. For a given inventory constraint and for linear demand and supply functions, the spread is narrowest at the preferred position.[31]

When the assumption of a specific stochastic order-flow process or an 'intrinsic value' of the stock is dropped as in O'Hara and Oldfield (1986) the market-maker faces order-flow and future-value (inventory) uncertainty. A risk-averse market-maker's spread can be decomposed into three parts: one for the known limit-orders, a risk-neutral adjustment for the expected market orders and a risk adjustment for market-order and inventory-value uncertainty. Under more realistic assumptions, in a market which is a combination between a call and a continuous auction,[32] since two types of uncertainty are considered, a risk-averse market maker may set a smaller spread than a risk-neutral specialist and may thus not be always driven out of the market. O'Hara and Oldfield (1986) also showed that, contrary to the finding of Ho and Stoll (1981), Zabel (1981) and Bradfield (1982),[33] the market maker's inventory affects both the placement and the size of the spread.[34] With both types of uncertainty the inventory was shown to affect both the placement and the magnitude of the spread. The pricing policy changed to reflect this dual uncertainty and the spread size was allowed to vary for more flexibility. This complexity had not been introduced in earlier work.[35]

The size of the optimal spread depends on the parameters of the supply and demand functions. With market order variability these functions are not known, as

[31] The Amihud and Mendelson (1982) model, together with Mendelson (1982), may be useful in comparing various methods of trading and exchange.
[32] As discussed in Cohen, Maier, Schwartz and Whitcomb (1981) the traders' optimal order strategies may differ between call and continuous markets.
[33] Zabel (1981) and Bradfield (1982) also developed discrete-time models of dealer behaviour.
[34] The spread equation had no explicit inventory terms but included the marginal utility term which implicitly contains the inventory variables.

they are in Demsetz (1968) and the risk-averse market-maker adjusts the spread to compensate for what could be large divergences in either supply or demand. The placement of the spread was shown to depend on both today's supply and demand and on the future expected value of the security. If the latter is variable, the market maker adjusts bid and ask prices to compensate for what could be large movements in tomorrow's security prices.[36]

All of the models described above have the common feature that inventory imposes some cost on the dealer and this is reflected in market bid and ask prices. It is also assumed that the dealer acts as the sole provider of liquidity, an assumption which is dropped in later models.

2.9.3.2 Multiple-Dealer Models

The following discussion distinguishes between theoretical and empirical work.

Theoretical models of dealer behaviour under competitive conditions like Ho and Stoll (1980, 1983) have attempted to analyze the problem of determining the market bid-ask spread. Earlier research like Garbade and Silber (1979a,b) or Cohen, Maier, Schwartz and Whitcomb (1978a,b and 1979) have either tackled isolated issues of this process or have not attempted to fully specify the determinants of the market spread. The transition from individual spreads to the market spread is not a simple aggregation process (like from individual to aggregate market demand and supply functions). The market spread is the outcome of a dynamic interaction involving many market participants. Thus market-spread models assume that the action of each dealer

---

[35] Inventory imposes two types of risk on the dealer: variability in market orders i.e. the size of inventory at day-end is uncertain, and price is a random variable i.e. the value of the inventory is also unknown.

in each period depends not only on his own characteristics and inventory position but also on the inventory position of his competitors as well as his competitor's other characteristics and attempt to determine optimal reservation selling and buying fees[37] for each dealer. However, due to the complexity of this analysis only two dealers have been considered. They also derive equations for the market spread and the market reservation spread (the minimum spread that could exist without either dealer being worse off).

Another issue that arises in multiple-dealer markets is that of inter-dealer trading which arises if a dealer would rather trade at the existing market bid or ask price than remain with his own bid or ask and take a chance on trading with an incoming market order.

Using the very restrictive assumption about homogeneous expectations, Ho and Stoll (1980) showed that when there is no inter-dealer trading and dealers have identical inventory positions, the less risk averse dealer can offer the lower buying or selling fee. Because inventory positions may differ, the better (less risk averse) dealer does not always quote the better buying and selling fee. The dealer with the lower reservation fee has no incentive to quote that fee but will instead quote his competitor's reservation fee (less a small amount). As a result the market spread is not the inside quote of the reservation prices but the outside quote of the reservation prices. While the reservation spread can be negative, the market spread is always positive. The lower bound on the market spread in the two dealer case is the reservation spread of the 'worst' dealer (the one who is the most risk averse).

---

[36] Zabel (1981) and Bradfield (1982) have shown that in the absence of price variability the placement of the spread depends on the specialist's book and the time of the day but the latter effect has not been considered in the O'Hara and Oldfield (1986) model.

[37] The reservation fee is the minimum fee such that the dealer's expected utility of terminal wealth would not be lowered were he to trade at that fee i.e. it represents the cost to the dealer of entering into a transaction that makes his overall portfolio non-optimal and/or moves him to a less desirable level of

Ho and Stoll (1980) also showed that in a multi-dealer environment, there is a gravitational pull effect, similar to that of Cohen, Maier, Schwartz and Whitcomb (1981) that limits the divergence of inventories and thereby limits the divergence of reservation prices of dealers.

In Ho and Stoll (1983)[38] where heterogeneous expectations are used (but homogeneous opinions about the true value of each stock) two sources of uncertainty were assumed: the random arrival of transactions and the random returns.[39]

The position of the market-spread depended on the spreads quoted by each dealer and in particular it was shown to be determined by the second best dealers. An equilibrium market- spread was defined to exist when the distribution of inventories across dealers is such that no dealer wishes (at that moment) to trade with any other dealer. The equilibrium market spread must be non-negative.

All dealers were shown to have the same reservation spread under heterogeneous as well as homogeneous opinions. The reservation spread of each dealer was independent of his inventory and therefore there is a tendency, in both cases, for the observed market spread to be the reservation spread of any dealer. An interesting insight of the model was that inventory should affect the placement of the spread differentially among dealers, however its drawback was that it did not include limit orders per se but allowed the dealers to trade either directly with the public or among themselves in an inter-dealer market. There was also a trade-off between trading with the public (uncertain execution but higher prices) versus trading with other market-

---

risk (bid price is true price minus reservation buying fee; ask price is true price plus reservation selling fee).

[38] The model is similar to those of Stoll (1978a) and Ho and Stoll (1981) in that the dealer cares about his overall portfolio and not about his trading activity only.

[39] It is important to note that the pricing and spread equations did not include expectations of the other dealers' actions or even the size of the other dealers' inventories. Pagano (1989) and Biais (1993) have addressed these effects in more detail.

makers (certain execution but lower prices). Thus the dealer's price may reflect more than the simple order-balancing issues analyzed in earlier models.[40]

The main difficulty encountered by early empirical researchers has been the non-availability of appropriate databases from multiple-dealer markets. Early empirical literature, which has dealt with inventories or dealership markets, [41] had not tested the central implications of inventory models of competitive dealership markets.[42] Other papers which tested some of the implications of inventory models of dealership markets[43] lacked more detailed data and consequently they could not explore the relationships between quote placement and individual dealer inventories which is the main idea behind the inventory model of dealership markets. Moreover, the nature of the NASDAQ database precluded the identification of the inventory positions of dealers and did not allow direct tests of the inventory model of dealership markets by the early researchers[44] of the quote-setting behaviour of NASDAQ dealers.

Hansch, Naik, and Viswanathan (1998) conducted direct tests of the Ho and Stoll (1983) model, using data from the London Stock Exchange and presented the first set of empirical results on inter-dealer trading and the relationships between inter-dealer trading and dealer inventory levels.[45] Their dataset allowed them to clearly identify the origin and nature (whether transactor was trading on his own account or as an agent) of each transaction, to construct market makers' inventory positions and, unlike

[40] Pagano (1989) showed that, since the depth and liquidity of a market depend on the entry decisions of all potential participants, each trader assesses them according to conjectures about entry by others. If trade is equally costly across markets it will eventually concentrate on one market. If not, some trade will concentrate on one market and large traders may resort to a separate market.
[41] Ho and Macris (1984) for the AMEX options market, Hasbrouck and Sofianos (1993) and Madhavan and Smidt (1993) for NYSE.
[42] Most available databases came from the NYSE the nature of which does not allow the analysis of either the inventory models of dealership markets or the inventory model of a monopolistic dealership market (Garman (1976) and Amihud and Mendelson (1980)).
[43] Stoll (1976) and Neuberger (1992), Lyons (1995), Manaster and Mann (1996).
[44] Chan et al (1995), Christie and Schultz (1994a,b), Christie, Harris and Schultz (1994).
[45] Recent work on interdealer trading has been presented by Reiss and Werner (1998) and Naik and Yadav (1996a).

the indirect tests of market maker behaviour,[46] to test directly the predictions of inventory models of competitive dealership markets.

Analyzing inter-dealer trading they found that market-makers posting competitive quotes execute a significantly larger proportion of the public trades and their relative inventory position is significantly related to their ability to execute large trades. Changes in quotes and inventories were shown to be strongly correlated and standardized and relative inventories (with respect to the median inventory) were mean reverting with a mean reversion coefficient increasing in the inventory level, which supports inventory models.[47]

They found that inter-dealer trading is an important mechanism for managing inventory risks in dealership markets, especially when inventories diverge a lot[48] and this mechanism allows dealers to take large inventory positions that they would not be willing to take in an auction-type market where they could only unwound their positions against the public order flow. Moreover, because of order-flow preferencing it was argued that posting competitive quotes may have little value in dealership markets.

2.9.3.3 Specialist Models

Even though most models of the dealer are essentially models of the specialist, early research has dealt directly with the effect of the specialist's behaviour on prices.

[46] Glosten and Harris (1988), Hasbrouck (1988), and Chan et. al. (1995).
[47] The mean half-life from the mean reversion estimates is 2.5 trading days (in contrast to that obtained by Madhavan and Smidt (1993) of 7.3 days) which indicates that dealership markets like the LSE differ considerably from markets with specialists like NYSE.
[48] This finding is consistent with Ho and Stoll (1983), Naik, Neuberger and Viswanathan (1996) and Lyons (1996).

In the NYSE the specialist's income comes from his dual role as a broker[49] (earns his income by managing orders) and a dealer (assumes risk). In order that he takes the opposite side of transactions the specialist charges a fee, the bid-ask spread. He can also earn trading profits on his inventory while he trades market orders as well as brokerage fees from limit orders.[50] However, early researchers[51] either assumed that his only decision variable was the spread and for this reason focused on its determinants, or maintained that the specialist changes the underlying price of the security (midpoint of the spread) in response to temporary trading imbalances which cause his actual inventory position to fall outside his preferred range of inventory holding,[52] a behaviour which may be independent of the bid-ask spread (i.e. the spread may remain constant around a highly variable mid-point). Others researchers examined his behavior as an inter-temporal arbitrager who buys when there is a temporary excess supply of shares and sells when there is a temporary excess demand.[53]

Madhavan and Sofianos (1998) have found evidence that specialists participate more in less active stocks and those where there is less competition. Rather than controlling their inventories by adjusting their quotes their evidence suggested that they do so through the timing and direction of their trades. They were found to participate more when the bid-ask spreads are wide, and when previous price movements have been significant. Their participation decreases with trade size in individual stocks.

[49] An examination of the Stock Exchange specialist system of the NYSE has been made by Stoll (1985) who diagnosed a historical decline in the participation of the specialist as broker. He also analyzed his role as price stabilizer and has distinguished between ex-post and ex-ante stabilization.
[50] Baumol (1965) first described the specialist's role.
[51] Demsetz (1968), Tinic (1972), West and Tinic (1971), Bagehot (1971), and Barnea and Logue (1975).
[52] Smidt (1971), the S.E.C. (1971), and Barnea (1974).
[53] Friedman (1953), Baumol (1957), Telser (1959), Farrell (1966) and Mandelbrot (1971).

Even early theories of the market maker's behaviour presented by Barnea and Logue (1975) showed that the spread imposed by the specialist was determined either due to:

a) his provision of liquidity[54], which relates to the market and ownership characteristics of the stock as well as to the amount of competition (a passive role in that he reacts by simply adjusting his spread),

b) inventory risk which could be broken down into new information[55] and marketability risk (his ability to make inventory adjustments when the market for an issue is thin) and

c) inventory adjustment[56] since according to perceived changes in inventory he will determine the equilibrium price (defined as the average of the bid and ask prices) so as to reduce his exposure risk to the stock.

Under liquidity trading, Bradfield (1979) had showed that when the specialist approaches his target level of inventory variability around the equilibrium price is reduced and if he holds the target level variability will be reduced relative to a clerk system[57] but he may at times aggravate price variability compared to what it would be if he did not trade, owing to inventory considerations plus access to the book. His impact on price variability also depends on the time of day since his emphasis early in the day will be on exploiting shifts in the book while later in the day it is on moving toward the target level[58] in order to avoid overnight costs. This analysis supported Barnea (1974) in that prices should be more volatile toward the end of the day.

[54] *The liquidity theory*, first developed by Demsetz (1968) and modified and refined by Tinic (1972).
[55] First introduced by Bagehot (1971).
[56] Presented by Smidt (1971).
[57] In Logue's words "*the specialist reduces trading costs by increasing depth in the neighbourhood of the equilibrium price*".
[58] A measure of market efficiency is the size and variability of the difference between transaction prices and the equilibrium price.

While he can increase the rapidity of exchange, when bid-ask quotations are too far apart, by offering a narrower bid-ask spread the specialist also needs to follow the rules of the exchange by offering a fair and orderly market for the stock. The stronger the competitive conditions in the market,[59] the closer will be the bid-ask spread or markup to the cost of waiting and carrying inventory in which case the spread will measure the cost of making transactions without delay.[60]

The efficiency of the specialist can be evaluated by whether transactions can take place at low cost and whether there are opportunities for systematic profit as a result of serial correlation in price series, or market imperfections which prevent price from fully and immediately reflecting new information.[61] These criteria however contrast with those used by the NYSE and the SEC which focus on price continuity and stability and may in fact tend to cause inefficiency by leading to price dependencies.[62] For this reason Logue (1975) argued in favour of 1) allowing the specialist to solicit business or search for the other side of the market for large block transactions, 2) eliminate the criterion of price continuity and stability in his performance, 3) allow direct competition among market makers by reducing barriers to entry (a more competitive market may minimize the need of the above rules of price continuity and stability).[63] He therefore concluded that in markets with liquidity only or with both liquidity and information trading a competitive market-maker will perform more

[59] As perceived by a particular market-maker, competition is affected by factors such as: 1) the existence of close substitutes for the securities in which the market maker makes a market, 2) the presence of alternative trading markets for the asset, 3) the existence of limit orders, 4) a high volume of autonomous trading which could increase the likelihood of offsetting competing market orders and 5) the level of commission rates (which act as barriers to entry for competing market makers).

[60] Commission brokerage fees are not determined by the same procedures.

[61] At any time period, the difference between the asset's current price and its equilibrium value can reflect either (a) the impact of new information on the relative price of the security or (b) a trading impact i.e. a temporary imbalance between supply and demand for the security at the equilibrium price as a result of random arrival of buy and sell orders.

[62] The'Stabilization Percentage Test' ('Tick Test'), 'Twice Total Volume Test', and 'On Balance Test'.

[63] Black (1971a,b) and Smidt (1971) made such proposals.

satisfactorily than a monopolistic market-maker and that regulatory criteria are unnecessary.[64] [65]

Cohen, Maier, Schwartz, and Whitcomb (1981) showed that transaction costs cause investors to use order-placement strategies which result in a non-trivial market bid-ask spread. They defined an *equilibrium market spread* and demonstrated that it is greater for thinner securities and argued that there exists a gravitational pull effect of limit orders as traders place them close to market orders. This eventually tempts them to accept the market price which implies certainty of execution. Thus they showed that the size of the spread depends on the movement of trades between limit and market orders and this movement depends on the execution probability of a limit order.

2.10 The Role of Volume

The efficient markets hypothesis which has produced voluminous research has assumed that current price impounds all information which is relevant to the valuation of a security and contrasts with the beliefs of technical analysts as well as with the empirical evidence which indicates that some price-based strategies may yield positive returns.[66] Volume can only be informative when adjustment of prices to information is not immediate in which case other statistics related to the trading process could be used to extract information about it.

---

[64] In information trading markets the types of information that could cause an individual to trade are: 1) price dependencies, 2) divergence of the arbitraged price from the unchanging equilibrium price, and 3) the true equilibrium price of the asset (may be the principal motivation for information trading).

[65] Like Barnea (1974), Logue (1975) argued that the chief cost of dealing with a market-maker is the difference between the theoretical, but unobservable, equilibrium price and the transaction price, rather than the bid-ask spread.

[66] Fama (1970, 1991) has provided reviews of the Efficient Market Hypothesis literature. Neftci (1991) and Brock, Lakonishok and LeBaron (1992) have carried out empirical tests of technical trading rules. A succinct definition of Efficiency has been given by Tobin (1989).

The behaviour of trading volume is closely linked to the heterogeneity among the traders in a market. Different proportions of various types of traders give rise to different volume behaviour and return-volume dynamics.

After investors receive private information they trade many rounds so that current volume is related to contemporaneous information flow as well as to existing private information which was received previously. Therefore volume lags behind information flow and may reach its peak many periods after traders first receive their private information and therefore volume is serially correlated even when information flow is independent over time. As informed investors continue to trade prices reveal more of their private information so their expected gains from speculation decrease.

Volume is also related to the volatility of prices. When information, exogenous to the trading process, arrives at the market trading takes place which changes the price of the stock because new information changes both the expectations and the uncertainty about the value of the stock. High volume is thus generated which causes volatility in prices. Volume which is generated by new information is accompanied by significant price changes which are much higher than when volume is generated by existing private information.

Even though the relation between volume and prices has been studied empirically in detail, the process by which these two variables are linked together is not known because it is not clear what kind of information volume brings into the trading process. In general there are two lines of research, which, based on rational expectations approaches attempt to study the informational role of volume: a) those which examine how volume arises when traders with differential information trade and b) those which study how and what traders can learn by observing volume.

Through the first approach, Wang (1994) has shown that volume is positively correlated with absolute price changes and dividends and this correlation increases with information asymmetry.[67] Moreover, different dynamic relations arise between trading volume and stock returns when informational as opposed to non-informational trading occurs which is an indication that volume conveys important information about how assets are priced in the market. Using a similar framework, He and Wang (1995) have shown that the pattern of volume is related to the flow and nature of information and that current volume is not only related to contemporaneous information but also to existing private information received previously. Therefore it can reach its peak many periods after investors first receive private information and it is serially correlated even when information flow is independent over time.

In the second line of theoretical research Blume, Easley and O'Hara (1994) have investigated how the statistical properties of volume relate to the underlying value of the asset and to the behaviour of prices. They showed that volume is important in that it provides information relating to the quality of traders' information which cannot be extracted by observing prices which only give information about direction.[68] Moreover, sequences of prices and volume can be informative and traders who use such information could do better that those not using it, providing a justification for the use of volume in technical analysis.[69]

Conrad, Hameed and Niden (1994) used a form of contrarian strategy to measure the strength of the relation between lagged volume and returns and found strong

---

[67] Bamber, Barron and Stober (1999) have confirmed Bachelier's (1900) intuition that differential interpretations are an important stimulus for trading, have supported Kandel and Pearson's (1995) argument that trading coincident with small price changes reflects investors' differential interpretation of information and have also found that differential interpretations explain a significant amount of the trading occurring when volume is higher than the non-announcement period average.

[68] Volume picks up signal quality independent from price because it is not normally distributed. Only if its statistical properties are known can traders use it to update their beliefs during their learning process.

correlation between lagged changes in trading activity and returns patterns in individual securities. They also found negative autocovariances in returns for actively traded stocks and positive for less frequently traded ones. They concluded that volume is a valuable source of information and should be included in any model of short-horizon, time-varying returns processes.

Jones, Kaul and Lipson (1994) have shown that information pertinent to the pricing of securities is contained in the transactions data and not in those of the size of the transactions.

Models which have investigated factors such as heterogeneous agents and incomplete markets have been used to model volume but they do not confront the data in its full complexity and have not evolved to a level at which they could furnish an empirical model of stock market data like, for example, a model with dynamically optimising heterogeneous agents which could jointly account for major stylized facts like serially correlated volatility, contemporaneous volume-volatility correlation and excess kurtosis of price changes.[70]

Major reviews of the literature on volume have been presented by Tauchen and Pitts (1983) and Karpoff (1987). While early research[71], using weekly data, found no correlation between price changes and volume, a number of researchers like Ying (1966) later documented a significant correlation between price changes, per se and absolute price changes with volume, using daily data. Using transactions and daily data on common stocks Epps (1977) found that volume changes are larger for

[69] A prerequisite for obtaining these results has been the assumption that traders know aggregate supply by observing volume which leads to a price which is 'fully-revealing'.

[70] Some of the most important models in this area include those of:
Admati and Pfleiderer (1988,1989) which comprise of informed and liquidity traders to explore the implications for within-day and weekend volume and price movements; Huffman(1987) who developed a capital growth model with overlapping generations to obtain a contemporaneous volume-price relationship and Huffman(1988) and Kettere and Marcet (1989) who examined trading volume and welfare issues in various economies comprised of heterogeneous infinitely lived agents facing limited trading opportunities.

[71] Granger and Morgenstern (1963) and Godfrey, Granger and Morgenstern (1964)

increases in the absolute price change, compared to decreases, regardless of the general movement in prices.[72] Rogalski (1978) argued that the relation is only contemporaneous. Harris (1986) showed that the positive correlation between volume and absolute price change varies across securities and Richardson, Sefcik and Thompson (1986) found that volume increases in proportion to abnormal returns around announcements of dividend changes.

Karpoff (1987) presented a model of the price-volume relation, based on previous empirical observations, which shows that when price-changes are positive (negative) the correlation is positive (negative), that is the relation is different for positive and negative price changes (this is essentially an asymmetric price-volume hypothesis).[73] He also argued that the relation can be traced to their common ties to information flows.

Jain and Joh (1988) analyzed the lead and lag relations between volume and the absolute value of returns using a long time series of hourly data and showed that there is a strong positive correlation between contemporaneous trading volume and the absolute value of returns (or square of returns) which is consistent with the mixture of distributions hypothesis of Clark(1973), Epps and Epps(1976) and Tauchen and Pitts (1983)[74]. Trading volume was found to be positively correlated with returns lagged up to four hours. They also found evidence of an asymmetric relation (i.e that the volume-returns relation is different for positive and non-positive price changes) which

---

[72] Wood, McInish and Ord (1985) and Smirlock and Starks (1985) documented the same relation . The latter, however, argue that it holds only in periods when the arrival of information can be distinguished ex-ante.

[73] This hypothesis also has the following implications: a) Tests using data on volume and the absolute value of price changes will yield positive correlations and heteroscedastic error terms. b) Tests using data on volume and price changes (not absolute) will yield positive correlations (ranked by the price change, the residuals from a linear regression of volume on price changes will be autocorrelated).

[74] The mixture of distributions hypothesis states that when no information is available, trading is slow and the price process evolves slowly. When new information violates old expectations trading is fast with the price process evolving even faster. Thus contemporaneous absolute returns and volume should be positively correlated.

is consistent with the behavioural model of Epps (1975) and the one of Karpoff (1985) and Karpoff (1987) which is based on institutional rules.

Jain and Joh (1988) also found that average volume follows a U- pattern during the day[75] and an inverse U-pattern during the week.[76] Although some evidence was found of volume causing returns, it was weak compared to that of returns affecting volume which is to be expected in an informationally efficient market.[77]

Using a very large data-set Gallant, Rossi and Tauchen (1992) documented the positive price-volume relation but showed that it is non-linear and that changes in price lead to a symmetric movement in volume.

---

[75] For the London Stock Exchange Abhyankar, Ghosh, Levin and Limmack (1997) have documented a two-humped pattern which they attributed to the differential trading patterns among stocks with different liquidities. High activity stocks however, showed a more definite U-shaped pattern.

[76] Average volume was found to differ significantly during the day, being highest during the first hour, declining monotonically until the fourth hour but increasing again on the fifth and sixth. During the week, average volume was also found to differ being lowest on Monday, increasing monotonically till Wednesday and then declining on Thursday and Friday.

[77] Smirlock and Starks (1985) found evidence of a similar relation using a smaller time series.

Chapter 3

The Bid-Ask Spread

3.1 Introduction

Existing literature on a number of areas pertinent to the bid-ask spread is presented in this chapter. The main issues discussed include the behavior of the bid-ask spread, the influences of the spread itself on the price-formation process, the distinction between the posted (quoted) and the effective spread and finally the models which attempt to determine the types, sizes and the behavior of the components of the bid-ask spread.

3.2 Behaviour of the Bid-Ask Spread

The intraday behavior of the spread is first discussed in this section, followed by the theories which attempt to explain this behavior. Finally research on the seasonality of the spread is presented.

3.2.1 Intraday Behaviour

The pattern followed by the bid-ask spreads of NYSE stocks has been documented by Brock and Kleidon (1992), McInish and Wood (1992), and Lee, Mucklow and Ready (1993)[78]. According to these papers, intraday spreads follow a U-shaped pattern. McInish and Wood (1992) found that the spread, as a percentage of price, has a U-shape during the day with particularly high values in the first minutes of trading, declining for the next 15 minutes to a level which lasts until the last few minutes of the day and rising again immediately preceding the close. Brock and Kleidon (1992) provided additional evidence showing that the U-shaped pattern is not restricted to one subset of stocks when these are grouped into price (bid-ask average), market value of equity and beta quintiles. In their study of the relationship between

[78] Several other authors like Porter (1988), Jaffe and Patel (undated) and Brown, Clinch and Foster (1991) also found similar evidence.

spreads and depths Lee, Mucklow and Ready (1993) documented a similar pattern but with a smaller sample of NYSE stocks.

The intraday pattern of bid-ask spreads for NASDAQ stocks has been investigated by Chan, Christie and Schultz (1995) who also related the results to the institutional features of the dealer market.[79] They found that the bid-ask spread for NASDAQ stocks is relatively stable throughout the day but narrows significantly during the final hour of trading. Kleidon and Werner (1993) found a similar pattern for London stocks during mandatory trading hours. The London stock market, like NASDAQ, is structured as a multiple-dealer system and therefore the difference in the intraday patterns between NYSE and NASDAQ (or London) market seems to support the idea that the institutional features of markets can exert an influence on the determination of spreads over the day.

In contrast to NASDAQ spreads, which remain relatively constant during the first hour of trading, NYSE spreads decline. However, the most striking difference between the spreads in the two markets found by Chan, Christie and Schultz (1995) was that they widen near the close for NYSE while those of NASDAQ narrow during the final 30 minutes of trading.[80] The authors used individual-dealer trade and quote revisions for NASDAQ stocks. The stocks selected were large enough for NYSE listing so that meaningful comparisons could be made.[81]

[79] Papers that focus on transaction costs differences across markets are those of Blume and Goldstein (1992), Christie and Huang (1994), Christie and Schultz (1994a,b), Goldstein (1993), Lee (1993), Neal (1992), and Petersen and Fialkowski (1992).
[80] This cannot be explained by the Brock and Kleidon (1992) model.
[81] The approach followed by the authors was that used by Foster and Viswanathan (1993) in their study of intraday volume, volatility and trading costs for NYSE securities, i.e. they tested for intraday variations in spreads, volume and volatility and other variables using Hansen's (1982) GMM procedure.

Even though the average dealer spread showed practically no intraday variation, the inside spread[82] declined, gradually during the day and sharply at close, reflecting changes in the location, rather than in the width of individual dealer quotes. The authors tested whether the intraday patterns in volume and volatility for NASDAQ stocks were similar to the intraday pattern of the spread for the same stocks. They found that even though they follow a U-shape, similar to that of NYSE stocks, they do not follow the pattern of NASDAQ spreads. Thus the difference in bid-ask spreads for NASDAQ vs. NYSE stocks near the close cannot be attributed to differences in the intraday pattern of return variability or volume.

For the London Stock Exchange which is a multiple-dealer market (known as ISE or International Stock Exchange, London) Abhyankar, Ghosh, Levin and Limmack (1997) have shown that average bid-ask spreads follow a U-shaped pattern during the trading period which is similar across deciles of stocks partitioned on the basis of liquidity. This behavior is similar to that of NYSE and unlike that of NASDAQ stocks. In contrast to the latter, the trading session in the London Stock Exchange consists of a Mandatory Quote-Period (MQP) before and after which dealers are not obliged to quote prices and the high values of the U-shape pattern of the spread corresponds to those periods (that is outside the MQP). The ISE and the NASDAQ markets are not directly comparable for a number of reasons the most important of which is that most of the trading in the ISE is done by institutional traders thus making the order-flow to consist of large, infrequent trades whereas NASDAQ is mainly a retail market.

McInish and Van Ness (2002) have shown that for the NYSE, variables which explain the intraday behavior of bid-ask spreads, such as activity, competition, risk

[82] The inside spread is the difference between the inside ask and the inside bid price. The average spread is defined as the difference between the average ask and the average bid price. The latter bracket

and information also determine its intraday components and in particular the inventory-holding and the asymmetric-information components which are discussed in section 3.4.

To appreciate the potential explanations given for the intraday behaviour of the bid-ask spread it is important to gain an understanding of the various models which have been used in the attempt to explain it. These fall under three main classes: inventory, specialist market power and information models and are presented in the next section.

3.2.2 Theories of the Behavior of the Spread

Microstructure Models which have been proposed in order to explain the intraday behavior of bid-ask spreads of stocks are founded on the theories presented in Chapter two and in particular the inventory and information models of dealer behavior. Other researchers have also investigated the market power of the specialist as a factor affecting the behavior of the spread.

3.2.2.1 Inventory Models

Stoll (1978a), Amihud and Mendelson (1980) and Ho and Stoll (1981) have argued that the spread exists to compensate the market maker for bearing the risk of undesired inventory. Lee, Mucklow and Ready (1993), Hasbrouck and Sofianos (1993), and Madhavan and Smidt (1993) all found evidence on the relation of Bid-Ask spreads to dealer inventory control costs. In particular, the former found that for a sample of NYSE stocks, spreads became wider in response to higher trading volume. Hasbrouck and Sofianos (1993) found that trades in which the NYSE specialist participates tend to have a bigger and more rapid impact on the spreads

---

the inside spread.

than trades with no specialist participation. Madhavan and Smidt (1993) reported that bid-ask quote revisions are positively related to order imbalances. Consequently during the open and the close of the market, when volume tends to be higher, the magnitude of order imbalances and therefore the spreads would be greater than during the rest of the day.

3.2.2.2 Specialist Market Power Model

Based on an extension of Merton (1971), Brock and Kleidon (1992) developed a model of intraday spreads where the market-maker possesses monopolistic power and showed that transactions demand at the open and close is greater and less elastic than at other times of the trading day due to overnight accumulation of information, changes due to the imminent non-trading period near the close and the fact that fund managers tend to trade near the close in order that they achieve optimal portfolio positions or transfer overnight risk. Heavy and inelastic demand causes the specialist to price-discriminate by charging a higher price for trading at those times. The Brock and Kleidon (1992) model predicts periodic demand with high volume at the open and close and concurrent wide spreads which is consistent with the observed patterns for NYSE stocks. It also implies spill-over effects with optimal increases in both transactions demand and spreads in intervals immediately after the open and prior to the close.[83]

Stoll and Whaley (1990) also supported the hypothesis that the wider spreads at the open reflect the specialists' ability to profit from their privileged knowledge of the order imbalances. Even though specialists are responsible for maintaining a

[83] Therefore one would not expect spreads in other markets (e.g. for CBOE options) to display a U-shaped intraday pattern if specialist market power is the key for observed pattern in NYSE stocks.

continuous presence in the market, they can often give quotes inferior to those of the limit order book, especially when the latter is thick.

3.2.2.3 Information Models

Copeland and Galai (1983), Glosten and Milgrom (1985), Kyle (1985), Easley and O'Hara (1987), Admati and Pfleiderer (1988), Hasbrouck (1988), Foster and Viswanathan (1990), (1994), and Madhavan (1992) all focused on the adverse selection faced by the market maker, which was first proposed by Bagehot (1971) and assumed different types of agents like informed traders, liquidity traders (must trade at a given time during the day regardless of the cost), and in some models discretionary liquidity traders. In these models the specialist, being at an informational disadvantage, must keep spreads wide enough so that the gains from trading with the uninformed compensate for the losses to the informed. These models predict that higher volume is associated with lower spreads given the optimal equilibrium behaviour of the models' traders. In particular:[84]

Foster and Viswanathan (1994) developed a model where strategic trading with two asymmetrically informed traders is analyzed. This model of competition between two informed traders predicts high volume, variances and spreads at the start of trading, a result that is not explained by earlier information models of intraday trading.

Madhavan (1992) developed a model in which information asymmetry is gradually resolved during the trading day by observing trading prices. He considered traders with diverse information regarding the value of the asset at the beginning of trading. As trading continues, private information is impounded into prices, and

[84] This prediction should normally be valid for NASDAQ stocks or low-volume ones where there are few limit orders.

specialists narrow their spreads as their informational handicap declines. Hence, the bid-ask spread declines through the day.

Admati and Pfleiderer (1988) modeled the strategic decisions of liquidity traders who have discretion over the timing of their trades. By concentrating their trades on specific periods, discretionary liquidity traders minimize the adverse selection costs facing specialists, leading to the simultaneous occurrence of heavy trading volume and narrow spreads at specific times during the day. [85]

It is evident from the above discussion that two broad classes of factors could be thought responsible for the pattern of the bid-ask spread intraday. Those pertaining to the structural characteristics of the market and those related to information.

The relative importance of market-structure versus information–based factors in determining the intraday pattern of bid-ask spreads can be measured by comparing the results for NYSE versus NASDAQ stocks. The distinction between these spread determinants is that the flow of information may be similar for NYSE and NASDAQ stocks, but the institutional structures of the markets are dissimilar.

The primary institutional differences[86] between markets are :

- A single specialist for a stock on the organized exchanges versus multiple dealers on NASDAQ.
- A call-auction market at the open on the organized exchanges versus a quote-driven system that searches for equilibrium prices in a dealer market and

---

[85] In both the Admati and Pfleiderer (1988) as well as the Madhavan (1992) model intraday spreads are attributed to the variation in the costs of adverse selection.

[86] Early papers focusing on the transaction characteristics of alternative market structures have been those of Grossman and Miller (1988) and Glosten (1989) which show that a specialist can enhance liquidity when trading volume is low or when adverse selection problems are large. Also Kamara (1988), Stoll and Whaley (1990) and Vijh (1990) documented important differences between call-auction and specialist markets, between organised exchanges and OTC markets and between the liquidity of stock and options markets.

o The consolidation of order flow through the specialist versus the fragmentation of order flow across dealers. Since the structural differences across markets are most pronounced around market closures it is important that one focuses on open and close trading periods in order to study difference in the bid-ask spread patterns.

Three other important studies of the spreads of securities, other than stocks, are those of Chang, Chung and Johnson (1995), George and Longstaff (1993) and Neal (1992).

Chan, Chung and Johnson (1995) studied the intraday behaviour of bid - ask spreads for actively traded CBOE options and compared them with their NYSE - traded underlying stocks in order to distinguish among the above three classes of competing hypotheses regarding variations in spreads for these two related securities. They found high spread for both options and stocks at the open which reflects market makers' desire to protect themselves from greater uncertainty at the open and high stock spread but lower option spread at the close which they attributed to the difference in market making structure. They concluded that both the degree of competition in market making and the extent of informed trading are important for understanding the intraday behaviour of spreads.[87]

George and Longstaff (1993) examined the cross-sectional distribution of bid-ask spreads in the S&P 100 index options market. The advantage of this cross-sectional approach was that the market structure and underlying sources of risk were held fixed across options, so that differences in bid-ask spreads could be directly related to

---

Evidence that transaction costs are related to market structure has been considered by the SEC in bringing about rule changes.

differences in the costs faced by market makers across options. It also allowed the direct examination of the interrelation between bid-ask spreads and trading activity. The important conclusions of this analysis were that most of the variation in bid-ask spreads can be explained by the exchange rules and the market structure and that bid-ask spreads in such a dealer market are larger than those in specialist markets.

Neal (1992) provided additional evidence on the relation between market structure and transaction costs by comparing the bid-ask spread and some properties of transaction prices for equity call options in two market structures: the American Stock Exchange (AMEX) specialist structure and the Chicago Board Options Exchange (CBOE) competitive market maker structure. He found that at low trading volumes the specialist structure has lower bid-ask spreads than competitive market structures and that this difference diminishes as volume rises, consistent with Grossman and Miller (1988). Moreover, he found that the distributions of trade size and transaction prices differ between the two structures. The competitive market maker structure is associated with a greater fraction of trades at the ask which are of a smaller size (consistent with Vijh (1990)).

3.2.3 Seasonality

Clark, McConneli, and Singh (1992) found that there is a distinct seasonal pattern for bid-ask spreads in NYSE stocks (which was more pronounced among low-priced stocks where the January seasonal was also most pronounced) in which both absolute and relative spreads decline from the end of December to the end of the following January but found little evidence of significant correlation between changes in spreads at the turn of the year and January stock returns.

[87] These results were consistent with papers which investigated non-NYSE securities like those of Chan, Christie and Schultz (1995) for NASDAQ and Kleidon and Werner (1993) for London but did

Stoll and Whaley (1983) have argued that excess returns on small stocks are a result of higher proportional bid-ask spreads in low-priced stocks. In their view, the bid-ask spread itself is the 'cause' of the higher returns on small stocks. That is, because of the higher proportional cost of transacting in small stocks, investors demand a higher rate of return.

Keim (1983) and Schultz (1983) have pointed out that small stock excess returns are concentrated in January and a seasonal in stock returns cannot be explained by the bid-ask spread unless there is a seasonal in the bid-ask spread as well.

Schultz (1983) compared bid-ask spreads in December with those in June but found no significant differences in bid-ask spreads between the two dates. He thus concluded that high transaction costs in December cannot explain the larger excess returns earned by small firms in January.

3.3 Spread Influences on The Price Process

Apart from the influence of other factors on the formation of the bid-ask spread research has investigated the spread's influence on factors related to the price process such as asset pricing and the price itself.

Amihud and Mendelson (1986) studied the effect of the bid-ask spread on asset pricing by analyzing a model in which investors with different expected holding periods trade assets with different relative spreads. They confirmed their hypothesis that market-observed expected return is an increasing and concave function of the spread.

Amihud and Mendelson (1987) examined the effects of the mechanism by which securities are traded on their price behaviour by comparing the behaviour of open-to-open and close-to-close returns on NYSE stocks, given the differences in execution

---

not compare different market making structures like Chan, Chung and Johnson (1995).

methods applied in the opening and closing transactions. Opening returns were found to exhibit greater dispersion, greater deviations from normality and a more negative and significant autocorrelation pattern than closing returns. They studied the effects of the bid-ask spread and the price-adjustment process on the estimated return variances and covariances and discussed the associated biases. They concluded that the trading mechanism has a significant effect on stock price behaviour.

Ho and Macris (1984), utilizing dealer's 'trading book' information, showed that much of the variation in transaction prices may be explained by the specialist's optimal determination of his bid and ask quotes and demonstrated that the dealer's bid-ask spread is an important explanatory variable in the observed transaction return. They also indicated that the dealer's inventory level may affect his quotes and thus the transaction prices and order arrivals.

3.4 Posted and Effective Spreads

Unlike the simple securities markets described in standard financial models, where trades take place at the prices posted by the specialist and market orders pay the whole spread between the bid and the ask, in the U.S. equity markets trades inside the spread are quite frequent. Market orders may transact inside the spread if specialists do not always display the best public limit orders or when they are matched with other market orders. In such cases, the posted spread overstates an investor's expected trading costs. Since investors can expect to buy at prices lower than the ask and sell at prices higher than the bid, the *effective spread* is the relevant measure of trading costs.[88] Accurate estimation of transaction costs is important for tests of market

[88] It is not possible to directly estimate the effective spread and compare it to the posted spread without using quotes or orders. Glosten and Harris (1988) have estimated it indirectly using transaction prices. Lee (1993) and Blume and Goldstein (1992) compare transaction prices to the midpoint of the quoted spread but this method accurately estimates the effective spread only when market orders are matched

efficiency, asset pricing models as well as tests of the theories which explain how the spread should vary with characteristics of the stock such as trading activity and the size of the order (large vs. small). However, the usefulness of the effective spread as a measure of such costs depends on the accuracy of the methods which are employed for the classification of trades as originating from buy or sell orders. These methods are discussed in section 4.3.

Using two samples of market orders, one based on orders submitted by retail brokers and one based on orders submitted electronically to the NYSE, Petersen and Fialkovski (1994) documented a significant difference between the posted spread and the effective spread paid by investors. By matching each order's execution price to the quote at the time the order was submitted, they measured how far inside the posted quotes each order was executed. This measure is called "price improvement" and it allows the estimation of the effective spread.

In the US equity markets price improvement can arise when : 1) the specialist intervenes to stop or match a market order, 2) market orders execute against hidden limit orders or against orders from floor traders and 3) buy and sell orders arrive simultaneously (a rare case). Then price improvement can be defined as:

Price improvement (relative to the exchanges' own quotes) =

Ask price – Transaction price (buy orders) =

Transaction price – Bid price (sell orders)

Price improvement (relative to the best bid or offer (BBO)) =

Min (ask price) – Transaction price (buy orders) =

Transaction price – Max (Bid price) (sell orders)

---

solely with limit orders (the specialist or a public limit order). If they are matched with each other the method will overestimate the spread actually paid.

The difference between the two price improvements is a measure of the quality of an exchange's quotes. The effective spread can be defined as:

Effective spread = Posted spread – 2 x Price improvement

since a trader may receive price improvement both buying and selling and assuming that it is the same on both sides.

Petersen and Fialkovski (1994) found a systematic difference between the two types of spread and similar to that of Branch and Freed (1977) that more-actively traded stocks have smaller posted spreads. The effective spread also declined with trading activity but the magnitude of the decline was much smaller.

It was found that the posted spread provides most of the explanatory power for the expected price improvement and once it is known other firm-specific factors have only a slight marginal effect. Petersen and Fialkovski (1994) also found that not only do the sizes of the posted and the effective spreads differ in magnitude but their correlation is also small. They also provided evidence that the posted spread is a poor measure of the costs of liquidity and that the effective spread on the regional exchanges is significantly larger than on the NYSE. Results suggested that the posted spread cannot be used to forecast the effective spread. Orders sent to the regional exchanges were found to receive significantly less price improvement whereas those sent to the NYSE receive better than average execution prices. The above results were not consistent with the notion that the US equity markets are integrated since in an integrated equity market the execution price of an order does not depend on where the order was sent.

3.5 Models of the Bid-Ask Spread and its Components

To set the foundations for a discussion of the models of the bid-ask spread and its components we first present below two basic models of the spread which are based on economic principles and describe the decision rules followed by dealers in setting their bid and ask prices. The first is the Stoll (1978) model, already described in the previous chapter, which assumes no informed trading and bases the decision of the dealer on inventory considerations. The second is the Glosten and Milgrom (1985) model which considers both informed as well as uninformed trading affecting the dealer's behaviour. We then proceed with the presentation of the rest of the models distinguishing between the two categories which attempt to measure spread components, namely covariance spread models, which make use of the serial covariance of transaction price changes and trade indicator models which use an indicator variable which characterizes trades as buyer- or seller-initiated. Both covariance spread as well as trade indicator models, even though they are based on economic principles like the Stoll (1978) and the Glosten and Milrom (1985) ones, they are less theoretic in nature since they attempt to derive tools which could be used for the empirical analysis of spreads and their components.

3.5.1 The Stoll (1978) Model

According to Stoll the dealer attempts to maximize the expected utility of his terminal wealth which consists of his initial wealth W_0, representing his initial position in the optimal efficient portfolio, plus the true value of his position in his trading account Q_p and any other funds he may posses. He is assumed to have an exogenous and unchanging belief regarding the 'true' price of the asset as well as a 'true' rate of return for it.

In this two-period model the dealer's decision problem consists of setting prices for the next transaction in such a way that his utility after the trade will be at least as good as before it. Then following a trade of a true value Q_i on the asset his terminal wealth, $\tilde{W}$ is given by:

$$\tilde{W} = W_0\left(1+\tilde{R}^*\right)+\left(1+\tilde{R}_i\right)Q_i-\left(1+R_f\right)\left(Q_i-C_i\right) \tag{3.1}$$

where $\tilde{R}^*$ is the rate of return on his initial portfolio

$\tilde{R}_i$ is the rate of return on stock i, and

R_f is the risk-free rate at which the dealer can borrow to finance his inventory or at which he can lend excess funds.

C_i is the cost of trading amount Q_i which is positive or negative depending on whether he buys or sells the asset.

$Q_i - C_i$ represents the dealer's exposure cost of holding a sub-optimal inventory.

Within the short time-period considered and since there are no limits on his ability to borrow, his risk of becoming bankrupt is zero.

The dealer will enter any transaction which will leave his expected utility unchanged according to:

$$\mathrm{E}\left[U\left(W_0\left(1+\tilde{R}^*\right)\right)\right]=\mathrm{E}\left[U\left(\tilde{W}\right)\right] \tag{3.2}$$

Using $R_f = 0$ since the time horizon considered is short:

$$c_i = \frac{C_i}{Q_i} = \frac{z}{W_0}\sigma_{ip}Q_p + \frac{1}{2}\frac{z}{W_0}\sigma_i^2 Q_i \tag{3.3}$$

where

c_i is the percentage cost that the dealer requires in order that he takes position Q_i in stock i, that is his cost of providing immediacy,

z is the dealer's coefficient of relative risk aversion,

σ_{ip} is the correlation between the rate of return of stock i and that of the optimal efficient portfolio, and

σ_i^2 is the variance of the return of stock i.

From the above equation it is evident that:

- The larger the dealer's initial wealth the lower the percentage cost
- The larger his risk aversion the larger this cost
- The larger the dealer's inventory position the larger this cost since it becomes more risky to increase his inventory.
- The size of the trade also affects this cost since a larger value will remove him away from his optimal position.
- The characteristics of the stock as well as its relation to other stocks in the market also affect the dealer's decision.

The dealer then will set his bid and ask prices in such a way as to compensate himself for this cost. For the bid, its proportional distance below the true price, expressed as a percentage of it, will be set according to:

$$\frac{\left(P_i^* - P_b\right)}{P_i^*} = c_i\left(Q_i^b\right) \qquad \text{and} \qquad \frac{\left(P_a - P_i^*\right)}{P_i^*} = c_i\left(Q_i^a\right) \tag{3.4}$$

where P_i^* is the true price of stock i before the transaction takes place

P_b is the bid price, P_a the ask price and

Q_i^b is the true value of a sale to the dealer.

Q_i^a is the true value of a purchase from the dealer.

Thus the bid-ask spread relative to the true price (percentage spread) is given by:

$$\frac{\left(P_a - P_b\right)}{P_i^*} = c_i\left(Q_i^b\right) - c_i\left(Q_i^a\right) = \frac{z}{W_0}\sigma_i^2 \tag{3.5}$$

3.5.2 The Glosten and Milgrom (1985) Model

Glosten and Milgrom (1985) modeled the behaviour of a dealer who can revise his bid and ask quotes after each trade assuming unit trades take place and only market orders are available. They modeled two types of investors: informed and pure liquidity traders who all start with the same information regarding the random value V of an asset. Informed traders trade on the basis of their private information whereas the uninformed are assumed to participate due to the disparity of preferences or endowments which is their main motive for trading. All market participants are assumed to be risk-neutral.

Each trader is assigned a time preference parameter which determines how much he is willing to pay to buy or accept to sell a unit of stock. He assigns random utility to shares of stock, x, and current consumption, c_c as follows:

$$\rho_l x V + c_c \tag{3.6}$$

where ρ_l is a 'liquidity parameter' which represents the differential subjective assessments regarding the distribution of the random variable V and is independent of V as well as any other information about V. It is a parameter of the individual investor's utility function and represents his personal trade-off between current and future consumption derived from owning the asset. The higher the value of ρ_l the higher the desire to invest for the future. Its value for the dealer is assumed to be equal to one.

The risk-neutrality assumption implies that for trade to take place the value of the parameter must vary across the participants otherwise a 'no-trade' situation will be reached and the market will collapse.

Investors are assumed to arrive one-by-one, randomly and anonymously and the dealer has knowledge of the probability structure of the arrival process which implies that he can extract correct inferences from observing the data. Investors also wish to maximize their expected utility using their existing information. Glosten and Milgrom (1985) distinguish between the public information set H_t available at time t which comprises of all past transaction prices, the current bid and ask quotes and all available public information and the informed traders' private information set J_t. An informed trader arriving at the process will have an information set F_t which includes J_t, H_t and the information furnished by the quoted bid and ask prices.

The optimal decision rule of an investor is formulated in terms of:

$$\text{Buy if } Z_t \rangle A$$

$$\text{Sell if } Z_t \langle B \tag{3.7}$$

where

$$Z_t = \rho_i \mathrm{E}[V \mid F_t] = \rho_{it}(1-U_t)\mathrm{E}[V \mid H_t, J_t, A, B] + \rho_{it} U_t \mathrm{E}[V \mid H_t, A, B]$$

$U_t = 1$ if trader arriving at time t is uninformed and

$U_t = 0$ otherwise.

The dealer, knowing the behaviour of the traders chooses bid and ask prices on the basis of his own information set S_t in such a way that his expected profit from a trader arriving at time t is:

$$\mathrm{E}\left[(A-V)I_{\{Z_t \rangle A\}}\right] + (V-B)I_{\{Z_t \langle B\}} \mid S_t \tag{3.8}$$

The indicator functions of these events:

$I_{\{Z_t \rangle A\}} = 1$ if $\{Z_t \rangle A\}$ and =0 otherwise, and

$I_{\{Z_t \langle B\}} = 1$ if $\{Z_t \langle B\}$ and =0 otherwise

The above equations assume that 1) there are zero costs associated with all short positions in cash or stock, 2) there are no transaction and inventory costs (since the dealer is assumed to be risk-neutral), 3) all participants are risk-neutral and competitive and 4) trades take place only at the ask and bid prices. Therefore it is only information that affects the spread.

Prices are assumed to be set in such a way that the expected profit is equal to zero on every trade the reasoning being that competition and risk neutrality would remove any rents the specialist might earn because of the actions of competing dealers (if the market maker was risk averse he would have inventory concerns).

prices. Glosten and Milgrom (1985) also assumed that at equilibrium the prices set by the market maker are ex-post regret-free, that is following the trade the dealer knows that while the ask has been exceeded or the transaction price has been below the bid and updates his information set according to this and his assessment of the trader's information. Thus the ask and bid prices are reservation prices which are arrived at because of competition and the ask (bid) is what the revised expectation of the value of the asset would be if a trader buys (sells). Then the bid and ask prices will be equal to the conditional mean of V given the specialist's information set as:

$$A_t = \mathrm{E}[V \mid S_t, Z_t \rangle A_t]$$

$$B_t = \mathrm{E}[V \mid S_t, Z_t \langle B_t] \tag{3.9}$$

A useful contribution of this model is that unlike Copeland and Galai (1983) where all information is revealed at one trade, Glosten and Milgrom (1985) incorporate the effect that information is not revealed instantly but only through a series of trades and thus the behaviour of prices is affected.

The price-decision problem of the market maker is affected by the sequential arrival of informed traders who force him to continually upgrade his expectation of the price of the asset and eventually new information is incorporated into this price (a learning problem). It is thus the opposite of previous inventory theoretic models which assumed that the order flow as well as the uncertainty about the value of the asset was exogenous to the market maker's decision problem (Bayesian Learning Model).

The model is incomplete since it does not capture the interaction between the adverse information and the need of the market maker to account for other risks such as inventory/bankruptcy/risk-neutrality etc.

Glosten and Milgrom (1985) also proved that:

- under these conditions the ask is always higher than the bid
- that price changes follow a martingale relative to the dealers and the public information which indicates that spreads due to the adverse selection effect are quantitatively different from those due to transaction cost, risk-aversion or monopoly power which, contrary to adverse selection, lead to negative serial correlation.
- The spread consists of both transitory as well as adverse selection components (explained in section 3.5.3) and the latter is given by:

$$\operatorname{cov}(\Delta P_t) = -\frac{1}{2}cA - c^2 \tag{3.10}$$

  where c is the per trade cost of the dealer.

- Spreads decline as the number of trades increases which indicates that the insiders' information is assimilated as trading proceeds leading to an approximate consensus expectation value, V of the asset in the market.

3.5.3 Covariance Spread Models

Covariance Spread models, which use the serial covariance of transaction price changes, are based on the work of Roll (1984) which is presented below. The most important of these models are those of Stoll (1989), Choi, Salandro and Shastri (1988) and George, Kaul, and Nimalendran (1991).

The determination of the relative sizes of the spread components is crucial to gaining an understanding of the return-generating mechanism since these components arise from the different trading behaviour of market participants reflecting their varying interests and actions. To the extent that the size of the components could be determined in every trading situation would allow researchers to gain a better insight of the return-generating process.

Early research has distinguished between three types of components of the spread, the order-processing, the inventory holding and the adverse selection cost components. The first two constitute what has been termed the transitory cost component which has a transient effect on price, because its effect on the time series of prices is not related to the underlying value of the security. The adverse-selection component, which allows market-makers to earn from uninformed traders what they risk losing from trading with informed ones, has a permanent effect on prices and does not cause serial correlation. The different time series behaviour of the two components has been used to separate them.

The bid-ask spread plays a very important role in the formation of the price and has been shown to affect the time series properties of returns.[89] The fact that the random arrival of buy and sell orders can make prices fluctuate from the bid to the ask (bid-ask bounce) can cause spurious volatility and serial correlation in returns even when there is no change in the fundamental value of the asset. In his seminal paper Roll (1984) showed that in a frictionless market (without transaction costs) when

[89] Blume and Stambaugh(1983), Keim (1989) and Philips and Smith (1980).

there is no change in the fundamental value of the security, the spread can induce negative serial correlation in transaction price changes and one can infer the spread from the observed pattern of prices. Choi, Salandro and Shastri (1988) have shown that this is true even when the types of orders arriving (order flow) are serially correlated which will only affect the magnitude of the correlation.[90]

Roll (1984) derived an equation for the realized spread, S, in terms of the serial covariance of transaction price changes, when the market is informationally efficient and the probability distribution of the observed price changes is stationary. He denoted the change in price between times t and t-1 as ΔP_t, and showed that the spread is given by:

$$S = 2\sqrt{-\operatorname{cov}\left[\Delta P_{t-1}, \Delta P_t\right]} \tag{3.11}$$

Since this spread is inferred by the transactions data it is equivalent to the effective (realized) spread which differs from the quoted spread since trades often occur within the quoted spread for reasons related to the particular characteristics of the trading situation and mechanism (for example broker discounts). This spread equation also accounts for the time-series properties of asset returns.

Stoll (1989) has developed a model where the covariance of transaction returns and the covariance of quote (bid and ask) returns are related to the probability of a price reversal and the size of a price reversal and estimated the last two. The model also distinguishes among the three components of the quoted spread (order processing, inventory holding and adverse selection) but assumes that the effective spread consists of the inventory and the order processing components only. It also assumes that the quoted spread, S^Q, is constant and relates it to stock characteristics:

$$\operatorname{cov}(\Delta P_t, \Delta P_{t-1}) = S^{Q^2}\left[\delta_c^{\,2}(1-2\pi) - \pi^2(1-2\delta_c)\right] \tag{3.12}$$

---

[90] Lucas (1978) and Leroy (1973) have suggested that this need not be true.

$$\mathrm{cov}(\Delta M_t, \Delta M_{t-1}) = \delta_c^2 S^{Q^2}(1-2\pi) \qquad (3.13)$$

where M_t : is the mid-point between the bid and ask prices at time t, equal to:

$$M_t = \frac{1}{2}(P_{a,t} + P_{b,t})$$

π : is the probability of a price reversal, that is the probability that a trade at the ask (bid) will be followed by a trade at the bid (ask)

δ_c : is the magnitude of a price continuation as a proportion of the quoted spread, where a price continuation occurs when a trade at the ask (bid) is followed by a trade at the ask (bid). Its value lies in the range between 0 and 1 assuming the value 0 for the case when the spread is solely due to order-processing costs and 0.5 for the case when it is due to adverse or inventory-holding costs.

S^Q : the quoted spread assumed to be constant

The adverse selection cost arises as the difference between the quoted and the realized spread. The realized spread is defined as the expected price change after a dealer purchase less the expected price change after a dealer sale and is given by:

$$2(\pi - \delta_c)S^Q \qquad (3.14)$$

Then the holding cost component is calculated from the above equation for the realized spread using $\delta_c = 0.5$ and the estimated value of π from the regressions. These calculations are based in the author's arbitrary assumptions that in the order processing cost model $\pi = 0.5$ and $\delta_c = 0$ whereas under the inventory holding cost model $\pi > 0.5$ and $\delta_c = 0.5$.

This model was tested using only a limited set of intraday trade and quote data for NASDAQ/NMS stocks from three months in 1984 and estimated the three components using OLS estimation. His results showed that the quoted spread for his sample consisted of 43% adverse, 10% inventory and 47% order-processing cost. The average realized spread was found to be 57% of the quoted spread. However, his use of data at the close of the trading session may have biased his results due to the different trading behaviour of the participants at those times as discussed in section 3.1.

Using the Stoll (1989) model for estimating the components of the realized spread but with complete intraday data, instead of the prices near the end of the trading session which Stoll(1979) had used, Menyah and Paudyal (2000) estimated the components of the quoted spread for the London Stock Exchange stocks as being 47%, 30% and 23% for the adverse, order-processing and inventory-holding costs respectively.

Choi, Salandro and Shastri (1988) allowed for serial dependence in transaction type which can occur when large orders are broken down in smaller orders so that informed traders can disguise themselves. Aged (stale) orders in the book also have the same effect. The authors related the serial covariance of transaction price changes with the realized spread as :

$$\operatorname{cov}(\Delta P_t, \Delta P_{t-1}) = -\pi^2 S^2 \tag{3.15}$$

They also did not assume inventory adjustment of dealers but merely extended Roll's (1984) model.[91]

Chen and Blenman (2003) have extended the single-period Stoll (1989) model into a two-period one and have incorporated the serial correlation in successive trades which allows for the accommodation of price-reversal effects which can be caused by asymmetric-information and inventory-holding costs. In this way they derived a spread estimator that is always positive.

[91] In Roll's (1984) model π is equal to ½.

George, Kaul and Nimalendran (1991) argued that previous estimators of the bid-ask spread and its components, which are based on autocovariances of transaction returns, are biased and inefficient. They are biased because these returns contain positively autocorrelated components as a result of time-variation in expected returns which leads to a downward bias in the estimation of the realized spread. They are inefficient because transaction returns contain a large 'unexpected' component. To correct for this bias they used the difference between transaction returns and returns based on bid-to-bid quotes which are unaffected by positive covariance generated by time-varying expected returns. The model they developed assumed time-varying expectations, no inventory adjustment and no serial dependence in order flow so that the probability of a price reversal was 0.5 and arrived at:

$$\operatorname{cov}(\Delta P_t, \Delta P_{t-1}) = -(1-\alpha)\frac{S^2}{4} \tag{3.16}$$

where 1-α is the order processing component. They used daily data for their estimation. It is doubtful however whether changing expectations are important at the microstructure level.

Kim and Ogden (1996) showed that this estimator is also biased since George, Kaul and Nimalendran (1991) also assumed that the spread is constant over time. The size of this bias has been estimated by Affleck-Graves, Hegde and Miller (1994) to be around 4 percent while Brooks and Masson (1996) argued that it is only present in short time-series and small sample cross-section estimates.

3.5.4 Trade-Indicator Models

Trade-indicator models use an indicator variable to categorize trades as originating from a buy or a sell order. The most important models in this category are those of

Glosten (1987), Glosten and Harris (1988), Huang and Stoll (1994), Lin, Sanger and Booth (1995), Hasbrouck (1988, 1991), Huang and Stoll (1997) and Madhavan, Richardson and Roomans (1997).

A basic simple asymmetric-information model has been presented by Glosten (1987) which is based on the same economic principles as the Glosten and Milgrom (1985) model. He developed a two-component theoretical model of the spread which incorporated the correlation between the price of the security and the direction of the trade. This characteristic was used to distinguish between the different effects of the two components on the behaviour of transaction prices. His model starts with a basic equation for the formation of the bid and ask prices which is based on the true value of the security. This value is the price which would be realizable in the market, that is the price at which agents would all agree to exchange the asset, in the absence of frictions (trading costs) and provided all participants had the same information, that is no private information existed in the market.

Glosten distinguishes between this common (public) information value P from the full information value P^*. The former is the expected value of the latter given the common information set Ω according to:

$$P = \mathrm{E}\left[P^* \mid \Omega\right] \tag{3.17}$$

The public information value, P, is assumed to remain constant as long as no new information regarding the particular security arrives which could lead to a different valuation of the security by the market participants. On this value all costs are subtracted (added)[92] to arrive at the bid, P_b (ask, P_a) price set by the dealer. Glosten (1987) distinguished between the adverse selection costs at the bid, A_b and the ask, A_a

[92] Flood et al. (1998) have shown that search costs which arise when dealers search for quotes in the large exchanges also affect trading costs and should be incorporated into the bid-ask spread since their size can be around one-third of the effective spread.

and between costs C_b at the bid and C_a at the ask, which account for all other costs such as those of market-making, inventory costs and a normal rate of return for operating as a market-maker. All of these costs were collectively called the gross-profit component. When a purchase occurs at the ask, market participants are assumed to revise their estimate of the true price from P to (P + C_a) and when a sale occurs at the bid they revise their estimate of the true price from P to (P - C_b).

The adverse selection costs A_a and A_b are given as the difference between the true price and the revised expectation of P^* conditioning also upon the latest trade (that is whether an investor buys at the ask or sells at the bid):

$$A_b = \mathrm{E}\left[P^* \mid \Omega \cup B\right] \tag{3.18}$$

$$A_a = \mathrm{E}\left[P^* \mid \Omega \cup A\right] \tag{3.19}$$

where A and B stand for the events:

A : investor buys at ask

B: investor sells at bid

Following the above discussion the bid and ask prices are formed according to:

$$P_b = P - A_b - C_b \tag{3.20}$$

$$P_a = P + A_a + C_a \tag{3.21}$$

The asymmetric information model, which is also a hypothesis about the way in which private information is incorporated into prices has been introduced by Glosten and Milgrom (1985) who incorporated the anticipation in the revision in expectations, conditional on the type of order submitted, into the bid and ask prices of the dealer. This hypothesis has subsequently been validated by the empirical work of Glosten and Harris (1988) who could not reject the hypothesis that a significant part of the spread of NYSE common stocks is due to asymmetric information.

Glosten's model of the spread explains the difference between the quoted and the realized spread. If the price at time t, P_t, either occurs at the bid or the ask then:

$$P_t = P + C_t q_t \tag{3.22}$$

where q_t is the buy-sell indicator (or trade-indicator) variable for the trade occurring at time t which takes the values of :

+1 if the trade is buyer-initiated (purchase from the dealer) and

-1 if the trade is seller-initiated (sale to the dealer)

and C_t is the gross profit component at time t and is:

C_a if the trade is buyer-initiated and

C_b if the trade is seller-initiated

Moreover, unless the adverse selection cost at time t, A_t, is zero, P_t will be correlated with q_t since the possibility that a buyer (seller) initiated trade is information based will cause P to be revised upwards (downwards):

$$\text{cov}[P_t, q_t \mid P] = \text{E}[A_t \mid P] \tag{3.23}$$

where A_t : is equal to A_a if $q_t = 1$ and

equal to A_b if $q_t = -1$

Arrival of new public information will revise P_t according to:

$$P_t = P_{t-1} + \varepsilon_t + A_t q_t \tag{3.24}$$

where ε_t is the revision in the true price due to the passage of time and the arrival of public information between trades at times t and t-1. It is assumed to be an independently and identically distributed random variable. The assumed serial independence implies rationality on behalf of the market-makers because if there was serial correlation in the location of the spread a market-maker entering the market could profit by incorporating this information into his own quotes.

Therefore the change in price ΔP_t, between trades at t and t-1 will be given by:

$$\Delta P_t = P_t - P_{t-1} = A_t q_t + \varepsilon_t + \left(C_t q_t - C_{t-1} q_{t-1}\right) \qquad (3.25)$$

that is the adverse selection component, $A_t q_t$, is permanent whereas the gross-profit component, $(C_t\, q_t - C_t\, q_t)$, can reverse leading to serial correlation in returns.

Assuming that: a) continuously compounded returns are uncorrelated, b) the spread is symmetric around the true price and c) that the gross-profit component does not cause conditional drift in prices, then the serial covariance of continuously compounded returns is approximately given by:

$$\text{cov} = -0.25\ \pi_g s_p^2 \qquad (3.26)$$

where π_g : is the proportion of the spread due to factors other than adverse selection (gross profit)

and s_p: is the proportional spread which is defined as:[93]

$$s_p = \frac{2(C + A)}{P_a + P_b} \qquad (3.27)$$

where $C_a = C_b = C$ and $A_a = A_b = A$

Based on the same economic principles as the Glosten and Milgrom (1985) model, Glosten and Harris (1988) extended Glosten's (1987) model and developed an econometric model of the spread attempting to estimate the two components (the information or adverse-selection component and the order-processing plus inventory component which together are referred to as transitory) by exploiting their different time-series behaviour as explained in the previous section. They allowed the adverse selection component and the gross profit component to depend on order size but assumed that there is no inventory cost. They also assumed that the components do not change with time. This idea has been introduced in the theoretical models of Easley and O'Hara (1987) and Kyle (1985) suggesting that the adverse-selection cost should increase with the quantity traded since informed traders wishing to maximize the return from their private information will want to trade in large quantities before their information becomes common knowledge (public).[94] Moreover, they allowed the error term in equation (3.25) to depend on the elapsed time between trades. Since they did not have information about quotes and thus could not infer the trade indicator variable, they estimated the adverse and transitory components conditioning on the observed volume at time t. Using fourteen months of NYSE common stock price changes in the years 1981-1983 and maximum likelihood estimation they were unable to reject the hypothesis that the adverse selection component is positive. They estimated the model[95]

$$\Delta P_t = C_0(q_t - q_{t-1}) + C_1(q_t V_t - q_{t-1} V_{t-1}) + A_0 q_t + A_1 q_t V_t + \varepsilon_t + r_t - r_{t-1} \tag{3.28}$$

where C_0, C_1 : are the intercept and slope of the gross-profit (or transitory) spread component equation $C_t = C_0 + C_1 V_t$, assumed to be linear in the volume of trade at time t, V_t

A_0, A_1 : are the intercept and slope of the linear adverse selection spread component equation $A_t = A_0 + A_1 V_t$, assumed to be linear in the volume of trade at time t, V_t

ε_t : public information innovation between trades at times t and t-1, assumed to be an independently, identically distributed normal

---

93 This way Glosten has shown that the part of the spread due to the gross profit component induces biases in the measurement of mean (simple) returns and the variance of returns and also induces negative serial correlation in measured returns.

94 This idea is equivalent to the permanent effect of block trades on prices in Holthausen, Leftwich and Mayers (1987).

random variable with mean and variance dependent on the time between trades at t and t-1, denoted as T_t, that is

$$\varepsilon_t \sim \text{iidN}\left(f_1(T_t), f_2(T_t) | T_t\right)$$

ΔP_t : change in observed prices from time t-1 to t

q_t : buy - sell indicator variable for the trade at time t

r_t : round-off error (drift between observed discrete price and unobserved price when no discrete prices are used), assumed to be a zero mean, asymptotically uniformly distributed random variable.

V_t : number of shares traded at time t

The conditional normality assumption for the public information innovation term, e_t, has been suggested by the mixture of distributions hypothesis of Clark (1973) and Harris (1987) already discussed in section 2.10.

Glosten and Harris (1988) found that A_0 and C_1 should be equal to zero.

Using the Glosten and Harris (1988) model Neuberger (1992) was not able to decompose the realized spread into inventory and asymmetric information.

Apart from the work of Glosten and Harris (1988) referred to above, which is a trade indicator model, even though trade direction could not be observed directly due to the lack of quote data, Huang and Stoll (1997) and Madhavan, Richardson and Roomans (1997) have also developed similar models.

Huang and Stoll (1997) developed a regression model to estimate all three components of the spread. They exploited the behaviour of trade and quote prices after a trade to distinguish between the adverse selection component and the other

[95] This model is similar to the inventory-theoretic model of Ho and Macris (1984) the main difference being that in the latter volume has a lagged effect on the bid-ask prices.

components since quote adjustments for inventory reasons tend to be reversed over time while quote adjustments for adverse information are not. They also used the idea in the inventory models, like that of Ho and Stoll (1981), according to which changes in quotes affect trade flow because after a sale at the bid the dealer lowers his bid to invoke purchases and vice versa thus increasing the probability of an opposite transaction and creating negative serial correlation in quote changes and trades. This is distinct from that of price changes described by Roll (1984) and attributed to the bid-ask bounce. Using this principle, the fundamental value of the security should be updated only by the unexpected part (innovation) of the information furnished by the trade flow which is given by:

$$\mathrm{E}(q_{t-1} \mid q_{t-2}) = (1-2\pi)q_{t-2} \tag{3.29}$$

Where π reflects the probability that a trade at time t is opposite in sign to the trade at time t-1.

The model estimated was:

$$\Delta P_t = \frac{S}{2}q_t + (A + IV - 1)\frac{S}{2}q_{t-1} - A\frac{S}{2}(1-2\pi)q_{t-2} + \varepsilon_t \tag{3.30}$$

Where A is the adverse selection and IV the inventory cost component of the spread[96] and S the traded (realized) spread which is assumed to be constant (estimated from the data). The error term, ε_t, which as in the previous models accounts for the public information innovation, is assumed to be an independently, identically distributed random variable. Their model ignored mid-quote trades.

Huang and Stoll (1997) used an ISSM sample of 20 of the largest and most actively traded NYSE stocks utilizing their trades and quotes for all trading days in

1992, without as well as with clustering trades, that is grouping trades at one price into one order (when no adjustment in quotes occurs). Their analysis was confined to transactions which were BBO (Best Bid Offered), coded as regular trades. Without clustering the average adverse selection was found to be negative around 3.1% of the spread whereas the inventory component 18.7%. With clustering the average adverse selection component was found to be 9.6% whereas the average inventory 28.65% of the traded spread. They also found that the adverse-selection cost is smaller for larger trade size categories.

Madhavan, Richardson and Roomans (1997) also attempted a two-component decomposition of the spread in which the order flow was serially correlated and was expressed as a function of the probability of a continuation and the probability of midquote execution. They also assumed that the revision in beliefs is positively correlated with the innovation in the order flow as in Glosten and Milgrom (1985) who however assumed that the latter is uncorrelated. In their model quotes were only used to derive the trade indicator variable and they did not explicitly model the inventory effect like Huang and Stoll (1997). The model estimated was:

$$\Delta P_t = (\varphi + \theta)q_t - (\varphi + \rho\theta)q_{t-1} + \varepsilon_t + \xi_t - \xi_{t-1} \tag{3.31}$$

Where ξ_t : is independently, identically distributed with zero mean and models the effect of rounding errors due to price discreteness or time-varying returns

ρ : is the first-order autocorrelation coefficient of the trade indicator variable

φ : transitory effect per share which covers the dealer's compensation

[96] Bollen, Smith and Whaley (2002) have presented a model of the inventory-holding and the adverse-selection cost components of the spread as an option with a stochastic time to expiration and showed that it performs well for NASDAQ stocks.

for providing liquidity, the inventory holding cost and risk bearing

θ : the permanent impact of order flow innovation in the true price equivalent to adverse selection

ε_t : captures the effect of public information on the true price of the stock, that is it represents the innovation in beliefs regarding the true price of the stock between trades at times t and t-1 and is an independently, identically distributed random variable with zero mean.

The assumption of uncorrelated errors is reasonable since the rounding error for NYSE stocks is on average one sixteenth of a dollar which is half the minimum variation and thus unlikely to have a significant impact on the results.

The above equation is based on the assumption that the bid and ask prices set by the dealer are ex-post rational conditional upon the next trade type under the rational expectations hypothesis. The resulting prices are regret-free.

The model was estimated using the first 750 stocks in the ISSM database for 1990 for NYSE-listed common stocks of which only 274 were eventually considered. The variation of the components of the spread was studied over five intraday intervals. The adverse selection cost was found to vary from 4.2 to 2.7 cents and to follow a U-shaped pattern over the day whereas the transitory cost was found to range from 3.4 to 4.6 cents and to increase progressively during the day.

The model was also used to explain known patterns of the behaviour of transaction variables like bid-ask spreads, execution costs, price and quote volatility etc.

To avoid estimating more complicated models previous researchers like Glosten and Milgrom (1985), Roll(1984), Stoll (1989), Choi, Salandro and Shastri (1988) and George, Kaul and Nimalendran (1991) have assumed that the size of each order is fixed and usually assign it a unit measure. Other authors, however, like Kyle (1985), Easley and O'Hara (1987) and Glosten (1989) have developed models in which the adverse selection component increases with the order size. In the Easley and O'Hara (1987) model trade size causes adverse selection because given that they wish to trade, informed traders prefer to trade in large amounts at a given price. Therefore large trades convey more information and move quoted prices more than small trades. Moreover, Glosten and Harris (1988) have used trade size in their empirical analysis like Madhavan and Smidt (1991). The last two papers have modeled the revision in beliefs as a function of the net order imbalance for a particular period. In a similar fashion and according to the inventory models of the spread, the order processing costs should decrease with trade size.[97]

Lin, Sanger and Booth (1995) have examined empirically the relationship between the two components of the spread (adverse selection and order processing) and found that the adverse selection component increases significantly and monotonically with trade size, increasing the average effective spread whereas the order processing component decreases monotonically for all but the largest percentile of the relative trade size distribution reflecting the extra cost associated with negotiating large trades. Using a one-period horizon model and thus not being able to capture longer-term inventory rebalancing effects, they derived three equations, one for each change in quote midpoint, transaction price and effective half spread respectively. Trading volume was not entered directly into their equations but their dataset was split into seven size categories. They used ISSM transaction data for 150 NYSE stocks in 1988 and estimated each equation separately using OLS.

They estimated the following two equations:

[97] Copeland and Stoll (1990) have argued that order-processing costs are fixed for a particular transaction so that average order-processing costs per share should decrease as the number of stocks traded in a transaction increases.

$$\Delta M_{t+1} = \lambda_e z_t + e_{t+1} \tag{3.32}$$

$$\Delta P_{t+1} = -\gamma_e z_t + u_{t+1} \tag{3.33}$$

where

z_t is the effective spread

γ_e is the order-processing cost component of the effective spread, and

λ_e is the adverse-selection cost component of the effective spread

Their results showed that the adverse-selection component of the effective spread increases monotonically with trade size from 19.8 to 62.6% of the effective spread. The order-processing cost was found to lie in the range between 46 to 22.4% of the effective spread and to decrease in trade size for all but the largest (99th percentile) trades.

Chapter 4

Research Design and Preliminary Data Analysis

4.1 Introduction

The way in which the empirical research of this thesis has been designed is presented in this chapter followed by an analysis of the database used and the way data has been extracted. Finally summary statistics of the data employed are presented and discussed.

4.2 Research Design

This section presents the aims of the empirical work carried out for this thesis and briefly outlines the methodology followed as well as the models developed.

The main aim of this thesis has been the investigation of the influence of both the trading volume and the volume of the quotes upon the formation of the bid-ask spread in a market. The second, minor, aim has been to research the influence of the waiting time between trades on the same variable, the bid-ask spread.

The analysis of the influence of volume on the spread is carried out through the analysis of the measurable variables of the trading process. These are the volume (number of shares) of the particular trade, the price at which the exchange of shares is effected, the bid and ask prices quoted, the volume at the bid and ask prices quoted prior to a particular trade, the number of trades, the number of quotes and quote revisions, the size of the spread and the time elapsing between two successive trades.

Following the existing literature, the influence of all these variables on the spread and subsequently the price of a given share is a function of their influence on the components of which the spread is composed, namely the adverse selection, the inventory holding and the order processing cost components. To this end we develop models of the change in the bid-ask spread through the change in the bid and the ask prices and then a model of the change in price. The first model, which has been named the "Quoted Spread Model", describes the change in the spread in terms of the three components of the spread as well as the waiting time between trades. The second model, called the "Price-Change Model" contains the same variables but describes the change in price in terms of the three components of the spread, however, the sizes of the parameters in these two equations could be different before and after a trade takes place.

Therefore, the use of these two models aids in the analysis of the components of the spread before and after a trade has taken place.

Since in order to derive the Quoted Spread Model we use equations describing the change in the bid and the ask, some information is lost through the differencing process. To gain more insight into the mechanism of the formation of the ask and bid prices we estimate those components separately.

Thus the aim of our analysis is manifold. First we wish to estimate the components of the bid-ask spread when the volume of trade and the volume of the quotes are included into the analysis directly. Second we wish to investigate how these components differ before and after a trade occurs. Third, we examine the difference between the sizes of these components under two different trading mechanisms, namely the NYSE and the NASDAQ. Finally, we investigate the impact which the waiting time between trades has on the formation of the spread and eventually on the prices of common stocks.

With the above aims in mind we use transactions data from the NYSE and NASDAQ for a particular period of time (one month) and following a procedure which is described below we randomly select two hundred stocks from NYSE and six hundred stocks from NASDAQ [98].

To account for the possibility that the parameters of the models vary with the trading activity in the stock, we split the stocks in each sub-sample into ten deciles, each of equal number of stocks, according to the level of trading activity in each stock, as measured by the number of shares traded in the month considered. Thus we end up with ten sub-samples from each exchange and carry out our analysis on these.

Since our models are trade-indicator models, that is they depend on whether the particular trade considered has originated from a buy or a sell order we follow certain procedures for the classification of trades into these two categories. We discuss and contrast all these procedures below outlining their merits and drawbacks.

The estimation of the models is carried out through the Generalized Method of Moments estimation procedure which does not require particular distributional assumptions for the variables involved and which does not force one into using Linear Regression or Maximum Likelihood estimation. All estimation is carried out using the SHAZAM econometrics package and the models are tested for specification errors using techniques based on the current literature on GMM-testing.

4.3 Data Analysis

The various databases available for the NYSE and NASDAQ exchanges are presented and discussed in this section. Particular emphasis is given to the Trade and Quote (TAQ) database of NYSE which is used for our empirical work. The procedure which has been followed in order that data for analysis is selected is critically

[98] A larger number of NASDAQ stocks, compared to NYSE, has been selected since the database used does not report all quotes for NASDAQ but only those at the bid-ask spread.

discussed together with current literature on the methods used to classify trades as buyer- or seller-initiated.

4.3.1 Database Used

A number of transaction-level databases have been used to carry out empirical analysis of intraday data for the NYSE and NASDAQ exchanges. Of these the most frequently utilized have been the TAQ, the TORQ and the ISSM databases.

The Trades, Orders, Reports and Quotes (TORQ) database is available from the NYSE and covers a period, starting from November 1990. It consists of detailed information on transactions, quotes, the NYSE audit trail and contains information about the parties to each side of a transaction as well as a record of all orders submitted to the floor of the NYSE through the SuperDOT system.[99] It comprises of four files: 1) the consolidated trade file, 2) the consolidated quote file, 3) the system order database (SOD) file and 4) the consolidated audit file. The first two are similar to the respective files in TAQ. The database also contains limit orders providing a thorough picture of the limit order book.

Harris and Hasbrouck (1996) report that the SuperDOT (DOT stands for Designated Order Turnaround System) accounts for 53% of participants in all transactions and for only 30% of total buy and sell volume. Of these orders 50% are market, 45% limit, 2% stop and 2% market-at-close orders. The fact that larger orders and those requiring special handling are submitted directly to the NYSE specialist by the floor brokers constitutes a major drawback of the database since it is not representative of the whole population of orders. The TORQ has been used in many studies like Hasbrouck (1996), Sias and Starks (1997), Koski and Scruggs (1998), Angel (1998), Chung et al (1999) and Ready(1998) among others.

[99] Hasbrouck, Sofianos and Sosebee (1993) as well as Hasbrouck (1992) discuss the TORQ database.

Another widely used database has been the one from the Institute for the Study of Security Markets (ISSM) the main components of which are distributed by SIAC (Securities Industry Automation Corporation) in real time over a high speed line. It consists of time-stamped trades, trade size and bid-ask quotes from NYSE and AMEX and consolidated regional US exchanges which belong to the National Market System (NMS) and covers various periods starting from January 4th 1988. This database is very large and for this reason only a sample of it has been used in experimental analysis. It does not indicate whether orders are buyer or seller-initiated.

Available to us for this thesis has been the Trade and Quote data (TAQ) from U.S. exchanges for the month of October 1994. This dataset, which is published by the NYSE, includes all trades and quotes of stocks traded in the New York Stock Exchange (NYSE), The American Exchange (AMEX) and the National Association of Securities Dealers Automated Quotation system (NASDAQ). It consists of a number of files which provide detailed trade and quote data together with information about the stocks, whether they have paid dividend or split and a includes a number of condition and correction codes from which one can infer the eligibility of the trades and quotes for one's analysis. However, it does not identify trades as buyer- or seller-initiated and does not give information concerning the identity of the parties involved in each transaction. An attempt is made to overcome the first problem by appropriate trade identification methods which are given in the relevant literature. Even though the use of data which identify traders could provide more information on the trading mechanism and would be necessary in more sophisticated models, such is beyond the aims of this thesis.

4.3.2 Data Selection Procedure

The Trade and Quote data in TAQ constitute a huge database which makes the task of analyzing all available data onerous. To select smaller subsets of data one needs to be careful about the possibility of 'data-snooping' biases. Therefore a random selection procedure has to be followed.

The most likely sources of trouble from the data, most of which follow from the discussion in chapter 3, are the following:

- Large splits cause large changes in the price level which in turn affect the impact of price-discreetness on the results of the model.
- Samples from different time periods entail the risk of selection bias due to temporal dependence.
- Data from opening and closing periods have different statistical properties from intra-day data reflecting the different institutional features involved. For example, in NYSE the opening price is determined by following a call-auction mechanism whereas the rest of trading is a double continuous auction market. Moreover, overnight accumulation of information may lead market participants to wish to alter their inventory positions immediately at the opening of trading. Therefore opening prices for NYSE stocks should be discarded. A similar problem occurs with closing prices even though there is no change in the trading mechanism. Nevertheless the strategies of various types of traders may lead to different trading behaviour close to the end of the trading session. For example, to avoid the build up of undesired inventory dealers may be willing to accept less-than optimum prices at that period. Therefore it is preferable to discard closing prices as well. Brock and Kleidon (1992) have examined such trading behaviour in the NYSE.

- Stocks with a large number of trades are likely to yield more accurate parameter estimates than infrequently traded ones.

Since our aim is to study the effect of volume on the behaviour of the bid-ask spread components under the two different trading mechanisms, the huge amount of available data is split according to:

1) The exchange where stocks were traded. Two different sets of stocks are selected, one for each type of exchange mechanism (hybrid like the NYSE versus quote-driven like the NASDAQ).
2) The total trading volume in the month considered. Deciles of trading volume are formed from the trade-files for each category of exchange.

To verify that data used are as pure as possible from errors we follow a procedure by which available data is filtered so that only correct data are allowed to enter our estimation procedure. To this end a number of checks is carried out on the data. A brief description of the data selection procedure is given below whereas details on the programming procedures used are described in some detail in Appendix C for both NYSE and NASDAQ stocks. For convenience we will only refer to the NYSE stock-selection procedure which is similar to that for NASDAQ.

All common stock symbols traded in NYSE[100] (or traded in NASDAQ as a primary market) which have a positive number of outstanding shares are extracted from the Master Table file of TAQ. Off-NYSE trades of NYSE-listed stocks are included in the sample since they constitute a large portion of the trading activity. Bessembinder and Kaufman (1997) argue that this is more than 50% of small trades in large firms. The dividend file is examined so that only stocks which have not paid dividend in a month or have not split are selected. The reason for the exclusion of such stocks is that we

[100] A good description of the trading systems and procedures of the New York Stock Exchange has been provided by Hasbrouck, Sofianos and Sosebee (1993).

focus our study on periods where there is little possibility that an information event would occur which would influence normal trading behaviour and would require an event study methodology for analysis. Eventually 1,829 stocks are selected from NYSE (5,850 from NASDAQ).

Selected stock symbols are then used to extract Trade data from the pertinent consolidated file of the TAQ database. The total volume in the month considered, as measured by the total number of shares traded, as well as the total turnover in the month, expressed in USD is calculated for each selected stock. The stocks are then sorted in ascending order according to the total number of shares traded in the month. NYSE stocks with no trades in the month considered are excluded whereas NASDAQ stocks with less than 10,000 shares traded in the month considered are left out. The higher minimum number of trades for NASDAQ against NYSE stocks aims to make the data for the empirical analysis more comparable. The above procedure reduces the number of NYSE stocks to 1,815 and those of NASDAQ to 5,032. Remaining stocks of each exchange and then split into equally-numbered deciles. Twenty stocks are then selected randomly from each NYSE and sixty stocks from each NASDAQ decile leaving a total of two hundred stocks for NYSE and six hundred for NASDAQ for subsequent analysis. The larger number of NASDAQ stocks is necessary in order to obtain a satisfactory number of observations because, as already mentioned, fewer quotes are reported for NASDAQ trades in the TAQ database. Moreover this procedure results in obtaining equivalent deciles where the total number of trades of one decile in a month is close to that of the equivalent decile of the other exchange so that comparison is made easier. The symbols of these randomly selected stocks are used to extract both Trade and Quote data from the respective consolidated files of the database. To avoid errors resulting from reporting delays trades and quotes for each symbol in a particular day are sorted first alphabetically and then by time. Filtering also takes place at this stage so that more errors are removed as described below. Even though a complete description of condition and correction codes, according to which filtering has taken place, is provided in the TAQ manual of the New York Stock Exchange, details have also been provided in Appendix E so as to facilitate the presentation of the data-selection methodology.

For trades, the consolidated trade file is checked so as to make sure that the condition code indicates that no out-of-sequence trades are included when other trades have taken place in the mean time (condition code 'Z'). Moreover, trades with correction codes indicating that a trade has been cancelled, or there is an error, are excluded (only good trades that is those with correction codes 0, 1 and 2 are accepted). Finally trades whose price is zero or negative or their size is zero or negative are left out.

Regarding quotes, we leave out those for which the bid or the ask price are zero or negative as well as those with zero or negative sizes at the bid or the ask. We also make sure that asks are larger than bids. The condition codes in the consolidated quote file are also examined so as to exclude trades with codes indicating that there has been a regulatory or non-regulatory trading halt due to a number of reasons such as news dissemination, high activity, pending news for the particular stock or for others affecting it and due to severe order influx (that is quotes with condition codes 0, 4, 7, 9, 11, 13, 14 and 15). The filtering procedure for NYSE data results in curtailing the original database of 345,114 trades and 809,996 quotes down to 325,501 trades and 766,908 quotes. For NASDAQ data, trades are reduced to 389,155 from 434,998 and quotes to 55,526 from 72,643.

Subsequently a matching procedure is followed for the stocks in the sample according to which trades are matched with quotes which precede them by at least five seconds. This method, which is explained in more detail below, derives from Lee and Ready (1991) who used a five[101] second cut-off condition to allow for non-synchronous reporting. The first and last trade in a particular stock in a day is discarded since, as explained above, opening-period data are expected to have different statistical properties compared to those during the day. Then each trade is examined trying to identify a qualifying quote which precedes it by at least five seconds. Only the quote immediately preceding the particular trade is considered whereas quote revisions preceding a particular trade are ignored.[102] Unmatched trades are also discarded. During this process all the necessary variables for the models developed are calculated from the data.

To enable us to estimate the adverse-selection component of the spread properly and to avoid obtaining negative values we aggregate half of the trades which occur in-sequence at the same price while quotes remain unchanged. This problem, also discussed in Huang and Stoll (1997), is encountered if trades are not aggregated (clustered) since an order may be broken down into a number of consecutive, individual trades thus inducing positive autocorrelation in trade flow. The effect of this is to yield a negative adverse selection effect which is theoretically unacceptable. We do not cluster all such trades, but only half, since a number of these may be executed on behalf of other traders who may have private information which would cause adverse-selection by the dealer. Therefore our decision is a compromise

[101] This interval depends on the particular data set even though most researchers are content with the five second period. Bessembinder (2002) argued that trade direction is best assessed when no adjustment is made for trade report lags.

[102] Such quote revisions are not used in the development of our model but they could have been included had further analysis been pursued.

between obtaining a theoretically acceptable adverse selection cost and the danger of clustering trades from different traders.

The trade classification method used in our empirical work is based on that proposed by Lee and Ready (1991) and described in the next section. This method has been found to be either superior or close to the performance of the best alternative methods for both NYSE and NASDAQ stocks as suggested by empirical research which is presented in the next section as well. As in Lee and Ready (1991) our program for matching trades with quotes also employs a five second cut-off condition to allow for non-synchronous reporting of trades. According to this method the quote rule is used first, that is the price is compared to the mid-point of the prevailing quotes and a trade is classified as buy if the price is above it and as a sell if it lies below the midpoint. Mid-point trades are then classified using the tick test (or tick rule) which compares the transaction price to the previous price as explained below. Unmatched trades are also classified using the tick rule. If the tick rule cannot lead to definite conclusions the trade is classified as indeterminate and discarded from the sample.

Finally the data are split into activity volume sub-samples (deciles) to be used in the estimation of the models. At the end of the procedure the number of matched trade-quotes pairs for NYSE stocks was 92,330 and that for NASDAQ 4,262.

It is worth noting at this point that the various data filters which have been employed during the data selection procedure have led to the removal of a very large number of trades or quotes which have been wrongly reported and included in the database used. Even though those data could not have been used because the information contained in them is wrong they result in the rejection of a large number of trades from the sequence affecting the robustness of the results. This comment also refers to trades around trading halts or severe information inflows. The large number of trades which were ignored because they could not be matched also has the same effect.

4.3.3 Trade Classification Methods

Transactions databases, like the NYSE Trade and Quote (TAQ) database, which are widely available, do not provide information on trade direction necessitating the use of trade classification algorithms in order to classify trades as buyer- or seller-initiated. There are three rules reported in the literature for identifying trade direction: the *quote rule*, the *tick test (or tick rule)* and that proposed by Lee and Ready (1991) which we will denote by *LR* in the rest of the text.

The *quote rule* compares the trade price to the bid-ask prices at the prevailing quote and classifies as buyer-initiated those trades which occur above the mid-point of the quotes at the bid and ask, as seller-initiated those trades which occur below the mid-point and does not classify trades which occur at the mid-point.

The *tick classification rule* (or tick-test as it is usually called) examines the price movement relative to the previous trade and classifies a trade as buyer-initiated if the transaction price is above the previous price or if the previous tick-change was up. It classifies a trade as seller-initiated if the transaction price is below the previous price or if the previous tick-change was down. A variant of the tick test is the reverse tick test which is similar to the former, the main difference being that the next transaction price is used to classify the current trade. If the next trade occurs at a price higher than the current one (uptick) or at the same price which has followed a price increase (zero uptick) the current trade is classified as a sell whereas if it occurs at a price lower than the current one (downtick) or at the same price which has followed a price decrease (downtick) the current trade is classified as a buy.

The *LR rule* combines the above two rules by classifying each trade according to the quote rule first, then if a trade occurs at the mid-point the tick test is used. Lee and Ready (1991) noted that the tick test and the reverse tick test give the same classification when the current trade is bracketed by a price-reversal (i.e. when the price change before the trade is the opposite of the price-change after the trade) but do not give the same results when the trade is bracketed by a price continuation. In the latter case the tick test performs better than the reverse tick test if quotes are changed in the direction of the order-imbalance and the reverse tick test performs better if the quote change is in a direction opposite to that of the preceding trade.They also showed that the tick test outperforms the quote-based method.

To avoid the problem arising when quotes are reported ahead of the trades which triggered them they employed a method according to which they compared the trade to the quotes prevailing a few seconds before the trade.[103] They also considered a number of factors which lead to certain trades being ignored:

- Trades not preceded by a quote were ignored (opening trades)
- The simultaneous arrival of market buy and sell orders was not considered since it is an extremely rare occasion in the NYSE according to Hasbrouck (1988).[104]
- Trades transacted under unusual settlement conditions, as well as those reported out of sequence, were excluded (these are identified with special condition codes) as explained above. Details for the codes are provided in appendix E.

A number of studies have examined the accuracy of trade classification algorithms indirectly or directly.[105] Studies by Peterson and Fialkowski (1994), Harris and

[103] The time interval depends on the particular exchange and is period-specific. It can be derived using the analysis in Lee and Ready (1991).

[104] Hasbrouck (1988) argues that roughly 85 % of the transactions can be classified on the basis of their proximity to the current quotes. The rest of them, which consist of transactions occurring at the midpoint of the current quotes were classified according to three methods presented in his paper.

Hasbrouck (1996), Knez and Ready (1996) and Lightfoot et. al. (1999) have found that occasionally, buy orders are executed below the quote mid-point and sell orders above it. Lightfoot et. al. (1999) in particular have shown that the LR rule can lead to biased estimates of effective spreads and price improvement. Aitken and Frino (1996) have tested the tick rule on Australian data and found a 75% success rate.

Lee and Radhakrishna (1999) have used a TORQ sample of NYSE stocks smaller than that of Finucane (2000), to check the accuracy of LR and found that 93% of all trades are classified correctly using the LR rule. They also examined one factor that affects the accuracy of the trade classification algorithms, in particular the trade location relative to the quoted spread.

Odders-White (2000) used NYSE data from the TORQ database to test the accuracy of the three methods and found that LR performs better than the tick test (successful in 85% of the trades) and also examined the impact of trade location relative to the posted quotes, the trade size and the transactions in large and frequently traded stocks on the accuracy of the LR. She also showed that eliminating midpoint trades when using the LR may increase the accuracy and that the LR overestimates the order-processing costs and over-(under-) estimates the number of buys (sells) for small trades.

Finucane (2000) compared the LR rule, the tick test and the reverse tick test using NYSE data which identified trade direction and showed that for NYSE trades the LR and the tick test showed approximately the same performance whereas the reverse tick test performed consistently worse than the other methods. The finding that the performance of LR, relative to the tick test, is similar and not superior to the tick-test, as found by the studies of Ellis, Michaely and O'Hara (2000) for NASDAQ and

[105] Finucane (2000) talks about studies in which these algorithms are used in particular applications.

Odders-White (2000) for the TORQ database, can be explained by the presence of orders which are more likely to receive price improvement.[106] Such orders are crossed market orders, market orders stopped by the specialist and trades involving multiple parties on the same side of the trade. He also showed that trading volume, trade size, spread size and frequency of trades and quotes all affect the accuracy of the algorithms. In particular, trades resulting from market orders which allow price-improvement, larger times between trades and shorter times between quotes are associated with higher classification errors.

Finucane (2000) found no significant improvement in his results when adjusting for delays in reporting quotes and also showed that the tick test gives estimates closer to the true effective spread and both methods provide estimates of signed volume biased towards zero[107] but the biases for the tick test are smaller than those using the LR. Moreover, attempting to improve the accuracy by systematically eliminating certain types of trades can introduce additional biases.

Ellis, Michaely and O'Hara (2000) examined the validity and accuracy of trade classification methods using NASDAQ data. In particular they compared the quote rule, the tick rule and the Lee and Ready (1991) rule (LR) using a unique data set from NASDAQ from which they were able to identify trade direction and found that the success rates for the quote rule, the tick test and the LR rule were 76.4%, 77.66 and 81.05% respectively.[108] They also found that standard classification rules for sorting trades as buys or sells are as accurate for NASDAQ as for NYSE trades.

[106] Lee and Ready (1991) predicted that the tick rule would classify correctly at least 85% of midpoint and 90% of bid/ask trades but their predictions were based on a model which assumed constant quoted prices and independent Poisson processes for the arrival of market buy and sell orders. When these assumptions are not valid the precision of the tick test cannot be accurately predicted by their method.
[107] The effective spread can be calculated as being equal to twice the difference between the transaction price and the midpoint price times a trade indicator variable which takes the value of one (minus one) when the trade is a buy (sell). An alternative definition is that the effective spread is twice the absolute difference between the transaction price and the midpoint. The midpoint is the average of the bid and ask prices.

Classification success rate for trades inside the quotes was significantly lower, dropping to only 60% for mid-point trades and 55% for other trades inside the quotes. They also found that small trades are easier to classify which is opposite to the results of Odders-White (2000). They concluded that trades outside the quotes are difficult to classify, that better classification is achieved during periods of slow trading due to the higher incidence of trades at the quotes and that the substantial delay in the reporting of NASDAQ trades on the TAQ database does not affect the ability of the algorithms to classify trades.

The authors proposed an alternative trade classification rule that relies more on ticks than quotes. According to this rule trades at the ask (bid) are classified as buys (sells) and all other trades are classified according to the tick rule. This rule was shown to increase the classification accuracy for trades inside the quotes from 55 to 61% and a slight improvement for out-of-quote trades (81.9% against LR's 81%). However the algorithm performs significantly better when calculating effective spreads. Another important result was that the use of LR in NASDAQ overstates the size of the effective spread (a result which was also found by Finucane (2000) for NYSE stocks) whereas their algorithm improves this estimate by 10%. This method has been shown by Bessembinder (2002) to lead to smaller, albeit more accurate, estimates of trading costs, as compared to the Lee and Ready (1991) method.

Their investigation has also showed that trade classification rules have similar success in normal trading and Initial Public Offering (IPO) periods, that excluding Electronic Communications Networks (ECN) trades does not change classification accuracy and that reporting delays into the algorithms does not improve accuracy for NASDAQ data which implies that one can use TAQ data with no delay for

[108] Odders-White (2000) respective figures for NYSE were 75, 79 and 85%.

classification of NASDAQ trades from TAQ data.[109]

4.4 Summary Statistics

Tables AD.4.1 and AD.4.2 in Appendix D present summary statistics, split in deciles of trading activity, for the trades of the NYSE and NASDAQ sub-samples of the stocks selected randomly from the TAQ database for the month of October 1994. Tables AD.4.3 and AD.4.4, also in Appendix D, present summary statistics for the quotes of the same sub-samples. To facilitate the discussion of these statistics, these data have also been plotted in Appendix D.

The average daily volumes as measured in number of shares traded in a day for all stocks in a decile, which have been plotted in Figure AD.1, are only slightly higher for most NADAQ compared to NYSE deciles of the same category ranging from 70 thousand to sixteen million for NYSE and from 84 thousand to 20.7 million for NASDAQ. These volumes are a result of our data selection procedure during which an effort has been made so as to obtain same category deciles in both exchanges with stocks which have an approximately equal number of trades. This would make comparison easier and would allow more correct inferences to be made. The pattern in both sub-samples increases smoothly in the first nine deciles, then sharply till the tenth decile.

The average daily turnover of all stocks in a decile, measured in thousands of USD and plotted in Figure AD.2, follows the same pattern but figures are higher for NYSE stocks in all deciles, even though NASDAQ deciles contain 60 as opposed to 20 stocks for the NYSE deciles. This difference may be an indication that on average NYSE stocks are much higher-priced compared to those of NASDAQ. Differences

---

[109] For the London Stock Exchange Berhhardt, Dvoracek, Hughson and Werner (1999) have reported that large trades receive better execution than small trades.

become progressively larger as one moves from the low towards the high activity deciles except for the last one. The value of this variable ranges from 1 to 653 million USD for NYSE and from 0.8 to 614 million USD for NASDAQ stocks but the overall average for NYSE is approximately 37% higher compared to the NASDAQ figure.

The average change in price, as can be seen in Figure AD.3, is very close to zero for all deciles of both sub-samples and is negative for all but the first three of NYSE and four of the NASDAQ deciles. Overall maximum values are higher (minimum values are lower) for NASDAQ stocks and the absolute average value for this exchange is 4-5 times higher. A possible explanation for this observation may lie in the fact that the NYSE trading mechanism involves a larger number of participants which allows the price change process to proceed in smaller steps considering the ability to trade within the quotes.

The average time between trades, which is measured in seconds and has been plotted in Figure AD.4, follows a steeply decreasing pattern as one moves from low to higher activity deciles in both sub-samples and values are very close for the first and last NYSE and NASDAQ activity deciles. However, the pattern for NYSE is convex and that of NASDAQ concave leading to larger differences in the medium activity deciles. This difference in the two patterns could be attributed to the quicker matching of orders in NYSE due to inter-quote transactions and also to the fact that NASDASQ dealers place quotes, firm for a period of time during which they wait for traders to place orders and this may cause extra delay. For NYSE this time ranges from approximately 2,500 seconds (41.6 minutes) from the lowest to 59 seconds in the highest activity decile and for NASDAQ from 2,368 seconds (39.5 minutes) to 107 seconds respectively. On average it is approximately 891 seconds (14.9 minutes) for NYSE and 1,469 seconds (24.5 minutes) for NASDAQ. Maximum times can be very long and reflect the cases when a stock has not transacted for a few hours.

The average number of trades in a stock in a day, plotted in Figure AD.5, increases, as expected, smoothly till the ninth decile for both sub-samples and then steeply. Figures for NYSE are higher in most deciles compared to those in NASDAQ of the same activity category and range from 83 to 7,240 and from 75 to 11,250 for NASDAQ. However, in the last NASDAQ decile this value is much higher than that for NYSE and the overall average for NASDAQ is above that for NYSE by almost 20%. Maximum values as high as 9.4 thousand for NYSE and 16.8 thousand for NASDAQ are observed. The fact that the number of trades in a decile for NASDAQ can be higher than that for the equivalent NYSE decile reflects the fact already discussed, namely that NASDAQ deciles contain more stocks which in the highest activity deciles can have trading activity comparable to that of NYSE stocks.

Caution should be exercised when examining the quote statistics and attempt to make comparisons between those of NYSE and NASDAQ. The TAQ database does not report the full sizes of quotes for NASDAQ-traded stocks but only the Best Bid Offered (BBO). For issues with more than one market-maker only the size of the first market-maker to quote at that price is shown. Moreover, NASDAQ market-makers are obliged to quote for only a fixed number of shares. Therefore the sizes of the quotes, unlike in NYSE data, are not representative of the market size and this constitutes a major drawback of the database.

The average volume of the ask (offer), measured as the average number of shares offered in a day for all stocks in the decile, plotted in Figure AD.6, follows an upward pattern for NYSE but with increasing variation in the high-activity deciles and increases steeply from the seventh towards the last decile. Its value ranges from 288 thousand to 54 million in the highest decile. The same pattern, with a similar degree of variability, can be observed for the average daily volume of the bid, plotted in Figure AD.7, its range for NYSE being between 280 thousand and 53 million shares. This pattern in the low activity deciles may be caused by the fact that traders, including the specialist, are reluctant to commit to larger trades in low activity stocks since they will be unable to reverse their positions should unfavourable information arrives. The average figures for the total NASDAQ sample are much lower than those of NYSE for both the average daily volume of the ask (9.7 million for NYSE; 229 thousand for NASDAQ) and the bid (9.5 million for NYSE; 231 thousand for NASDAQ) owing to the fact that NASDAQ quotes in the TAQ database are only reported for the Best Bid Offered (BBO) whereas all quotes are reported for NYSE.

The average daily number of quotes has been plotted in Figure AD.15 and follows a rising pattern for NYSE (steeper from decile 6 towards the largest decile) ranging from 321 to 12,300 (average 3,445). These figures are significantly higher than those for NASDAQ which range from 89 to 864 (average 250). Overall, the average ask per quote and the average bid per quote are much higher for NYSE as compared to NASDAQ deciles, the difference increasing with increasing activity.

The average change in the ask (Figure AD.8) and bid (Figure AD.9) prices, which follow similar patterns, is close to zero for all NYSE deciles but varies a lot for NASDAQ deciles where it starts negative in the first decile, increases till the fifth, decreases abruptly in the sixth and then increases for the rest of the deciles. This erratic behaviour in the lower activity deciles probably indicates the greater uncertainty of the dealers who may attempt to adjust faster to new information or their tendency to overreact.

The average time between quotes, plotted in Figure AD.10, drops steeply as one moves from the low to the high activity NYSE deciles. As with the time between trades the pattern is concave for NASDAQ where inter-quote time is almost constant in the first six deciles and then decreases. However, this pattern arises from the fact that not all quotes are reported for NASDAQ stocks. For NYSE it is convex. NYSE times are much smaller than NASDAQ and range from 1,244 to 41 seconds for NYSE with an average of 373 and from 3,118 to 1,168 seconds for NASDAQ with an average of 2,754. As already mentioned, the larger number of participants and the ability to trade within the quotes may explain the pattern for NYSE.

The average change in the spread, as shown in Figure AD.11, is very close to zero for both sub-samples and the average change in the proportional spread, as shown in Figure AD.12, which is the spread expressed as a percentage of the midpoint price (bid plus ask divided by two), is also close to zero for both NYSE and NASDAQ.

The average spread, which has been plotted in Figure AD.13, decreases smoothly for NYSE and more steeply for NASDAQ as one progresses from the low towards the high activity deciles, probably owing to the higher uncertainty of NASDAQ dealers when trading in less liquid stocks and ranges from 0.46 to 0.31 USD for NYSE (average 0.4 USD) and from 0.68 to 0.30 USD for NASDAQ stocks (average 0.42 USD), both averages being close to 3/8ths. NASDAQ values are higher in all deciles the difference being higher in the low activity deciles which is in accordance with the remarks made earlier concerning the uncertainty of dealers.

The proportional spread, plotted in Figure AD.14, follows a similar, decreasing pattern with NASDAQ values being higher and ranging from 12.6 to 1.8 percent. NYSE values range from 4.45 to 0.84 % and their average value (2.21%) is much lower than that for NASDAQ (7.14%).

Chapter 5

The Quoted-Spread Model

5.1 Introduction

Based on the work of previous researchers, presented in chapter 3, we develop here a model of the quoted bid-ask spread from first principles but with two crucial modifications. The first is that we incorporate the volume factor explicitly into our equations both in the form of the volume of a particular trade as well as in the form of the volume of shares offered or bid prior to a particular trade. The second major difference between this model and the rest in the literature is the incorporation of the time factor between two successive trades, that is the waiting time between the two trades, through a more analytic equation for the Order Processing cost which incorporates both the time between trades and the volume at the bid or the ask (depending on the side of the spread considered) apart from the constant term which represents the fixed cost of a trade.

To arrive at the quoted spread equation we first develop the equations for the change in the bid and the ask separately and finally combine these two in the former. However, since certain terms, and thus some of the information in the data is lost during this procedure we estimate the bid- and ask-change equations separately so as to investigate the possibility of a difference in the behaviour of the parameters from one equation to the other.

Our discussion in this chapter follows the steps outlined below:

- First we develop the models for the Quoted Spread as well as the Ask- and Bid-Change Models from principles already presented in chapter 3.
- Next we provide summary statistics for the variables used in the estimation of the models.
- Results of the estimations are presented and discussed.

- The specification tests of the models and their results are provided and commented and
- Finally conclusions, based on the estimation and specification results for the above models, are discussed.

5.2 Development of the Models

The timing convention followed in the models is as follows: We denote by P_t the price at which the transaction at time t takes place. Prior to the transaction at time t it is assumed that, as in Glosten (1987), there exists a 'true', or common-information price of the asset at t-1, m_{t-1} which has been based on the trade preceding that time. Let us denote by κ and λ two very short intervals of time preceding transaction time t where $\kappa > \lambda$. We assume that, based on m_{t-1} and the trade at t-1 as well as on new public information which has arrived in the market from time t-1 till t-κ, market

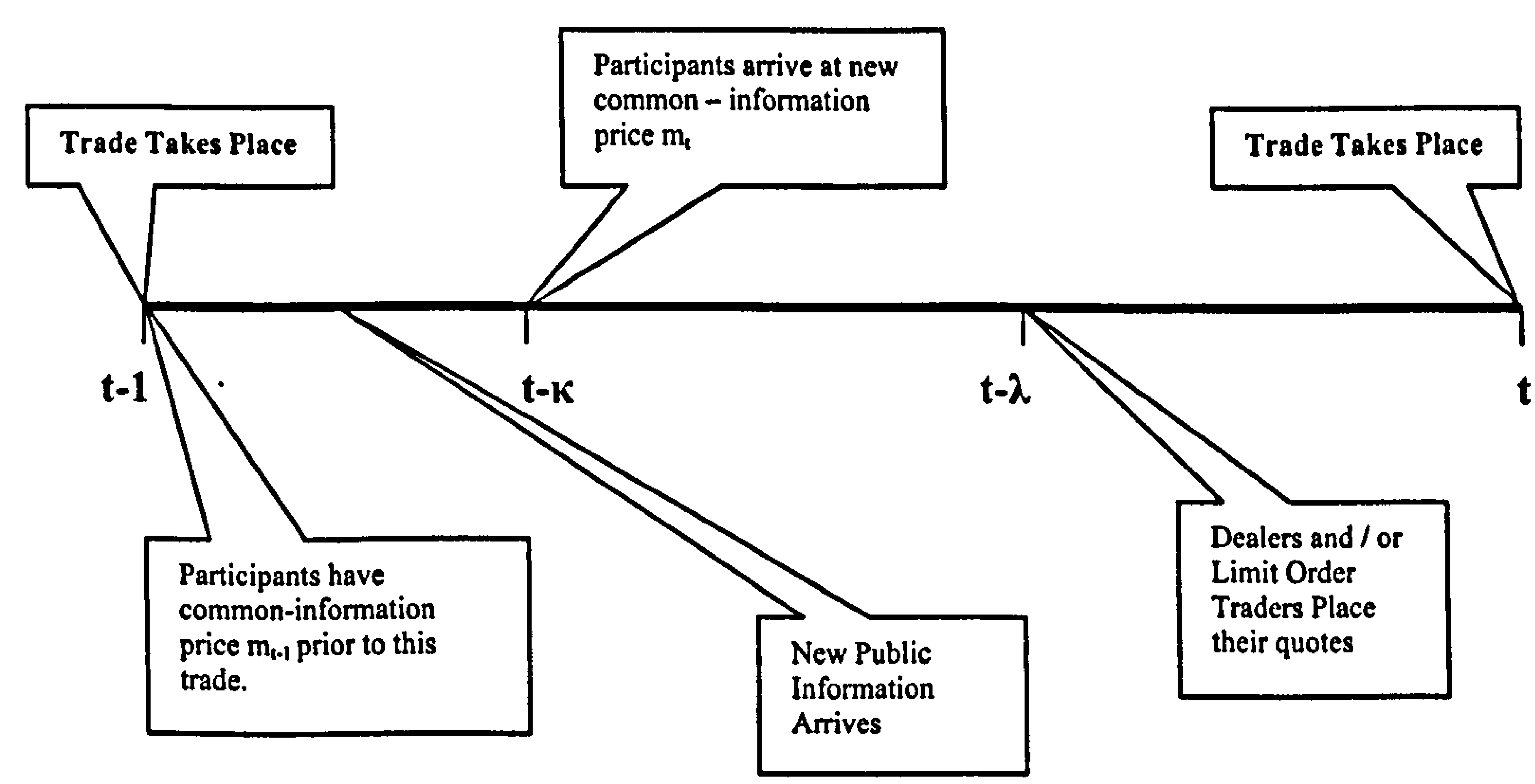


Figure 5.1 : Time Scale of Assumed Events for the Bid-Ask Spread Models

participants arrive at a new 'true' price m_t according to

$$m_t = m_{t-1} + A_{t-1}q_{t-1} + \varepsilon_t \tag{5.1}$$

where m_t is the true price of the stock at time t, A_{t-1} is the adverse-selection component at time t-1, q_{t-1} is the trade indicator variable (=+1 for trades which are buyer-initiated and –1 for seller-initiated) for the trade which occurred with price P_{t-1} at time t-1 and ε_t is a term which accounts for the public information signal which affects the fundamental value of the stock and is assumed to be an independent, identically distributed random variable. Based on this new price, dealers (and limit order traders) place their quotes (and limit prices) at time t-λ just before a trade occurs which leads to the new transaction price at t, P_t.

Equation (5.1) is identical to that used in Glosten and Milgrom (1985) and Huang and Stoll (1997) among others. There is a slight difference in the timing conventions used in the models of Huang and Stoll (1997) and Madhavan, Richardson and Roomans (1997) which is worth pointing out. In the latter, the 'true' price is formed after the trade and is the expected value of the stock following the trade at t. In the former paper, as well as in our model, the 'true' price is formed before the transaction at t. Even though both conventions are right when considered under the right perspective we believe that the pre-trade convention is more realistic since traders and / or dealers should base their quotes on the most recent common-information price before a trade takes place.

The term $A_t q_t$ accounts for the effect of adverse selection,[110] or information asymmetry or the permanent impact from the information furnished by the order flow as explained above. The adverse selection component is assumed to be equal for both sides of the spread (bid and ask) since it is assumed to depend only on the size and direction of the previous trade in anticipation of the unexpected information component of that trade. Models analyzing the spread and the intraday price formation describe the term A_t in equation (5.1) as the innovation in trade flow either as a deviation between the observed and the expected change in trade flow like Madhavan, Richardson and Roomans (1997), or as a fraction of the half spread like Huang and Stoll (1997). We follow the former which is based on Glosten (1987) in that A_t represents the total size of the adverse selection component which is equivalent to the permanent effect of the previous trade on the price. However we extend the model to incorporate the effect of volume in the above equation, as in Glosten and Harris (1988) and Lin, Sanger and Booth (1995) since the size of a particular order executed probably signals to the dealer the degree of the information carried by the trade flow. Assuming that the adverse-selection cost is symmetric around the mid-point of the spread, so that this cost for the ask, A_t^a, is equal to the same cost for the bid, A_t^b, expressing this component as the product of an adverse selection parameter, θ, and the volume of the previous trade, leads to:

$$A_t^a = A_t^b = \theta V_{t-1}$$

where θ indicates the change in the commonly accepted price per unit volume of shares traded and V_{t-1} is the volume of the trade at t-1 expressed as number of shares traded. Trading volume is assumed to be distributed[111] according to

$$V_t \sim \left(\bar{V}, \sigma_V^2\right)$$

and uncorrelated with the trade indicator variable, i.e.

$$\operatorname{cov}\left(q_{t+k}, V_{t+l}\right) = 0$$

for all k and l where k,l = 0, 1, 2, ...

[110] Information asymmetry is assumed to depend upon the volume of the previous trade, the volume difference between the bid and ask quotes and the difference between supply and demand as expressed by the difference between buy and sell volumes over a given period.

[111] We avoid the assumption that volume is normally distributed since research cited in section 2.8 has indicated that it is not (for example Blume, Easley and O'Hara (1994)).

To distinguish between the inventory and the adverse selection effects one needs to include relations which describe the probability of a particular type of trade (purchase or sale) occurring and this is usually accomplished by modeling the serial correlation in trade flows because the dealer is tempted to revise his quote to the opposite direction of the last trade in order to balance his inventory. This has the effect of introducing negative serial correlation in trades. This effect is different from the negative serial correlation in transaction price changes owing to the bid-ask bounce as described in Roll (1984). Moreover we subtract the anticipated effect of the trade to include in the adverse selection component only the innovation in the trade flow which is that part of the new information likely to influence the price. As already discussed above we use the relationship for the conditional expectation of the trade indicator variable q_t developed by Madhavan, Richardson and Roomans (1997)

$$E(q_t \mid q_{t-1}) = \rho q_{t-1}$$

where ρ is the first order autocorrelation coefficient of the trade-indicator variable, q_t. Therefore, the unanticipated effect of the trade on the 'true' price of the stock which is not known until after the trade at time t has taken place and which is to be substituted in equation (5.1) is given by

$$\theta q_{t-1} V_{t-1} - E(\theta q_{t-1} V_{t-1} \mid q_{t-2}) = \theta\left(q_{t-1} V_{t-1} - \rho \bar{V} q_{t-2} \right)$$

and equation (5.1) becomes

$$m_t = m_{t-1} + \theta\left(q_{t-1} V_{t-1} - \rho \bar{V} q_{t-2} \right) + \vartheta_t \qquad (5.3)$$

where ϑ_t is an i.i.d random variable.

Therefore, whereas actual trades affect quote adjustment which is motivated by inventory-holding concerns, it is innovations in trades which affect adverse selection and this principle is used to separate the two effects.

Based on the expected 'true' price following the trade at time t-1 and prior to the trade at time t the dealer sets his bid, b_t and ask, a_t quotes by adding (subtracting) the inventory holding cost at time t, IV_t, and the order-processing cost at time t, O_t, to (from) the expected true price of the stock which will be arrived at by traders after the trade at time t has taken place and this expected true price is conditional upon the direction of trade (whether buyer- or seller-initiated) as shown below:

$$a_t = E(m_{t+1} \mid q_t = +1) + IV_t + O_t^a + \eta_t \tag{5.4}$$

$$b_t = E(m_{t+1} \mid q_t = -1) - IV_t - O_t^b + \nu_t \tag{5.5}$$

where

$$E(m_{t+1} \mid q_t = +1) = m_t + \theta \bar{V}(1 - \rho q_{t-1}) \tag{5.4a}$$

$$E(m_{t+1} \mid q_t = -1) = m_t - \theta \bar{V}(1 + \rho q_{t-1}) \tag{5.5a}$$

In the above two equations η_t and ν_t account for the effect of price discreteness and are assumed to be asymptotically uniformly distributed random variables with a zero mean, as in the Glosten and Harris (1988) model described in section 3.4 and uncorrelated with each other. Price discreteness arises from the fact that transactions occur at prices rounded to the nearest $1/16^{th}$ whereas the true price is a real number. The terms O_t^a and O_t^b carry into the model the effect of risk bearing of the market maker as well as the cost of providing immediacy and order processing. Following Glosten and Milgrom (1985) and Madhavan, Richardson and Roomans (1997), equations (5.4) and (5.5) assume that the bid and ask prices quoted are ex-post rational in the sense that they are set in such a way so that the dealer is compensated conditional on whether a trade occurs on the bid or ask, that is they are regret-free. For this reason, in the above two equations we use the expected value of m_{t+1} which is the expected 'true' price of the stock given that a transaction occurs at the bid or ask at time t. In this way, the unanticipated part of the innovation in trade flow, equivalent to the adverse selection component, is incorporated into our model.

The order-processing costs O_t^a and O_t^b are usually taken to be a constant proportion of the spread but in our model we assume they are different for the bid and the ask sides of the spread. It is further assumed that this component depends on a constant factor α, which accounts for a minimum fixed cost charged by the provider of immediacy, the time which has elapsed between the last two trades at times t-2 and t-1, denoted as τ_{t-1} and the volumes (depth) of the existing quotes at the bid, V_t^b and the ask, V_t^a, at time t (the standing quotes) according to the following equations:

$$O_t^a = \alpha + \beta\tau_{t-1} + \gamma V_t^a \tag{5.6}$$

$$O_t^b = \alpha + \beta\tau_{t-1} + \gamma V_t^b \tag{5.7}$$

V^a and V^b are the number of shares quoted at the ask and bid respectively and α and β are parameters. Glosten and Harris(1988) also model this cost as a linear function of the size of the trade and their specification search has shown them that γ is equal to zero.[112] Since in markets like the NYSE certain participants like the specialist have access to the book of orders and are aware of the price schedule of buy and sell orders we assume that quotes are set based on the volumes at the bid or the ask. That is the dependence of the order-processing cost on the number of shares quoted is likely to be due to the response of the specialist to what appears to be an information event. This can also be justified by the fact that if a long period elapses since the last trade and the depth of quotes builds up it is more likely that a trade will consume more shares than before. In his analysis of the stock-exchange specialist system Stoll (1985) also adopts a linear formulation for the order processing costs which includes a constant term, a term which varies with the number of transactions and one which varies with the size of transactions.

The intuition behind the use of the time between trades, τ_t, comes from real trading situations where the absence of trading may itself add to the information set. Glosten and Harris (1988) have also used it to describe the process of public information innovation by expressing the mean and variance of ε_t as functions of it. We incorporate time between transactions into the equation since we wish to model this effect separately and not consider it merely as part of the public information innovation. Moreover, it is natural to assume that it is an additional factor affecting the placing of quotes, directly associated with the cost of waiting of the dealer. For this reason only the latest quotes, the ones closest in time to the transaction should be used in the analysis. It is worth noting that according to our interpretation the likely time-effect on the spread may serve both as a compensation of the dealer for the extra waiting time as well as an information event.

As in Huang and Stoll (1997) the inventory effect at time t, IV_t is modeled as :

$$IV_t = \delta \sum_{i=1}^{t-1} q_i V_i \qquad (5.8)$$

where the summation sign shows the net inventory accumulation in a stock since the beginning of trading in a particular day. When substituted into equations (5.4) and (5.5) parameter δ indicates the effect of this inventory accumulation on the quoted ask

---

[112] Only 20 firms were used in this study for specification search and only 800 successive price changes for each stock. The dataset used was much smaller than the one used in this thesis and only for NYSE

and bid prices respectively. Inventory models of the spread, like Stoll (1978) and Ho and Stoll (1981), assume that the inventory effect affects the midpoint of the bid ask spread in the sense that the dealer adds it on the true price of the stock to arrive at the midpoint, that is the dealer adjusts the midpoint between the bid and ask in such a way as to compensate for a deviation from his desired inventory level, based on the inventory imbalance from previous trades. Such theories however also assume that the spread set by dealers is constant in the sense that they change its position and not its size whenever they react to changes in their inventory with the aim of maintaining a balanced fixed inventory at least within a certain time horizon. Since we attempt to model the asymmetry around the true price we add or subtract the inventory holding cost component directly on the expected true price of the stock after the trade in the ask and bid quote equations (5.4) and (5.5) respectively. The effect of the inventory component in these equations will therefore depend on the nature of the inventory imbalance up to and including time t-1. If the dealer has a positive inventory, a surplus which has probably resulted from a dominance of seller-initiated trades which have a negative trade-indicator variable, q_t, he will wish to lower his ask and his bid so as to discourage sales to him due to his low bid and at the same time encourage purchases from him due to his low ask. If the dealer's inventory accumulation is negative due to a dominance of buyer-initiated trades the dealer will wish to increase both his ask and bid prices so as to encourage sales to him by traders due to his high bid and at the same time discourage purchases from him due to his high ask. It follows that in both of the above cases the inventory parameter, δ, in equation (5.4) should be negative whereas that in equation (5.5) should be positive.

---

stocks.

Taking the first differences of equations (5.4) and (5.5), substituting for O_t^a from equation (5.6) into equation (5.4) , substituting for O_t^b from equations (5.7) into equation (5.5) and substituting for Δm_t from equation (5.3) we obtain:

$$\Delta a_t = \theta q_{t-1} V_{t-1} - \theta \bar{V} \rho q_{t-1} + \delta q_{t-1} V_{t-1} + \beta \Delta \tau_{t-1} + \gamma \Delta V_t^a + \Delta \eta_t + \vartheta_t \quad (5.9)$$

$$\Delta b_t = \theta q_{t-1} V_{t-1} - \theta \bar{V} \rho q_{t-1} - \delta q_{t-1} V_{t-1} - \beta \Delta \tau_{t-1} - \gamma \Delta V_t^b + \Delta \nu_t + \vartheta_t \quad (5.10)$$

where

$$\Delta V_t^a = V_t^a - V_{t-1}^a \quad \text{and} \quad \Delta V_t^b = V_t^b - V_{t-1}^b$$

and $\Delta \eta_t$ and $\Delta \nu_t$ are assumed to be , uniformly distributed random variables with a zero mean. The inventory term in equations (5.9) and (5.10) is the difference between the cumulated inventories at times t-1 and t-2.

Since the quoted spread is expressed as the difference between the bid and ask quotes:

$$S_t^Q = a_t - b_t$$

the expression for the change in the *quoted* spread becomes:

$$\Delta S_t^Q = \Delta a_t - \Delta b_t = 2\delta q_{t-1} V_{t-1} + 2\beta \Delta \tau_{t-1} + \gamma\left(\Delta V_t^a + \Delta V_t^b\right) + \Delta \xi_t \quad (5.11)$$

where

$$\Delta \xi_t = \Delta \eta_t - \Delta \nu_t$$

is assumed to be a uniformly distributed random variable with a zero mean.

There is a large number of stocks used in this study which, apart from the fact that they trade at different exchanges, which implies that the different mechanisms will probably have a different impact on the behavior of both quotes and prices, they also have prices and spreads which exhibit large variation. For this reason, in order to be able to compare the parameters of the components of the spread as well as the other parameters, it is more meaningful to study them relative to some other variable such as the quoted spread. Therefore by dividing the left-hand side of equation (5.11) by the quoted spread at time t-1 we obtain the change in the quoted spread between times t and t-1 relative to the quoted spread at time t-1 which implies that since the same variables are used for estimation in the right-hand side the parameters of the components of the spread, that is δ, β and γ will now be given relative to the quoted spread at time t-1. Thus equation (5.11) becomes:

$$\Delta \bar{S}_t^Q = 2\delta_p q_{t-1} V_{t-1} + 2\beta_p \Delta\tau_{t-1} + \gamma_p \left(\Delta V_t^a + \Delta V_t^b\right) + \Delta \upsilon_t \qquad (5.11b)$$

where $\Delta \bar{S}_t^Q$ denotes the change in the quoted spread at time t expressed as a fraction of the spread at time t-1 and the subscript p under the parameters δ, β and γ denotes parameters expressed as a fraction of the spread at time t-1 so that:

$$\delta_p = \frac{\delta}{S_{t-1}}, \quad \gamma_p = \frac{\gamma}{S_{t-1}}, \quad \beta_p = \frac{\beta}{S_{t-1}} \quad \text{and} \quad \theta_p = \frac{\theta}{S_{t-1}}$$

It is worth pointing out that, owing to our data-selection procedure, the spread at time t-1 cannot be zero.

To make comparison between the models easy we also divide the left-hand sides of equations (5.9) and (5.10) by the quoted spread at time t-1 so as to obtain:

$$\Delta \bar{a}_t = \theta_p q_{t-1} V_{t-1} - \theta_p \bar{V} \rho q_{t-1} + \delta_p q_{t-1} V_{t-1} + \beta_p \Delta\tau_{t-1} + \gamma_p \Delta V_t^a + \Delta\chi_t + \varpi_t \qquad (5.9b)$$

$$\Delta \bar{b}_t = \theta_p q_{t-1} V_{t-1} - \theta_p \bar{V} \rho q_{t-1} - \delta_p q_{t-1} V_{t-1} - \beta_p \Delta\tau_{t-1} - \gamma_p \Delta V_t^b + \Delta\lambda_t + \pi_t \qquad (5.10b)$$

where $\Delta \bar{a}_t$ and $\Delta \bar{b}_t$ denote the change in he quoted ask and bid respectively at time t, expressed as a fractions of the spread at time t-1 and the subscript p under the parameters θ, δ, β and γ denotes the same parameters also expressed as a fraction of the spread at time t-1.

Since there is no reason to support that the random variables $\Delta\eta_t$, Δv_t and ϑ_t in equations (5.9), (5.10) as well as the random variable $\Delta\xi_t$ in equation (5.11) are correlated with the spread at time t-1 we assume that random variables $\Delta\chi_t$ and $\Delta\lambda_t$ in equations (5.9b), (5.10b) are also uniformly distributed random variables with a zero mean and that ω_τ and π_τ are i.i.d random variables. Moreover the random variable $\Delta\upsilon_t$ in equation (5.11b) is also an i.i.d random variable.

In selecting the spread at time t-1 to express the parameters to be estimated, instead of that at time t, we aim to study the changes in the quoted spread, the ask and bid prices in terms of explanatory variables which lag in time behind the dependent variable in order to investigate whether these can in a way explain the formation of the dependent variables. In other words we try to answer the question as to how the variables at time t-1 as well as the change in the standing quotes prior to the trade affect the setting by the dealer of the variables which affect the quoted spread.

5.3 Data

The variables used for the estimation of the models in this chapter have been obtained following the procedures described in chapter four. The technical details for extracting the data and estimating the variables are given in appendix C. In this section we present and discuss these variables and their correlations for the two exchanges considered.

5.3.1 Description

Tables AG.5.1 and AG.5.2 in Appendix G present summary statistics for the variables of the quoted spread and the bid and ask change models for the NYSE and the NASDAQ samples of stocks. It is important to note that the proportional change in spread is expressed as a fraction of the spread at time t-1. The average change in spread is very small and fluctuates around zero for NYSE but for NASDAQ it has its highest value in the lowest activity decile, then decreases as one moves towards the high activity deciles.

Whereas the change in spread from time to time is very small, the average change in the proportional spread is 23 percent for NYSE and 13 for NASDAQ. The pattern for NYSE is increasing as one moves towards the higher activity deciles, then slightly decreases at the last two deciles. For NASDAQ it starts high in the first decile, then follows a U-shape till the ninth decile after which it drops.

The average number of shares traded increases with the size of the decile but on average is of the order of 1.65 thousand shares for NYSE (1.5 for NASDAQ) even though the maximum number can reach a few million. The average spread is of the order of 0.4 USD for both NYSE and NASDAQ stocks and has a distinct trend decreasing progressively as one moves from the lowest to the highest activity deciles.[113]

The change in the number of shares offered fluctuates around zero but after the eighth decile becomes increasingly negative for NYSE. The magnitude of this variable for NYSE is however around ten times that of NASDAQ. For NYSE the same pattern is observed for the change in the number of shares bid and the magnitude of this variable is almost four times larger for NASDAQ compared to that for NYSE. The pattern for NASDAQ however is one of variation around the mean.

As expected, the change in the waiting time between NYSE trades decreases with higher activity volume deciles. The pattern for NASDAQ is not representative of the true trading conditions since only a few quotes are available and therefore not all trades (only those matched with quotes at the BBO) have been included in the analysis. On average however these changes are larger for NYSE. The average change in waiting time is about 215 seconds for NYSE and 35 for NASDAQ.

The average change in the offer, expressed as a proportion of the spread at time t-1 is positive and that in the bid is negative for both NYSE and NASDAQ - for which absolute values are about ten times higher implying more abrupt quote adjustment. For NYSE it decreases for the first three deciles then increases to zero till the seventh decile and remains around this value in the highest activity range. For NASDAQ, it also starts from a high value, fluctuates around zero till the sixth decile, then remains close to zero. For NYSE the pattern is repeated for the average change in the bid price where the average value is generally larger compared to that for the offer price.[114] This is also true for NASDAQ but the pattern is one of a low value at the first decile, increasing till the fourth and remaining around zero thereafter.

5.3.2 Correlation of variables

The correlations among the variables used to estimate the models developed in this chapter are discussed separately for NYSE and NASDAQ in the following subsections.

5.3.2.1 NYSE

There is negative serial correlation between the proportional change in the quoted spread and the change in the offer size or bid size which is of the order of 0.2 to 0.3,

---

[113] NYSE spreads exhibit slightly higher variation around the mean as compared to NASDAQ spreads probably owing to the intervention of the specialist who makes the process more dynamic. Extremely low values appear in NASDAQ stocks and extremely high in NYSE.
[114] Note that this discussion refers to absolute values.

except for the last three deciles where it is around 0.15. There is also positive correlation between it and the proportional change in the offer price (negative for that in the bid price) of the order of 0.6 which drops to 0.4 for the proportional change in offer price only in decile 9.

The magnitude of the correlation between the proportional change in spread with the spread drops after decile 8.

The volume of trading in the previous trade shows low correlation (less than 0.05) with all variables.

There is also positive correlation between the change in the size of the offer and that of the bid of the order of 0.3 to 0.4 which increases to around 0.45 after the fifth decile. The change in the size of the offer is also negatively correlated with the proportional change in the offer price (0.10 to 0.20) but positively correlated with both the proportional change in the bid price (0.2 but negligible after decile nine) and the previous spread (0.15 but negligible after decile 6; decile 7 is however 0.12).

The change in the size of the bid shows a similar pattern with negative correlation (0.2) with the proportional change in the offer price, but positive correlation with both the proportional change in the bid price (0.12 to 0.23 but negligible at decile 9) and the previous spread (0.15 but negligible after decile 5, seven excluded).

The change in the waiting time between trades shows no significant correlation with all variables.

The trade indicator variable is positively correlated with the proportional change in the offer price (0.2 to 0.3 but drops after decile seven to 0.12 and then becomes negligible) and its correlation with the proportional change in the bid price shows a similar pattern.

The proportional change in the offer price is negatively correlated with both the proportional change in the bid price (0.20 to 0.40) and the previous spread (0.30 to 0.50) and the proportional change in the bid price is positively correlated with the previous spread (0.36 to 0.62 attaining the highest values at the large activity deciles).

5.3.2.2 NASDAQ

The proportional change in spread is positively correlated with the spread (average value of coefficient is 0.84), the change in the offer size (0.1 to 0.4 in all but the two highest activity deciles), the change in the bid size (0.01 to 0.44), the proportional change in the offer price (strong correlation ranging from 0.86 to 0.45) and negatively correlated with the proportional change in the bid price (-0.24 to – 0.69) and the previous spread (- 0.21 to – 0.62).

The change in the proportional spread is negatively correlated with the volume of the previous trade only in the lowest activity decile. It is positively correlated with both the change in the size of the offer (0.50 to 0.12) and the change in the size of the bid (0.32 to 0.01) with the larger values encountered in the low activity deciles. It is also positively correlated with the proportional change in the offer price (0.40 to 0.64) but negatively with the proportional change in the bid price (-0.36 to –0.53). Its correlation with the waiting time between trades is only significant in the three lowest activity deciles and almost non-existent with the trade indicator variable. Significant negative correlation of this variable with the previous spread is however observed in all deciles and ranges from –0.25 to – 0.49.

The volume of the previous trade is only significantly correlated with the change in the waiting time between trades (four deciles), the trade indicator variable (two deciles), the proportional change in the offer price (one decile), the proportional change in the bid price (0.19 and 0.53 in the first two deciles respectively) and with the previous spread (weakly negative in six deciles).

The change in the size of the offer and that of the bid are significantly negatively correlated in four deciles (- 0.10 to - 0.18). The former is also positively correlated with the proportional change in the offer price (0.10 to 0.35) in all deciles whereas the change in the size of the bid is negatively correlated with the proportional change in the bid price in six deciles (-0.13 to –0.21) and with the previous spread in four deciles (-0.48 to –0.11).

The change in the waiting time between trades is not significantly correlated with any of the variables except weakly negatively with the trade indicator variable in the four lowest activity deciles and the proportional change in the bid price in the second and third deciles.

The trade indicator variable is positively correlated with the proportional change in the offer price in all deciles (0.46 to 0.16), with the proportional change in the bid price (0.15 to 0.40).

Finally the proportional changes in the offer and bid prices are positively correlated (0.17 to 0.38) in all deciles. The proportional change in the offer price is negatively correlated with the previous spread (- 0.01 to - 0.47) in all deciles but the proportional change in the bid price is positively correlated with the previous spread in only five of the highest activity deciles (0.19 to 0.33).

5.4 Estimation of the Models

The procedure for estimating the parameters of the models developed in this chapter is described in this section and the results of the estimations are provided in appropriate tables. Details on the theoretical underpins of the Generalized Method of

Moments (GMM) methodology which has been used for estimation are given in Appendix A.

5.4.1 The Quoted Spread Model

The basic model to be estimated is that described by equation (5.11b) which is the difference between equations (5.9b) and (5.10b) and is recited below for easier reference:

$$\Delta \bar{S}_t^{Q} = 2\delta_p q_{t-1} V_{t-1} + 2\beta_p \Delta\tau_{t-1} + \gamma_p \left(\Delta V_t^{a} + \Delta V_t^{b}\right) + \Delta\upsilon_t \qquad (5.11b)$$

Equations (5.9b) and (5.10b) describe the changes in the ask and bid prices between times t and t-1, relative to the spread at time t-1. However, since some information is lost in the differencing process we estimate both of these equations separately. This allows us to study the bid and offer change process separately and provides more insight into the quoted spread change process.

Equation (5.11b) involves three parameters, namely δ_p, β_p and γ_p and five variables.

In selecting the GMM estimation method we attempt to avoid making strong assumptions regarding the price-generating process. We also do not impose strong distributional assumptions on the error terms in our equations which account for a number of different stochastic information events as well as rounding errors. Moreover, during estimation we can account for various types of serial correlation of the residuals as well as any unknown form of conditional heteroscedasticity. Finally the latest literature on the testing of the results from GMM estimation, details of which are given in appendix B, can be used for specification testing.

In estimating this equation using the Generalized Method of Moments (GMM), apart from the above parameters which we can denote by the vector φ, we require a set of orthogonality restrictions which can be represented by the following vector function of population moments, h_t :

$$f(\varphi, h_t) = \begin{Bmatrix} e_t V_{t-1} \\ e_t q_{t-1} \\ e_t \Delta \tau_{t-1} \\ e_t \Delta V_t^a \\ e_t \Delta V_t^b \end{Bmatrix}$$

where e_t is the error term in equation (5.11b) which corresponds to $\Delta \upsilon_t$.

The set of the above restrictions correspond to strict exogeneity conditions in the sense that

$$E[e_i \mid X_1, X_2, ..., X_n] = 0 \qquad \text{(i=1,2,...,n)}$$

where X_j is the vector of all the observations of regressor j.

Thus the expected value of the error term, conditional on the regressors for all observations is equal to zero. This implies that the unconditional mean of the error term is zero:

$$E(e_i) = 0 \qquad \text{(i=1,2,...,n)}$$

and also that the above five regressors are orthogonal to the error term for all observations:

$$E(x_{jk} \cdot e_i) = 0 \qquad \text{(i,j=1,......,n; k=1,.........,K)}$$

or

$$E(X_j \cdot e_i) = \begin{Bmatrix} E(x_{j1} e_i) \\ E(x_{j2} e_i) \\ \cdot \\ \cdot \\ \cdot \\ E(x_{jk} e_i) \end{Bmatrix} = \mathbf{0} \text{ (for all i,j)} \quad \text{and} \quad \mathbf{0} \text{ is (Kx1)}$$

Therefore the regressors are assumed to be orthogonal to the past, current and future error terms (or the error terms are assumed to be orthogonal to the past, current and future values of the regressors).

Even though the above conditions are difficult to be satisfied in finite samples in a time series model, they are derived from the assumptions made in constructing the model. We therefore investigate the validity of these assumptions, during the specification tests on the results derived from our estimations, in order to identify which particular conditions appear not to be valid in the particular samples used.

This vector valued function exactly identifies the parameter vector φ by minimizing the function

$$Q_T(\varphi) = g_T(\varphi)' S_T g_T(\varphi)$$

where $g_T(\varphi)$ is the sample mean of $h(\varphi, x_T)$, where $Q_T(\varphi)$ is the function to be minimized and S_T is the Newey-West (1987) estimator as explained in appendix A.

The implied moment conditions imposed for estimation are:

$$E[f(\varphi, h_t)] = 0$$

There are three parameters to be estimated and five moment conditions, therefore this model is over-identified by a degree of two.

In our estimations we express volumes in thousands of shares and the change in the waiting time between trades in hours.

5.4.2 The Offer-Change and the Bid-Change Models

The models to be estimated are those described by equations (5.9b) and (5.10b). In the last two equations $\Delta \bar{a}_t$ and $\Delta \bar{b}_t$ describe the changes in the ask and bid prices respectively as a fraction of the spread at time t-1 and are re-written below for ease of reference:

$$\Delta\bar{a}_t = \theta_p q_{t-1} V_{t-1} - \theta_p \bar{V} \rho q_{t-1} + \delta_p q_{t-1} V_{t-1} + \beta_p \Delta\tau_{t-1} + \gamma_p \Delta V_t^a + \Delta\chi_t + \varpi_t \quad (5.9b)$$

$$\Delta\bar{b}_t = \theta_p q_{t-1} V_{t-1} - \theta_p \bar{V} \rho q_{t-1} - \delta_p q_{t-1} V_{t-1} - \beta_p \Delta\tau_{t-1} - \gamma_p \Delta V_t^b + \Delta\lambda_t + \pi_t \quad (5.10b)$$

Each equation involves five parameters, namely ρ, θ_p, δ_p, β_p and γ_p, which make up the parameter vector φ, and five variables. Since ρ and the average volume can be calculated in advance from the data and then treated like constants for the particular decile we can reduce the estimation problem into a four parameter- four variable one. The vector function of population moments, h_t, is now:

$$f(\varphi, h_t) = \begin{Bmatrix} e_t^a V_{t-1} \\ e_t^a q_{t-1} \\ e_t^a \Delta\tau_{t-1} \\ e_t^a \Delta V_t^a \end{Bmatrix}$$

for the Offer-Change model and

$$f(\varphi, h_t) = \begin{Bmatrix} e_t^b V_{t-1} \\ e_t^b q_{t-1} \\ e_t^b \Delta\tau_{t-1} \\ e_t^b \Delta V_t^b \end{Bmatrix}$$

for the Bid-Change model, where e_t^a and e_t^b are the combined error terms in equations (5.9b) and (5.10b) respectively, that is:

$$e_t^a = \Delta\chi_t + \varpi_t \qquad \text{and} \qquad e_t^b = \Delta\lambda_t + \pi_t$$

The above error terms have been used to model both random information events as well as rounding-off errors due to price-discreteness.

The comments regarding the strict exogeneity conditions of the regressors which were made in section 5.4.1 are also valid here, that is the

$$E(X_j \cdot e_i) = \begin{Bmatrix} E(x_{j1}e_i) \\ E(x_{j2}e_i) \\ \vdots \\ E(x_{jk}e_i) \end{Bmatrix} = \mathbf{0} \quad \text{(for all i,j)} \qquad \text{and} \qquad \mathbf{0} \text{ is } (\text{Kx1})$$

The validity of the above strict exogeneity conditions The for the four regressors in each model is examined in the specification tests which follow the estimation of the models.

This vector valued function exactly identifies the parameter vector φ by minimizing the function

$$Q_T(\varphi) = g_T(\varphi)' S_T g_T(\varphi)$$

where $g_T(\varphi)$ is the sample mean of $h(\varphi, x_T)$, where $Q_T(\varphi)$ is the function to be minimized and S_T is the Newey-West (1987) estimator as explained in appendix A.

The implied moment conditions are:

$$E[f(\phi, h_t)] = 0$$

There are four parameters to be estimated and four moment conditions, therefore this model is exactly identified.

In our estimations we express volumes in thousands of shares and the change in the waiting time in hours.

5.5 Discussion of Results

The results from the GMM estimation of the three models in this chapter are discussed in turn in the subsections which follow.

5.5.1 The Quoted Spread Model

Tables AG.5.3 and AG.5.4 in Appendix G show the results of the GMM estimation for the Quoted Spread Model using the NYSE and NASDAQ random samples of common stocks respectively. These data are plotted in figures 5.2 to 5.4 where each graph shows the variation of one parameter for both NYSE and NASDAQ in order to facilitate the comparison between the two exchange mechanisms.

As can be seen from figure 5.2 the size of the Quoted Volume Parameter γ_p follows completely distinct patterns for the two types of exchange. In NYSE it is negative and increases as one moves from the low (1) towards the highest (10) activity deciles. Its value is significant for all deciles and ranges from -0.153 to -0.011 total sample value is -0.015). That is the per thousand share impact of the total volume of shares quoted at the bid and the ask accounts for 15 to 1 percent of the proportional change in the spread, which is calculated using the spread at time t-1 in the denominator. For NASDAQ stocks it is positive and follows a U-shape with its lowest value in the fifth decile while those at the two extremes are the largest values. NASDAQ values range from 1.23 to 0.24 (total sample value is 0.4852). Overall, absolute values for NASDAQ are much larger compared to NYSE.

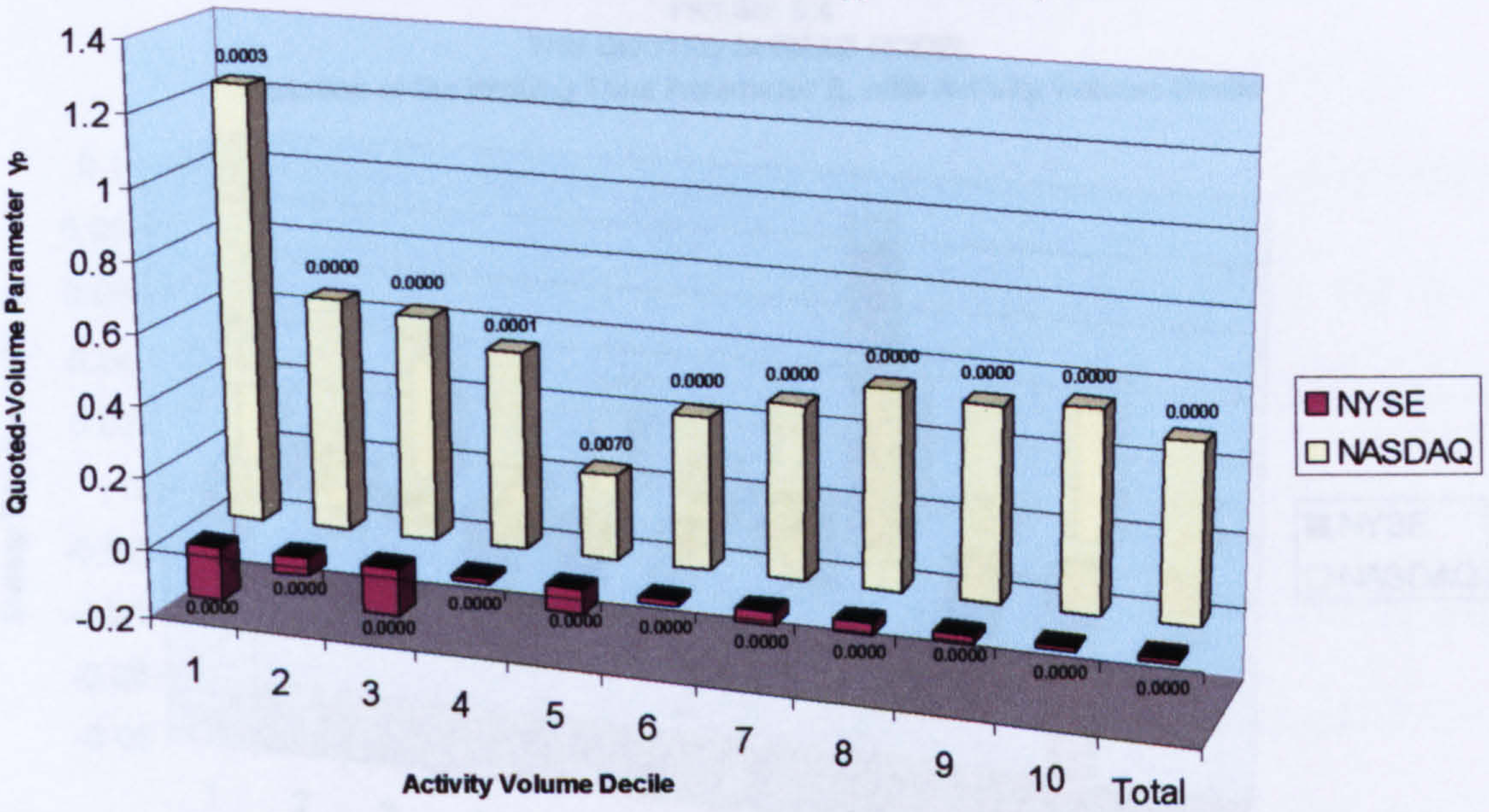
The inventory cost parameter δ_p, plotted in figure 5.3 is only significantly different from zero in 3 out of ten deciles for NYSE and for only two of the NASDAQ deciles (1 and 10)[115]. For NYSE stocks it is negative and exhibits an increasing pattern from decile 5 to seven ranging from -1.67 to -0.9 percent of the previous spread for every thousand shares traded in the previous time interval total sample value is not significant). For NASDAQ it is significant in decile 10 where it is around

[115] Data of the first NASDAQ decile are not believed to give valid estimates of the parameters since their very small number of observations makes GMM estimation inefficient.

1.9 percent and negative whereas its value for decile 1 is 8.7% and positive (total sample value is -0.66%). In general it could be argued that the inventory effect appears to be valid mainly for NYSE stocks and present only in the medium activity deciles.

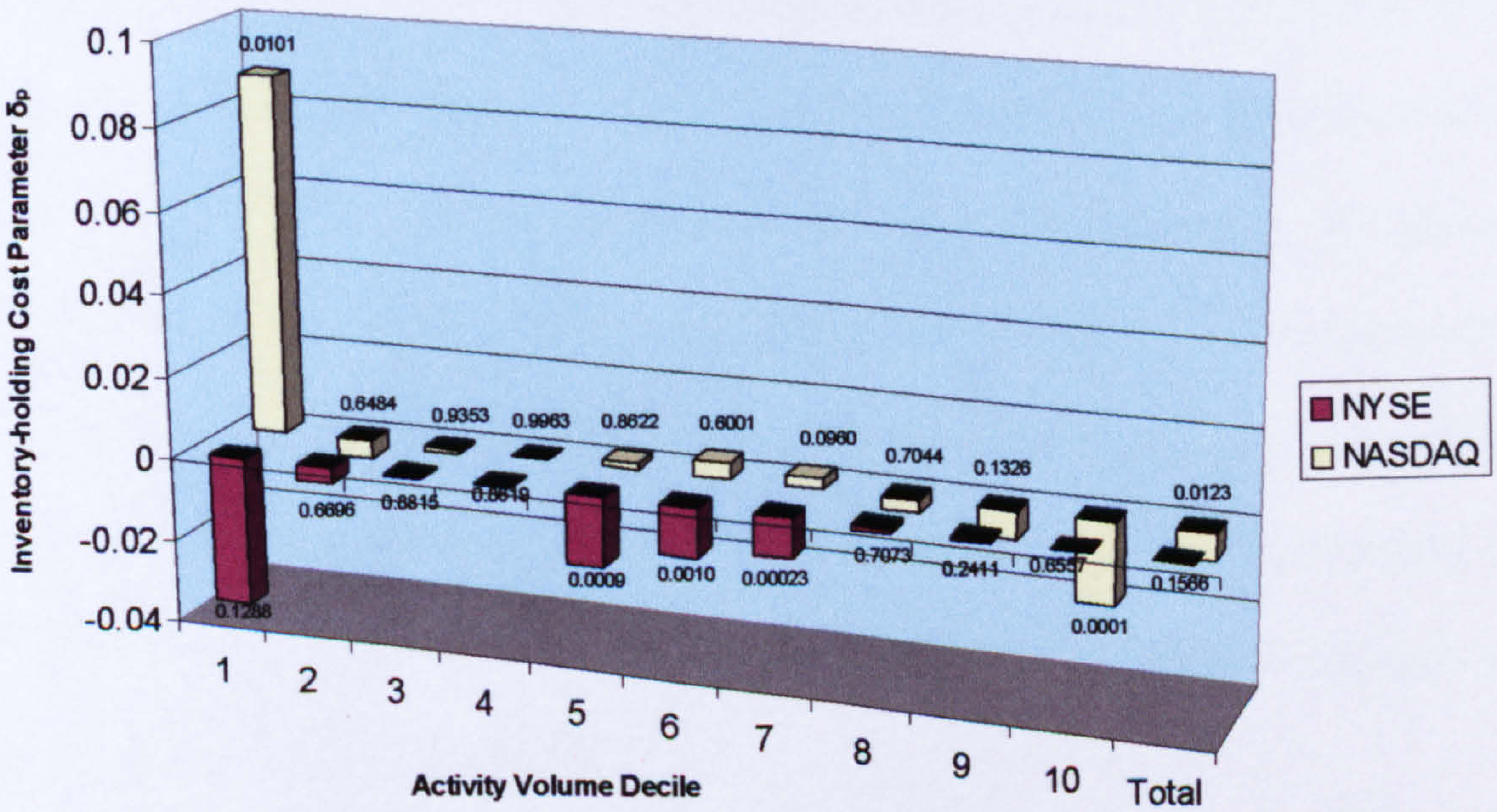
The time component parameter β_p, plotted in figure 5.4, is only significantly different from zero in three of the NYSE deciles namely 3 (at 10%), 8 (at 5%) and 10 (at the 10% level). For NASDAQ it is significantly different from zero at deciles 3 and 5 at the 10% level. Overall there is little evidence of it being different from zero and its sign appears to be positive for NASDAQ and of mixed sign for NYSE. Total sample values for both NYSE and NASDAQ are not significant.

FIGURE 5.2
THE QUOTED SPREAD MODEL
Variation of the Quoted-Volume Parameter γ_p with Activity Volume Decile



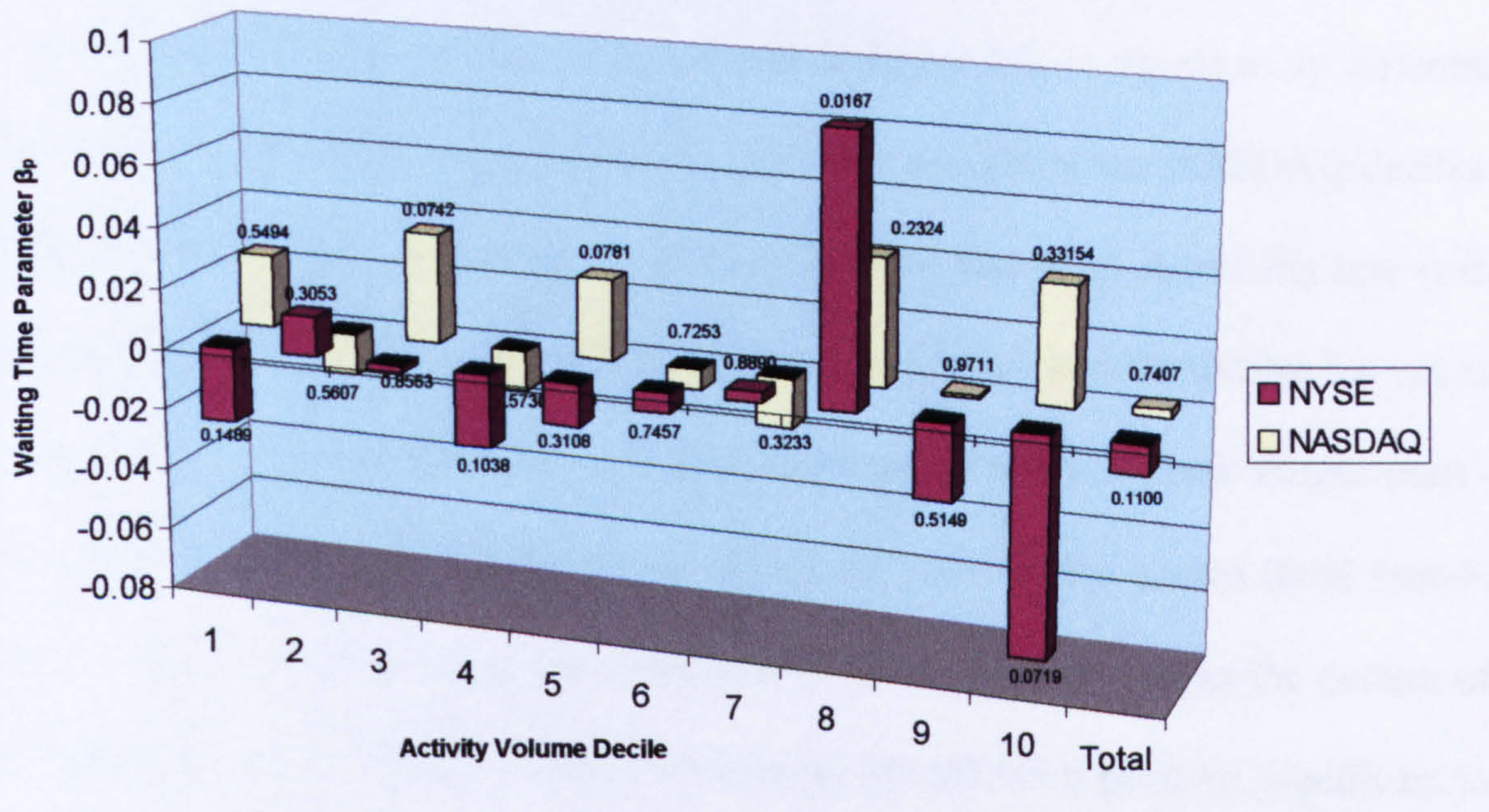
Numbers appearing on top of the bars refer to probability values.

FIGURE 5.3
THE QUOTED SPREAD MODEL
Variation of the Inventory-holding Cost Parameter δ_p with Activity Volume Decile



Numbers appearing on top of the bars refer to probability values.

FIGURE 5.4
THE QUOTED SPREAD MODEL
Variation of the Waiting Time Parameter β_p with Activity Volume Decile



Numbers appearing on top of the bars refer to probability values.

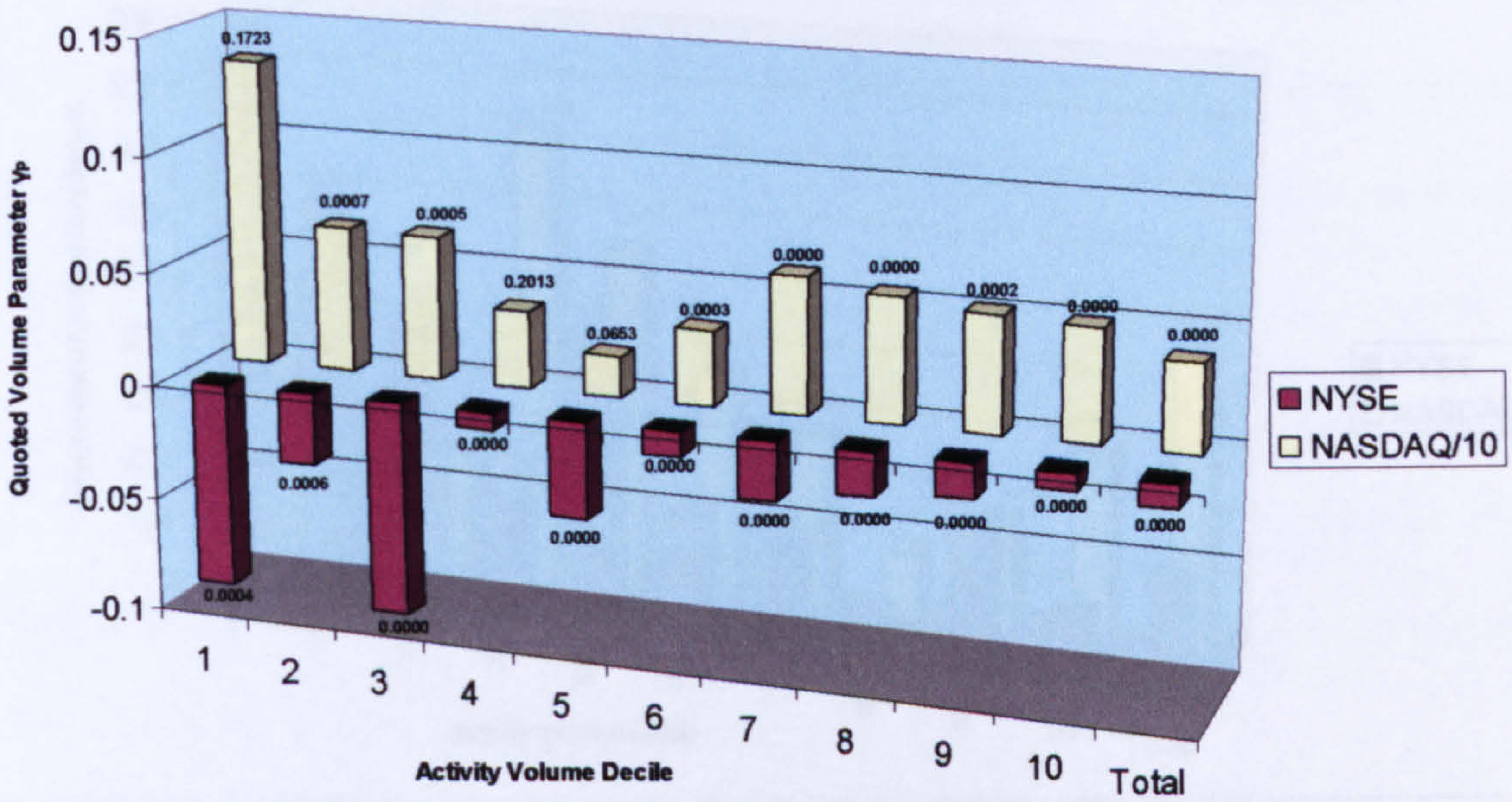
5.5.2 The Offer-Change Model

The estimates of the parameters of the Offer-Change model are exhibited in tables AG.5.5 and AG.5.6 in Appendix G and have been plotted in figures 5.5 to 5.7.

The Quoted Volume Parameter γ_p, plotted in figure 5.5, is significantly different from zero in all NYSE deciles at the 1% level and in eight of the NASDAQ deciles. The pattern for NYSE, as shown in figure 5.5, is still that of an increasing size with negative values as one moves from the low towards the high activity deciles but not as smooth as in the quoted spread model. The magnitude of this parameter ranges from -9.3 to -0.7 % of the previous spread for NYSE for 1000 shares quoted (total sample value is -1%). There are significant differences in the size, as well as the pattern of this parameter for NASDAQ deciles. Its value for NASDAQ is positive, significant in eight deciles and its absolute value larger than NYSE (5-6 times) ranging from 63 to 18 % of the previous spread per 1000 shares quoted total sample value is 38.5%). For this reason, NASDAQ values in figure 5.5 have been scaled down by a factor of 10.

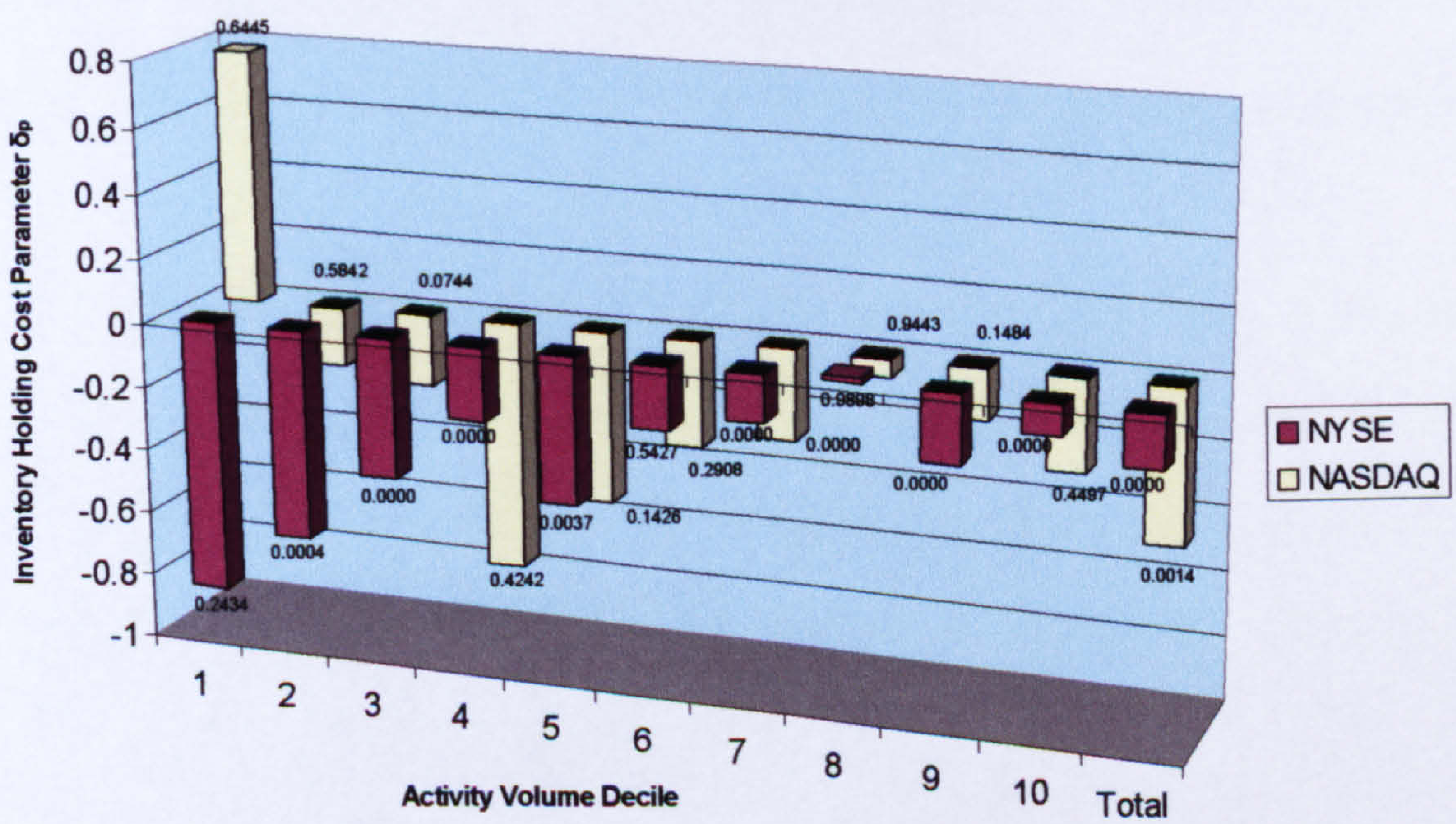
Our model does not allow us to estimate separately the inventory parameter δ_p and the adverse selection parameter θ_p since their estimation depends on the volume of the previous trade. The second term in equation (5.9b) which involves θ_p, the average volume and the correlation coefficient ρ, has a smaller effect in the calculation of θ_p because the product of the other two quantities is smaller compared to the size of the volume of the trade at time t. Thus one should be cautious when interpreting the results for θ_p and δ_p from this equation. The same comments should also apply for the case of the bid-change equation.

FIGURE 5.5
THE OFFER-CHANGE MODEL
Variation of the Quoted-Volume Parameter γ_p with Activity Volume Decile



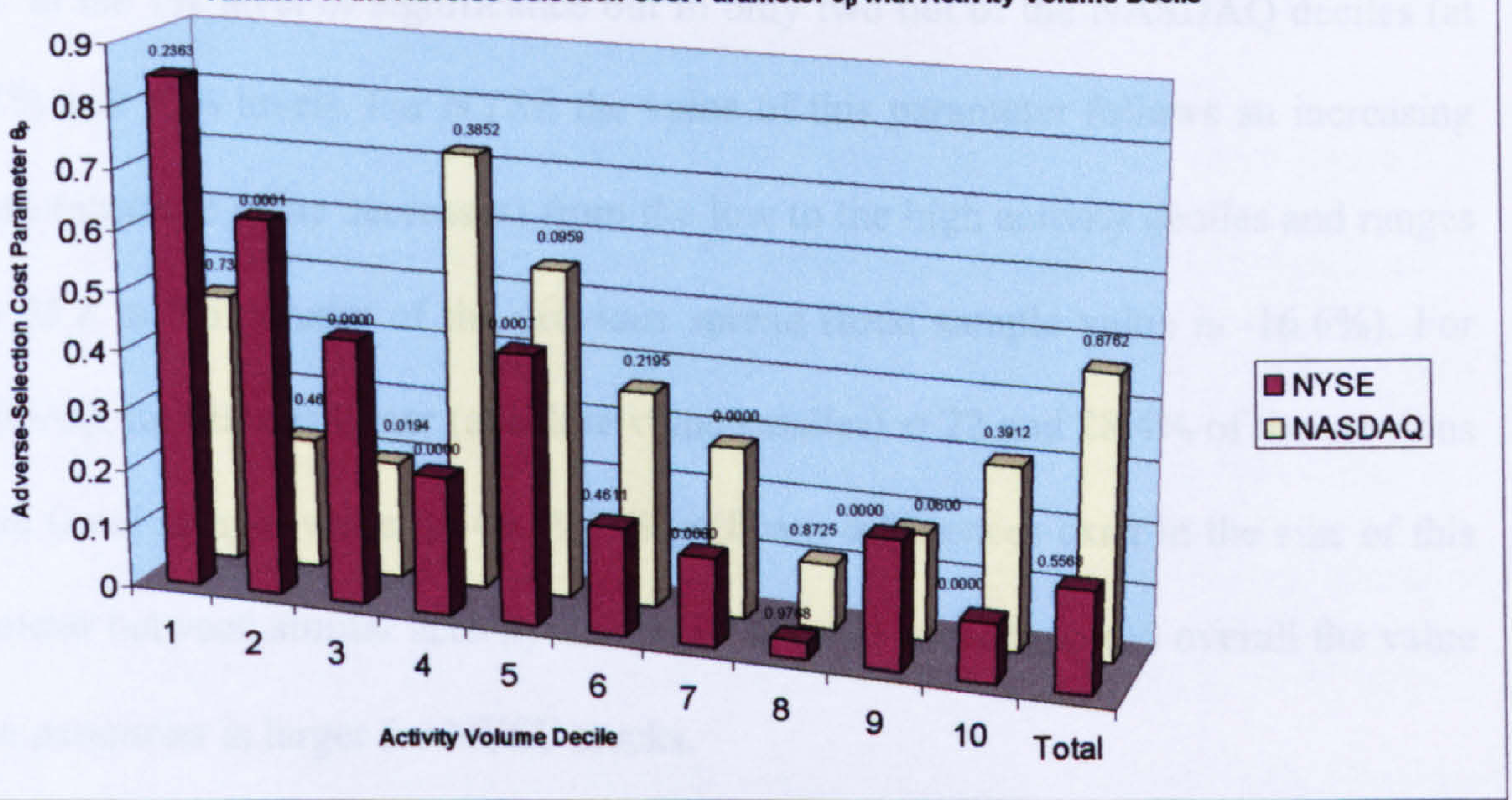
Numbers appearing on top of the bars refer to probability values.
The value for decile 1 of NASDAQ has been scaled down by a factor of 10 since it is large (1.33) and would distort the graph.

FIGURE 5.6
THE OFFER-CHANGE MODEL
Variation of the Inventory-Cost Parameter δ_p with Activity Volume Decile



Numbers appearing on top of the bars refer to probability values.
The value for decile 1 of NYSE has been scaled down by a factor of 10 since it is large (-8.49) and would distort the graph.

FIGURE 5.7
THE OFFER-CHANGE MODEL
Variation of the Adverse-Selection Cost Parameter θ_p with Activity Volume Decile



Numbers appearing on top of the bars refer to probability values.
The value for decile 1 of NYSE has been scaled down by a factor of 10 since it is large (8.46) and would distort the graph.

As can be seen from figure 5.6 the inventory cost parameter δ_p is negative for both NYSE and NASDAQ and significantly different from zero in seven NYSE deciles, most at the 1% level of significance but in only two out of the NASDAQ deciles (at the 1% and 10% level). For NYSE the value of this parameter follows an increasing pattern (absolute value decreases) from the low to the high activity deciles and ranges from 65.8 to 9.6 percent of the previous spread (total sample value is -16.6%). For NASDAQ, its value is larger (absolute value smaller) at 22 and 28.4% of the previous spread (total sample value is -48.1%). Significant differences exist in the size of this parameter between similar activity deciles of the two exchanges and overall the value of the parameter is larger for NYSE stocks.

The adverse selection parameter θ_p, plotted in figure 5.7, decreases from the low towards the higher activity deciles for NYSE where its value is significantly different from zero in seven deciles at the 1% level and ranges from 62.5 to 9.7% (total sample is 16.7%) of the previous spread. It is also significantly different from zero in four NASDAQ deciles (one at the 1% and 5% and two at the 10% level) and its value ranges from 54.8 to 17.6% of the previous period spread following an inverse-U pattern (total sample value is 46.4%).

The time component parameter β_p is negative and significantly different from zero in only one of the NYSE deciles (at the 5% level) and in only one of the NASDAQ deciles (decile three at the 10 % level) where it is negative. For this reason this parameter has not been plotted. Total sample values are not significant for both the NYSE and the NASDAQ sub-samples.

5.5.3 The Bid-Change Model

Tables AG.5.7 and AG.5.8 in Appendix G show the results of the GMM estimates of the Bid-Change Model which are plotted in figures 5.8 to 5.10.

For NYSE the quoted volume parameter γ_p, which has been plotted in figure 5.8, follows a similar pattern as in the case of the offer-change model and is significantly different from zero in all deciles (in eight of these at the 1% level). Its size ranges from -9.2 to -0.5 percent of the previous spread for NYSE (total sample value is -0.97%). The pattern for NASDAQ is irregular and the parameter value, which is positive and significantly different from zero in seven deciles, ranges from 68 to 18% (total sample is 43.68%), that is the absolute values are larger for NASDAQ stocks by five to six times. For this reason, NASDAQ values in figure 5.5 have been scaled down by a factor of 10.

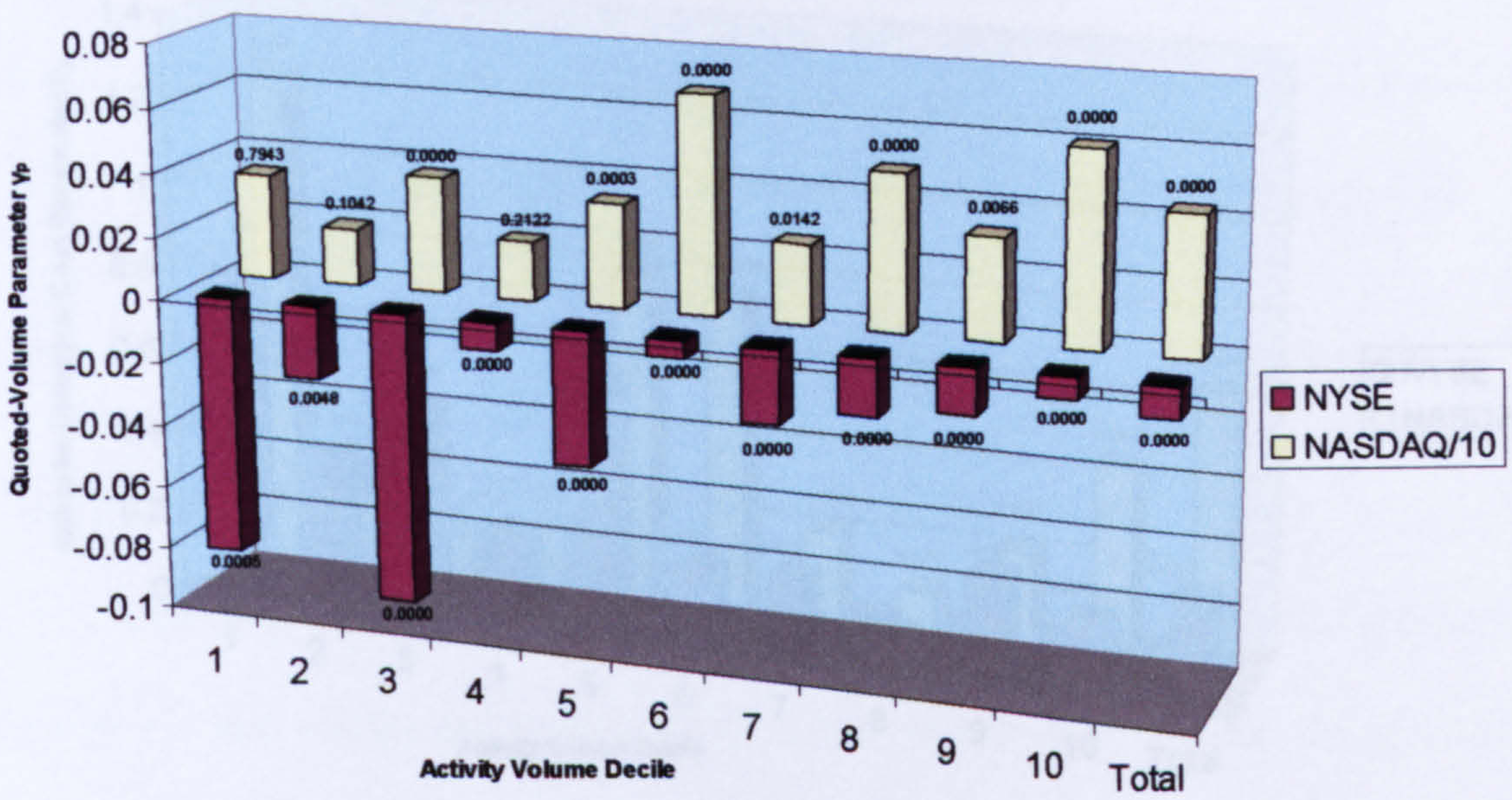
The inventory cost parameter δ_p, plotted in figure 5.9, is significantly different from zero in four NYSE deciles, three at the 1% level of significance. Its value for NYSE is positive and ranges from 60.8 to 17.4 percent of the previous spread following an inverse-U shape (total sample value is 23.93%). For NASDAQ stocks this parameter decreases with increasing activity deciles. Its value is significant in also five deciles (two at the 1% level) and ranges from 83 to 19% of the previous spread (total sample value is 67.82%).

As mentioned in the discussion of the parameter for the offer-change model the value of the adverse selection parameter θ_p, plotted in figure 5.10, is close to that of the inventory cost parameter δ_p. It is statistically significantly different from zero in five of the NYSE (four at the 1% level) and five of the NASDAQ deciles (two at the 1% and two at the 5% level). Its value for NYSE ranges from 59.8 to 17.7 % of the previous spread (total sample value is 24%) whereas for NASDAQ it ranges between

80.5 and 21.6% (total sample value is 64.5%). The pattern is an inverse-U shape for both exchanges while NASDAQ values are in general larger than NYSE ones.

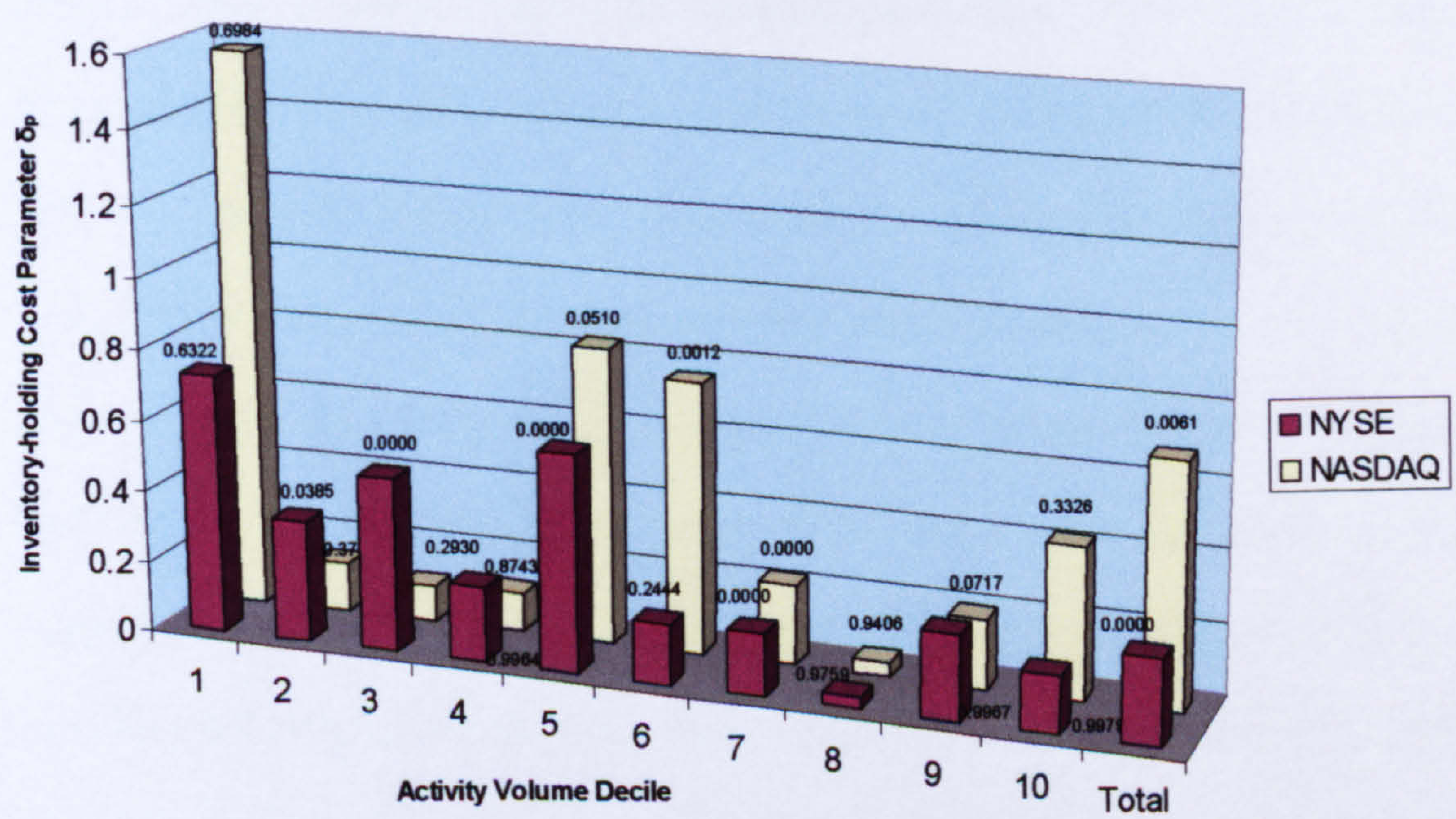
The time effect parameter β_p is only statistically significantly different from zero in three out of the ten NYSE deciles and in none of the NASDAQ deciles whereas total sample values are not significant for both exchanges.

FIGURE 5.8
THE BID-CHANGE MODEL
Variation of the Quoted-Volume Parameter γ_p with Activity Volume Decile



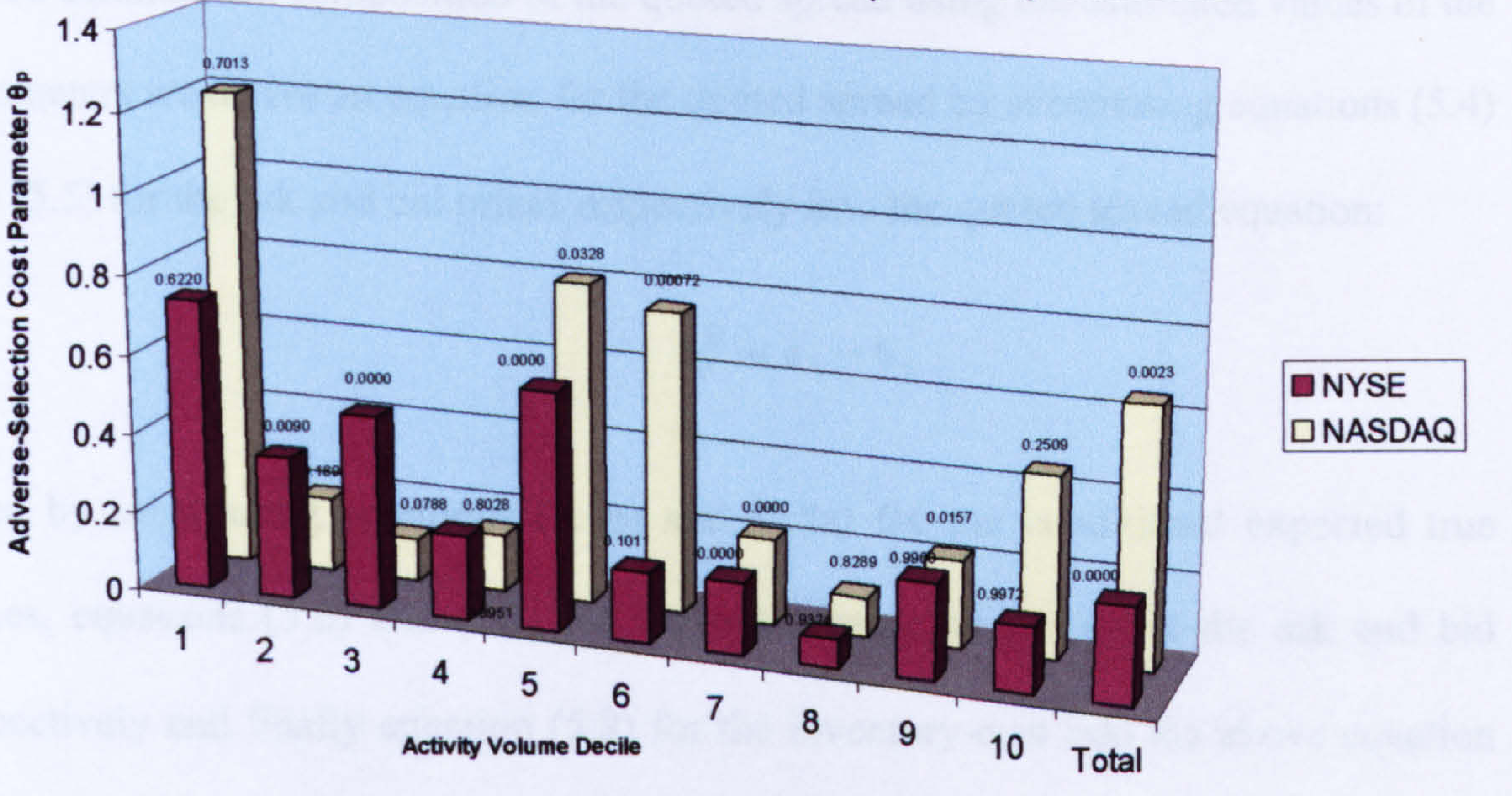
Numbers appearing on top of the bars refer to probability values.

FIGURE 5.9
THE BID-CHANGE MODEL
Variation of the Inventory-Holding Cost Parameter δ_p with Activity Volume Decile



Numbers appearing on top of the bars refer to probability values.
The value for decile 1 of NYSE has been scaled down by a factor of 10 since it is large (7.26) and would distort the graph.

FIGURE 5.10
THE BID-CHANGE MODEL
Variation of the Adverse-Selection Cost Parameter θ_p with Activity Volume Decile



Numbers appearing on top of the bars refer to probability values.
The value for decile 1 of NYSE has been scaled down by a factor of 10 since it is large (7.34) and would distort the graph.

A summary of the results of the Quoted-Spread and the Offer- and Bid-Change Models is provided in table 5.1.

To estimate the composition of the quoted spread using the estimated values of the parameters we derive an equation for the quoted spread by substituting equations (5.4) and (5.5) for the ask and bid prices respectively into the quoted spread equation:

$$S_t^Q = a_t - b_t$$

Then by substituting equations (5.4a) and (5.5a) for the conditional expected true prices, equations (5.6) and (5.7) for the order processing costs at the ask and bid respectively and finally equation (5.8) for the inventory-cost into the above equation we arrive at:

$$S_t^Q = 2\theta\bar{V} + 2\delta\sum_{i=1}^{t-1} q_i V_i + 2\alpha + 2\beta\Delta\tau_{t-1} + \gamma\left(\Delta V_t^a + \Delta V_t^b\right) \qquad (5.12)$$

Using appropriate spreadsheets in which we estimate the above parameters from the estimated parameter values which were expressed as fractions of the spread at time t-1, we are able to calculate the components of the quoted spread for each decile for the NYSE and NASDAQ sub-samples. These results have been plotted in figures (5.1) and (5.2) for NYSE and NASDAQ respectively where the adverse-selection and the inventory-holding cost components are presented as percentages of the total quoted spread explained by our results. Depending on the errors of the estimation the total explained spread can vary from zero to even 150 percent of the actual spread but such large values are rare. The adverse selection cost for NYSE varies from 68.9 to 1.23% of the total explained quoted spread (27.18 to 1.14 for NASDAQ) whereas the inventory-holding cost for NYSE varies from 0.006 to 1.2% (0.12 to 0.42% for

NASDAQ). Since the latter component is very small compared to the adverse-selection one it has been scaled up by a factor of 10 in figures 5.11 and 5.12.

Table 5.1 Summary of the Results from the Estimations of the Quoted-Spread, the Offer-Change and the Bid-Change Models, NYSE and NASDAQ TAQ Data: October 1994.

Table 5.1 presents a descriptive summary of the results from the GMM estimation of the Quoted-Spread Model of equation (5.11b), the Offer-Change Model of equation (5.9b) and the Bid-Change Model of equation (5.10b) for both the NYSE and NASDAQ subsamples from the TAQ database for October 1994.

MODEL		NYSE	NASDAQ
Quoted-Spread			
	γ_p		
		Significant in all deciles (at 1% level)	Significant in all deciles (at 1% level)
		15.3 to 1.1 % of previous spread	123 to 24 % of previous spread
		Negative; Increases from lowest to highest activity deciles	Positive; U-shape / lowest at fifth activity decile Larger than NYSE
	δ_p		
		Significant in three deciles (at 1% level)	Significant in two deciles (at 1% level)
		1.7 to 0.9 % of previous spread	1.9 % of previous spread
		Negative; Increases from fifth to seventh decile	Very weak evidence
	β_p		
		Significant in three deciles (at 10% level)	Significant in two deciles (at 10% level)
		Two negative values and one positive	Positive
Offer-Change			
	γ_p		
		Significant in all deciles (at 1% level)	Significant in eight deciles (in 7 at the 1% level)
		9.3 to 0.7 % of previous spread	63 to 18 % of previous spread
		Negative; Increases from lowest to highest activity deciles	Positive; U-shape; 5 to 6 times larger than NYSE
	δ_p		
		Significant in seven deciles (most at 1% level)	Significant in two deciles(at the 1 and 10% level)
		65.8 to 9.6 % of previous spread	22 and 28.4 % of previous spread
		Negative; Increases from lowest to highest activity deciles but not smoothly	Negative; Decreases with activity decile Smaller than NYSE
	θ_p		
		Significant in seven deciles (at the 1% level)	Significant in four deciles (at 1, 5, 10 and 10% level)
		62.5 to 9.7 % of previous spread	54.8 to 17.6 % of previous spread
		Decreases from lowest to highest activity deciles	Inverse-U shape
	β_p		
		Significant in one decile (5% level); Negative	Significant in one decile (10% level); Negative
Bid-Change			
	γ_p		
		Significant in all deciles (most at the 1% level)	Significant in seven deciles
		-9.2 to -0.5 % of previous spread	68 to 18 % of previous spread
		Negative; Decreases from lowest to highest activity deciles	Irregular pattern; 5 to 6 times larger than NYSE
	δ_p		
		Significant in four deciles (3 at the 1% level)	Significant in four deciles (2 at the 1%, 2 at the 10% level)
		60.8 to 17.4% of previous spread	Positive; 83 to 19% of previous spread
		Positive; Inverse-U pattern	Decreasing with increasing activity decile
	θ_p		
		Significant in five deciles (4 at the 1% level)	Significant in five deciles (2 at the 1%, 2 at the 5% level)
		59.8 to 17.7 % of previous spread	80.5 to 21.6 % of previous spread
		Inverse U-pattern	Inverse U-pattern
	β_p		
		Significant in three deciles; Negative	Nowhere significant

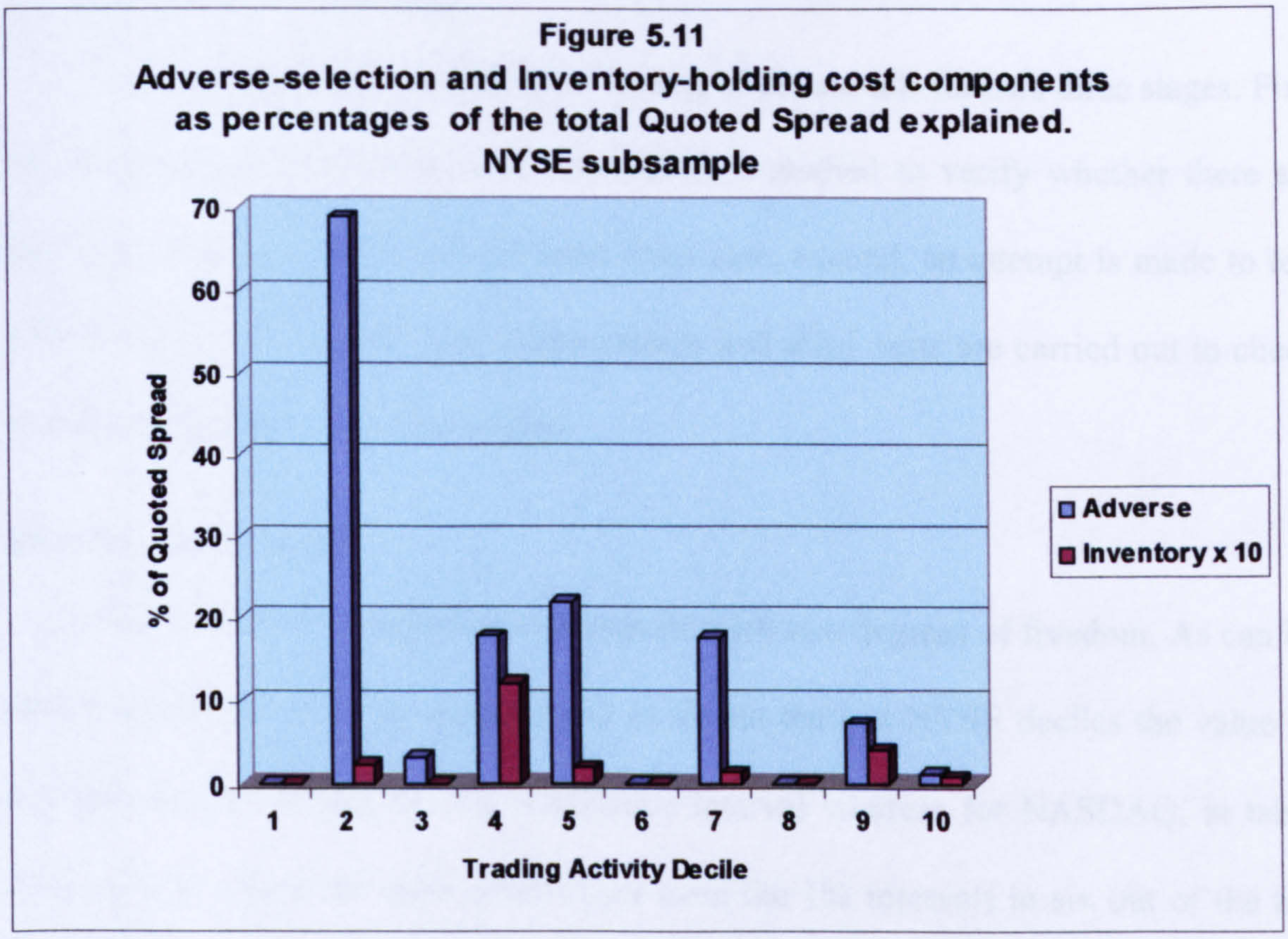


Figure 5.11
Adverse-selection and Inventory-holding cost components
as percentages of the total Quoted Spread explained.
NYSE subsample
70
60
50
40
30
20
10
0
% of Quoted Spread
1
2
3
4
5
6
7
8
9
10
Trading Activity Decile
Adverse
Inventory x 10

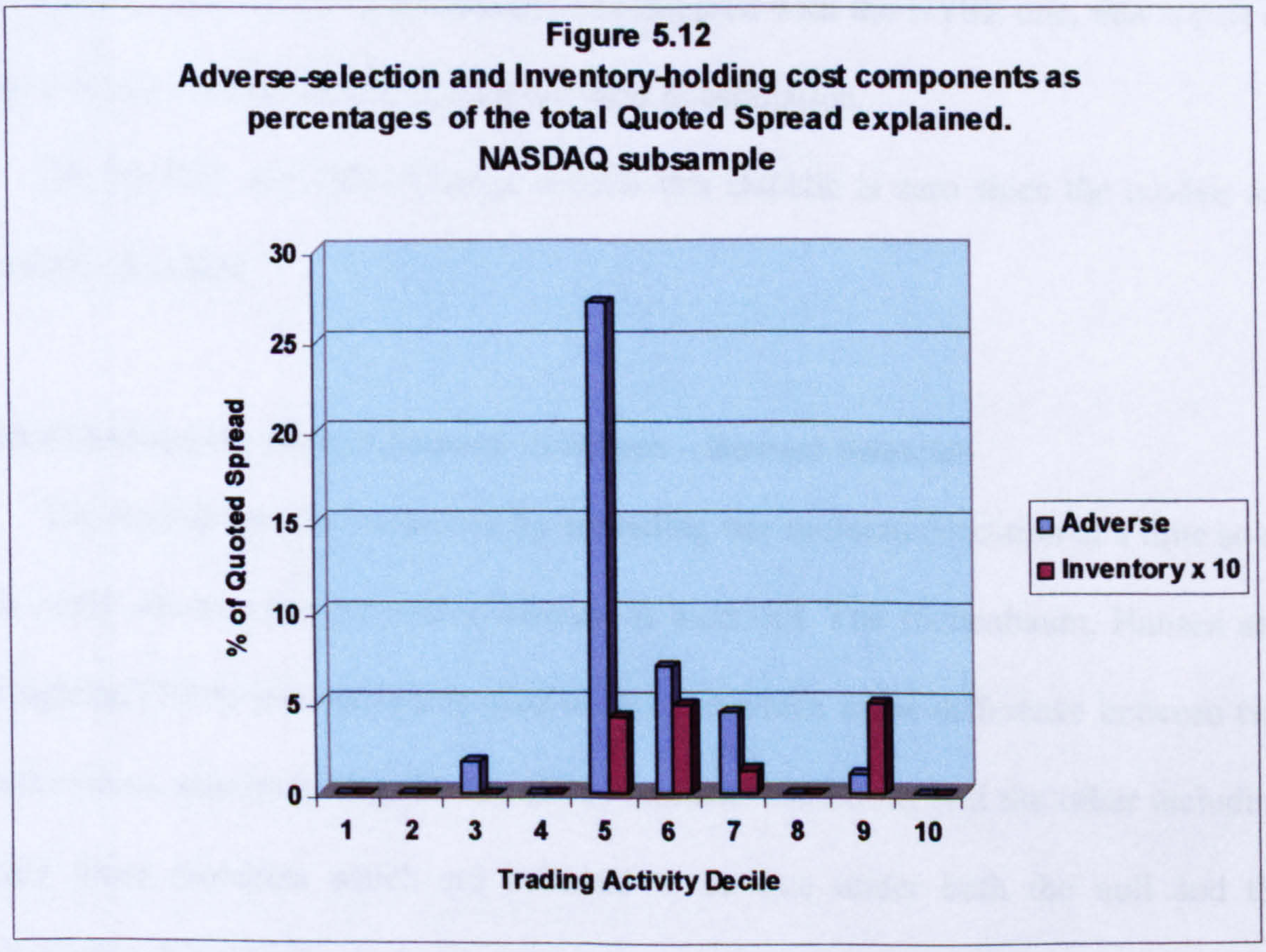


Figure 5.12
Adverse-selection and Inventory-holding cost components as
percentages of the total Quoted Spread explained.
NASDAQ subsample
30
25
20
15
10
5
0
% of Quoted Spread
1
2
3
4
5
6
7
8
9
10
Trading Activity Decile
Adverse
Inventory x 10

5.6 Specification Testing

The procedure for the specification testing of the models follows three stages: First the J-Statistic of over-identifying restrictions is studied to verify whether there are moment conditions which are different from zero, second, an attempt is made to test which moment conditions hold in the sample and third, tests are carried out to check whether parameters are equal to zero.

5.6.1 The J-test statistic

This should be distributed as a chi-square with two degrees of freedom. As can be seen from table AG.5.3 in Appendix G in all but the last NYSE deciles the value of this statistic lies within the 5% confidence interval whereas for NASDAQ, in table AG.5.4, it is within the same interval (or even the 1% interval) in six out of the ten deciles. This indicates that there are in general more moment conditions which are not satisfied in the NASDAQ sub-sample as compared with the NYSE one, that is part of the original moment conditions are not used in estimation.

For the Bid- and Offer-Change models this statistic is zero since the models are exactly identified.

5.6.2 Checking for correct moment conditions – Moment Selection

The models are then estimated by excluding one moment condition at a time so as to verify whether that particular moment is incorrect. The Eichenbaum, Hansen and Singleton (1988) test statistic is used to this end which is the difference between two estimations, one including the full set of moment conditions and the other including only those moments which are believed to be true under both the null and the alternative hypothesis. A detailed description of this and the other relevant tests is given in Appendix B but the fundamental ideas of the test are described here in brief.

Assuming that the vector of all the moment conditions is given by $E[f(x_t, \theta_0)]$ where x_t are the variables in the model and θ_0 is the true parameter vector we can express our null hypothesis as:

$$H_O^S : E[f_1(x_t, \theta_{01})] = 0 \qquad \text{and} \qquad E[f_2(x_t, \theta_0)] = 0$$

and the alternative as:

$$H_A^S : E[f_1(x_t, \theta_{01})] = 0 \qquad \text{and} \qquad E[f_2(x_t, \theta_0)] = \mu_2 T^{-1/2}$$

$$\text{where} \quad \theta_0' = (\theta_{01}', \theta_{02}') \qquad \theta_{0i} : (p_i \times 1)$$

$$f(x_t, \theta_0)' = [f_1(x_t, \theta_{01})', f_2(x_t, \theta_0)'] \qquad f_i(\bullet) : (q_i \times 1)$$

that is we have partitioned both the parameter vector θ_0 as well as that of the moment restrictions $E[f(x_t, \theta_0)]$ into two sub-vectors.

The C_T statistic is then given by

$$C_T = T\left\{Q_T\left(\hat{\theta}_T\right) - Q_{1T}\left(\tilde{\theta}_{1T}\right)\right\}$$

where $\tilde{\theta}_{1T}$: value of θ which minimizes the objective function Q_{1T} using part of the moment conditions

and is distributed according to:

$C_T \sim \chi^2_{\nu_1}$ under H_0^S $\quad \nu_1 = q_2 - p_2$ (degree to which θ_{01} is overidentified by $E(f_2(x_t, \theta_0)) = 0$ if θ_{01} is known).

that is ν_1 is the degree to which the first parameter sub-vector is over-identified by the second sub-vector of moment conditions.

In tables AG.5.9 and AG.5.10 an attempt is made to diagnose the source of the problem, that is to determine which moments are candidates to be characterized as over-identifying. At the 5% level it is evident from table AG.5.9 for NYSE that the C_T statistic is not significant in seven deciles for all moments. In deciles 5 and 6 the null hypothesis that all moment conditions in the sample are satisfied can be rejected for the variables related to the change in the size of the bid and ask and for the change in the size of the bid and the waiting time respectively. In decile 10 it is rejected for the volume, the change in waiting time as well as the trade indicator variable related moments. In this decile, the C_T statistic has a very high value which indicates that the particular moment conditions are probably not satisfied in the sample. For the NASDAQ sub-sample from table AG.5.10 the C_T statistic is nowhere significant at the 5% (or even the 10%) level providing no further evidence as to which moments are not satisfied in this sub-sample.

The next problem we are faced is to choose which moment restrictions are valid in the sample. Andrews (1999) has proposed moment selection criteria (MSC) analogous to the model selection criteria used when choosing among competing models. We use his GMM-BIC (Bayesian Information Criterion) statistic which is based on the J statistic of over-identifying restrictions of Hansen and is given by:

$$MSC_{BIC,n}(c) = J_n(c) - h(|c|)k_n$$

where c is a moment selection vector (selects only certain moments), |c| is the number of moments selected by the vector, p is the number of parameters estimated by GMM and $J_n(c)$ stands for the J-test statistic for over-identifying restrictions constructed using the moment selection vector c. The J-test statistic is distributed as a chi-square with [|c| - min(p,|c|)] degrees of freedom under the null hypothesis that all moment conditions in c are correct.

The second term in the above equation is a bonus term which rewards the selection of vectors which use more moment conditions and is required in order that the increase in J, which occurs when more moment conditions are added, even if they are correct, is counterbalanced. For this reason $h(\bullet)$ is defined as a strictly increasing function with $k_n \rightarrow \infty$ and $k_n = o(n)$ where n is the sample size. Andrews (1999) has stated that both $h(\bullet)$ and k_n should be specified by the researcher. As the latter increases the bonus given for more moment conditions employed increases without bound. For our purposes we define

$$h(|c|)k_n = (|c| - p)\log n$$

Based on this statistic we adopt a downward testing procedure that is a procedure which starts from the most restrictive model and progresses towards the least restrictive one. According to this procedure we start with moment vectors containing the largest number of moments and estimate our models progressively decreasing the number of moments until the J-statistic (i.e. the null hypothesis that all moment conditions considered are correct) is not rejected.

In table AG.5.9 in Appendix G for NYSE the exclusion of one of the initial moment conditions from the estimating procedure led to J-statistics which were not significant in seven out of ten deciles at the 5% level whereas for the NASDAQ sub-sample in table AG.5.10 this statistic was nowhere significant. Referring to tables AG.5.11 and AG.5.12 in Appendix G which tabulate the GMM-BIC moment-selection statistic of Andrews (1999) which in turn indicates which moment conditions are most likely to be over-identifying one can see that in NYSE such moment conditions can be the change in the size of the bid and offer, which give a low MSC value in six out of the ten deciles. The change in waiting time moment also gives very low MSC values in five deciles while the moment conditions related to the volume and the trade indicator variables seem to hold well in most activity deciles, especially the higher ones. For NASDAQ one can see from table AG.5.12 that the same three moment conditions appear to be over-identifying in almost half of the deciles but in this case the volume and trade indicator moment conditions appear to be violated more often that in the NYSE sub-sample. It should be noted that one cannot proceed further with this test by excluding more moment conditions in this model since excluding two moment conditions will result to the model becoming just identified which will produce a J-statistic equal to zero.

5.6.3 Testing the Parameters of the Model

The second stage involves testing the parameters of the model. The null hypothesis is

$H_0^R : r(\theta_0) = 0$ versus $H_A^R : r(\theta_0) = T^{-1/2}\mu_R$ where $\mathrm{r}(\bullet)$ is a $(\mathrm{s} \times 1)$ vector of continuous differentiable functions of the form $\mathrm{R}(\theta) = \partial r(\theta)/\partial\theta'$

The number of restrictions, s, should not exceed the number of parameters, p. Since the SHAZAM program, which has been used to estimate the model, calculates the Wald statistic for each parameter separately, we carry out a Likelihood Ratio (LR) test for the joint hypothesis that all parameters in the model are equal to zero. The reason for using the LR statistic is that the Wald statistic has two disadvantages. First it is not invariant to a re-parametrization of the model or the restrictions and second it tends to be less well approximated by the chi-square in finite samples. Rejection of the null however can occur either because the identifying restrictions are satisfied at the true parameter vector and restrictions are different from zero or because restrictions are equal to zero and the identifying restrictions are not satisfied at the true parameter vector.

As can be seen from tables AG.5.13 and AG.5.14 in Appendix G for the Quoted Spread Model the null hypothesis that all parameters are equal to zero should be rejected for all deciles of the NYSE sub-sample and for most of the NASDAQ deciles with the exception of the first three lowest activity ones. Especially for the case of the first and second NASDAQ deciles the LR statistic is very low while decile 3 is close to the 5 % confidence region. This observation could probably be attributed to our limited number of observations for NASDAQ in these deciles. A similar observation can be obtained from tables AG.5.15 and AG.5.16 for the Offer- and Bid-Change models. While for NYSE the null hypothesis should be rejected for all deciles in the NASDAQ sub-sample it should be accepted for the first decile whereas the statistic for the second decile is close to the 5% interval value.

5.6.4 Testing for Structural Stability

A third stage of the specification testing procedure could be to test for instability in the model. However due to limitations in our package we are not in a position to carry out tests such as those of Sowell (1996a,b) presented in Appendix B, which could investigate the potential existence of structural breaks at known or unknown points in the sample.

5.6.5 Conclusions from the Specification Tests

The above specification tests indicate that in the NYSE sub-sample the assumptions made in the Quoted-Spread Model about the moments related to the change in the size of the ask and the bid as well as those related to the waiting time between trades do not hold well in the medium activity deciles violating the strong exogeneity assumptions made regarding the corresponding regressors which appear to be correlated with the error term. It appears that for stocks in those deciles the change in the spread at some point in time has an effect on the above variables which is not described by our model.

Similarly, in the highest activity decile of NYSE the strong exogeneity assumption is violated for the trading volume, the change in waiting time and the trade indicator variable regressors. This indicates that in the highest liquidity stocks these variables are influenced by the change in the spread an effect not captured by the model.

The strict exogeneity conditions appear to hold better in the NASDAQ sub-sample even though the much smaller number of observations does not help to uncover their likely violation.

5.7 Conclusions

The evidence from the Quoted Spread Model indicates that there exists a relatively strong quoted-volume parameter, γ_p, which, as described by equations (5-6) and (5-7) affects the order processing cost the dealer charges when setting the bid-ask quote. In NYSE this should be interpreted either as imposed by the specialist, or as the cost of placing an order by the investor. The evidence in this chapter documents that the quoted spread is affected by the number of shares quoted prior to a trade. However, the sign of this parameter for NYSE is negative indicating that it tends to reduce the order-processing cost and thus the spread and this effect is the same for both the ask and the bid side of the spread as can be seen from tables AG.5.5 and AG.5.7 in Appendix G as well as from figures (5.5) and (5.8). Three important aspects of this parameter should be taken into consideration in this analysis. Its magnitude, its pattern across the deciles and the degree of symmetry in its size between the ask and the bid sides of the spread. In NYSE the size of this effect tends to diminish as one moves from the low to the high activity deciles. Even though it is small throughout it is very small in the highest activity decile. What is more interesting to point out, however, is the degree of the symmetry of the parameter which is almost equal for the ask and bid sizes thus justifying our use of the same parameter to model both sides. The question remains as to what causes the negative value which is due to the negative relationship between the changes in the values of the quotes and the depths at the quotes. A probable explanation may be based on the exchange-mechanism for NYSE. The specialist has knowledge of the orders in the book and sets the quotes according to the limit-orders placed but can charge what he believes to be an appropriate order-processing cost. However as orders build up, increasing the quoted volume at both sides, this initial cost, which may be quite conservative need not remain the same since with increased volume the specialist will be able to recover his clerical costs from more transactions by charging less per transaction. Thus when volume at the quotes builds up quoted prices are reduced slightly. The effect is larger in the low activity deciles where it is likely that a larger initial fixed order processing cost would be charged which would be corrected sooner as volume picks up. These arguments are supported by the fact that we do not observe this negative effect in NASDAQ where the dealer is not alone in making a market for a particular stock.

Caution should be exercised when interpreting this component for NASDAQ. Since the TAQ database only gives the prices and volumes of the Best Bid Offered (BBO) and not the total size of the quotes of all dealers in a stock at a given price and since dealers in a quote-driven market like NASDAQ are only obligated to quote for a certain size , which is usually 10 lots (1000 shares), this size cannot be used to make inferences on the effect of the volume quoted since it is likely that it will not vary much. NASDAQ dealers can use this limited size to protect themselves against private-information holders as well as a means for discovering prices gradually. They thus set an order-processing cost proportional to the volume they quote since by positioning themselves on the BBO they have already offered a competitive price and should attempt to recover their clerical costs through the increased volume. For this reason the size of this parameter is larger compared to NYSE. The U-shape for the offer-change model may be attributed to the attempt of dealers to recover their clerical costs quickly in less liquid stocks and to take advantage of public demand for transactions in the most liquid ones. In contrast with the symmetry in the size of this parameter for the ask and bid sides in NYSE we observe a large asymmetry for NASDAQ stocks. This asymmetry is probably due to the nature of our database since the ask and bid prices of the BBO quoted do not originate from the same dealer and thus this finding cannot support the idea that the behavior of NASDAQ dealers might influence this parameter differently for each side of the spread.

The inventory cost parameter δ_p, introduced through equation (5-8) of the model, appears to be significant for NYSE in the middle deciles of the Quoted-Spread Model in figure (5.3). Looking into figures (5.6) and (5.9) for the Offer- and Bid- Change models we observe that even though the net effect of this parameter on the spread is nil for most deciles, which should be expected since empirical evidence has indicated that inventory adjustments occur over long time intervals, its absolute value follows a diminishing pattern in the offer and bid change equations. This can be explained by the fact that the specialist over-charges in low liquidity stocks to compensate for his adverse inventory position because these stocks are more likely to bring him into such a position and also because it is easier to recover from an undesired position when trading in high liquidity stocks where recovery could be accomplished through a large number of small steps. For NASDAQ there is weaker evidence for this effect in all of the models even though the size of the parameter appears to be smaller for the ask side but larger for the bid side of the spread compared to that for NYSE. A possible explanation for the weaker evidence is that the NASDAQ dealer, as opposed to the NYSE specialist, is partially protected from extreme undesired inventory positions by both his obligation to quote for a limited number of shares only, which of course applies also to the specialist, but also from his ability to post quotes which are away from the BBO whenever he feels he does not wish to take additional risk. The large number of dealers in a NASDAQ stock implies that the collective inventory risk is shared among them unlike the specialist who is obliged to maintain a 'fair and orderly market'. Moreover, as is also discussed in Bessembinder (1999) for NASDAQ as well as in Menyah and Paudyal (2000) for the ISE, preferencing (sending an order to a market-maker who has not posted the best quote) and internalization (sending an order to a dealer who belongs to the same firm but does not have the best quote) which are possible in these trading mechanisms can affect the sizes of the adverse-selection and the inventory-holding cost. Another reason why evidence for NASDAQ is weaker is the nature of the data for this exchange where the total quoted volume is not reported, therefore a complete picture of the collective inventory imbalance is not available.[116] Finally it is worth pointing out the asymmetry in the absolute sizes of this component between the bid and ask sides of the spread. Even though NYSE absolute values of this component at the bid can be larger or smaller than values of the equivalent decile at the ask NASDAQ values at the bid are larger compared to those at the ask. Finally the sign of the parameter is as expected from theory and discussed in section 5.2, that is it is negative for the ask and positive for the bid.

[116] For example a dealer with excess inventory might offer more shares for sale.

The quoted spread model does not provide information on the adverse selection parameter θ_p since the pertinent term cancels out during differencing when deriving equation (5.11b). The Offer- and Bid-Change models, however, do so and the parameter has been plotted in figures (5.7) and (5.10) respectively. Due to the nature of our model the pattern of this parameter is similar to that of δ_p for the reasons given below. Examining equations (5.9b) and (5.10b) it is evident that the difference between the adverse-selection and the inventory-holding terms, apart from the different parameters, is that in the adverse-selection term the product of the average volume[117] and the autocorrelation coefficient of the trade indicator variable is subtracted. This raises the question as to how strong this term is in differentiating between the two components. Its differentiating power depends on the size and sign of the autocorrelation coefficient, the value of the average volume, which however does not change very much from decile to decile and finally on the sign of the trade-indicator variable. An analysis of the values which the above parameters can take reveals that negative adverse-selection parameters can arise from small trades when the autocorrelation coefficient is positive or from large trades when the change in the ask or bid price is to the opposite direction to that of the last trade. The problem of negative adverse-selection parameters has been overcome by clustering trades at the same price when there is no change in the quotes as was explained in section 4.3.2 which discusses the data-selection procedure. However, when the autocorrelation coefficient is small the term subtracted in the adverse-selection component part of the above equations (the product of the average volume and the autocorrelation coefficient of the trade indicator variable) is also small and thus the estimated

[117] Estimations have been carried out using the median as well as the mode of the volume but there has been no significant difference in the results obtained. The median was used because it minimizes the effect of extreme observations whereas the mode is the most likely statistic to be close to the impression the trader has of an expected volume (the one most frequently occurring).

adverse-selection component is close to the estimated inventory-holding cost component. Unfortunately in our dataset the autocorrelation coefficients calculated and used for estimation are the averages of the values of all the stocks in each decile because we do not run the estimations for each stock separately but for all the trades of all stocks in a given decile. A stock-by-stock estimation would be an onerous task requiring the use of specially written computer programs. The effect of using average autocorrelation coefficients is that positive values are averaged with negative ones resulting in small values. This is a limitation of our procedure.

The value of the adverse-selection parameter for the NYSE subsample follows a diminishing pattern for the ask side of the spread and an inverse-U pattern for the bid side. In general, for both the ask and bid sides high activity deciles have smaller values which can be attributed to the fact that specialists can recover this cost through a large number of trades in the high activity deciles so the cost recovered per trade is small in those deciles. The value of the parameter is larger for the ask, compared to that for the bid which indicates that our assumption about the symmetry of the adverse-selection component may be wrong. For NASDAQ, where evidence is weaker, there appears to be an inverse-U pattern for both the ask and bid sides of the spread. We also observe that some bid-side values are very large and that there is asymmetry between the ask and bid sides of the spread for this component.

There is some evidence, albeit weak, from the Quoted Spread model for the significance of the waiting time parameter β_p, plotted in figure (5.4). However there is little evidence elsewhere to support the hypothesis that it is different from zero and we therefore have to either reject the hypothesis or assume that our models, methodology or our database, are not capable of yielding any sensible values.

The average sizes of the components as percentages of the explained quoted spread for stocks in the NYSE and the NASDAQ deciles have been plotted in figures (5.11) and (5.12) respectively. For NYSE we can observe from figure (5.11) that the adverse-selection cost is large ranging from 68.9 to 1.23 percent of the quoted-spread, wherever it is significantly different from zero and is larger in the medium-activity deciles following an inverse-U shape (excluding decile 2 which has the highest value). The inventory-holding cost is small, mainly present in the medium-activity deciles and ranging from 0.006 to 1.2 % of the quoted spread. It also follows an inverse-U pattern.

For NASDAQ the adverse-selection cost ranges from 27.18 to 1.14 percent of the quoted spread and is larger in the medium activity deciles. The inventory-holding cost is very small and present in the higher-activity deciles and ranges from 0.12 to 0.42% of the explained quoted spread. Stoll's (1989) value for the adverse-selection of the quoted spread in NASDAQ (43%) is larger and outside our range.

Compared with the values for the adverse-selection cost for NYSE estimated by Huang and Stoll (1997), which however refers to the realized spread, our values for the higher activity deciles, which correspond to the large stocks they use, compare well with their average of 9.6% (range from 1.3 to 21%). We also find much smaller values for the inventory-holding cost component compared to the above authors. However their results are not comparable to ours since they examine fewer stocks the trades of which however span a period of twelve months compared to the one month of our data. This should allow them to uncover the inventory-adjustment effect which may not be as evident in the month we consider.

Overall the contribution of the models in this chapter has been to:

1. Document the significance of the quoted-volume parameter γ_p for both NYSE and NASDAQ, providing evidence that the volume quoted at the bid and ask directly affects the size of the spread.
2. Document the existence and estimate parameter values for an adverse selection component which depends directly on the number of shares traded, unlike previous studies which only use dummy variables for ranges of size. This has been accomplished for both NYSE and NASDAQ but within certain limitations which have been described above and have been imposed by the models employed.
3. Document the existence and estimate parameter values for an inventory-holding component which varies with the trading activity of the stock for both NYSE and NASDAQ.
4. Show that the waiting time between trades may account for a very small part the spread but further research is required to uncover a possible relationship.
5. All of the above have been accomplished for deciles of trading activity which help determine the patterns of the parameters.

Chapter 6

The Price-Change Model

6.1 Introduction

Following the development of the Quoted-Spread Model we will now attempt to develop a model which describes the change in price of a security based on the idea that it is the outcome of the dealer's attempt to be compensated both for his order processing as well as for his inventory holding costs. Moreover the adverse selection component should have been incorporated into his idea of the true price of the stock, immediately preceding a trade which is expressed through m_t in equation (5.1). Therefore the development of the price-change equation draws from previous concepts but this does not imply that the dealer will be able to incorporate all his costs into the spread since, as is true especially in the New York Stock Exchange market, trades can and do occur inside the spread due to many reasons among which the most important is the ability of the trader to place limit-orders inside the spread.[118] Placing limit orders was not possible in NASDAQ in 1994 where our data come from. Implementation of a limit order book has been introduced in NASDAQ in 1998 and customer limit orders should now be displayed to the public as NASDAQ quotations if they improve the prices of the market-makers allowing limit-order traders to compete with them. Discussion in this chapter follows the same procedure as that for chapter 5 and an outline of the steps taken is given below:

- First a model is developed for the change in price of a stock following on from principles and equations presented in chapter 5.

[118] Barclay, Christie, Harris and Schultz (1999) provide a description of the new rules for NASDAQ. The change in the behaviour of this market, as far as the components of the spread are concerned, after the implementation of the book is a topic of future research. Bessembinder (1999) has found that trading costs for NASDAQ still remain larger than in NYSE probably owing to order-preferencing even though the bid-ask spread decreases as shown by McInish, Van Ness and Van Ness (1998).

- Next summary statistics for the variables used in the estimation of the above model are tabulated and commented.
- Results of the estimations using the two datasets from NYSE and NASDAQ are presented and discussed.
- The specification tests of the model developed and their results are provided and commented and
- Finally conclusions drawn, based on the estimation and specification results for this model, are discussed.

6.2 Development of the Model

At time t a trade takes place at the bid, the ask or inside the spread and the price at t is assumed to result from the common information price after the trade at time t, m_{t+1}, (notation of figure 5.1) after the inventory-holding as well as the order-processing costs have been added on this price :

$$P_t = m_t + IV_t + q_t O_t^a I_t^a + q_t O_t^b I_t^b + \zeta_t \tag{6.1}$$

Where ζ_t is a zero mean error term which is an asymptotically uniformly distributed random variable which accounts for the effect of price discreetness. I^a (I^b) is an indicator variable which takes the value of 1(0) when a trade is buyer-initiated and the value of 0 (1) when it is seller-initiated. The above equation is constructed in such a way that when the trade is seller-initiated and q_t is thus equal to -1, I^a is equal to 0 and I^b equal to 1. The order-processing cost component is thus subtracted from the 'true' price of the stock m_t. The opposite effect takes place when the trade is buyer-initiated. In that case where q_t is equal to 1, I^a is equal to 1 and I^b equal to 0.

The order-processing cost component is thus added to the 'true' price. The relationship between I_t and q_t is therefore the following:

$$I_t^a = \tfrac{1}{2}(q_t + 1) \qquad\qquad I_t^b = 1 - I_t^a \tag{6.2}$$

The asymmetric (adverse-selection) cost of the dealer is contained into the common information price, m_{t+1}.

Taking first differences of equation (6.1) we obtain:

$$\Delta P_t = \Delta m_{t+1} + \Delta IV_t + q_t O_t^a I_t^a - q_{t-1} O_{t-1}^a I_{t-1}^a + q_t O_t^b I_t^b - q_{t-1} O_{t-1}^b I_{t-1}^b + \Delta\zeta_t \tag{6.3}$$

where $\Delta P_t = P_t - P_{t-1}$

and $\Delta\zeta_t = \zeta_t - \zeta_{t-1}$ is also a zero mean asymptotically uniformly distributed random variable.

Substituting for the components of the spread from equations (5.3), (5.8), (5.6) and (5.7) into equation (6.3) leads to:

$$\begin{aligned}\Delta P_t = {} & \left[\alpha(I_t^a + I_t^b) + \beta\tau_{t-1}(I_t^a + I_t^b) + \gamma(I_t^a V_t^a + I_t^b V_t^b) + \theta V_t\right] q_t - \\ & \left[\alpha(I_{t-1}^a + I_{t-1}^b) + \beta\tau_{t-2}(I_{t-1}^a + I_{t-1}^b) + \gamma(I_{t-1}^a V_{t-1}^a + I_{t-1}^b V_{t-1}^b) + \rho\bar{V}\theta - \delta V_{t-1}\right] q_{t-1} + \\ & \Delta\zeta_t + \vartheta_{t+1}\end{aligned} \tag{6.4}$$

Substituting equations (6.2) into equation (6.4) we obtain:

$$\begin{aligned}\Delta P_t = {} & \left[\alpha + \beta\tau_{t-1} + \frac{\gamma}{2}\left(V_t^a + V_t^b + q_t(V_t^a - V_t^b)\right) + \theta V_t\right] q_t - \\ & \left[\alpha + \beta\tau_{t-2} + \frac{\gamma}{2}\left(V_{t-1}^a + V_{t-1}^b + q_{t-1}(V_{t-1}^a - V_{t-1}^b)\right) + \rho\theta\bar{V} - \delta V_{t-1}\right] q_{t-1} + \\ & \Delta\zeta_t + \vartheta_{t+1}\end{aligned} \tag{6.5}$$

In order that we allow for easier comparison of the results from the estimation of the parameters of this model with those of the Quoted-Spread, the Ask- and the Bid-

Change models we express, as was done in the case of the latter models, the change in price as a fraction of the spread at time t-1. Moreover, owing to the diversity of the stocks used in this study, which trade at different exchanges (different mechanisms) with prices and spreads which exhibit large variation, it would be more meaningful to study the parameters of the components of the spread relative to some other variable such as the quoted spread.

Dividing the left-hand side of equation (6.5) by the quoted spread at time t-1 we obtain the change in the price between times t and t-1 relative to the quoted spread at time t-1 which implies that since the same variables are used for estimation in the right-hand side the parameters of the components of the spread, that is δ, β and γ will now be given relative to the quoted spread at time t-1. Thus equation (6.5) becomes:

$$\Delta \bar{P}_t = \left[\alpha_p + \beta_p \tau_{t-1} + \frac{\gamma_p}{2}\left(V_t^a + V_t^b + q_t\left(V_t^a - V_t^b\right)\right) + \theta_p V_t\right] q_t -$$
$$\left[\alpha_p + \beta_p \tau_{t-2} + \frac{\gamma_p}{2}\left(V_{t-1}^a + V_{t-1}^b + q_{t-1}\left(V_{t-1}^a - V_{t-1}^b\right)\right) + \rho\theta_p \bar{V} - \delta_p V_{t-1}\right] q_{t-1} -$$
$$+ \kappa_t \qquad (6.5b)$$

where $\Delta \bar{P}_t$ denotes the change in price between times t and t-1, expressed as a fraction of the spread at time t-1 and the subscript p under the parameters α, δ, β, θ and γ denotes parameters expressed as a fraction of the spread at time t-1, so that:

$$\alpha_p = \frac{\alpha}{S_{t-1}}, \quad \delta_p = \frac{\delta}{S_{t-1}}, \quad \gamma_p = \frac{\gamma}{S_{t-1}}, \quad \beta_p = \frac{\beta}{S_{t-1}} \quad \text{and} \quad \theta_p = \frac{\theta}{S_{t-1}}$$

Owing to our data-selection procedure, the spread at time t-1 cannot be zero.

Moreover, we assume that

$$\kappa_t = \frac{\Delta\zeta_t}{S_{t-1}} + \frac{\vartheta_{t+1}}{S_{t-1}}$$ is a zero mean, independently, identically distributed random variable since there is no reason to support that $\Delta\zeta_t$ and ϑ_{t+1} are correlated with the spread at time t-1.

The difference between equations (5.11b) and (6.5b) is that whereas (5.11b) describes the change in the quoted spread in terms of the components of the spread as set by the market maker in anticipation of the impact which the size and direction of the order flow will have on the true price, equation (6.5b) describes the change in the transaction price in terms of the actual sizes of the components of the spread following the trade at time t. It can thus be argued that (5.11b) gives us the ex-ante values of the components whereas (6.5b) gives the ex-post values. When trades occur at the quoted bid or the ask the spread components in the two equations should be equal. However when trades occur inside the spread the components will have different sizes and it is not certain that all cost will be recovered at the full value set by the market-maker.

By estimating the parameters of the two equations separately one can therefore infer to what extent each component of the spread is reduced following a trade inside the quotes. It is expected that by choosing to trade inside the spread, if that is possible, a trader can reduce his adverse selection cost as well as that related to the dealer's inventory holding cost.

6.3 Data

Following the procedures discussed in chapter 4 and appendix C the variables which are required for the estimation of the parameters of the Price-Change Model have been calculated and are presented and discussed in this section together with the correlation between them.

6.3.1 Description

Tables AG.6.1 and AG.6.2 in Appendix G present summary statistics for the variables of the price change model for the NYSE and the NASDAQ samples of stocks. The change in price is expressed as a proportion of the spread at time t-1 so as to derive equivalent values for the parameters of the model since those of the previous models have been expressed in the same way. The mean proportional change in price for the NYSE sample follows a decreasing pattern till the third decile then increases remaining close to zero till the largest activity decile. For NASDAQ it starts from a negative (lowest) value and increases in the second decile, then fluctuates around zero. Its value ranges from 0.0068 to –0.0017 for NYSE and from -0.0059 to 0.068 for NASDAQ.

The mean number of shares traded in the previous time period increases uniformly with the larger activity volume deciles and ranges from 949 to 2690 for NYSE and from 938 to 1974 for NASDAQ the mean in both samples being around 1600.

The waiting time between trades at t and t-1 and between t-1 and t-2 decreases uniformly for both NYSE and NASDAQ samples ranging from 2890 to 65 seconds for NYSE and from 3,100 to 116 seconds for NASDAQ. Average waiting times are almost twenty percent lower for NYSE compared to NASDAQ samples with NASDAQ values exhibiting much higher variability.

The mean offer size for NYSE stocks increases irregularly till the seventh decile then smoothly till the largest activity decile ranging from 693 to 4770 shares, its average being 2180 shares. For NASDAQ it follows a smoother pattern increasing till the third, remaining almost constant till the sixth decile and finally increasing towards the last one. It ranges from 794 to 969 shares with an average of 869.

The mean bid sizes follow the same patterns as those followed by the mean offer sizes of each sample. The mean bid size for NYSE ranges from 571 to 4630 (average 2150) and that of NASDAQ from 688 to 4970 with an average of 973.

The trade indicator variable for NYSE follows an increasing pattern with troughs at deciles 4 and 8 and has an average value of 0.148. That of the previous time period is also increasing but with more troughs (average 0.139) whereas that at time t-2 is constant till decile 4 then increases smoothly (average 0.32). For NASDAQ a smoother increasing pattern is observed with troughs at the fourth and seventh deciles. First decile values are negative. Average values are 0.064 for the current variable, 0.061 for that at time t-1 and 0.425 for t-2.

The average quoted spread decreases smoothly with increasing activity decile for the NYSE and NASDAQ samples. For NYSE it ranges from 0.497 to 0.287 USD with an average of 0.394 whereas for NASDAQ it ranges from 0.69 to 0.32 USD but the average spread is almost the same as for NYSE. NASDAQ spreads however exhibit lower variability.

Finally, the realized spread, as a percentage of the quoted spread, for NYSE and NASDAQ stocks shown in tables 6.1 and 6.2 has been calculated using the method of Petersen and Fialkovski (1994) described in section 3.4, that is for buy orders we calculate the difference between the ask price and the transaction price, for sell orders we calculate the difference between the transaction price and the bid price and each time we subtract each difference from half the quoted spread. NASDAQ effective spreads are larger than NYSE ones in all deciles and spreads in both subsamples follow the same pattern, that is they are larger in the medium-activity deciles and have large values in the lowest and highest activity deciles. NYSE effective spreads range from 32.2 to 42.63% of the quoted spread whereas NASDAQ effective spreads range from 49.6 to 71.7 % of the quoted spread.

6.3.2 Correlation of variables

The correlations among the variables used to estimate the models developed in this chapter are discussed separately for NYSE and NASDAQ in the following subsections.

6.3.2.1 NYSE

The change in price, relative to the quoted spread at time t-1, is positively correlated with the trade indicator variable at time t (0.65 to 0.50) and negatively with that at time t-1 (-0.19 to –0.35 increases with decile).

The volume traded at time t shows little correlation with any variables.

The same behaviour is exhibited by the waiting time between times t and t-1 except that it is slightly correlated with that between times t-1 and t-2 from decile 3 and higher and this correlation becomes stronger with increasing activity decile (ranges from 0.12 to 0.38).

The waiting time between times t-1 and t-2 is slightly positively correlated with the bid and offer sizes in decile 1, with the previous bid size in deciles 2 and 3 and slightly negatively correlated with the with the trade indicator variable at time t-2 in deciles 2, 5, 7 and 8.

The offer size is positively correlated with the bid size (0.24 to 0.55) and the offer size in the previous time period (0.32 to 0.66). It also exhibits a weaker positive correlation with the bid size in the previous time period (0.10 to 0.36) and a weak negative correlation with the spread in the previous time period (-0.08 to 0.27).

The bid size is positively correlated with the offer size in the previous time period (0.095 to 0.36) but shows a rather strong positive correlation with the bid size in the previous time period (0.4 to 0.71). It is also negatively correlated with the quoted spread in the previous time period (-0.07 to –0.26) with the weakest values in most of the highest activity deciles.

The offer size and bid size at time t-1 exhibit stronger negative correlation with the quoted spread in the same time period ranging from -0.16 to -0.39 and from –0.15 to –0.41 respectively (smaller values are generally encountered in the higher activity deciles).

The trade indicator variables at times t, t-1 and t-2 show little correlation with any other variables.

6.3.2.2 NASDAQ

The change in price, relative to the quoted spread at time t-1, is positively correlated with the trade indicator variable at time t (0.48 to 0.58) and negatively with that at time t-1 (-0.01 to –0.36).

The volume traded at time t shows little correlation with any variables except for the first decile where it is weakly positively correlated with the offer and bid sizes at time t and t-1 and negatively with the quoted spread at time t-1 in six of the other deciles.

The waiting time between trades between times t and t-1 is weakly positively correlated with that between trades between times t-1 and t-2 and the strength of the relationship increases with the activity decile reaching 0.29 at decile 10.

The waiting time between times t-1 and t-2 is weakly negatively correlated (-0.14 max) with the trade indicator variable at time t-2 in six of the deciles.

The offer size is positively correlated with the bid size (0.10 to 0.58) and strongly correlated with the offer size at time t-1 (0.26 to 0.82 weaker in the high deciles). It also exhibits a weaker correlation with the bid size of the previous time period (0.22 to 0.58, in general weaker in the high activity deciles).

The bid size is positively correlated with the offer size in the previous time period (0.002 to 0.52), shows a stronger positive correlation with the bid size in the previous time period (0.35 to 0.70) and a weak negative correlation with the quoted spread in the previous time period in four of the deciles, with higher values at the low activity deciles.

The offer size in the previous time period is positively correlated with the bid size in the same period (0.16 to 0.51) and positively correlated with the quoted spread in the same time period in seven of the deciles. The bid size in the previous time period is positively correlated with the spread in the same time period in six of the deciles.

The trade indicator variable at time t is weakly correlated with that at t-1 in five deciles and positively correlated with the same variable at time t-2 in three deciles. The same variable at time t-1 is positively correlated with that in the previous period in only three of the deciles.

6.4 Estimation of the Model

The procedure for estimating the parameters of this model is described in this section and the results of the estimations are provided in appropriate tables. Details on the theoretical underpins of the Generalized Method of Moments (GMM) methodology which has been used for estimation are given in Appendix A.

The model to be estimated is that described by equation (6.5b) which is recited below:

$$\Delta\bar{P}_t = \left[\alpha_p + \beta_p \tau_{t-1} + \frac{\gamma_p}{2}\left(V_t^a + V_t^b + q_t\left(V_t^a - V_t^b\right)\right) + \theta_p V_t\right] q_t -$$

$$\left[\alpha_p + \beta_p \tau_{t-2} + \frac{\gamma_p}{2}\left(V_{t-1}^a + V_{t-1}^b + q_{t-1}\left(V_{t-1}^a - V_{t-1}^b\right)\right) + \rho\theta_p \bar{V} - \delta_p V_{t-1}\right] q_{t-1} -$$

$$+ \kappa_t \qquad (6.5b)$$

This model describes the change in the transaction prices between times t and t-1, relative to the quoted spread at time t-1.

Equation (6.5b) involves six parameters, namely ρ, θ_p, δ_p, β_p, α_p and γ_p and eleven variables. As in the offer-change and the bid-change models, ρ and the average volume can be calculated in advance from the data and then treated like constants for the particular decile allowing us to reduce the estimation problem into a five parameter - ten variable one. In estimating this equation using the Generalized Method of Moments (GMM), apart from the above parameters which we can denote by the vector φ, we require a set of orthogonality restrictions which can be represented by the following vector function of population moments :

$$f'(\varphi, h_t) = \left[e_t^p \tau_{t-1}\ ,\ e_t^p \tau_{t-2}\ ,\ e_t^p q_t\ ,\ e_t^p q_{t-1}\ ,\ e_t^p V_t\ ,\ e_t^p V_t^a\ ,\ e_t^p V_{t-1}^a\ ,\ e_t^p V_t^b\ ,\ e_t^p V_{t-1}^b\ ,\ e_t^p V_{t-1}\right]$$

where e_t^p is the error term in the equation (6.5b) which corresponds to κ_t

The set of the above restrictions correspond to strict exogeneity conditions in the sense that

$$E\left[e_i^p \mid X_1, X_2, \ldots, X_n\right] = 0 \qquad (i=1,2,\ldots,n)$$

where X_j is the vector of all the observations of regressor j.

Thus the expected value of the error term, conditional on the regressors for all observations is equal to zero. This implies that the unconditional mean of the error term is zero:

$$E\left(e_i^p\right)=0 \qquad \text{(i=1,2,...,n)}$$

and also that the above ten regressors are orthogonal to the error term for all observations:

$$E\left(x_{jk}\cdot e_i^p\right)=0 \qquad \text{(i,j=1,......,n; k=1,.........,K)}$$

or

$$E\left(X_j\cdot e_i^p\right)=\begin{Bmatrix} E\left(x_{j1}e_i^p\right) \\ E\left(x_{j2}e_i^p\right) \\ \vdots \\ \cdot \\ E\left(x_{jk}e_i^p\right) \end{Bmatrix}=\mathbf{0} \ \text{(for all i,j)} \quad \text{and} \quad \mathbf{0} \text{ is (Kx1)}$$

Therefore the regressors are assumed to be orthogonal to the past, current and future error terms (or the error terms are assumed to be orthogonal to the past, current and future values of the regressors).

As in chapter 5, the validity of the above strict exogeneity conditions is examined in the specification tests which follow the estimation of the model.

The implied moment conditions are:

$$\mathrm{E}[f(\varphi,h_t)]=0$$

There are five parameters to be estimated and ten moment conditions, therefore this model is over-identified by a degree of five.

In our estimations we express volumes in thousands of shares and the change in the waiting time between trades in hours.

6.5 Discussion of Results

Tables AG.6.3 and AG.6.4 in Appendix G present the Generalized Method of Moments Estimates of the Parameters of the Price Change Model described by equation (6.5b) for the NYSE and NASDAQ samples of stocks respectively. The parameter values are also plotted in figures 6.1 to 6.5.

The fixed-cost-of-trade parameter α_p, plotted in figure 6.1, is significantly different from zero in eight of the ten deciles for NYSE (seven at the 1% level of significance) and in only four of the NASDAQ deciles (one at the 1% and two at the 5% level). Its value for NYSE follows an inverse-U shape and ranges from 4 to 21 percent of the previous period spread (total sample value is 4%). The pattern for NASDAQ stocks is different, the parameter being significantly different from zero mainly in the higher activity deciles and decreases with increasing activity, its value ranging from 60 to 25 percent of the previous spread (it is 10% for the total sample).

The time parameter β_p, plotted in figure 6.5, is significant in two of the NYSE deciles (at the 1% level of significance). In the NASDAQ sub-sample it is also significant in two of the deciles (one at the 1% and one at the 10% level). For NYSE its values are 99 and 264 percent of the previous period spread and the total sample value is significant at the 1% level and equal to 222%. For NASDAQ its values are negative at 20 and 30% of the previous spread and the total sample value is not significant.

The quoted volume parameter γ_p has been plotted in figure 6.2. For NYSE it is significant in eight of the deciles, in seven of these at the 1% level. Its value decreases steadily towards the highest activity decile and ranges from 3 to 0.2 percent of the previous spread (total sample value is 0.4%). For NASDAQ, where the value of this parameter is significant in only the second lowest and the highest activity deciles at the 1 and 5% level respectively, its values are 61.6 and 26.1% of the previous spread and the total sample value is not significant.

The inventory cost parameter δ_p, plotted in figure 6.3, is significant in six of the NYSE deciles (two at 1% and two at the 5% level), its value ranging from 3 to 0.2% of the previous spread (total sample value is 0.06%). Its pattern follows a U-shape (with the exception of decile 9 which has a low value) being insignificant in the medium-activity deciles. For the NASDAQ sub-sample this parameter is significantly different from zero in also five deciles, in four at the 1% level and shows an increasing pattern as trading activity increases. Its values range from 5.3 to 8.6 % of the previous spread being much larger than NYSE values (total sample value is not significant).

The adverse selection parameter θ_p, plotted in figure 6.4, which is significant in four NYSE deciles (at 1%), has values ranging between 6.3 and 0.3% of the previous spread (the value for the total sample is 0.1%), decreasing as activity increases. For NASDAQ stocks it is significant in three medium-activity deciles (two at the 1 and one at the 10% level) ranging from 8.8 to 3.3% of the spread and follows the same decreasing pattern as for NYSE (total sample value is 9.3%).

A summary of the results of the Price-Change Models is provided in table 6.1.

Using the parameter values estimated and presented in tables AG.6.3 and AG.6.4 in Appendix G together with the values for the average realized spreads which are presented in tables AG.6.1 and AG.6.2 in the same Appendix the adverse-selection and the inventory-holding costs present in the realized spread have been calculated for the NYSE and NASDAQ subsamples as percentages of the realized spread and are given in table 6.2. As is evident from this table only a very small part of the above costs is recovered through trading. For NYSE the adverse-selection cost ranges between 0.11 and 2.59 percent of the realized spread whereas the inventory-holding cost ranges from 0.12 to 0.86 percent. For NASDAQ stocks the adverse-selection cost varies from 1.18 to 2.34 percent whereas the inventory-holding cost from 0.16 to 3.61 percent of the realized spread. However, as has been shown by Ellis, Michaely and O'Hara (2000) and discussed in section 4.3, the effective spread we have calculated may be over-estimated due to our use of the Lee and Ready (1991) rule for trade-classification in which case the estimated components as percentages of the realized spread could be even larger.

Table 6.1 Summary of the Results from the Estimations of the Price-Change Model, NYSE and NASDAQ TAQ Data: October 1994.

Table 6.1 presents a descriptive summary of the results from the GMM estimation of the Price-Change Model of equation (6.5b) for both the NYSE and NASDAQ sub-samples from the TAQ database for October 1994.

PRICE-CHANGE MODEL	NYSE	NASDAQ
α_p		
	Significant in eight deciles (7 at 1% level)	Significant in four deciles (1 at 1%; 2 at 5%)
	4 to 21 % of previous spread	25 to 60 % of previous spread
	Inverse U Shape	Decreases with increasing activity decile
γ_p		
	Significant in six deciles (four at 1% level)	Significant in two deciles (1 and 5% level)
	1.8 to 0.6 % of previous spread	61.6 and 26.1 % of previous spread
	Decreases from lowest to highest activity deciles	Highest activity decile has lower value
δ_p		
	Significant in six deciles (2 at 1% and in 2 at 5% level)	Significant in five deciles (in four at 1% and in one at 5% level)
	Positive; 3 to 0.2 % of previous spread	Positive; 5.3 to 8.6 % of previous spread
	U-shape with highest value in low activity deciles (but lowest value in highest)	Increases with increasing activity deciles
θ_p		
	Significant in four deciles (at 1% level)	Significant in 3 medium – activity deciles (2 at 1% and 1 at 10% level)
	6.3 to 0.3 % of previous spread	8.8 to 3.8% of previous spread
	Decreases with increasing activity decile	Shows decreasing trend as activity increases
β_p		
	Significant in two deciles (at 1%)	Significant in two deciles (1 and 10% level)
	99 to 264 % of previous spread	20 to 30 % of previous spread
	Positive	Negative

Table 6.2 Adverse-selection and Inventory-holding costs as Percentages of the Realized Spread, NYSE and NASDAQ Common Stocks, TAQ Data: October 1994.

Table 6.2 presents the adverse-selection and inventory-holding costs present in the realized spread, as percentages of this variable, for NYSE and NASDAQ subsamples from the TAQ database for October 1994. Figures for the adverse-selection parameter, θ_p and the inventory-holding cost parameter, δ_p, from tables AG.6.3 and AG.6.4 in Appendix G, which have been estimated by the Price-Change Model, together with values for the average realized spreads, S_R, presented in tables AG.6.1 and AG.6.2 in the same Appendix have been used for the calculations.

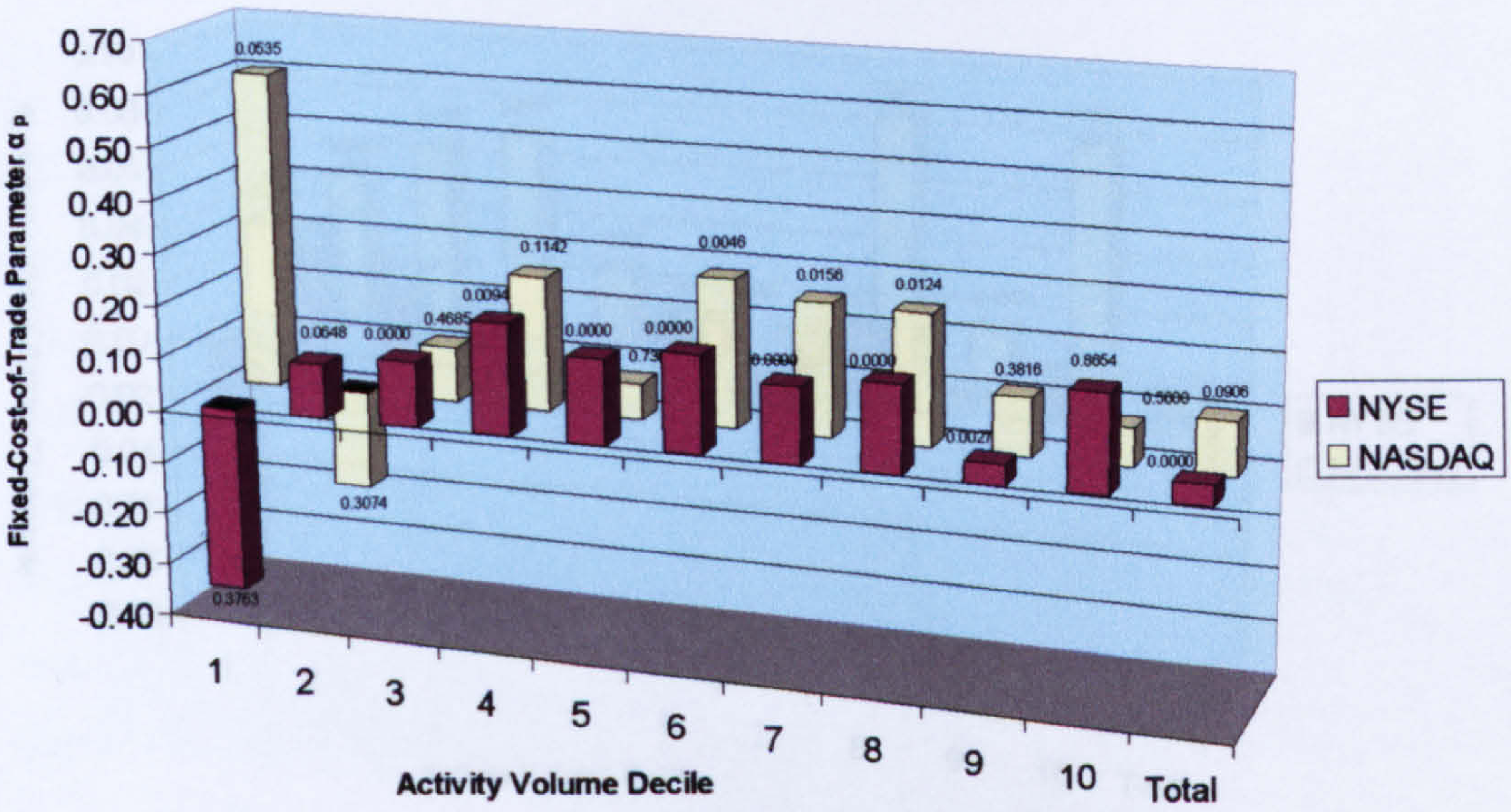
NYSE

Decile	1	2	3	4	5	6	7	8	9	10
Averse-Selection Cost	0.00	0.00	2.59	0.00	2.16	0.00	1.37	0.00	0.12	0.00
Inventory-Holding Cost	0.00	0.12	0.23	0.27	0.12	0.57	0.00	0.86	0.12	0.00

NASDAQ

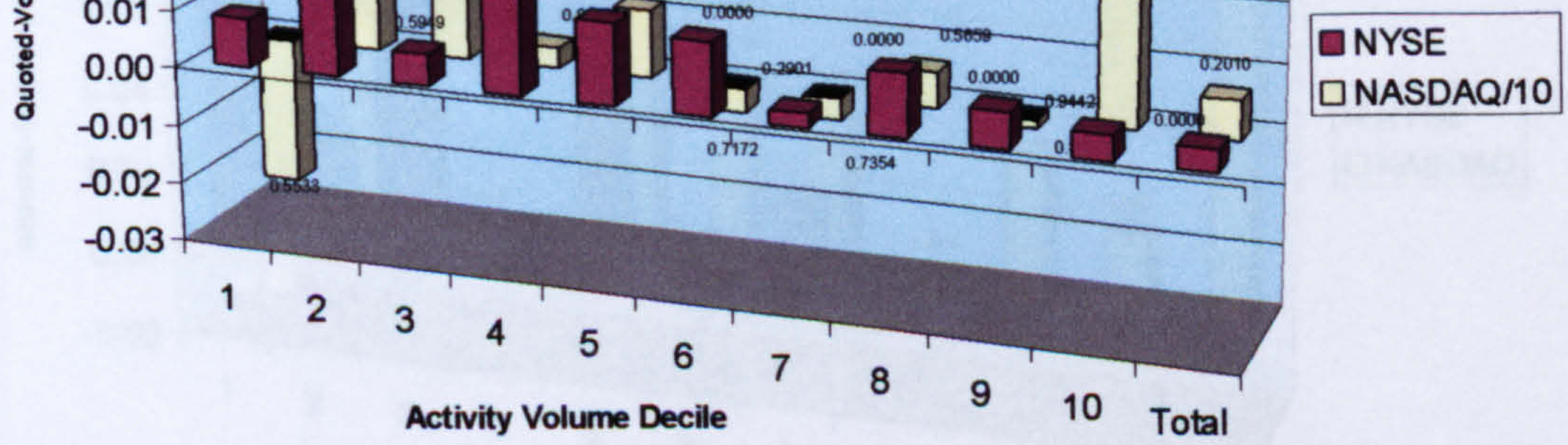
Decile	1	2	3	4	5	6	7	8	9	10
Averse-Selection Cost	0.00	0.00	0.00	0.00	2.33	1.18	1.73	0.00	1.78	0.00
Inventory-Holding Cost	0.00	0.16	0.00	2.05	0.00	0.00	0.00	2.03	0.00	3.61

FIGURE 6.1
THE PRICE-CHANGE MODEL
Variation of the Fixed-Cost-of-Trade Parameter α_p with Activity Volume Decile



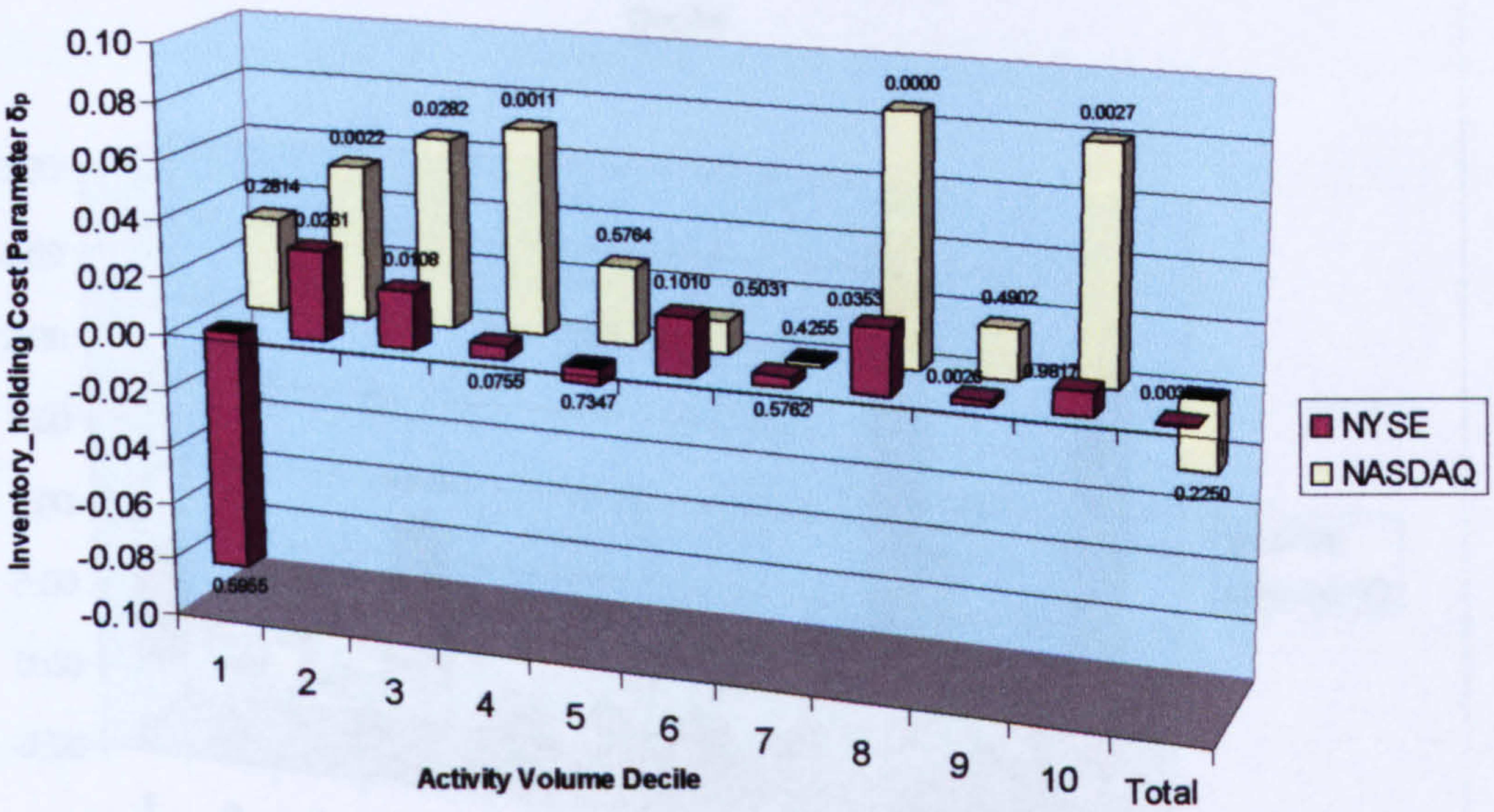
Numbers appearing on top of the bars refer to probability values.

FIGURE 6.2
THE PRICE-CHANGE MODEL
Variation of the Quoted-Volume Parameter γ_p with Activity Volume Decile



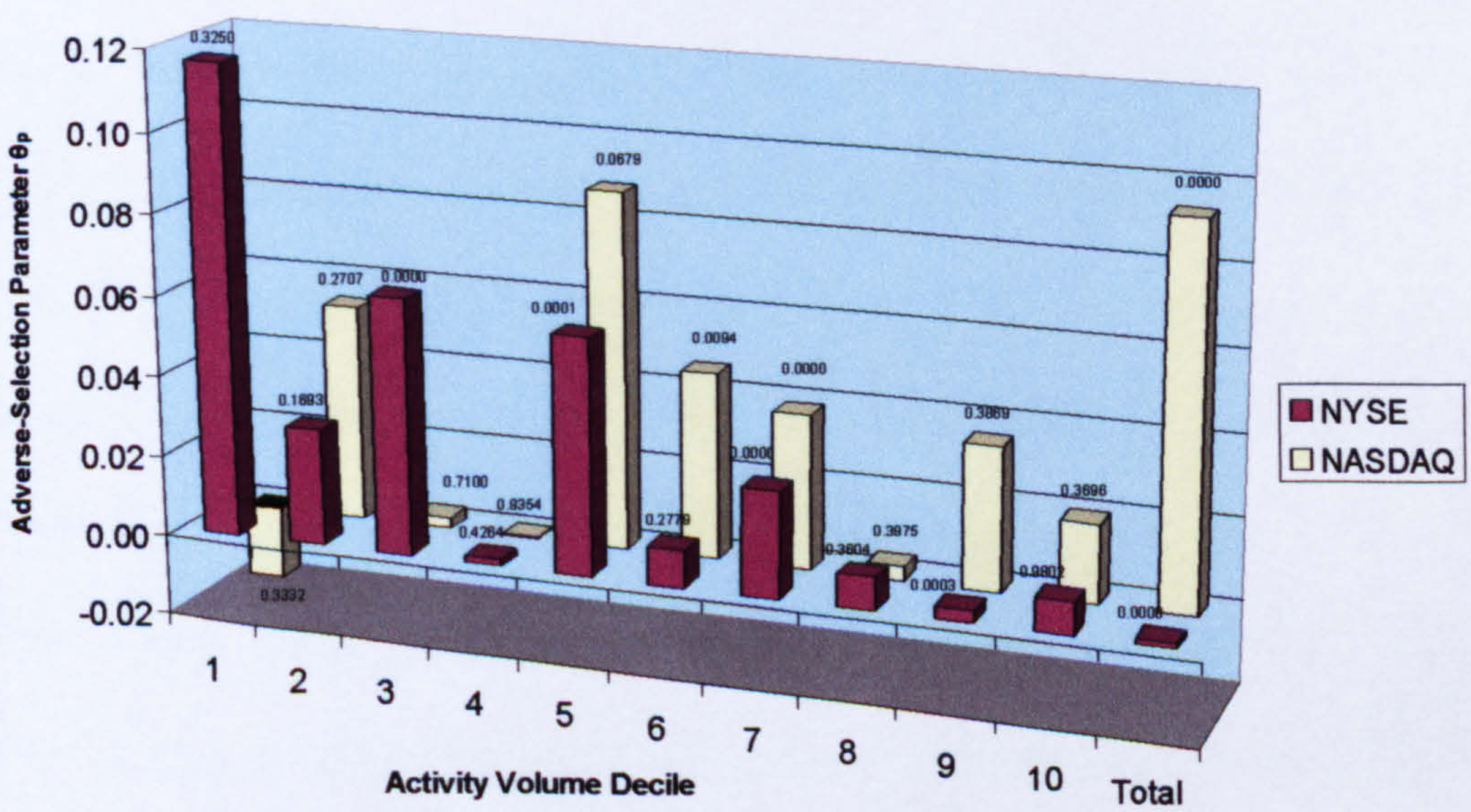
Numbers appearing on top of the bars refer to probability values.

FIGURE 6.3
THE PRICE-CHANGE MODEL
Variation of the Inventory-Holding Cost Parameter δ_p with Activity Volume Decile



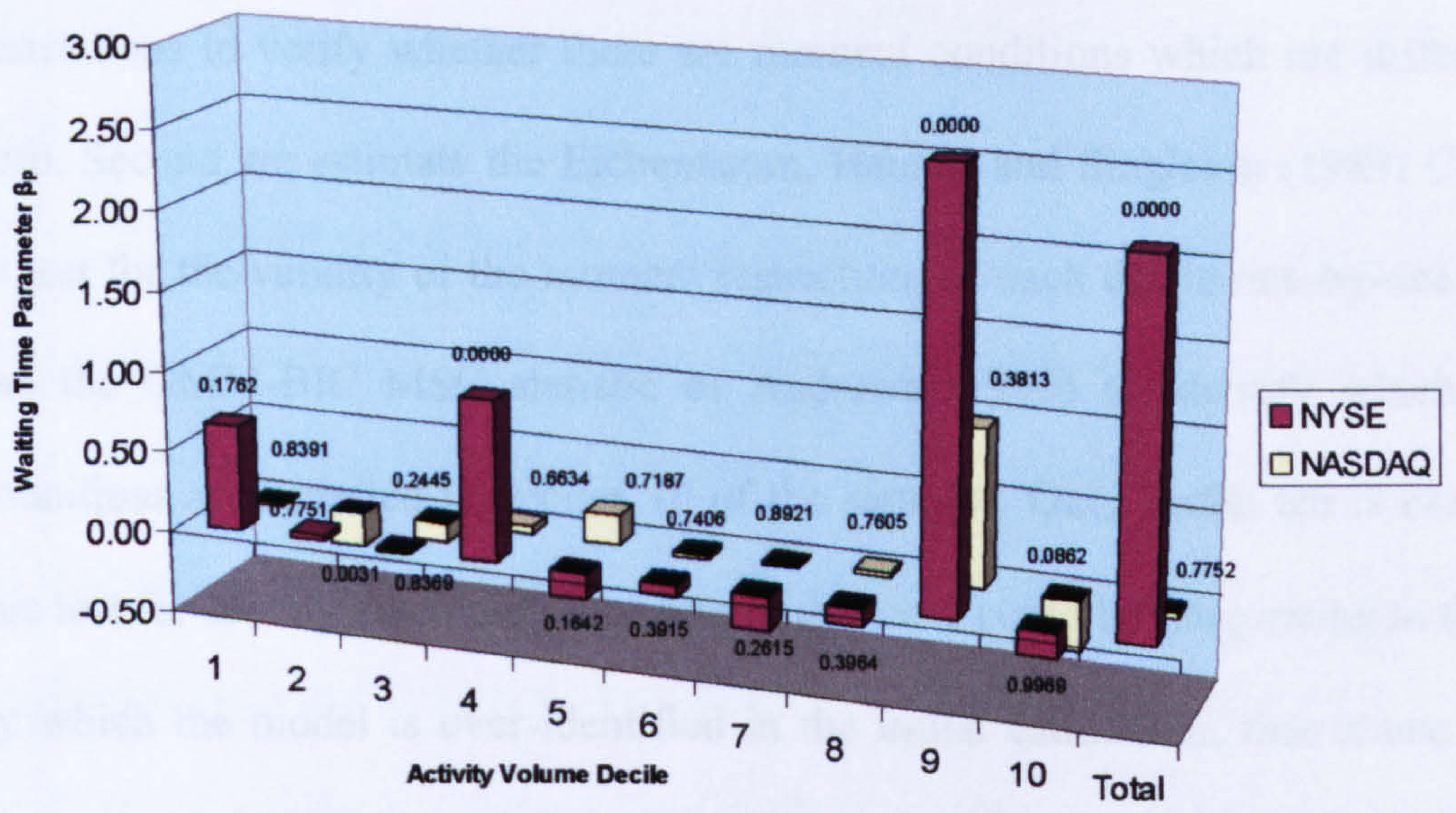
Numbers appearing on top of the bars refer to probability values.

FIGURE 6.4
THE PRICE-CHANGE MODEL
Variation of the Adverse-Selection Parameter θ_p with Activity Volume Decile



Numbers appearing on top of the bars refer to probability values.

FIGURE 6.5
THE PRICE-CHANGE MODEL
Variation of the Waiting Time btn Trades Parameter β_p with Activity Volume Decile



Numbers appearing on top of the bars refer to probability values.

6.6 Specification Testing

We follow the same procedure for the specification of this model as for the Quoted Spread Model in chapter five. We first examine Hansen's J-statistic of overidentifying restrictions to verify whether there are moment conditions which are different from zero. Second we estimate the Eichenbaum, Hansen and Singleton (1988) C_T statistic to test for the validity of the moment restrictions in each decile one-by-one. We also use the GMM-BIC MSC statistic of Andrews (1999) to identify which moment conditions are satisfied in deciles 10 of the samples. Only decile ten is examined in this test for brevity since the procedure in this case could be long owing to the degree by which the model is over-identified in the initial estimation, that is one could go down to a series of steps, each time omitting one more variable. Subsequently we test for the validity of the parameter restrictions using the Likelihood Ratio (LR).

6.6.1 The J-test statistic

This should be distributed as a chi-square with five degrees of freedom. As is evident from table AG.6.3 in Appendix G for the NYSE sub-sample the statistic lies within the 10% interval in four and within the 5% interval in two deciles. It is only for deciles 3, 7, 9 and 10 that we can infer that part of the moment conditions used in the initial estimation are over-identifying. The picture is different for the NASDAQ sub-sample as can be seen from table AG.6.4 where the J statistic lies outside the 1% interval in only one decile (decile 7) and between the 1 and 5% critical values in deciles 5, 6 and 10. Overall it can be argued that the moment conditions used in the original estimation of the NASDAQ sub-sample hold better compared to the NYSE one.

6.6.2 Checking for correct moment conditions – Moment Selection

To determine which moment conditions could be characterized as over-identifying the Eichenbaum, Hansen and Singleton (1988) C_T statistic has been presented in tables AG.6.5 and AG.6.6 in Appendix G for all deciles. As is evident from table AG.6.5 for NYSE at the 5 percent level the statistic is not significant for the majority of moment conditions in all except in the ninth and tenth deciles where it is significant for most moment conditions. At the same level of significance the null hypothesis that all moment conditions in the NYSE sub-sample are satisfied can be rejected in most deciles for the trade indicator variable-related moments at times t and t-1 as well as the difference in waiting time-related moment at time t-1. As can be seen from table AG.6.6 for NASDAQ for most deciles and the majority of the moment conditions the null hypothesis cannot be rejected at the 5% level. It is only in three out of the ten deciles that the hypothesis can be rejected for the trade indicator–related moments at time t and t-1 as well as for the trading volume at t-1 – related moment condition (and for one decile for the volume at time t).

To choose which moment conditions are valid in the sample we use the GMM-BIC (Bayesian Information Criterion) statistic of Andrews (1999) as in chapter 5. This statistic is presented for the NYSE and the NASDAQ sub-samples in tables AG.6.7 and AG.6.8 in Appendix G respectively where two moment conditions are excluded from the full set of moment conditions which had been used in the initial estimation. The degree to which the Price-Change model is over-identified in the initial GMM estimation in tables AG.6.3 and AG.6.4 is five which implies that by following the same downward testing procedure, as in chapter 5, that is by excluding one moment condition in turn until the J-statistic is not rejected, could lead to a very large number of estimations. For this reason the procedure has only been carried out for decile 10 and conclusions on which moment conditions hold in the whole sub-sample are drawn from this decile. As can be seen from table AG.6.7 for NYSE by excluding the moment conditions related to the couples of variables change in volume of ask at time t with the same change in time t-1 as well as with the change in the volume of bid at time t-1 one obtains the lowest values of the BIC statistic together with J-values which are well within the 5% confidence interval. Values of the BIC statistic very close to the above are also obtained for the moment conditions related to a) the change in waiting time at t-2 with the change in the volume of the bid at time t-1, b) the change in the volume of the ask at time t-1 with the change in the volume of the bid at t-1, c) the trade indicator variable at time t with that at time t-1 and d) the change in the volume of the bid at time t with both the volume at time t and at time t-1. From table AG.6.8 for NASDAQ, the couples of moment conditions giving the lowest BIC values, as well as J-values well within the 5% confidence interval are those related to a) the waiting time at t-1 with that at t-2 as well as the trade indicator variable at t-2 and b) the waiting time at t-2 with both the change in the volume of the ask at t as well as the trade indicator variable at time t.

6.6.3 Testing the Parameters of the Model

As was done in chapter 5 we test for the joint hypothesis that all parameters in the model are equal to zero by carrying out a Likelihood Ratio (LR) test. It is evident from tables AG.6.9 for NYSE and AG.6.10 for NASDAQ that at the five percent level of significance the null hypothesis should be rejected for all deciles of the NYSE subsample and for all but the first NASDAQ decile. It is possible that the small number of observations in the latter decile is not adequate for carrying out the test.

6.6.4 Conclusions from the Specification Tests

The above specification tests indicate that many of the strong exogeneity assumptions regarding the regressors used in the Price-Change model appear to be violated for the NYSE sub-sample, especially in the case of very liquid stocks. In particular, the changes in the size of the bid and ask as well as the trade indicator variable related assumptions seem not to be valid in the highest activity decile a fact which indicates that for highly liquid stocks these variables appear to be correlated at some points in time with the change in price. This effect has not been accounted for in the development of the Price-Change model where the particular variables have been assumed to be able to explain the change in price and not vice-versa.

Overall the exogeneity assumptions appear to hold better in the NASDAQ sub-sample where the ones related to the change in waiting time and the trade indicator variable appear to be violated more often than the rest. However, as discussed in section 5.6.4 the stronger validity of the assumptions for NASDAQ may be due to the smaller number of observations.

Finally, the specification tests in this section indicate that at least for NYSE the change in price should be modelled in such a way as to better capture the effect of lagged price changes on variables such as the change in the volume of bid and ask as well as on the trade indicator variable.

6.7 Conclusions

The results from the GMM estimation of the Price Change Model indicate that there is a significant Fixed-Cost-of-Trade parameter, α_p for NYSE stocks which, apart from the second and third decile where we obtain two smaller values, is larger in the medium activity deciles and becomes smaller as liquidity increases. Even though for NASDAQ the evidence is weaker, this component appears to be large in illiquid stocks and decreases slightly with increasing activity. The trend of a decreasing fixed cost of trade as liquidity increases, which is evident in figure (6.1), accords well with the idea of a dealer who charges a larger fixed order-processing cost for low liquidity stocks so as to deal with the danger of failing to recover his costs due to the small expected number of trades. In liquid stocks the dealer sets a lower fixed cost knowing that he could recover his costs from a larger number of trades.

The Quoted Volume parameter γ_p, plotted in figure (6.2), is only a fraction of the fixed cost and follows a decreasing pattern for NYSE while it is only present in the highest and lowest activity deciles of NASDAQ. The decreasing pattern as one moves from the low to the higher activity deciles for NYSE can be attributed to the rate at which the dealer attempts to recover his order-processing cost according to the liquidity in the stock and can be combined with the evidence for the fixed-cost-of-trade component. Evidence for the first and third deciles of NYSE is weak implying that there is probably only a fixed cost-of-trade component for the order-processing cost in the very low activity deciles. This is also true for NASDAQ as evidenced by the large difference between the two values which indicates that the procedure for the recovery of clerical costs probably follows the same mechanism. A point which requires an explanation though is the positive sign of this parameter for NYSE which was found to be negative for the equivalent parameter in the Quoted-Spread and Ask- and Bid-Change models in chapter 5. A possible explanation for this seemingly conflicting evidence for NYSE is that the increase in volume at the quotes forces traders to place limit orders inside the quotes so that they stand a better chance of having their orders carried out ahead of the queue. This has the effect of decreasing the size of the quoted spread. The quoted spread in NYSE however is only representative of the state of the book and the depth of the inside spread may be small. Anxious traders who wish to transact at high volumes and willing to pay a higher price may place market orders which will quickly exhaust the depth at the posted ask and bid and will eventually lead to an increase in the effective spread which would not have taken place had such orders not been placed.

It appears that both the Inventory-holding Cost Parameter, δ_p, in figure (6.3) and the Adverse Selection Parameter, θ_p, in figure (6.4) are close to zero in most deciles of both exchanges. The most likely explanation for this finding, is that, even though the two components are present in the spread quoted prior to the trade, the Order-Processing cost component is the main constituent of the realized spread which eventually leads to the formation of the price. Thus our model which, contrary to previous research like Huang and Stoll (1997) and Madhavan et.al (1997), uses trading volume directly into the equations for the quoted spread and the price does not find serious evidence that these two components are recovered through the price which from our evidence appears to be driven mainly by the order-processing cost. Note that equation (6.5b) described the change in price taking into account the realized spread and in this sense it refers to the outcome of the trade while equation (5.11b) for the quoted spread refers to the time before the trade.

For the inventory-holding cost this may imply that even though it is used in the quote-generating mechanism a very small part is recovered through price during the limited time-horizon (one month) which our dataset spans. NASDAQ values are much larger than NYSE and no evidence is found of values which are significantly different from zero in the medium-activity deciles. The values for the total inventory-holding cost relative to the realized spread which are given in table 6.6 reveal that only a very small proportion of the cost implicit in the quoted-spread is eventually recovered in the realized spread. This cost, compared to that estimated by Huang and Stoll (1997) for NYSE which ranges from 7.5 to 58 percent for the sample of the twenty stocks they used, is much smaller and apart from the different datasets used can only be attributed to the period of one year they examine which is longer than our one month. Nevertheless their assumption of a constant realized spread may also influence their results. Madhavan, Richardson and Roomans (1997) do not model this cost in their equations. Stoll(1989) estimated the quoted spread and the realized spread and from this he inferred the components of the realized spread which he argued should exclude adverse-selection. Using his results the average inventory-holding cost in the realized spread should drop to 17.5 percent.

The adverse-selection parameter, θ_p, in the three deciles of NYSE where it is significant also appears to be small, around 6 to 2 % of the previous spread whereas for NASDAQ it is also significant in three medium-activity deciles. The pattern for both exchanges is one of decreasing values as activity increases. The values for the total adverse-selection cost relative to the realized spread which are given in table 6.2 indicate that, as in the case of the inventory-holding cost, only a very small proportion of the cost implicit in the quoted-spread is eventually recovered in the realized spread. Both Huang and Stoll (1997) and Madhavan et. al (1997) use only NYSE stocks for their estimations. Our values are of the same order as those found in both of the above papers. Whereas our values range from 0.12 to 2.6 percent of the realized spread from decile to decile, Madhavan et. al. (1997) values range from 4.15 to 2.87 percent during the day. Huang and Stoll (1997) find an average value of 9.6 percent but the values for individual stocks range from 1.4 to 21 percent. Lin, Sanger and Booth (1995) with their statistical model find very large values (ranging from 19.8 to 62.6 percent) but spread components are not estimated together as in the other two models. It is possible that if a stock-by-stock estimation was carried out in this thesis the results obtained could lie between those of Huang and Stoll(1997) and Madhavan et. al. (1997).

The average adverse-selection cost in the realized spread for NASDAQ which has been estimated in this thesis is 12.5 % higher than that for NYSE. This finding is in agreement with the findings of Van Ness, Van Ness and Warr (2002) who showed that the adverse-selection component for NASDAQ stocks is higher than for NYSE and AMEX stocks, a result which is opposite to that found by Affleck-Graves, Hegde and Miller (1994).

Of course it is possible that our model cannot properly uncover this effect from the data and that a better equation may be required to be developed which would model this effect in a different way. Moreover, since this model is much more complex, compared to those in chapter 5, it may be severely affected by the averaging of the autocorrelation coefficients of the trade-indicator variables already discussed in section 5.7. For the NASDAQ dataset there is the additional problem that there are fewer matched trades which render estimation from this model more difficult.

The waiting time between trades parameter β_p, plotted in figure (6.5) does not appear to play a significant role in the formation of the price in the NYSE but its very large value in the ninth decile together with a similar and significant value for the total sample is worthy of further investigation perhaps by modeling its relation to prices and quotes differently or using a larger and more accurate database. The idea that the time between trades may serve as an information signal has already been presented in the work of Diamond and Verrecchia (1987) and Easley and O'Hara (1992a). As the latter paper argues, the absence of trade may serve as a signal for the existence of new information. An explanation for our evidence may thus be that in high activity deciles where a large number of trades take place, the existence of a long period of no-trade, which would not normally be expected, adds to the uncertainty regarding information and results in an increase in the price of the stock. How this increase is brought about, given that the evidence of a similar effect in the quoted spread models is weak, is not however clear and requires further investigation.

Even though our data for NASDAQ do not provide us with adequate information on the volume of shares quoted, they give us full information regarding the trades taking place. However, during the matching procedure a number of trades, which could not be matched with quotes (including mid-quote trades) have been discarded. Perhaps, if a different matching procedure is followed, during which trades would be matched with older quotes, assuming those still prevail, better results could be obtained. Judging from the results from NYSE, which come from a fuller database, it is anticipated that these components will not be large. Nevertheless, the weak evidence for the adverse-selection and inventory-holding costs from the Price-Change model necessitates the development of an improved model to study the components of the effective spread before accepting the idea that volume does not influence these components.

Overall the contribution of the model in this chapter has been to:

1. Document that there is a significant fixed-cost-of-trade component in both NYSE and NASDAQ stocks which decreases as trading activity increases.
2. Show that quoted volume affects the price-formation process of NYSE stocks its effect decreasing with increasing trading activity in the stock.
3. Indicate that trading volume does not seem to exert a significant influence on the adverse-selection component of the realized spread even though there are doubts regarding the effectiveness of the model in uncovering this effect.
4. Indicate that there is weak evidence of the existence of an inventory-holding cost affecting the realized spread and thus the price but it is doubtful that such could be uncovered in the short horizon which our datasets span.
5. Show that there is some evidence of the significance of the time between the trades for the price-formation process of NYSE stocks which have very high liquidity but this factor requires more extensive research.

Chapter 7

Conclusions

7.1 Introduction

The two models which have been developed in chapters five and six of this thesis have attempted to measure the components of the bid-ask spread of common stocks when the volume of shares traded as well as the volume quoted at the bid and the ask prices (depth) are taken into account explicitly and not merely considered as ranges of volume in the analysis. The role played by the waiting-time between trades in the price-generating process, through the formation of the bid-ask spread, has also been investigated by building appropriate variables and parameters into the models and the fixed cost of trading has been incorporated into the order-processing cost equation. Finally the behaviour of the spread has been investigated for two different exchange mechanisms as well as across deciles of varying trading activity within each exchange.

It should be pointed out that the main equations of the two models, in particular the Quoted-Spread and the Price-Change model, deal with two separate stages of the price-generating process. The former model covers the first stage in the process where the bid and ask quotes are formed by the market-makers or the traders, just before trading takes place and examines the components of the quoted-spread that arise from this procedure. The latter model examines the components of the spread which have led to a change in price after a trade has taken place, following the quotation procedure. The interpretation of the results of the two models together aids in the analysis of the spread components highlighting their differences before and after a trade. The main difference between the data used in the estimation of the two models is the use of prices which the Price-Change model employs leading to the particular results of the estimation of the parameters.

In contrast to other papers, which employ the same estimation methodology as the one in this dissertation, we have also examined the validity of the moment conditions in the samples used for the GMM estimations of our models in order that a clearer picture is obtained regarding the accuracy of the estimation results.

The two papers most relevant to this research are those of Huang and Stoll (1997) and Madhavan, Richardson and Roomans (1997). Neither of these papers considers the effect of the quoted volume of shares on the spread. Moreover, the latter uses quotes only in order to determine the trade indicator variable and does not consider the inventory cost effect. The former model assumes that the spread is constant and estimates it using all the data. None of these studies incorporates the volume of shares traded explicitly like we have done and none attempts to study the behaviour of these components based on activity deciles. They also do not contrast the behaviour of the two exchanges in the two separate trading mechanisms. Therefore our results, especially those pertaining to the quoted spread, are not directly comparable to the findings of the above papers since in our models a larger number of parameters is involved with fewer assumptions.

7.2 Main Findings

The main findings of this thesis are:

1. Trading volume affects the formation of the adverse selection-cost implicit in the quoted spread. In the NYSE the adverse-selection parameter becomes smaller the higher the trading activity in the stock imposing a similar pattern on the adverse-selection cost. This is probably due to the ability of specialists to recover it through a large number of trades in more liquid stocks. It is not

symmetric for the bid and ask sides of the spread (larger for the ask) that is the assumption of symmetry in our as well as most other empirical work is wrong. For NASDAQ stocks its effect is in general larger compared to NYSE probably showing that the information asymmetry in this exchange mechanism is stronger than in NYSE and demands the placement of a stronger charge for this cost on behalf of the dealers when it comes to quoting. We also find an asymmetry between the bid and ask sides in NASDAQ stocks but this effect is believed to be the outcome not of the quote-setting behavior of the dealer but rather results from the nature of our data which only report quotes at the BBO. Evidence indicates that the adverse-selection cost component in the realized spread which depends on the volume of shares traded is very small. This shows that the largest part of these components, build by dealers during the quoting procedure, is practically not recovered on average during the transactions. Our results which show that the average adverse-selection cost in the realized spread for NASDAQ is 12.5 % higher than that for NYSE agree with the findings of Van Ness, Van Ness and Warr (2002) who showed that the adverse-selection component for NASDAQ stocks is higher than that for NYSE and AMEX stocks. Prior evidence by Affleck-Graves, Hegde and Miller (1994) had indicated the opposite result.

2. The absolute value of the inventory-holding cost parameter in the quoted spread is found to decrease as trading activity in a stock increases. The fact that, for both NYSE and NASDAQ, this component is found to decrease as one moves from the low to the high activity deciles, contrary to volume which of course exhibits the opposite pattern, indicates that in higher activity deciles the total inventory holding cost is recovered through the increased number of shares

traded, even though at a slower per share rate to comply with the faster adjustment in inventory for frequently traded stocks which has been found empirically by Hasbrouck and Sofianos (1993). This implies that dealers can possibly recover this cost through a large number of trades in high activity deciles. Nevertheless the pattern found for the total inventory-holding cost in the quoted spread is found to be present mainly in medium-activity deciles and to follow an inverse-U shape. The value of the inventory holding cost parameter is not symmetric for both the NYSE and NASDAQ implying that the assumption of symmetry built into our models as well as all other models in the literature is probably erroneous. There is weaker evidence for this effect in NASDAQ stocks probably owing to the particular trading mechanism which allows dealers to share the collective inventory risk or avoid exposing themselves to this risk by quoting away from the BBO. The inventory-holding cost parameter however appears to have a smaller effect on the quoted spread compared to the impact of the quoted-volume parameter. Even though the size of this component is large, as evidenced by the results from the Bid- and Ask-Change models, the net effect is quite small since the two effects apparently cancel out. Evidence also indicates that this cost is almost completely recovered through the effective spread. Our values for this cost are smaller than those found by other researchers probably due to the short horizon of our database (one month). Inventory changes, which according to Madhavan and Smidt (1993)[119] and Hasbrouck and Sofianos (1993) take place during a particularly long period of time, cannot be captured by our intraday data which span a period of one month only, thus the parameter in our models may not be able to reflect this effect

[119] As mentioned in footnote 22 in chapter 2.

since there may be a small number of large changes, beyond the horizon of our database, which affect the inventory position of dealers. Another explanation for the small size of this effect is the fact that we have used in our models the average inventory imbalance of all stocks in a decile to estimate the inventory-holding component and it is possible that part of the effect could cancel out due to this averaging procedure.

3. There is strong evidence that the volume of shares quoted at the ask and bid prices (depth) affects the size of the quoted spread of both NYSE and NASDAQ stocks and the impact of this effect appears to be larger for NASDAQ as compared to NYSE stocks. The parameter values for this effect in NYSE are negative since, on average, there is negative correlation between the changes in the prices of the quotes and the depth at the quotes. Our explanation for this effect on NYSE is that the specialist has the power to correct his initial fixed-cost-of-trade in low activity deciles when he realizes that quoted volume is building up. For NYSE stocks the effect is symmetric at both sides of the spread justifying our assumption of modeling it in the same way for both sides. For NASDAQ stocks however the effect is not symmetric probably owing to the nature of our data as explained above. It is found that quoted volume has a positive effect on the price of both NYSE and NASDAQ stocks and this effect for NYSE decreases with increasing activity in the stock exerting a significant but small impact on the formation of the price of a stock. One explanation for this finding is that the large impact that this component has on the quotes is reduced as prices alternate between the bid and the ask prices set by the dealers, nevertheless it affects price-formation slowly but steadily. The evidence that this component affects the prices of NASDAQ stocks is weak but, wherever it is

significantly different from zero, it has a large value. Our finding regarding this component is in agreement with the findings of Chakravarty, Harris and Wood (2001) who have shown that new information is first reflected in the depths of the bid and ask quotes, rather than in the updates in the spreads, implying that size leads prices as far as the information brought into the market is concerned. While they have used a statistical model to arrive at this result we have built this effect into a comprehensive model of the bid-ask spread which is based on sound theoretical principles and have estimated it for both NYSE and NASDAQ stocks.

4. There is weak evidence that the waiting time between trades affects the quoted spread which indicates that it plays practically no role in the formation of the quoted spread. There is also weak evidence that it affects the price-formation process of NYSE (not NASDAQ) stocks with its effect increasing with increasing activity in the stock. The mechanism of how this takes place is worthy of further investigation.

5. There is strong evidence for the existence of a fixed-cost-of-trade parameter for both NYSE and NASDAQ stocks which becomes smaller with increasing trading activity in a stock. This is attributed to the fact that dealers in less liquid stocks set a larger fixed cost since they are less likely to recover their order-processing costs from a large number of trades.

7.3 Contribution

The main contribution of the research work in this thesis has been to show that:

1. There is an adverse-selection component implicit in the quoted-spread which depends directly on the number of shares traded, unlike previous studies which

only use dummy variables for ranges of size. It has been shown that this effect is not symmetric for the ask and bid sides of the spread implying that empirical models which assume that it is are wrong. Parameter values for this effect have been estimated for both NYSE and NASDAQ stocks but within certain limitations which are described below and have been imposed by the models employed. Evidence has been found that trading volume has a small effect on the adverse-selection component of the realized spread. The evidence in this thesis also reinforces recent evidence which indicates that the adverse-selection costs recovered through the realized spread are larger for NASDAQ compared to NYSE stocks.

2. There is an inventory-holding component implicit in the quoted spread which varies with the trading activity of both NYSE and NASDAQ common stocks. It has been shown, as for the adverse-selection component, that it is not symmetric for the ask and bid sides of the spread. Evidence has also been found of the existence of a small inventory-holding cost affecting the realized spread and thus the price but it is doubtful that such could be completely uncovered in the short horizon which our datasets span.

3. Even though the adverse-selection cost constitutes a large part of the quoted spread it appears to be a very small part of the realized spread. The results of this thesis for both the adverse-selection and the inventory-holding cost show that they are smaller compared to those estimated in other papers.

4. The volume of shares quoted at the bid and ask (depth) plays a significant role in the formation of the quoted spread for both NYSE and NASDAQ common stocks, a feature not considered in earlier work. This effect is found to be symmetric for the ask and bid sides of the spread. Depth also affects the price-

formation process of NYSE stocks its effect decreasing with increasing trading activity in the stock. This component, even though it has been presented here as part of the Order Processing Cost, it could also be considered as a component of the Adverse Selection cost. The reason for extending this argument is that it too acts as a signal to the trader / dealer of the expected change in price and the market participant accordingly places his order so as to minimize the effect of this expected demand / supply of the stock in line with the original idea of Bagehot (1971). The evidence in this thesis is in line with recent findings which indicate that new information first enters the price-process through the depths of the quotes and not the size of the spread, however we provide a thorough analysis of this effect for two separate market mechanisms and also for stocks of varying trading activity.

5. There is a significant fixed-cost-of-trade component in both NYSE and NASDAQ stocks which decreases as trading activity increases.
6. The waiting-time between trades may account for a very small part of the quoted spread but may also play a significant role in the price-formation process of those NYSE stocks which have very high liquidity. Further research is required to uncover a possible relationship and further research is required to investigate the mechanism through which this effect takes place.
7. The above analysis has been carried out for two different trading mechanisms, at both the pre-trade and post-trade stages, using deciles of trading activity which help determine the patterns of the parameters of the above components.
8. A thorough analysis of the components of the spread cannot be carried out without explicitly incorporating the effect of volume whether it is in the form of shares traded or shares quoted prior to a particular trade.

9. The results in this thesis have demonstrated that, in spite of the competitive nature of the multiple-dealer market of NASDAQ, its dealers overreact to adverse information extracting higher rents from traders compared to what a specialist in NYSE would charge. This behaviour could lead to a liquidity crisis easier than under the specialist system of NYSE where the monopoly power of the specialist and his privileged position, owing to the consolidation of the order flow for a stock through him, could help him moderate his losses. Even though quotes in NASDAQ are competitive, the lack of a limit order book, as was the case in the period where our data originate from, leads to higher effective spreads. However, the ability of NASDAQ dealers to share the risks of inventory imbalances allows them to charge lower inventory costs as has been found by our analysis.

7.4 Policy Implications

A comparative examination between the results of this thesis relating to the NYSE and NASDAQ sub-samples can shed some light into their implications to design issues regarding the two exchange mechanisms as well as to likely policies which could be adopted to the advantage of certain types of market participants.

As mentioned in chapter three the two major differences between the two exchange mechanisms are the monopoly power of the specialist in NYSE and the consolidation of the order flow for a particular stock through him against the competitive multiple dealer system of NASDAQ.

Even though, average quoted spreads in NYSE and NASDAQ are very close, as is evidenced from tables AG.6.1 and AG.6.2 in Appendix G, the realized spreads of NASDAQ stocks are significantly larger implying that traders in that market bear higher trading costs. Combined with the results for the adverse selection component which indicate that it is also larger in NASDAQ stocks, one can conclude that uninformed traders in this multiple dealer market eventually pay higher information-related costs. This implies that competition is effective in shaping the quotes of the dealers but price improvement under this mechanism is lower relative to the specialist system of NYSE. For the latter, consolidation of the order flow through the specialist brings him to a better position in assessing the level of informed trading and he can thus set a more reasonable adverse selection cost. In NASDAQ however, dealers only experience part of the order flow and they probably overreact to informed trading extracting excessive rents from uninformed traders. Another advantage of consolidating the order flow is that, as the probability of orders arriving simultaneously at the exchange increases with liquidity, the probability of matching opposite orders also increases and the specialist has the tendency to decrease his adverse selection cost to the advantage of traders. Moreover, he is not obliged, or even allowed, to match all orders himself since many of these transact against limit orders in the book. As liquidity increases, together with competition from limit order traders, the monopolist position of the specialist can help him extract some rents from traders and thus continue to offer a market in the stock, otherwise he would not risk deviating from his optimum inventory position during large imbalances of buy or sell orders in which case a liquidity crisis would emerge. Thus our results indicate that the inefficiency of NASDAQ dealers in handling adverse information without extracting excessive rents from uninformed traders could easily lead to liquidity crises when in the face of excessive volatility they overreact increasing their spreads very much thus driving liquidity traders out of the market.

On the other hand the specialist, as indicated by our results, charges higher inventory holding costs since he is the only one to bear the risks of undesired inventory during order imbalances compared to the risk-sharing capabilities of the dealers in the quote-driven NASDAQ market who handle part of the order flow.

The effect of the quoted volume parameter, which from our results appears to be larger for NASDAQ and which also does not decrease with trading activity, contributes to our conclusion regarding the overreaction of dealers in NASDAQ to increased information trading.

Whereas increasing liquidity can decrease costs in both exchanges it is believed that it is only through the introduction of a limit order book that the adverse selection component in the effective spread could be reduced in NASDAQ since trading could then take place inside dealers' quotes who then have to compete with limit order traders. It should be noted that the NASDAQ data used for the empirical analysis in this thesis come from a period of time when limit orders were not allowed.

Since as was expected, empirical evidence in this thesis indicates that adverse selection as well as inventory holding costs decrease with trading activity regulators in all markets should aim at increasing the scale of trading. The problem with thin markets, that is those where the number of transactions per unit time is low, is that they are not able to absorb large imbalalances of buy and sell orders without increasing price volatility. Then, risk averse traders who face increased transaction costs have the tendency to leave the market aggravating the situation.

The number of traders and the number and size of orders play an important part in arriving at a clearing price but depending on the choice between monopoly power, market transparency and types of orders allowed certain market participants will be worse off. So designers have to strike a balance between the satisfaction of the desires of all types of traders while at the same time keeping liquidity high and allow the hedging of price risks.

Assessing which type of market organization is appropriate or which policies should be implemented however demands a clarification as to whose advantage the design is aiming at. Exchanges normally wish to have high commissions or lower them and increase volumes. Dealers certainly wish to maximize their profits if competition would allow them and at the same time to abide by the rules of the exchange. Traders wish to carry out their orders at the minimum cost while at the same time minimize the effects of their trading on price. Finally, regulators wish to keep the market as stable and liquid as possible. The design of a market should attempt to compromise on the above motives and satisfy what Domowitz (1990) has defined as the three most important requirements of the market, namely: 1) reliable price discovery, 2) broad price dissemination and 3) effective hedging against price risks.

A reliable price-discovery implies that the market is always capable of arriving at a market clearing price. A monopolistic market like NYSE contributes to this end but this comes usually at the expense of the traders since consumer surplus goes to the specialist. Moreover, the price is not close to the underlying value of the security rendering prices inaccurate signals of value. However removing the monopoly power of the specialist could lead to liquidity crises as discussed above.

When prices are widely disseminated markets become transparent and traders can obtain information from market prices without losing to certain types of privileged traders. However, removing some of the advantages of informed traders may tempt them to move to other markets where they might have those advantages.

Uninformed traders will also be able to protect themselves from price risks when market liquidity is high and there is no anonymity as long as the lack of anonymity does not deter informed traders from trading since they will be unable to conceal their trades and thus take advantage of their information resulting in an illiquid market.

Market designers should at all times strive to create an efficient market in the sense that new information should be quickly aggregated and impounded into prices. The speed of assimilation of information is important since when assimilation is fast there is large price volatility and the ability of uninformed traders to hedge against price risk decreases. The speed of adjustment of prices in turn depends on the extent of informed trading which affects the losses of the uninformed but this conflicts with the task of the minimization of trading costs. The question then arises as to how a given surplus should be allocated among market participants. The answer is not straightforward and will depend on which types of traders' motives are the most important to be satisfied by the designer.

7.5 Limitations

The limitations of our models and database should be considered when interpreting the results of the estimations. In particular there are two types of limitations or drawbacks to our work. Those related to the models and those related to the data used.

1. The ability of our models to thoroughly describe the price-change mechanisms of the two exchanges is hindered by a number of simplifying assumptions which have been put forward during their development. The first is that the adverse-selection and the inventory-holding cost components have been assumed to be symmetric, that is to behave similarly for the bid and ask quotations which as evidenced by our results of the Offer and Bid-Change models is probably not true. It has also not been possible to estimate the adverse-selection parameter completely independently from the inventory holding cost one since we have used average autocorrelation coefficients of the trade-indicator variable so as to avoid a stock-by-stock estimation which would be difficult. Another assumption is that it is only

standing quotes which affect the components and not the quote revisions preceding them. As Bessembinder (2002) has shown estimated trading costs increase if, during trade classification, trade prices are compared to earlier, rather than contemporaneous, quotes reflecting adverse quote movements prior to trade report times. If these movements occur after order-submission but before trade execution then they impose a cost to traders which cannot be captured if trade prices are compared to quotes which are in effect at trade report times. Thus utilizing the information in the quote revisions could possibly improve the estimation of this component. Another limitation of our method is that we have used the average inventory imbalance of all stocks in a decile to estimate the inventory-holding component. It is possible that part of the effect cancels out due to this averaging procedure.

2. Data limitations include the fact that the TAQ database does not provide information regarding the origin and nature (buy or sell orders) of the trades as well as that NASDAQ quotes in this database do not disclose the full depth of the number of shares at the bid or ask across all the dealers who make a market in the stock. The first problem has necessitated the use of one of the methodologies presented in Chapter Four in order that trades can be classified as buyer- or seller-initiated and this results in introducing a number of errors due to the potentially erroneous classification of some of the trades and / or the inability of the methods to classify certain trades, especially mid-point ones. The second problem unavoidably leads to the evidence obtained from the NASDAQ subsample being poorer compared to that from NYSE. The short time-horizon of our dataset does not allow to properly uncover the inventory-holding effect since, as explained above, inventory imbalances have been shown to reverse over long time intervals.

Another problem during the preparation of the data for estimation has been that a large number of it is lost during the filtering and matching procedures described in chapter four. During these procedures significant information is lost which could otherwise improve estimation.

Finally, had intra-day seasonal patterns been removed from the variables in the models better estimation results would have been obtained.

7.6 Recommendations for Further Research

The above mentioned limitations and drawbacks of our work suggest the way into refining and improving potential future research. Our models, even though they carry the burden of more complexity, could furnish improved results by using more detailed databases (which unfortunately are not easily publicly available) and by making certain improvements on them provided that the new data can support these.

Assuming that improved data were available[120] which a) could identify the origin of a trade, that is whether it came from a buy or a sell order, the nature of the participants in the transaction, that is whether market orders were matched with market or limit orders, or with trades from the specialist, and b) would indicate the full depth of the quotes, it would not be necessary to use any trade classification methods with a resulting positive effect on the accuracy of the data and better results would be obtained for NASDAQ. Moreover, information on the inventory positions of dealers and traders would aid in the exact estimation of the size of the inventory holding cost component allowing for a better estimation of the adverse selection component as well.

As improvements to our models it is suggested that one should attempt to model the effect of quote revisions on the formation of the bid and ask through its effect on the adverse-selection and order-processing cost equations. The adverse-selection component should also be modeled independently from the inventory-holding cost component either by the use of improved data, as described above, or analytically probably using an alternative method if one could be devised. Different adverse-selection and inventory-holding cost parameters should be used in the ask- and bid-change equations to model the asymmetry in these parameters as indicated by the evidence in this thesis. The fixed-cost-of-trade parameter in the order-processing cost equation could also be modeled separately for the bid and ask since it may not exhibit the same behaviour when dealers buy as compared to when they sell.

Repeating the estimations in this dissertation using deciles constructed on the basis of turnover, instead of volume, might throw more light into the price-generating mechanism. It is expected that the use of data for more than one month in a period, from different months in a year or from two similar calendar months in two separate years would improve our understanding of the mechanisms involved in the formation of the spreads and prices of common stocks. Data for periods longer than the one month considered in this thesis could also allow for the proper estimation of the inventory-holding cost. NASDAQ data after 1998, when a limit-order book was implemented could also be used to investigate the differences in the particular exchange mechanism brought about by the new rules. Removing intra-daily seasonal patterns from the variables used for estimation in the models would probably lead to better results. Finally, by developing special programs which would allow for the stock-by-stock estimation at a fast pace would certainly improve the results obtained in this thesis.

[120] As in the papers of Hansch, Naik and Viswanathan (1998) and Hasbrouck and Sofianos (1993)

Appendix A
GMM Methodology

To avoid imposing strong assumptions about the distribution of the parameters of interest we choose to estimate equations (5.9b), (5.10b), (5.11b) and (6.5b) by the Generalized Method of Moments as developed by Hansen (1982). This method requires one to specify only a certain number of moment conditions and not the full probability density. However, this may have the drawback that it only uses some of the information in the sample.

Hansen (1982) offered an extension of the earlier minimization methods based on the minimum chi-square estimators of Cramer (1946), Ferguson (1958) and Rothenberg (1973) as well as the minimum distance estimator of Malinvaud (1970). He provided the most general characterization of this approach and derived the asymptotic properties for serially dependent processes.

For an (hx1) vector of variables w_t observed at time t, and an (ax1) vector of unknown coefficients (parameters) θ, an (rx1) vector valued function $h(\theta, w_t)$ of instruments is selected so that the true value of θ, θ_o, is characterized by the condition

$$E\{h(\theta_o, w_t)\} = 0$$

Which is a set of orthogonality conditions.

If

$$y_T \equiv \left(w'_T, w'_{T-1}, \ldots, w'_1\right)'$$

is a (Thx1) vector containing the observations in a sample of size T, and $g(\theta, y_T)$ is a (rx1) vector-valued function denoting the sample average of $h(\theta, w_t)$:

$$g(\theta; y_T) \equiv \frac{1}{T}\sum_{t=1}^{T} h(\theta, w_t)$$

The method consists of choosing θ so as to make the sample moment of $g(\theta, y_T)$ as close as possible to the population moment of zero.

The GMM estimator is the value of θ which minimizes the scalar

$$Q(\theta; y_T) = \left[g(\theta; y_T)\right]' W_T [g(\theta; y_T)]$$

where

$$\{W_T\}_{T=1}^{\infty}$$

is a sequence of (rxr) positive definite weighting matrices (may be a function of the data y_T). The minimization problem is non-linear and requires numerical methods.

The first order condition is

$$D_T\left(\hat{\theta}_T\right)' W_T g_T\left(\hat{\theta}_T\right) = 0$$

Where $D_T(\theta)$ is a matrix of partial derivatives given by

$$D_T(\theta) = \partial g_T(\theta) / \partial \theta$$

which converges asymptotically to D_o and $\hat{\theta}_T$ is asymptotically distributed according to

$$\sqrt{T}\left(\hat{\theta}_T - \hat{\theta}_o\right) \xrightarrow{d} N(0, V)$$

$$V = \left(D_o' W D_o\right)^{-1} D_o' W S W D_o \left(D_o' W D_o\right)^{-1}$$

$$S = \lim_{T \to \infty} \operatorname{Var}\left[T^{1/2} \sum_{t=1}^{T} f_t(\theta_o)\right] = \lim_{T \to \infty} \operatorname{Var}\left[T^{-1/2} g_T(\theta_o)\right]$$

which is the asymptotic variance matrix whose inverse times any positive scalar gives the optimal value for the weighting matrix W_T. The variance covariance matrix of the parameters is thus given by

$$V^{*} = \left(D_o' S^{-1} D_o\right)^{-1}$$

The minimized objective function times T is distributed as chi-square with $(N_f - N_\beta)$ degrees of freedom where N_f is the number of orthogonality conditions and N_β is the number of parameters to be estimated. Estimation starts with an initial value for W_T (usually the identity matrix).

To avoid problems related to the autocovariances of the function of the instruments (the number of estimated autocovariances cannot increase rapidly with the size of the sample) and to ensure that the estimator of S will be positive definite, the most common method is to use the Newey and West (1987) estimator which guarantees positive definiteness by downweighting higher-order autocovariances. The method is also consistent since this downweighting disappears asymptotically. Thus the estimator used is

$$S_T\left(q, \hat{\theta}_T\right) = \Gamma_{o,T}\left(\hat{\theta}_T\right) + \sum\left(\frac{q-j}{q}\right)\left(\Gamma_{j,T}\left(\hat{\theta}_T\right) + \Gamma_{j,T}'\left(\hat{\theta}_T\right)\right)$$

q-1 is the maximum lag length which receives a non-zero weight.

Other estimators which can be applied are those proposed by Gallant (1987), Andrews (1991), Andrews and Monahan (1992).

The non-linear regression command (NL) in SHAZAM version 7.0 has been used with the GMM = BARTLETT option for estimation and a maximum lag length of 2 (with AUTOCOV=2) for computing the weighting matrix in the Newey-West estimator. For this estimator, when BARTLETT is selected, the weights, w_j, used in the estimator are given by:

$$w_j = 1 - \frac{j}{L+1}$$

where L is the lag length and j the number of weights used (maximum=L).

Appendix B

GMM Testing Methodology

B.1 Testing Hypotheses about the set of moment conditions $E[f(x_t,\theta_0)]$

For a vector of observed variables x_t of size T and a vector of unknown parameters θ_0 (px 1) the model implies that there is a (q x 1) set of population moment conditions:

$$E[f(x_t,\theta_0)]=0$$

The null hypothesis in this case is $H_0: V^{-\frac{1}{2}}E[f(x_t,\theta_0)]$

and V is a non-singular matrix where it is also true that:

$$\frac{\partial Q_T\left(\hat{\theta}_T\right)}{\partial \theta}=F_T\left(\hat{\theta}_T\right)' V_T^{-1} f_T\left(\hat{\theta}_T\right)=0$$

where $\hat{\theta}_T$ is the value of θ which minimizes

$$Q_T(\theta)=f_T(\theta)' V_T^{-1} f_T(\theta)=0$$

$$f_T(\theta)=\frac{1}{T}\sum_{t=1}^{T} f(x_t,\theta)$$

and V_T is a consistent estimator of $V=\lim_{T\to\infty} \operatorname{var}\left[T^{1/2} f_T(\theta_0)\right]$

The null hypothesis cab be partitioned into that hypothesis pertaining to the set of the identifying and to the set of the over-identifying restrictions as follows:

$$H_0=H_0^I \cap H_0^O$$

$$H_0^I: PV^{-\frac{1}{2}}E[f(x_t,\theta_0)]=0$$

$$H_0^O: (I_q-P)V^{-\frac{1}{2}}E[f(x_t,\theta_0)]=0$$

Since $\hat{\theta}_T$ satisfies $\frac{\partial Q_T\left(\hat{\theta}_T\right)}{\partial \theta}=0$ it must also satisfy $\hat{P} V_T^{-1/2} f_T\left(\hat{\theta}_T\right)=0$ where

$$\hat{P}=\hat{M}\left(\hat{M}'\hat{M}\right)^{-1}\hat{M} \quad , \quad \hat{M}=V_T^{-1/2}F_T\left(\hat{\theta}_T\right) \text{ and } F_T(\theta)=\frac{\partial f_T(\theta)}{\partial \theta'}$$

Therefore estimation is actually based on the population analogue of

$\hat{P} V_T^{-1/2} f_T\left(\hat{\theta}_T\right) = 0$, that is $PV^{-1/2}E[f(x_t,\theta)] = 0$ and $P = M(M'M)^{-1}M'$,

$M = V_T^{-1/2}E[F_t(\theta_0)]$, $F_t(\theta) = \dfrac{\partial f(x_t,\theta)}{\partial \theta'}$.

i.e. P is a projection matrix (of rank p) which sets only p unique linear combinations on the (q x 1) vector $E[f(x_t,\theta_0)]$ equal to zero.

The first sub-hypothesis cannot be tested since the sample analogue to the identifying restrictions are satisfied by the estimated sample moments. The second however can be tested because the overidentifying restrictions are not imposed during estimation.

Hansen (1982):

$$J_T = TQ_T\left(\hat{\theta}_T\right)$$
$$J_T \overset{A}{\sim} \chi^2_{q-p} \text{ under } H_0^O$$

This is the well-known *Test of Overidentifying Restrictions.*

The local alternatives to the above null hypotheses are:

$$H_A^I : PV^{-1/2}E[f(x_t,\theta_0)] = T^{-1/2}\mu_I$$
$$H_A^O : (I_q - P)V^{-1/2}E[f(x_t,\theta_0)] = T^{-1/2}\mu_0 \qquad \text{where } \mu_I \text{ and } \mu_0 > 0$$

The main question posed is whether the J_T statistic has power against violations of the overidentifying restrictions.

Assuming that the overidentifying restrictions are satisfied by the data, i.e.

$H_0^I \cap H_A^O$ are satisfied, it is shown that $J_T \overset{A}{\sim} \chi^2_{q-p}(\mu_0'\mu_0)$ and since $\mu_0'\mu_0 > 0$, J_T has power against this alternative.

Assuming that the overidentifying restrictions are satisfied but the identifying restrictions are not, i.e.

$H_A^I \cap H_O^O$ is satisfied, then it is shown that $J_T \overset{A}{\sim} \chi^2_{q-p}$

Regardless of whether H_O^I or H_A^I holds the statistic has the same distribution under H_O^O, that is it has no power to discriminate between the latter two hypotheses. However, failure to reject H_O^O does not imply the validity of H_O because H_O^I may be violated.

B.2 Testing Hypotheses about Subsets of the moment conditions $E[f(x_t,\theta_0)]$

Local alternatives can be expressed as:

$$H_O^S : E[f_1(x_t,\theta_{01})]=0 \qquad \text{and} \qquad E[f_2(x_t,\theta_0)]=0$$

$$H_A^S : E[f_1(x_t,\theta_{01})]=0 \qquad \text{and} \qquad E[f_2(x_t,\theta_0)]=\mu_2 T^{-1/2}$$

where $\theta_0' = (\theta_{01}', \theta_{02}') \qquad \theta_{0i} : (p_i \times 1)$

$$f(x_t,\theta_0)' = [f_1(x_t,\theta_{01})', f_2(x_t,\theta_0)'] \qquad f_i(\bullet):(q_i \times 1)$$

Eichenbaum, Hansen and Singleton (1988) proposed a statistic, C_T which is based on two estimations, the first including the full set of moment conditions and the second only those population moment conditions which are held to be true under both H_0^S and H_A^S

$$C_T = T\left\{Q_T\left(\hat{\theta}_T\right) - Q_{1T}\left(\tilde{\theta}_{1T}\right)\right\} \qquad \text{where } \tilde{\theta}_{1T} : \text{value of } \theta_1 \text{ which minimizes } Q_{1T}$$

$$\text{and } Q_{1T} = f_{T1}(\theta_1)' \hat{V}_{11}^{-1} f_{T1}(\theta_1)$$

$C_T \sim \chi^2_{v_1}$ under H_0^S $\quad v_1 = q_2 - p_2$ (degree to which θ_{01} is overidentified by $E(f_2(x_t, \theta_0)) = 0$ if θ_{01} is known).

This statistic offers a method for the detection of $E[f_2(x_t, \theta_0)] \neq 0$ via its impact on the overidentifying restrictions.

Under certain conditions it is possible to develop complementary statistics which are sensitive to violations in the identifying restrictions.

If $p_2 = 0$ and $q_1 \geq p$ we can use $E[f_1(x_t, \theta_0)] = 0$ to obtain an alternative estimator of θ_0, θ_T which, by construction, does not satisfy the identifying restrictions. This test is sensitive to alternatives which satisfy both H_A^S and H_A^I. A convenient Hausman (1978) type statistic, expressed in terms of the parameter estimators of two separate estimations is:

$$H_T = T\left(\hat{\theta}_T - \tilde{\theta}_T\right) V_h^- \left(\hat{\theta}_T - \tilde{\theta}_T\right)$$

where $\hat{V}_h$ is a consistent estimator of V_h which results from partitioning the sample moments.

This statistic allows testing H_0^{SI}. Newey (1985a) has shown that it converges to a $\chi^2_{\nu 2}$ under H_0^S where ν_2 is equat to the rank of V_h.

All of the above statistics are conditioned on the validity of $E[f_1(x_t, \theta_0)] = 0$

If this does not hold then the statistics can be significant even though $E[f_2(x_t, \theta_0)]$ is actually zero.

B.3 Testing Hypotheses about the Parameter Vector

The general framework for testing these types of hypotheses is:

$H_0^R : r(\theta_0) = 0$ versus $H_A^R : r(\theta_0) = T^{-1/2}\mu_R$ where $r(\bullet)$ is a $(s \times 1)$ vector of continuous differentiable functions of the form $R(\theta) = \partial r(\theta)/\partial\theta'$

Assuming that the unrestricted estimator is denoted by $\hat{\theta}_T$ and the restricted by $\bar{\theta}_T$ the three statistics are given by :

Wald Test: $$W_T = Tr\left(\hat{\theta}_T\right)'\left[R\left(\hat{\theta}_T\right)\left(\hat{M}'\hat{M}\right)'R\left(\hat{\theta}_T\right)'\right]^{-1} r\left(\hat{\theta}_T\right)$$

LR Test: $$LR_T = T\left[Q_T\left(\bar{\theta}_T\right) - Q_T\left(\hat{\theta}_T\right)\right]$$ and

LM Test: $$LM_T = Tf_T\left(\bar{\theta}_T\right)' V_T^{-1/2'}\bar{P}V_T^{-1/2}f_T\left(\bar{\theta}_T\right)$$ where $\bar{P} = \bar{M}\left(\bar{M}'\bar{M}\right)^{-1}\bar{M}'$ and $\bar{M} = V_T^{-1/2}F_T\left(\bar{\theta}_T\right)$

The number of restrictions, s, should not exceed the number of parameters and these restrictions must form a coherent set of equations in such a way so that given the value of p-s elements of θ_0, it should be possible to solve uniquely for the remaining s values using the null hypothesis.

Newey and West (1987) have proposed three main statistics which are extensions of the Wald, Likelihood Ratio (LR) and Lagrange Multiplier (LM) tests to the GMM framework.

All three statistics were shown to be asymptotically equivalent under both the null and the alternative hypothesis. Under the null, the statistics converge to a chi-square distribution with s degrees of freedom.

LR is the most difficult statistic to calculate since it requires both the restricted and the unrestricted estimation to be carried out. However the Wald statistic has two serious disadvantages, first that it is invariant to a reparametrization of the model or the restrictions and second that it tends to be less well approximated by the chi-square in finite samples.

Sowell (1996a) has shown that the above null hypothesis can be rejected either because the identifying restrictions are satisfied at θ_0 but $r(\theta_0) \neq 0$

or when $r(\theta_0) = 0$ because the identifying restrictions are not satisfied at the true parameter vector θ_0. The two main issues which arise from the above discussion are first that specification tests must precede inferences about the parameters and second that the trio of the tests are functions of the identifying restrictions and are thus asymptotically independent of the previous statistics which are based on the overidentifying restrictions under the null or any other local alternatives.

B.4 Testing Hypotheses about Structural Stability

Structural stability may occur even when the expected value of the moment conditions is the same at all t. On the contrary, structural instability occurs when the expected value of the moment conditions is not equal to zero for only part of the sample.

Assume that there exists a breakpoint πT where some aspect of the model changes. This point may be known or unknown but in the latter case the null hypothesis is a much broader one since it postulates that there is no instability at any point. The two sub-samples thus formed are:

$$T_1 = \{1,2,\ldots,[\pi T]\} \text{ and } T_2 = \{[\pi T]+1,\ldots,T\}$$

B.4.1 Andrews and Fair (1988) tests for parameter variation

To allow for the possibility that the usual moment conditions are satisfied at different parameter values before and after the breakpoint, they introduced an augmented population moment condition:

$$E[g(x_t,\phi_0)] = \begin{bmatrix} d_t(\pi)f(x_t,\theta_1) \\ (1-d_t(\pi))f(x_t,\theta_2) \end{bmatrix} = 0$$ where $d_t(\pi)$ is a dummy variable equal to 1 for $t \le \pi T$ and $\phi_0' = (\theta_1', \theta_2')$

The procedure followed is :

1) Estimate ϕ_0 using GMM
2) Calculate W, LR or LM for the hypothesis : $[I_p, -I_p]\phi_0 = 0_p$
3) All of the above statistics have a limiting χ^2_p distribution if the restriction holds.

B.4.2 Ghysels and Hall Test

They test :

$H_0 : E[f(x_t, \theta_0)] = 0 \quad \text{for } t \in T_1 \cap T_2$ against the local alternative

$H_A^1 : E[f(x_t, \theta_0)] = 0 \quad \text{for } t \in T_1$ and

$E[f(x_t, \theta_0)] = T^{-1/2}\mu_2 \quad \text{for } t \in T_2 \quad \text{where } \mu_2 \neq 0$

i.e. they evaluate the sample moments from T_2 at $\hat{\theta}_1(\pi)$

Therefore under the null these moments should (in probability) converge to zero.

The *Predictive Test Statistic* they derived is:

$$PR_T(\pi) = T^{-1/2} \sum_{t \in T_2} f\left(x_t, \theta(\hat{\pi})\right)' \hat{V}_{PR}^{-1} T^{-1/2} \sum_{t \in T_2} f\left(x_t, \theta(\hat{\pi})\right)$$

$\hat{V}_{PR}$ is a covariance matrix defined in Ghysels and Hall (1990b).

The above statistic converges to a χ_q^2 under the null. Since this limiting distribution has more degrees of freedom than the Wald statistic, the Predictive Statistic clearly tests something more than parameter variation, as is explained below.

To understand the difference between the above two types of tests it is useful to restate the hypotheses distinguishing between the identifying and the over-identifying restrictions. Since the latter are imposed in estimation some parameter values must satisfy them in both subsamples and therefore these restrictions will be structurally stable if they are satisfied by the same parameter values in both. Three types of null hypotheses can thus be stated. For the identifying restrictions these are:

$$H_0^I(\pi) = PV^{-\frac{1}{2}}E[f(x_t,\theta_0)] = 0 \quad \text{for} \quad t \in T_1$$

$$PV^{-\frac{1}{2}}E[f(x_t,\theta_0)] = 0 \quad \text{for} \quad t \in T_2$$

For the overidentifying restrictions structural stability implies that they hold before and after the sample, namely:

$$H_0^{O1}(\pi) = [I_q - P_1]V^{-\frac{1}{2}}E[f(x_t,\theta_1)] = 0 \quad \text{for} \quad t \in T_1 \quad \text{and}$$

$$H_0^{O2}(\pi) = [I_q - P_2]V^{-\frac{1}{2}}E[f(x_t,\theta_2)] = 0 \quad \text{for} \quad t \in T_2$$

where $P_i = M_i(M_i'M_i)^{-1}M_i'$ and $M_i = V_i^{-1/2}E[F_t(\theta_i)]$

V_i is defined on T_i for $i = 1,2$.

Therefore the two null hypotheses for the overidentifying restrictions imply that they are satisfied at different values.

For instability to occur any one of the three null hypotheses must be violated. The null hypothesis for the identifying restrictions is equivalent to that of no parameter variation and therefore the trio of test statistics (Wald, LR and LM) can be used to test it. However, since the overidentifying and the identifying restrictions are orthogonal, these statistics have no power against any local alternatives on the null hypotheses for the overidentifying restrictions but Ghysels, Guay, and Hall (1997) have shown that their Predictive Test has such power (hence the increased degrees of freedom).

To overcome the weaknesses of the above tests, Hall and Sen (1996a) have proposed testing the two types of restrictions separately. To test the stability of the

$O_T(\pi) = O1_T(\pi) + O2_T(\pi)$ where $O1_T(\pi)$ and $O2_T(\pi)$ are the overidentifying restrictions tests based on the sub-samples T_1 and T_2 respectively.

overidentifying restrictions one can use:

Hall and Sen (1996a) have shown that under the null hypothesis $O_T(\pi)$ converges to a $\chi^2_{2(q-p)}$. As they have shown, a Wald (or any of the trio) test has power against local alternatives to $H_0^I(\pi)$, denoted as $H_A^I(\pi)$ but no power against local alternatives to $H_0^O(\pi)$, denoted as $H_A^O(\pi)$. However, $O_T(\pi)$ has power against $H_A^O(\pi)$ but none against $H_A^I(\pi)$. Since these two statistics are asymptotically independent they can be used together to diagnose the source of the instability.

B.4.3 Unknown Breakpoint

When the position of the breakpoint is not known the composite null hypothesis to be tested is: $H_0^I(\pi)$ for all $\pi \in \Pi \subset (0,1)$

Suppose we use the Wald statistic, then we calculate it for each possible π to produce a sequence of statistics and inference is based on some function of this sequence which is chosen is such a way so that the power of the test is maximized against a local alternative in which a weighting distribution is used to indicate the relative importance of departures from $H_0^I(\pi)$ in different directions at different breakpoints.

Andrews and Ploberger (1994) have presented a general framework for the derivation of these optimal tests which has been generalized to GMM by Sowell (1996a). The local alternatives to $H_0^I(\pi)$ to be tested are:

$$H_A^I(\pi): PV^{-1/2}E[f(x_t,\theta_0)] = T^{-1/2}\mu_{11} \quad \text{for } t \in T_1 \quad \mu_{11} = 0$$

$$PV^{-1/2}E[f(x_t,\theta_0)] = T^{-1/2}\mu_{12} \quad \text{for } t \in T_2$$

Andrews (1993), Andrews and Ploberger (1994) and Sowell (1996a) have proposed statistics for testing the above alternatives and based on the same principles Sowell (1996b) has derived optimal tests for testing the null hypothesis for the overidentifying restrictions, i.e. that $H_0^O(\pi)$ holds for all $\pi \in \Pi$. His alternative hypothesis is that the overidentifying restrictions are violated before the breakpoint i.e. $H_A^{O1}(\pi) \cap H_0^{O2}(\pi)$. Hall and Sen (1996b) have shown that his statistics are suboptimal against an alternative where the overidentifying restrictions are violated after, instead of before, the breakpoint, i.e . $H_0^{O1}(\pi) \cap H_A^{O2}(\pi)$ However, the more general alternative, where the overidentifying restrictions can be violated either before or after the breakpoint is more interesting, i.e.

$$\{H_A^{O1}(\pi) \cap H_0^{O2}(\pi)\} \cup \{H_0^{O1}(\pi) \cap H_A^{O2}(\pi)\}$$

Hall and Sen (1996a) have derived statistics to test the above alternative.

Sowell's (1996b) test, apart from being optimal for local alternatives that imply a one-time structural jump in the moment conditions, it is also consistent for any testable alternative. These tests are independent of his optimal tests for parameter instability presented in Sowell (1996a) and (1994). He presented tests which simultaneously test for the stability of both types of restrictions but he further expanded these by separating them into statistics appropriate for each type of restriction separately, as described below:

Optimal test for alternatives with one time jumps of the *identifying* restrictions when the level of the jump is extremely large:

$$E_A = \frac{1}{T}\sum_{t=1}^{T}\exp\left\{\frac{1}{2}TF_t\left(\hat{\theta}_T\right)' W_T\hat{M}\left(\hat{M}' W_T\hat{M}\right)^{-1}\hat{M}' W_T F_t\left(\hat{\theta}_T\right)\right\}$$

Optimal test for alternatives with one time jumps of the *overidentifying* restrictions when the level of the jump is extremely large:

$$E_B = \frac{1}{T}\sum_{t=1}^{T}\exp\left\{\frac{1}{2}TF_t\left(\hat{\theta}_T\right)'\left(W_T - W_T\hat{M}\left(\hat{M}' W_T\hat{M}\right)^{-1}\hat{M}' W_T\right)F_t\left(\hat{\theta}_T\right)\right\}$$

Optimal test for alternatives with one time jumps of the *identifying* restrictions when the level of the jump is extremely small:

$$L_A = \sum_{t=1}^{T}F_t\left(\hat{\theta}_T\right)' W_T\hat{M}\left(\hat{M}' W_T\hat{M}\right)^{-1}\hat{M}' W_T F_t\left(\hat{\theta}_T\right)$$

Optimal test for alternatives with one time jumps of the *overidentifying* restrictions when the level of the jump is extremely small:

$$L_B = \sum_{t=1}^{T} F_t\left(\hat{\theta}_T\right)'\left(W_T - W_T\hat{M}\left(\hat{M}'W_T\hat{M}\right)^{-1}\hat{M}'W_T\right)F_t\left(\hat{\theta}_T\right)$$

Sowell (1996b) also provides critical values for the distributions of E_A, $\log(E_B)$, L_A and L_B. Apart from their optimality properties these statistics only require the single standard GMM optimization to be carried out and are easily computed from terms typically calculated during GMM estimation.

Appendix C

Data-Selection Procedure

C.1 Procedure for Selecting Data

Chapter 4 describes the methodology used to for processing raw data in such a way as to obtain appropriate inputs for the GMM estimation of the models. However technical details have been omitted in order to make it easier to read the text.

The detailed technical procedure for selecting data for the GMM estimation of the models is similar for both the NYSE and the NASDAQ deciles and is described below in detail while an outline is given in table AC.1. The following steps have been carried out so as to select data which are as free from errors as possible:

Step 1

The first step is to select from the master file the symbols of all common stocks traded in NYSE and all common stocks traded in NASDAQ as a primary market (see appendix E for details of the codes used). This is accomplished by running program MAST of the TAQ package which retrieves user-selected symbols by security type and / or exchange from the master file of the particular month. Extracted symbols are stored in file MAST9410.IN

Step 2

The file containing the extracted symbols in step 1 is renamed as SELECT.IN to be used to extract more detailed data in step 3.

Step 3

Symbols in file SELECT.IN are used to extract more detailed data for the above stocks from the MAST9410.IN file using program TAQSEL.PAS which has been written for this purpose. Detailed data for each stock include the symbol, the number of issued and outstanding shares, the tick size, and the number of shares in a round lot (UOT number). These details are stored in file NEWSYMB.DAT whereas symbols only are stored in file SELSYMB.DAT for further use. The TAQSEL.PAS program also attempts to filter errors at this stage by checking whether the particular stock is actually traded in the particular exchange and also ensures that the number of outstanding shares is not equal to zero.

Step 4

File SELSYMB.DAT is renamed as SELECT.IN so as to be used to extract trade and quote data from the Consolidated Trade and the Consolidated Quote binary files of the TAQ database using the TAQ program SELECT. The extracted trade and quote data are automatically stored in files SELECT.T and SELECT.Q respectively.

Step 5

Program TAQSTAT2.PAS, written specifically for this purpose, is used to classify symbols in file SELECT.IN by trading activity, based on the trading volume of each stock in the particular month examined, which is extracted and calculated from file SELECT.T. Classified stocks are then grouped into deciles, using a routine ('heapsort') which sorts stocks by descending volume and then 20 stocks are selected randomly (using a random selection routine) from each NYSE decile (60 from each NASDAQ decile so as to obtain an adequate number of observations). Selected stocks for NYSE are then stored in file NYRAND.DAT (NASRAND.DAT for NASDAQ) and detailed data for NYSE are stored in file RANDSTK.DAT (RANDNAS.DAT for NASDAQ). These detailed data for each selected stock include the symbol, the volume as number of shares traded in the month, the turnover as total value in USD for the month and the decile which the particular stock has been grouped into.

Step 6

File NYRAND.DAT (NASRAND.DAT for NASDAQ) is renamed to SELECT.IN and is used again to extract final trade and quote data from the respective Consolidated binary files by running the TAQ program SELECT. Trade and quote data are stored in files SELECT.T and SELECT.Q and are renamed as NYALL.T and NYALL.Q for NYSE (NASALL.T and NASALL.Q for NASDAQ) to be used in later steps.

Step 7

For each of the files NYALL.T and NYALL.Q for NYSE (NASALL.T and NASALL.Q for NASDAQ) the following procedure is used to sort symbols alphabetically using program SortSymbols which has been written specifically for this purpose:

1) For each calendar day trades and quotes are sorted alphabetically and are output in files NYALLN1.T and NYALLN1.Q respectively for NYSE (NASALLN1.T and NASALLN1.Q for NASDAQ).

2) For each symbol in a particular day trades and quotes are sorted by time and are output in files NYALLN.T and NYALLN.Q respectively for NYSE (NASALLN.T and NASALLN.Q for NASDAQ).

The above procedure is necessary so as to avoid errors during matching of trades with quotes which may occur due to the specific structure of the TAQ trade and quote files.

Further errors are filtered using the specially written programs KILLTRAD.PAS and KILLQUOT.PAS for trades and quotes respectively and results are output in files NYALLN2.T and NYALLN2.Q respectively for NYSE (NASALLN2.T and NASALLN2.Q for NASDAQ). The above filtering programs ensure that certain critical fields like the price, size of trade, bid price, ask price and sizes of quotes are not less or equal to zero and also checks certain condition and correction codes which are described in Chapter 4 and tabulated in Appendix E.

Step 8

Trade and quote data from step 7 are read and matched using the specially written program NY20GMM.PAS for NYSE (NAS20GMM.PAS for NASDAQ) and are output in files NYMOD1.DAT and NYMOD2.DAT for the Quoted Spread Model and the Price-Change Model respectively for NYSE (NASMOD1.DAT and NASMOD2.DAT for NASDAQ). These files are later used for the GMM estimation of the models by the SHAZAM package after their data are split into deciles of trading activity.

Step 9

For each exchange-related subsample the specially written program MAKEDECS.PAS is used to split the data in the above files of matched trades and quotes into deciles of trading activity and the results are output separately for each decile and for each model to be used for GMM estimation by the SHAZAM package.

Summary Statistics

Summary statistics for the trade and quote data of each sub-sample are extracted using programs NYSUMT.PAS and NYSUMQ.PAS respectively for NYSE (NASUMT.PAS and NASUMQ.PAS for NASDAQ). These programs have been written specifically for this purpose.

TABLE AC.1 : PROCEDURE FOR SELECTING DATA

STEP	NYSE	NASDAQ
1	Run : **MAST N C.** Symbols are stored in *MASTyymm.IN.*	Run : **MAST T C.** Symbols are stored in *MASTyymm.IN.*
2	Rename *MASTyymm.IN* as *SELECT.IN*	Rename *MASTyymm.IN* as *SELECT.IN*
3	Program *TAQSEL.PAS:* Use *SELECT.IN* as input to *TAQSEL.PAS* to select data which are saved in two files: *NEWSYMB.DAT* (symbol / shares / tick / UOT) and *SELSYMB.DAT* (contains symbols only).	Program *TAQSEL.PAS:* Use *SELECT.IN* as input to *TAQSEL.PAS* to select data which are saved in two files: *NEWSYMB.DAT* (symbol / shares / tick / UOT) and *SELSYMB.DAT* (contains symbols only).
4	Rename *SELSYMB.DAT* to *SELECT.IN* and use it to extract trade and quote data by running **SELECT**. Quote and trade data is then stored in *SELECT.Q* and *SELECT.T.*	Rename *SELSYMB.DAT* to *SELECT.IN* and use it to extract trade and quote data by running **SELECT**. Quote and trade data is then stored in *SELECT.Q* and *SELECT.T.*
5	Program *TAQSTAT2.PAS:* Classify stocks in *SELECT.IN* by trading activity. Based on trading volume classify stocks by trading activity, group into deciles and select 20 stocks from each group. Output stock symbols into file *NYRAND.DAT* and detailed data (symbol / volume / turnover / decile) in file *RANDSTK.DAT.*	Program *TAQSTAT2.PAS:* Classify stocks in *SELECT.IN* by trading activity. Based on trading volume classify stocks by trading activity, group into deciles and select 20 stocks from each group. Output stock symbols into file *NASRAND.DAT* and detailed data (symbol / volume / turnover / decile) in file *RANDNAS.DAT.*
6	Rename *NYRAND.DAT* to *SELECT.IN* and use it to extract final trade and quote data by running **SELECT**. Rename *SELECT.T* and *SELECT.Q* as *NYALL.T* and *NYALL.Q* for future use.	Rename *NASRAND.DAT* to *SELECT.IN* and use it to extract final trade and quote data by running **SELECT**. Rename *SELECT.T* and *SELECT.Q* as *NASALL.T* and *NASALL.Q* for future use.
7	For both *NYALL.T* and *NYALL.Q* (one at a time) follow the procedure below using program ***SortSymbols*** in Visual Basic: 1) Sort symbols alphabetically each day and output in file *NYALLN1.T(Q)* 2) Sort symbols by time each day and output in file *NYALLN.T(Q)* Using program ***KILLTRAD.PAS (KILLQUOT.PAS)*** remove quotation marks at beginning of line and filter; then output in file *NYALLN2.T(Q)*	For both *NASALL.T* and *NASALL.Q* (one at a time) follow the procedure below using program ***SortSymbols*** in Visual Basic: 1) Sort symbols alphabetically each day and output in file *NASALLN1.T(Q)* 2) Sort symbols by time each day and output in file *NASALLN.T(Q)* Using program ***KILLTRAD.PAS (KILLQUOT.PAS)*** remove quotation marks at beginning of line and filter; then output in file *NASALLN2.T(Q)*
8	Read trade and quote data using ***NY20GMM.PAS*** to prepare data for GMM estimation of the B/A spread model. Output results in *NYMOD1.DAT* and *NYMOD2.DAT*	Read trade and quote data using ***NAS20GMM.PAS*** to prepare data for GMM estimation of the B/A spread model. Output results in *NASMOD1.DAT* and *NASMOD2.DAT*
9	Using Program **MAKEDECS.PAS** split data in *NYMOD1(2).DAT* into deciles and output data in directory C:\taq\phdata\nyse\model(1/2)	Using Program **MAKEDECS.PAS** split data in *NASMOD1(2).DAT* into deciles and output data in directory C:\taq\phdata\nasdaq\model(1/2)
10	Use programs *NYSUMT.PAS* and *NYSUMQ.PAS* to extract summary statistics for the trade and quote data of the NYSE sub-sample.	Use programs *NASUMT.PAS* and *NASUMQ.PAS* to extract summary statistics for the trade and quote data of the NASDAQ sub-sample.

Appendix D

Tables and Graphs of Summary Statistics

Tables AD.4.1, AD.4.2, AD.4.3 and AD.4.4

Summary Statistics of the Trades and Quotes of the Random Samples of NYSE and NASDAQ Common Stocks, TAQ data: October 1994.

Tables AD.4.1 and AD.4.3 present Summary Statistics of the Trades and Quotes respectively for the sample of 200 NYSE Common Stocks and Tables AD.4.2 and AD.4.4 present similar statistics of the Trades and Quotes respectively for the sample of 600 NASDAQ Common Stocks. Stocks for each sample have been randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares of each stock traded in October 1994.

For the variables: Daily Volume of Shares traded, Daily Turnover, Daily Number of Trades, Daily Volume of Ask and Bid and the Daily Number of Quotes the averages are given by:

$$\text{Average}(X) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}}{n}$$

where n is the number of trading days in the month (equal to 21 for our samples)
m is the number of trades of all stocks in a decile, and
X is the variable for which the average is calculated

The maximum and the minimum for each variable are given by :

$$\text{Maximum}(X) = \underset{i=1}{\overset{n}{\text{Max}}}\left(\underset{j=1}{\overset{m}{\text{Max}}}\left(X_{ij}\right)\right)$$

$$\text{Minimum}(X) = \underset{i=1}{\overset{n}{\text{Min}}}\left(\underset{j=1}{\overset{m}{\text{Min}}}\left(X_{ij}\right)\right)$$

where Max and Min denote the maximum and Minimum of a set of values respectively.

For the other variables, namely: the Change in Price of transaction, Change in Ask and Bid Prices, the Time between Trades, Time between Quotes, Change in Spread, Change in Proportional Spread, the Spread and the Proportional Spread the average the maximum and the minimum are given by:

$$\text{Average}(X) = \frac{\sum_{i=1}^{m} X_i}{m-21}$$

and

$$\text{Maximum}(X) = \underset{i=1}{\overset{m}{\text{Max}}}\left(X_i\right)$$

$$\text{Minimum}(X) = \underset{i=1}{\overset{m}{\text{Min}}}\left(X_i\right)$$

The Averages, Maximum and Minimum values of the total sample have been calculated as the average, maximum and minimum values of the individual deciles respectively, that is:

$$\text{Average}(X_{Total}) = \underset{k=1}{\overset{10}{\text{Avg}}}\left(\text{Avg}(X_k)\right)$$

$$\text{Maximum}\left(X_{\text{Total}}\right) = \underset{k=1}{\overset{10}{\text{Max}}}\left(\text{Max}\left(X_k\right)\right)$$

$$\text{Minimum}\left(X_{\text{Total}}\right) = \underset{k=1}{\overset{10}{\text{Min}}}\left(\text{Min}\left(X_k\right)\right)$$

where

Avg denotes the average of a set of values and

k is the number of the decile the prorerty of which is calculated by the above formula.

Table AD.4.1 Summary Statistics of the Trades of the Random Sample of Common Stocks, NYSE TAQ data: October 1994.

DECILE	1	2	3	4	5	6	7	8	9	10	Total Sample
Daily Volume (no of Shares Traded)											
Average	70,024	167,700	291,295	451,462	668,352	963,767	1,546,576	2,865,152	4,924,243	16,006,690	2,795,526
Maximum	140,400	316,800	521,000	792,800	1,044,200	1,648,000	2,516,700	5,476,000	7,666,600	27,513,200	27,513,200
Minimum	38,800	130,900	215,000	284,900	461,400	732,900	939,900	1,871,900	2,713,100	10,757,100	38,800
Daily Turnover ('000 of USD)											
Average	992.92	2,409.98	6,592.22	9,191.20	17,609.72	18,255.52	43,717.29	103,189.53	181,263.21	653,194.54	103,641.61
Maximum	2,701.46	4,441.14	11,818.39	16,377.88	29,252.05	29,013.01	69,595.68	171,487.01	275,793.01	1,037,333.99	1,037,333.99
Minimum	657.71	1,755.53	4,601.66	5,241.95	11,259.78	11,527.26	26,100.66	68,410.23	99,907.28	443,161.94	657.71
Change in Price (USD)											
Average	0.000363	0.000545	0.000292	-0.001634	-0.001586	-0.000411	-0.000651	-0.000246	-0.000051	-0.000010	-0.000339
Maximum	0.500000	0.500000	0.500000	0.875000	1.000000	2.000000	0.875000	1.000000	1.000000	1.625000	2.000000
Minimum	-0.500000	-0.375000	-0.750000	-1.250000	-1.000000	-2.000000	-0.875000	-1.250000	-0.875000	-1.375000	-2.000000
Time between Trades (seconds)											
Average	2,526	1,710	1,270	1,187	677	671	393	248	164	59	891
Maximum	23,402	22,127	20,893	22,020	17,650	17,997	11,904	17,557	7,366	9,708	23,402
Minimum	0	0	0	0	0	0	0	0	0	0	0
Daily Number of Trades											
Average	83	187	286	289	600	602	1,007	1,716	2,574	7,240	1,458
Maximum	120	288	423	428	842	793	1,391	2,292	3,450	9,408	9,408
Minimum	73	138	232	255	540	543	761	1,367	1,845	6,048	73

Table AD.4.2 Summary Statistics of the Trades of the Random Sample of Common Stocks, NASDAQ TAQ data: October 1994.

DECILE	1	2	3	4	5	6	7	8	9	10	Total Sample
Daily Volume (no of Shares Traded)											
Average	84,610	173,429	294,743	486,557	754,014	1,086,505	1,629,643	2,681,610	4,982,548	20,686,719	3,286,038
Maximum	133,800	267,800	533,900	836,400	1,200,400	1,656,400	2,445,600	4,075,200	7,214,900	33,011,300	33,011,300
Minimum	57,000	114,300	175,500	324,800	652,200	699,700	895,900	2,020,700	3,684,900	14,846,600	57,000
Daily Turnover ('000 of USD)											
Average	787.67	1,638.57	2,516.87	3,281.30	5,627.79	8,344.12	17,589.71	33,345.31	70,487.88	614,411.54	75,803.08
Maximum	1,721.49	3,380.58	4,768.05	5,153.83	8,684.98	15,982.17	31,840.90	51,919.69	114,653.13	1,219,151.38	1,219,151.38
Minimum	294.57	796.48	1,324.88	2,422.12	4,606.03	5,385.83	9,699.31	23,964.63	51,209.30	448,424.86	294.57
Change in Price (USD)											
Average	0.007863	0.004487	0.003054	0.001866	0.001400	-0.000970	-0.001679	-0.000232	-0.000253	0.000106	0.001564
Maximum	2.000000	1.000000	1.250000	1.500000	1.500000	1.125000	3.062500	1.000000	1.000000	4.250000	4.250000
Minimum	-2.750000	-2.000000	-1.250000	-1.750000	-1.750000	-1.250000	-3.250000	-1.000000	-1.250000	-4.250000	-4.250000
Time between Trades (seconds)											
Average	2,368	2,438	2,210	1,990	1,669	1,543	1,129	736	502	107	1,469
Maximum	23,264	22,394	23,200	22,834	24,361	23,451	23,210	23,340	19,486	16,399	24,361
Minimum	0	0	0	0	0	0	0	0	0	0	0
Daily Number of Trades											
Average	75	136	195	293	441	521	814	1,446	2,287	11,250	1,746
Maximum	104	181	264	406	713	754	1,061	2,199	3,864	16,804	16,804
Minimum	59	96	149	242	366	393	577	1,317	1,840	8,262	59

Table AD.4.3 Summary Statistics of the Quotes of the Random Sample of Common Stocks, NYSE TAQ data: October 1994.

PART A

DECILE	1	2	3	4	5	6	7	8	9	10	Total Sample
Daily Volume of Offer (No of Shares)											
Average	288,114	655,305	885,676	2,057,819	2,061,114	4,015,067	4,581,538	10,024,743	17,383,129	54,816,929	9,676,943
Maximum	512,800	1,263,900	1,423,900	3,475,300	3,271,500	6,155,600	6,868,600	15,811,000	25,205,300	75,681,900	75,681,900
Minimum	174,900	394,700	753,400	1,438,600	1,566,700	2,373,100	2,790,900	8,267,500	12,297,100	42,236,100	174,900
Daily Volume of Bid (No of Shares)											
Average	279,457	734,314	818,467	1,776,029	2,050,890	4,046,610	4,819,095	10,935,557	16,139,862	52,931,852	9,453,213
Maximum	495,000	1,221,600	1,371,700	2,891,600	2,905,000	6,202,800	7,814,900	20,477,200	26,816,200	74,388,800	74,388,800
Minimum	203,000	489,200	636,100	1,385,200	1,579,200	3,186,300	2,979,400	8,323,400	9,419,500	44,790,500	203,000
Change in Bid Price (USD)											
Average	-0.000946	0.000072	-0.000604	-0.000722	-0.001058	-0.000292	0.000554	0.000627	-0.000104	0.000551	-0.000192
Maximum	0.625000	1.375000	2.125000	1.375000	4.000000	3.250000	4.500000	9.125000	10.500000	30.500000	30.500000
Minimum	-0.875000	-0.625000	-2.250000	-1.125000	-4.000000	-3.375000	-4.125000	-9.125000	-10.500000	-17.750000	-17.750000
Change in Offer Price (USD)											
Average	-0.000260	0.000100	0.000255	-0.000164	-0.000160	-0.000541	-0.001076	-0.000735	0.000084	-0.000144	-0.000264
Maximum	1.375000	1.875000	1.125000	6.250000	1.500000	1.375000	6.875000	6.625000	6.750000	17.500000	17.500000
Minimum	-1.750000	-2.000000	-1.250000	-5.625000	-2.250000	-8.125000	-6.875000	-6.875000	-6.750000	-17.500000	-17.500000
Time between Quotes (secs)											
Average	1,244	710	408	489	219	305	143	97	71	41	373
Maximum	23,755	23,794	23,372	23,843	17,332	19,911	13,262	22,775	7,221	14,647	23,843
Minimum	0	0	0	0	0	0	0	0	0	0	0
Daily Number of Quotes											
Average	321	625	1,070	874	2,130	1,510	3,240	5,300	7,080	12,300	3,445
Maximum	493	968	1,449	1,182	2,925	2,077	4,562	7,826	10,580	17,392	17,392
Minimum	270	521	911	772	1,873	1,217	2,659	4,071	5,364	9,431	270

Table AD.4.3 - PART B

DECILE	1	2	3	4	5	6	7	8	9	10	Total Sample
Change in Spread (USD)											
Average	0.000687	0.000029	0.000859	0.000559	0.000898	-0.000249	-0.001630	-0.001361	0.000188	-0.000578	-0.0000
Maximum	1.750000	1.750000	2.250000	6.250000	4.000000	4.750000	7.500000	9.875000	13.500000	35.250000	35.2500
Minimum	-2.000000	-2.250000	-2.250000	-5.375000	-4.000000	-8.250000	-7.250000	-9.875000	-13.375000	-35.500000	-35.5000
Change in Proportional Spread (USD)											
Average	0.0195	-0.0187	0.0029	0.0008	0.0043	-0.0089	-0.0053	-0.0035	0.0015	-0.0022	-0.00
Maximum	41.3793	46.5839	16.6464	12.7815	18.1774	21.1584	42.4028	28.6004	41.3793	41.9322	46.58
Minimum	-46.4368	-46.5839	-16.4729	-10.9302	-18.1774	-20.6222	-24.0283	-28.6004	-40.9925	-41.7812	-46.58
Spread (USD)											
Average	0.4634	0.4491	0.4343	0.4219	0.4064	0.3748	0.4026	0.4288	0.3515	0.3149	0.40
Maximum	2.5000	2.7500	2.6250	7.1250	4.1250	8.3750	7.8750	10.0000	14.0000	35.8750	35.87
Minimum	0.1250	0.0625	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.06
Proportional Spread (USD)											
Average	4.45	4.87	1.88	2.05	1.70	2.35	1.60	1.26	1.08	0.84	2.
Maximum	53.33	60.87	19.18	14.67	18.70	27.03	43.11	34.48	42.91	43.41	60.
Minimum	0.43	0.16	0.24	0.26	0.22	0.27	0.17	0.10	0.15	0.13	0.

Table AD.4.4 Summary Statistics of the Quotes of the Random Sample of Common Stocks, NASDAQ TAQ data: October 1994.

PART A

DECILE	1	2	3	4	5	6	7	8	9	10	Total Sample
Daily Volume of Offer (No of Shares)											
Average	64,833	91,062	94,076	110,262	155,167	146,038	202,043	252,752	325,457	849,562	229,125
Maximum	106,200	152,500	168,000	177,200	228,000	215,700	308,400	400,600	526,000	1,312,700	1,312,700
Minimum	57,600	77,500	80,000	90,200	121,500	124,400	163,000	205,400	238,000	612,400	57,600
Daily Volume of Bid (No of Shares)											
Average	65,200	89,981	95,657	113,186	159,619	150,476	205,281	250,376	328,424	854,995	231,320
Maximum	107,600	150,000	167,500	177,900	250,000	218,200	307,700	402,500	530,000	1,309,200	1,309,200
Minimum	58,200	76,500	81,500	92,400	127,500	126,100	167,200	202,800	241,000	618,000	58,200
Change in Bid Price (USD)											
Average	-0.006163	-0.002366	-0.003643	-0.000643	0.003308	-0.005167	-0.001197	0.002780	0.000771	0.003212	-0.000911
Maximum	0.625000	0.500000	0.750000	0.750000	0.750000	0.500000	0.500000	1.000000	0.500000	1.625000	1.625000
Minimum	-1.000000	-1.125000	-0.750000	-0.750000	-0.750000	-1.000000	-0.500000	-0.500000	-0.750000	-1.625000	-1.625000
Change in Offer Price (USD)											
Average	-0.004371	-0.002853	-0.002747	-0.001630	0.003412	-0.004503	-0.001904	0.002445	0.000912	0.003223	-0.000802
Maximum	0.500000	0.500000	0.625000	0.750000	0.750000	0.750000	0.750000	1.000000	0.750000	1.000000	1.000000
Minimum	-0.500000	-0.750000	-0.500000	-0.750000	-0.750000	-1.500000	-0.750000	-0.500000	-0.500000	-1.000000	-1.500000
Time between Quotes (secs)											
Average	3,118	2,888	3,100	3,134	2,960	3,228	2,841	2,615	2,484	1,168	2,754
Maximum	27,453	27,266	27,755	27,638	27,522	27,479	27,373	27,756	27,716	27,383	27,756
Minimum	0	0	0	0	0	0	0	0	0	0	0
Daily Number of Quotes											
Average	89	111	107	135	187	175	226	271	337	864	250
Maximum	146	184	189	216	276	248	346	441	549	1,318	1,318
Minimum	82	96	89	111	146	153	188	225	247	601	82

Table AD.4.4 - PART B

DECILE	1	2	3	4	5	6	7	8	9	10	Total Sample
Change in Spread (USD)											
Average	0.001574	0.000108	0.001598	-0.000942	-0.000024	0.000664	-0.000416	-0.000335	-0.000230	-0.000155	0.0001
Maximum	0.750000	1.125000	0.750000	0.750000	0.750000	0.750000	0.500000	0.500000	0.500000	1.625000	1.6250
Minimum	-0.500000	-0.750000	-0.750000	-0.750000	-0.750000	-1.500000	-0.500000	-0.500000	-0.500000	-1.625000	-1.6250
Change in Proportional Spread (USD)											
Average	0.0613	0.0094	0.0473	-0.0324	-0.0011	0.0025	0.0101	-0.0089	-0.0002	0.0000	0.00
Maximum	36.6667	45.4887	49.3151	27.7778	21.8182	27.7778	45.3441	15.9637	27.7778	43.2127	49.31
Minimum	-21.4286	-27.2446	-23.6364	-21.5321	-21.9608	-27.7778	-21.4286	-13.2942	-27.7778	-18.0870	-27.77
Spread (USD)											
Average	0.6792	0.5086	0.4509	0.4066	0.4290	0.3412	0.3725	0.3693	0.3258	0.2972	0.41
Maximum	2.7500	2.0000	2.0000	1.7500	3.2500	1.5000	1.5000	1.2500	1.0000	1.0000	3.25
Minimum	0.0313	-0.7500	-0.1250	0.0000	-0.1250	-0.2500	0.0000	0.0000	-0.2500	-2.5000	-2.50
Proportional Spread (USD)											
Average	12.60	8.90	9.07	9.19	7.82	8.24	5.72	4.14	3.96	1.79	7.
Maximum	57.14	58.82	61.22	58.82	54.55	50.00	57.14	34.48	54.55	52.63	61.
Minimum	0.51	-2.30	-1.12	0.00	-3.17	-0.39	0.00	0.00	-2.30	-13.33	-13.

Figure AD. 1 Variation of Average Daily Trading Volume with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.1 and AD.4.2 have been used in Figure AD.1 to plot the variation of the average value of daily trading volume, expressed as the average number of shares traded in a day, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

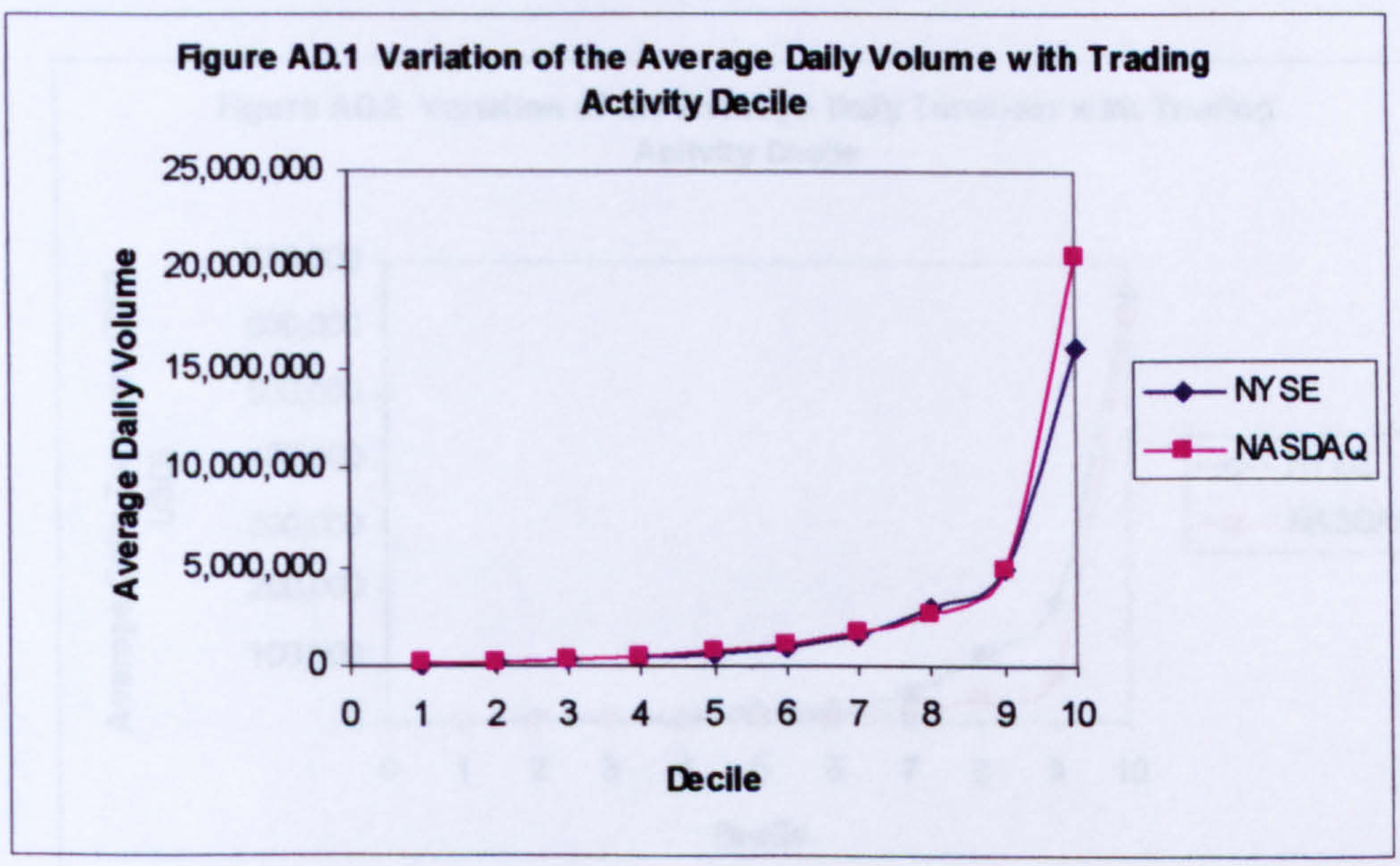


Figure AD. 2 Variation of Average Daily Turnover with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.1 and AD.4.2 have been used in Figure AD.2 to plot the variation of the average value of daily turnover, expressed in thousands of US Dollars traded in a day, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

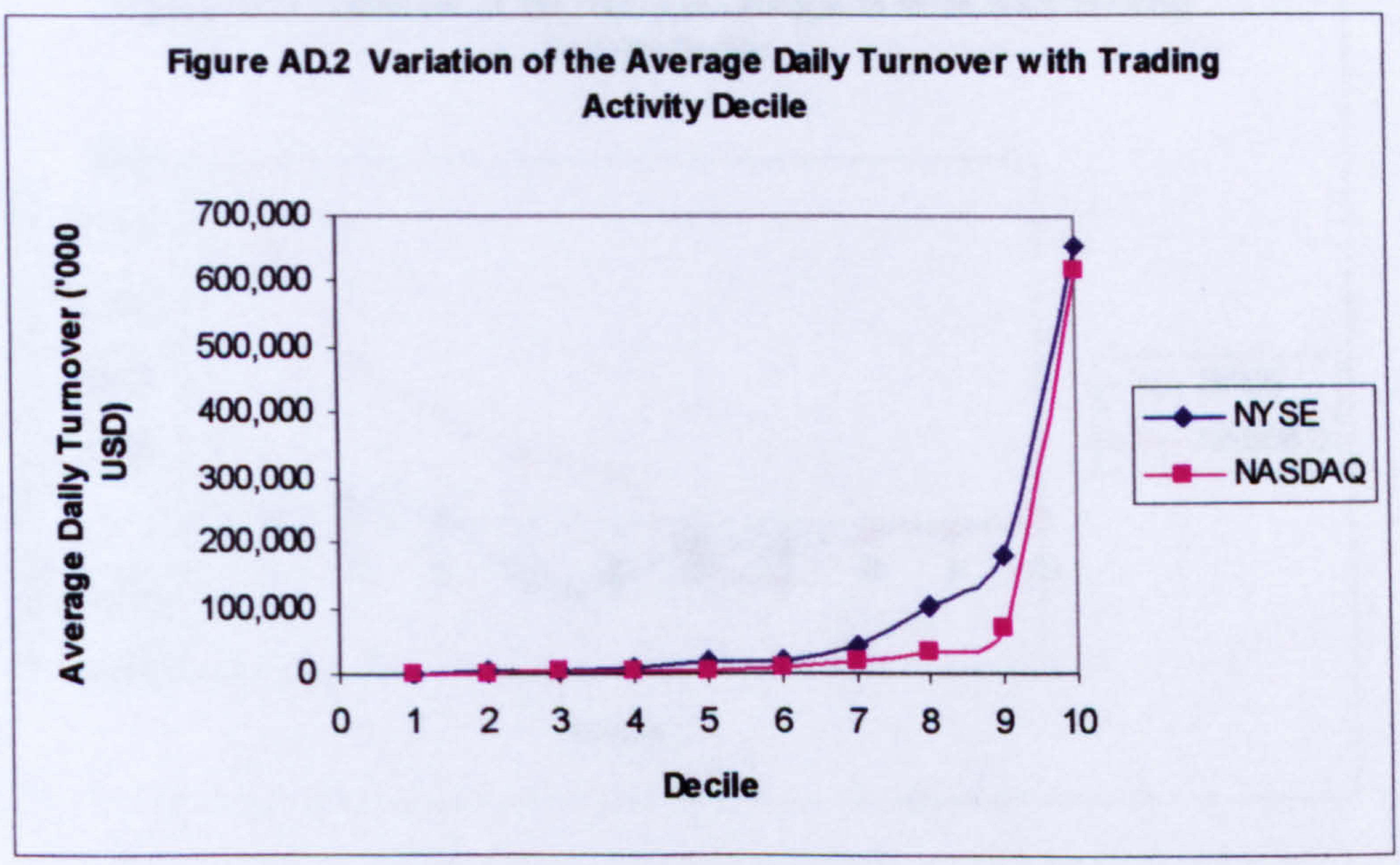


Figure AD. 3 Variation of the Average Change in Share Price with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.1 and AD.4.2 have been used in Figure AD.3 to plot the variation of the average change in the price of shares, expressed in U.S. Dollars, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

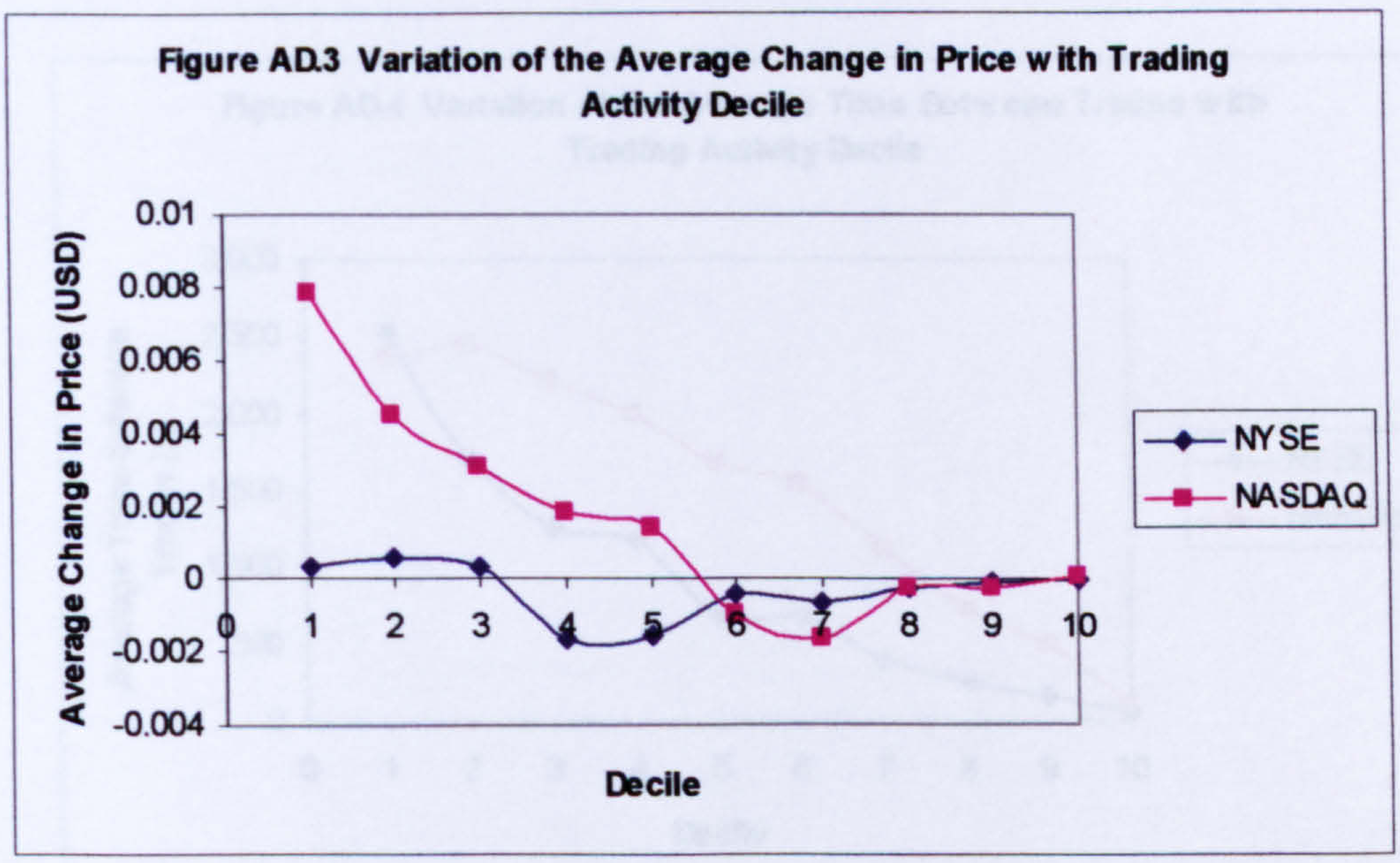


Figure AD. 4 Variation of the Average Time between Trades with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.1 and AD.4.2 have been used in Figure AD.4 to plot the variation of the average time between trades, in seconds, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

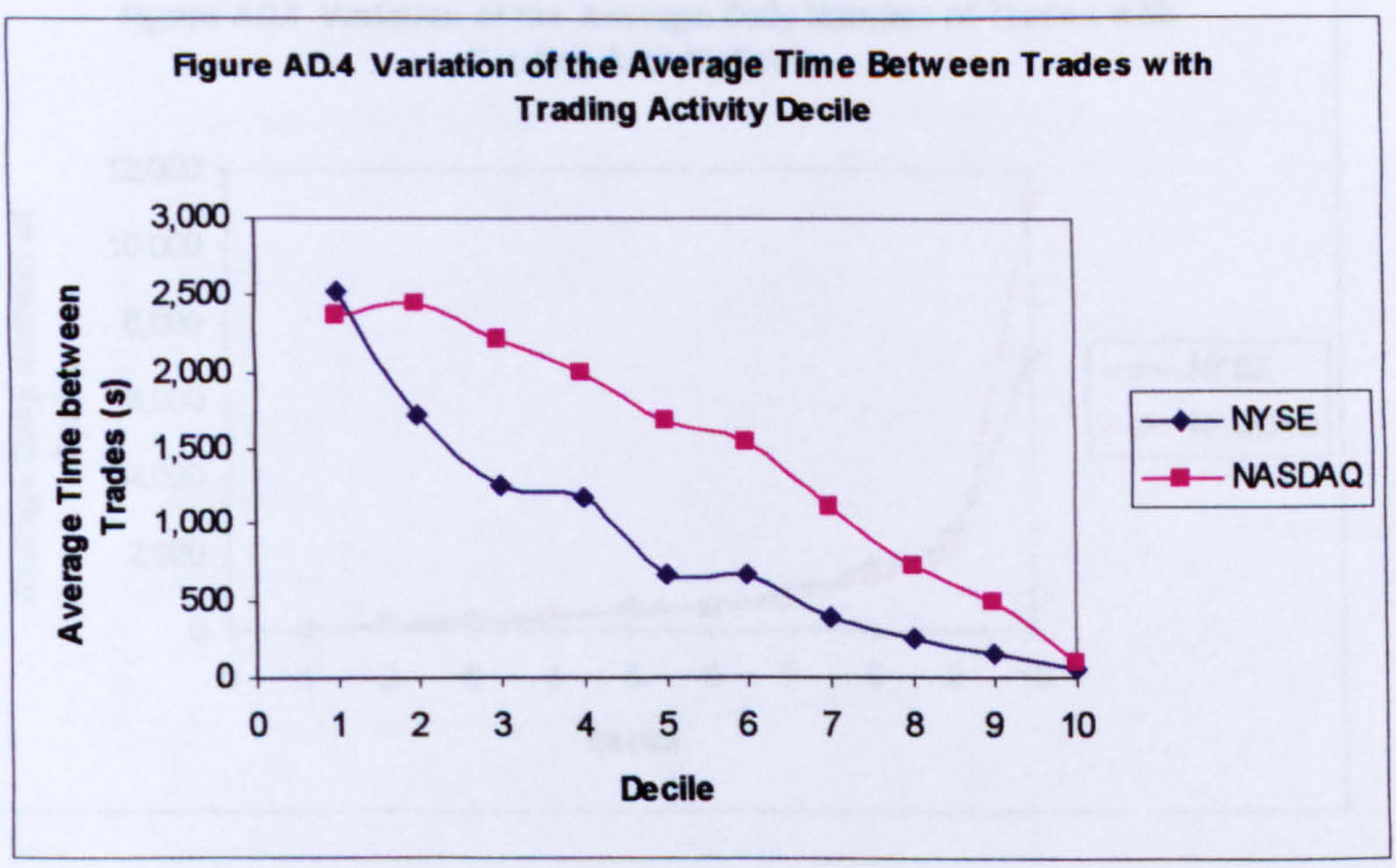


Figure AD. 5 Variation of the Average Daily Number of Trades with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.1 and AD.4.2 have been used in Figure AD.5 to plot the variation of the average daily number of trades, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

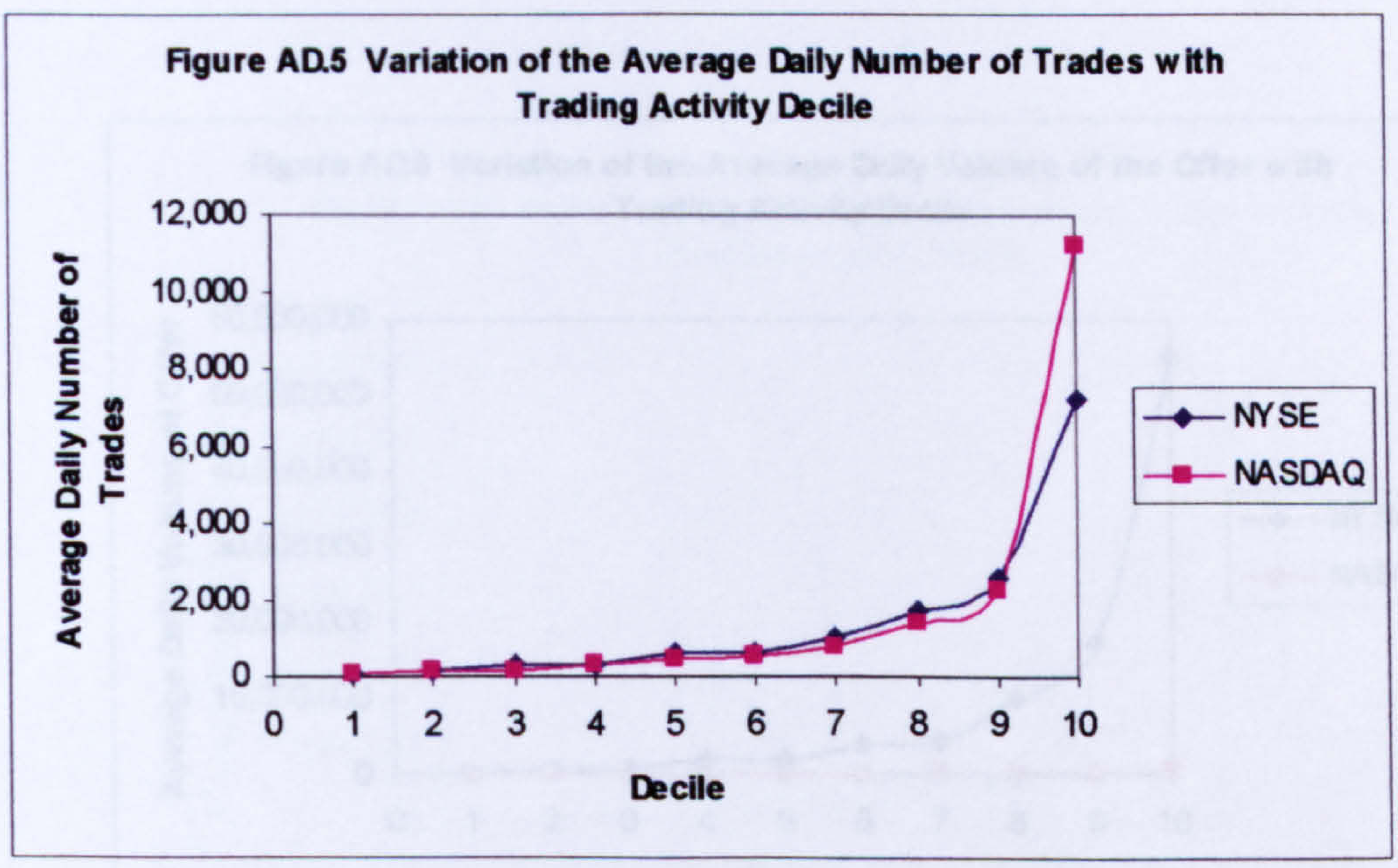


Figure AD. 6 Variation of the Average Daily Volume of the Offer (ASK) Price with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.6 to plot the variation of the average daily volume of the offer (ask) price, expressed as the average number of shares offered for all stocks in a decile in a day, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

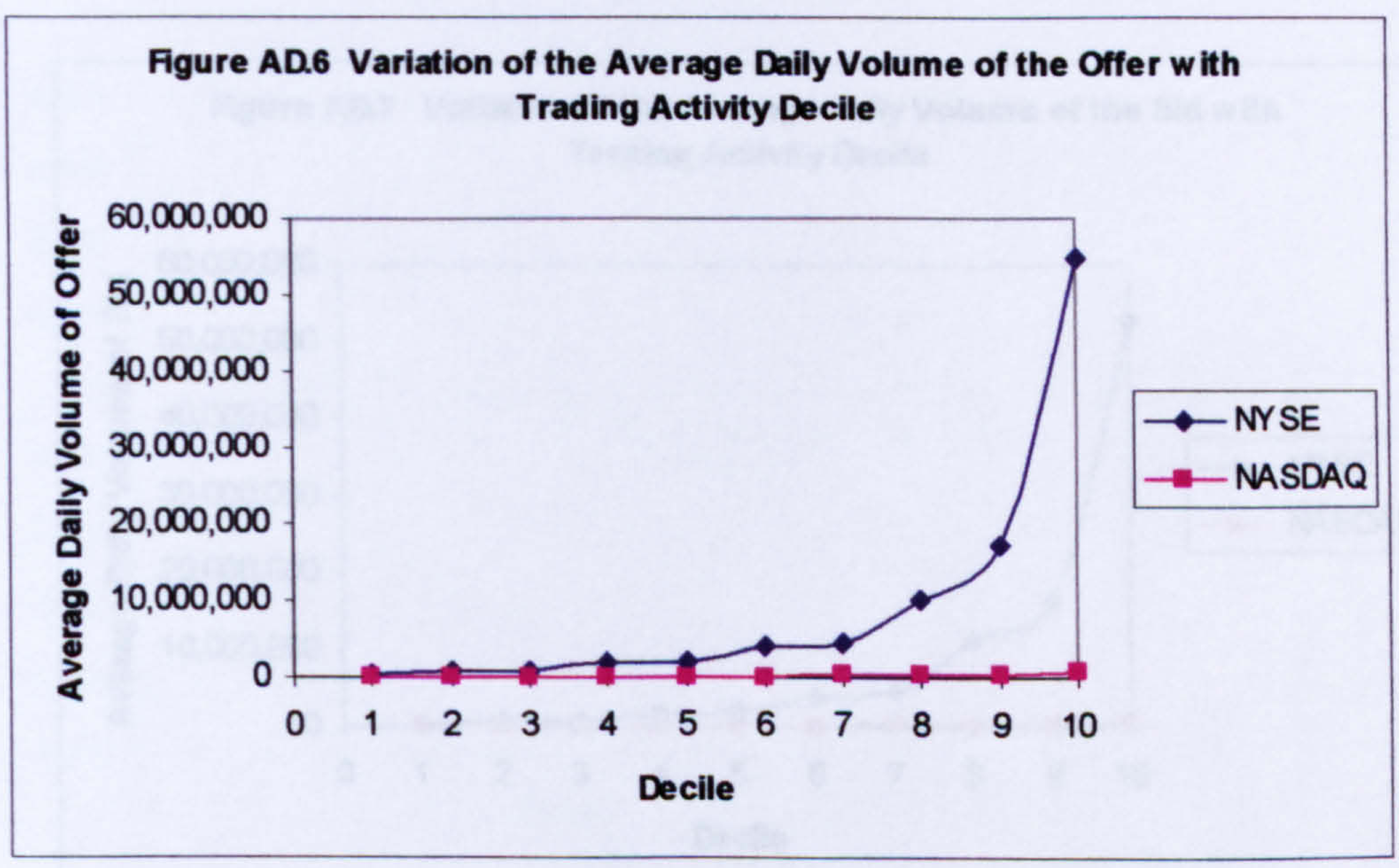


Figure AD. 7 Variation of the Average Daily Volume of the Bid Price with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.7 to plot the variation of the average daily volume of the bid price, expressed as the average number of shares bid for all stocks in a decile in a day, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

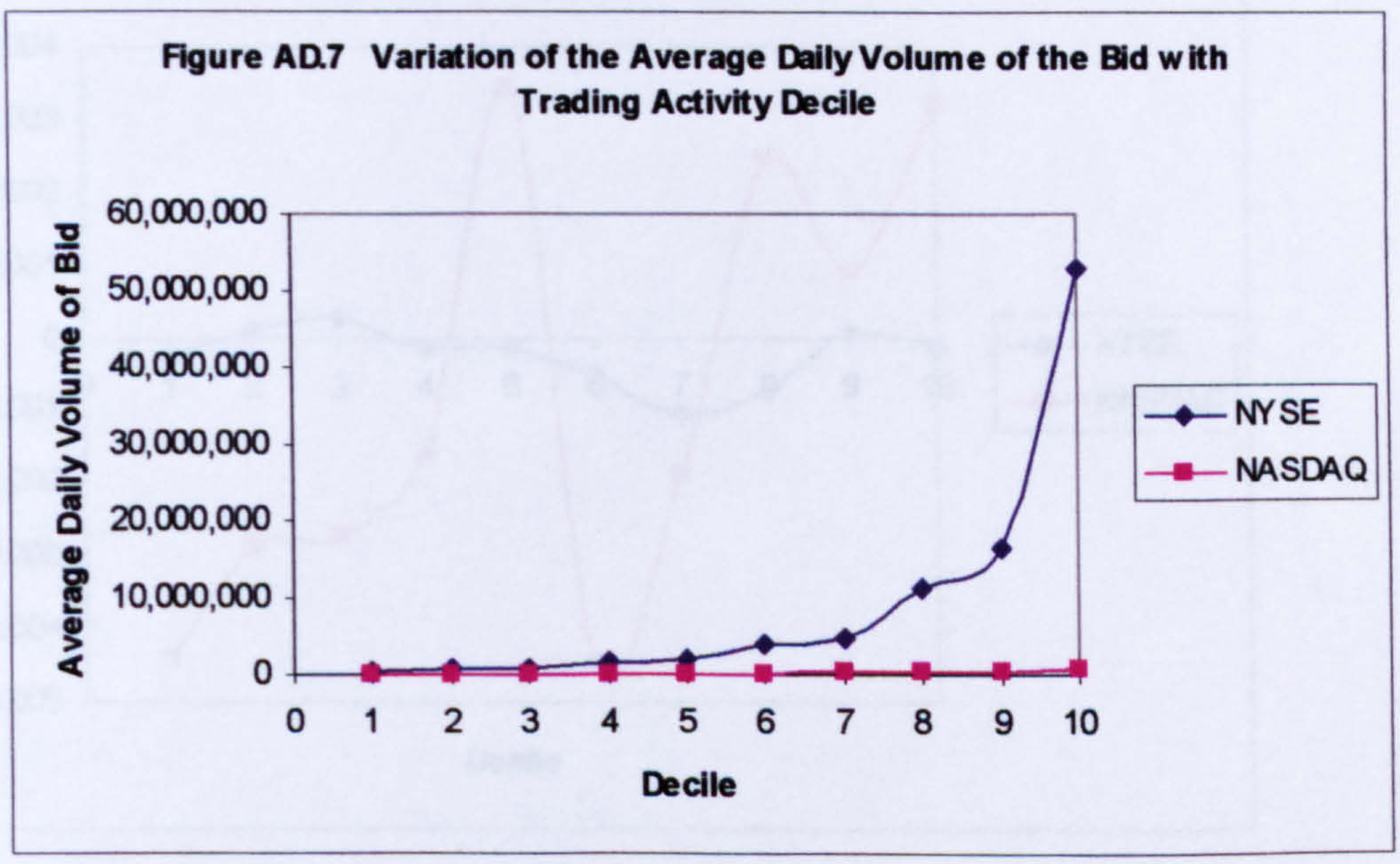


Figure AD. 8 Variation of the Average Change in the Offer (Ask) Price with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.8 to plot the variation of the average change in the Offer (Ask) Price, in U.S. Dollars, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

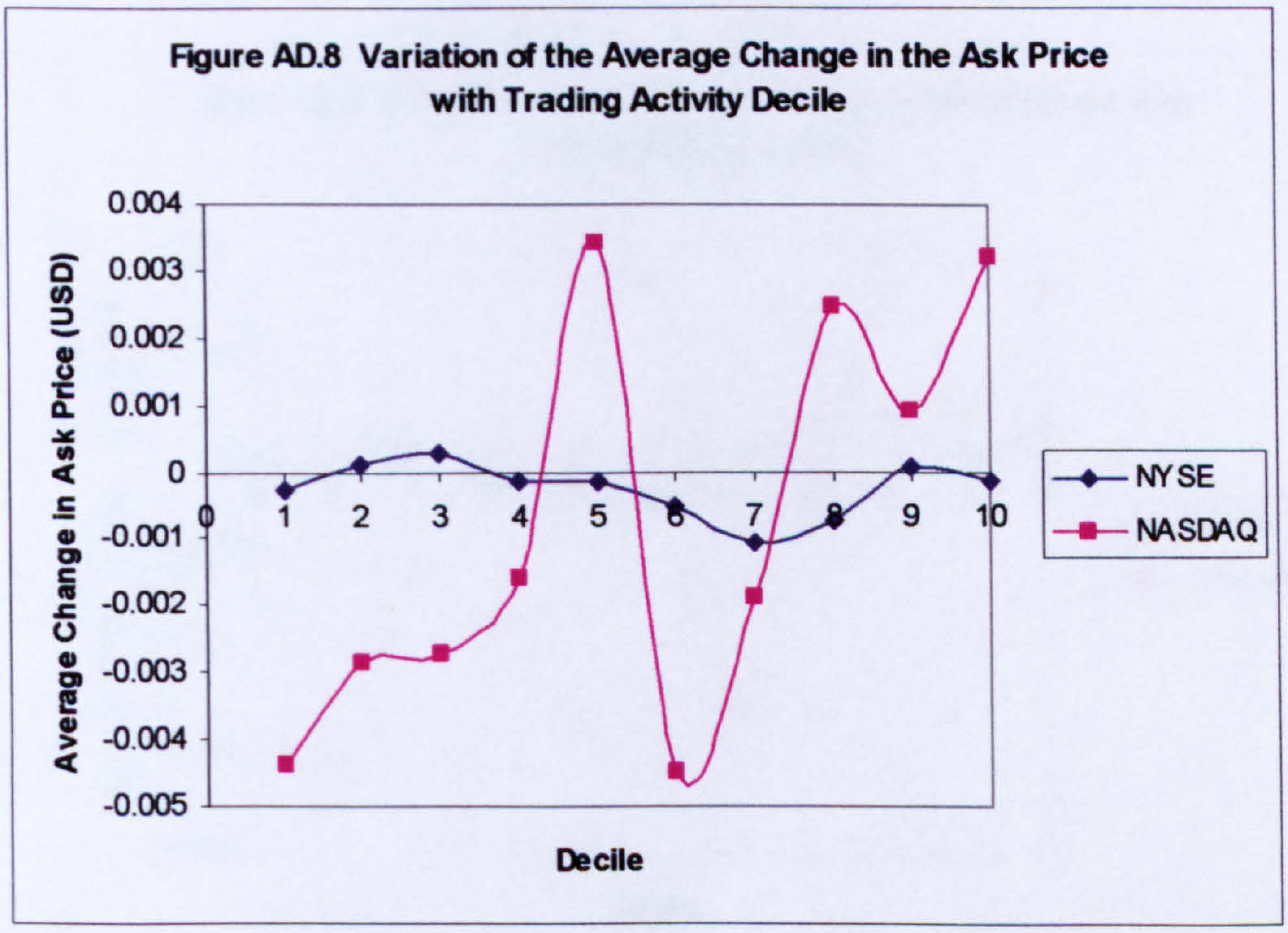


Figure AD. 9 Variation of the Average Change in the Bid Price with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.9 to plot the variation of the average change in the bid price, in U.S. Dollars, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

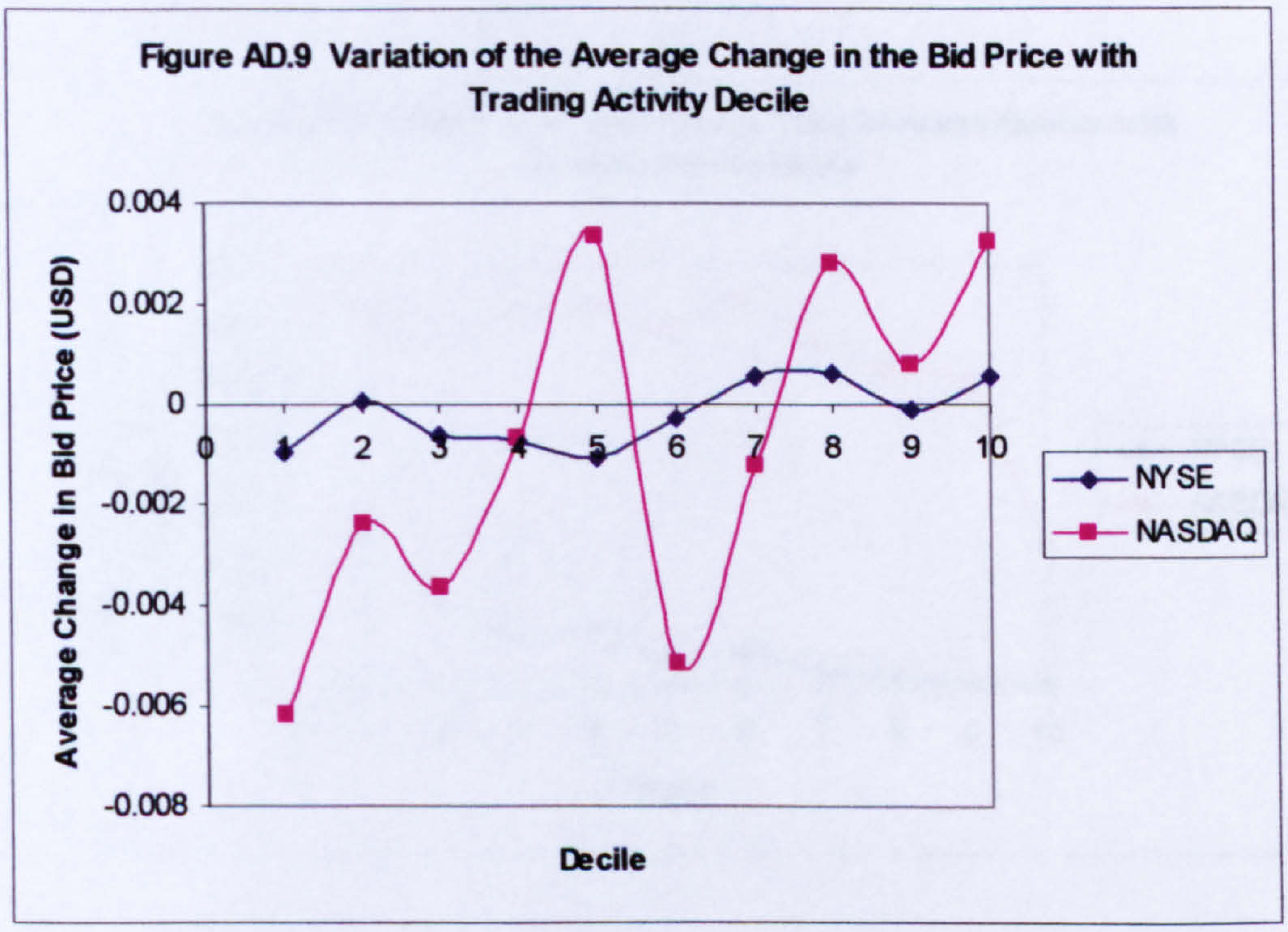


Figure AD. 10 Variation of the Average Time between Quotes with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.10 to plot the variation of the average time between quotes, in seconds, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

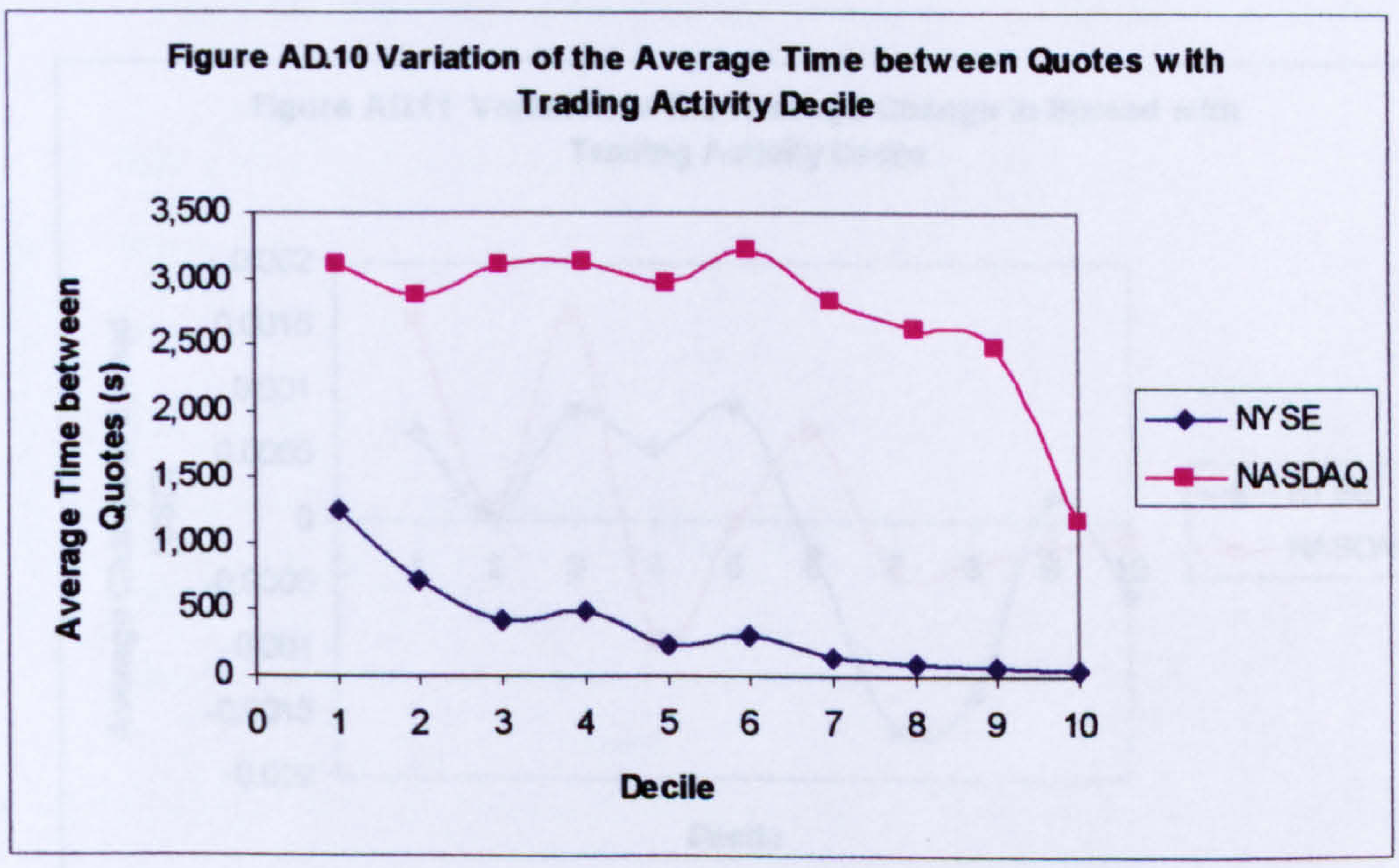


Figure AD. 11 Variation of the Average Change in Spread with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.11 to plot the variation of the average change in spread, in U.S. Dollars, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

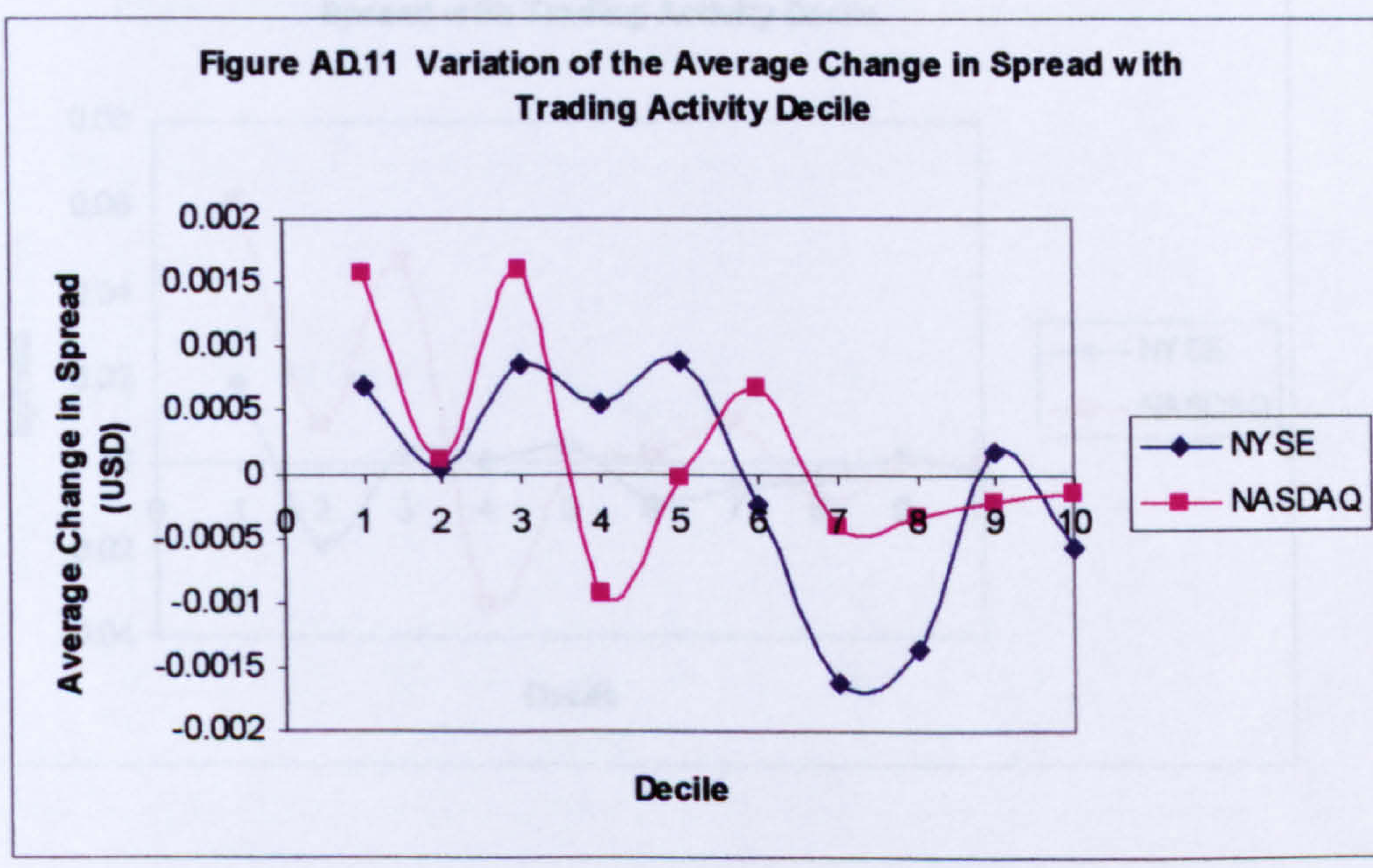


Figure AD. 12 Variation of the Average Change in Proportional Spread with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.12 to plot the variation of the average change in the Proportional Spread, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

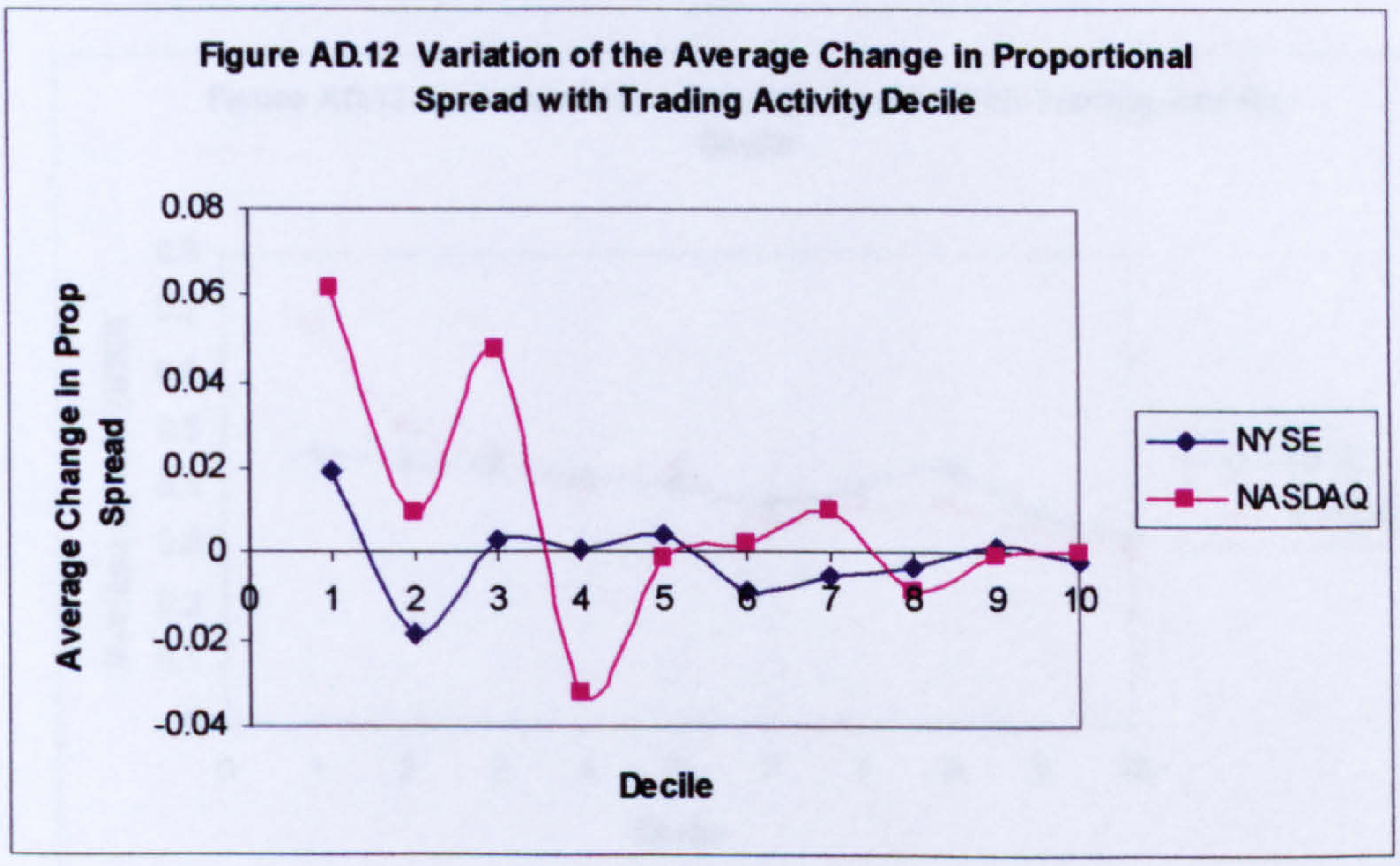


Figure AD. 13 Variation of the Average Spread with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.13 to plot the variation of the average spread, in U.S. Dollars, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

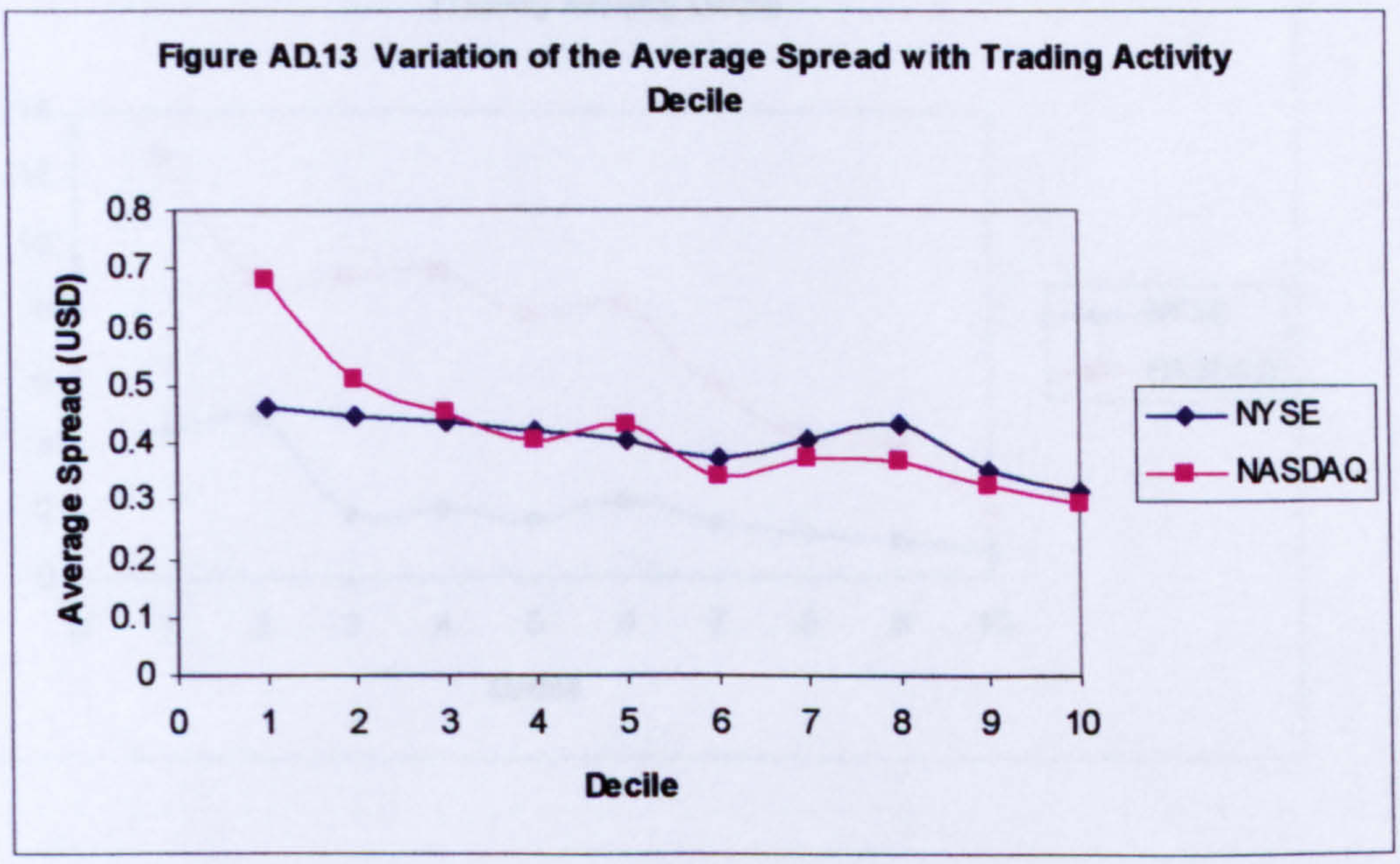


Figure AD. 14 Variation of the Average Proportional Spread with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.14 to plot the variation of the average proportional spread, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

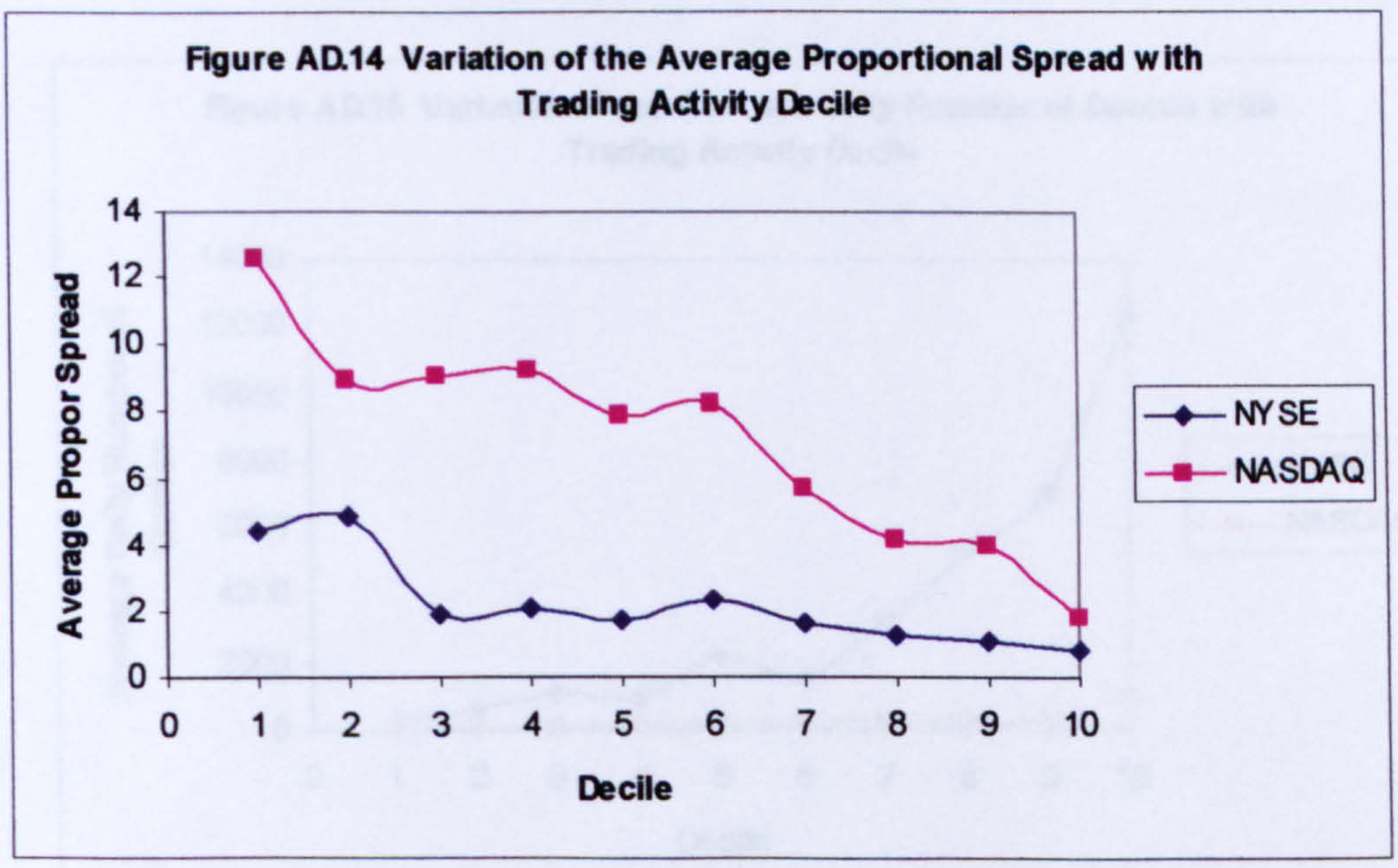


Figure AD. 15 Variation of the Average Daily Number of Quotes with Trading Activity Volume Decile, NYSE and NASDAQ TAQ data: October 1994.

Data presented in tables AD.4.3 and AD.4.4 have been used in Figure AD.15 to plot the variation of the average daily number of quotes for all stocks in a decile in a day, with activity volume decile for NYSE and NASDAQ samples of data.

Following appropriate filtering procedures, described in chapter four, a total number of 200 NYSE and 600 NASDAQ common stocks are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. One is the lowest activity and ten the highest activity volume decile.

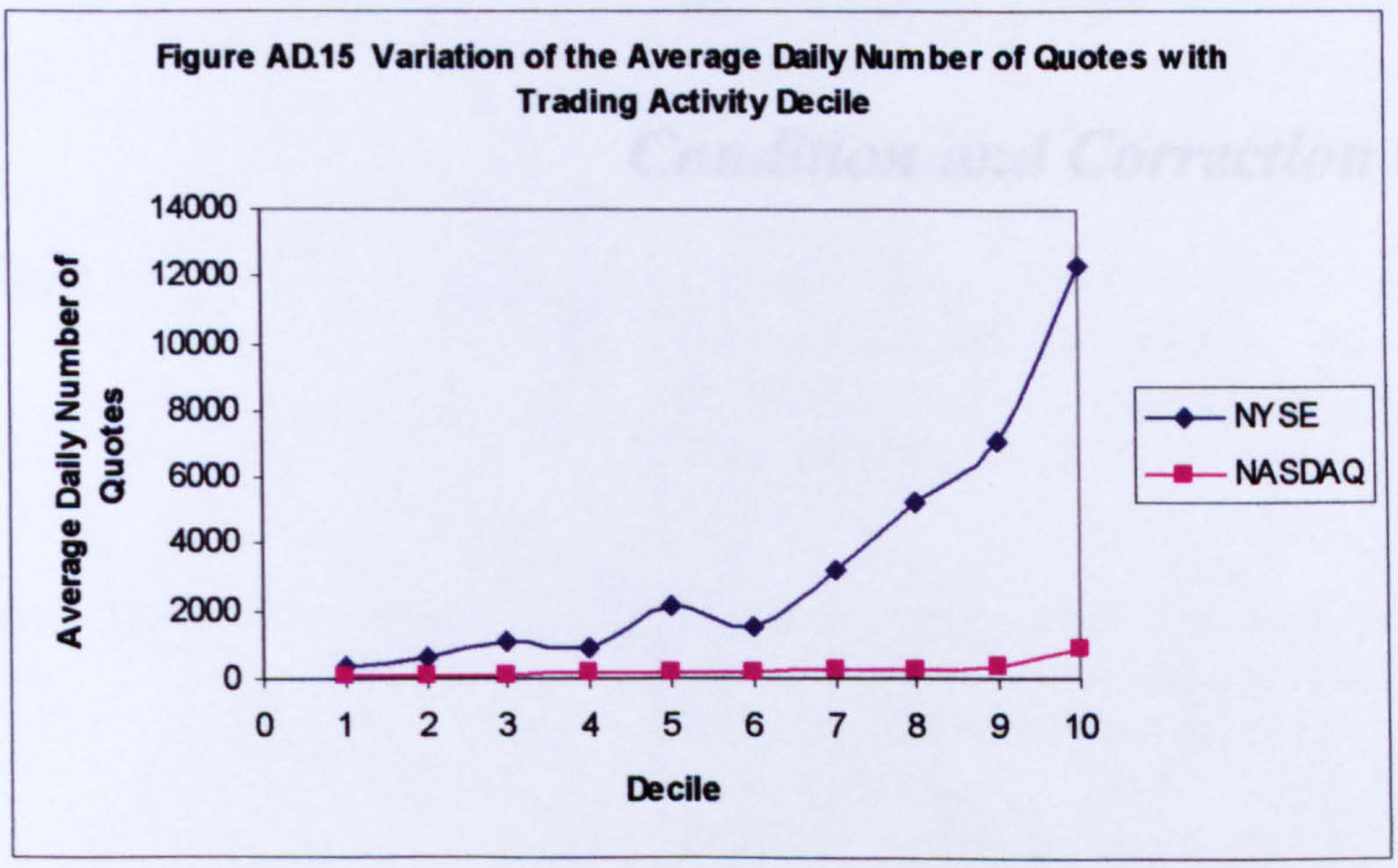


Figure AD.15 Variation of the Average Daily Number of Quotes with Trading Activity Decile

Appendix E

Trade and Quote (TAQ) Database

Condition and Correction Codes

AE.1 Introduction

A number of condition and correction codes have been used for the creation of the Trade and Quote (TAQ) database. These provide useful information which aids researchers in evaluating the validity of the data reported.

The following information has been obtained from the TAQ Database Manual of the New York Stock Exchange, Version 3.1, 1995.

The main files included in the TAQ database are the following:

Name of File	Description
Date File	Provides information on the dates on which transactions have taken place and the positions in the trade and quote files where data can be found for a particular date.
Consolidated Trade File	Contains one record for each trade reported and contains fields for the time, price, size and exchange of each trade. Also contains condition and correction codes.
Consolidated Quote File	Contains one record for each quote reported and contains fields for the time, bid price, bid size, offer price, offer size and exchange of each quote. Also contains condition codes as well as information identifying market makers for NASDAQ trades.
Master Table File	Contains reference information about the stocks in the trade and quote files, including, among other: symbol, name, CUSIP (Committee on Uniform Security Identification Procedure) number which uniquely identifies each stock, exchange which trades stock, number of outstanding shares, size of round lot, trading denomination (tick size), type of stock (i.e. common, preferred etc).
Dividend File	Contains one record for each symbol that either paid a dividend or redistributed

	stock during the month. Fields: symbol, CUSIP number, cas dividend, stock dividend or split and date of split or dividend.

AE.2 Condition and Correction Codes

The following conditions and correction codes have been used for filtering data in this thesis:

Code	Type	File	Description
Z	Condition	Consolidated Trade File	Transaction has been reported to the tape at a time later than it occurred and out of sequence. There are other trades between the time the transaction occurs and the time it is reported.
0	Correction	Consolidated Trade File	Good Trade – No Correction
1	Correction	Consolidated Trade File	Good Trade – Corrected Trade
2	Correction	Consolidated Trade File	Good Trade – Corrected Symbol
0	Condition	Consolidated Quote File	Invalid Field for this issue
4	Condition	Consolidated Quote File	A regulatory halt used because relevant news influencing the stock is disseminated. Trading has been suspended until the resultant impact has been assessed.
7	Condition	Consolidated Quote File	A non-regulatory halt used because there is severe buy or sell order imbalance. To prevent a disorderly market, trading is temporarily suspended.
9	Condition	Consolidated Quote File	A regulatory halt used because the level of trading activity in a security is such that the Exchange cannot collect, process and disseminate quotes that accurately reflect market conditions. The specialist, with Floor Official approval, may

			switch to 'non-firm mode' for 30 minutes.
11	Condition	Consolidated Quote File	Indicates a regulatory trading halt or delayed opening due to an expected news announcement which may influence trading in the stock. A trading halt or opening delay may be reversed once the news has been disseminated.
13	Condition	Consolidated Quote File	A non-regulatory halt used because events relating to one security will affect the price and performance of another security.
14	Condition	Consolidated Quote File	A non-regulatory halt used because matters that affect the common stock of a company may affect the performance of other classes of securities for the same company.
15	Condition	Consolidated Quote File	A non-regulatory halt used because there is a severe influx of buy and sell orders. To prevent a disorderly market, trading is temporarily suspended.

Appendix F

Limit Orders

F.1 Introduction

This appendix discusses a number of issues relating to limit orders. In particular, existing literature is presented on the following research areas:

1) The effect of the existence of limit orders on the trading process.
2) The decision process followed by traders in choosing among different types of orders.
3) Comparison between market and limit orders
4) Components of limit orders and the limit-order book, and
5) Types of limit orders available and proposed

F.2 How the existence of limit orders affects the trading process.

The difficulty in modeling limit-orders[121] has led early researchers to either ignore them completely or involve both types of order but not include quote-driven dealers. However, even in models of the latter type[122] bid and ask quotes are either infinite[123] or are exogenously specified and never exhausted by the incoming order flow.[124] The presence of the limit-order book[125] has been shown by Cohen, Meir, Schwartz, and Whitcomb (1981) to introduce significant negative serial correlation into both

[121] Kyle (1989), Rock (1996), Angel (1995), Kumar and Seppi (1994), Foucault (1993) Glosten (1994). One of the earliest models to use them was Kyle (1989) who defined the schedule of limit orders as a linear demand schedule covering all quantities. O'Hara and Oldfield (1986) and Hamao and Hasbrouck (1995) also used limit orders in their analyses. Biais, Hillion and Spatt (1995), Hamao and Hasbrouck (1995) and Harris and Hasbrouck (1996) described empirical properties of limit orders in the Paris, Tokyo and New York Stock Exchanges respectively. Hollifield, Miller and Sandas (1996) and Sandas (1998) carried out structural estimation on Swedish data.

[122] Easley and O'Hara (1991), Chakravarty and Holden (1995), Harris (1994), Parlour (1998), Seppi (1997).

[123] Harris (1994), Parlour (1998).

[124] Easley and O'Hara (1991), Chakravarty and Holden (1995). Seppi (1997) allows the depth of a monopoly specialist to be exhausted but his analysis focuses on tick size.

[125] Pure limit order book markets have been examined by Niemeyer and Sandas (1993) for Stockholm, Hedvall (1994) for Helsinki, Lehmann and Modest (1994) for Tokyo, Niemeyer (1994) for Helsinki and Stockholm, Biais, Hillion and Spatt (1995) for Paris, Frino and McCorry (1995) for Australia,

transaction and quotation returns for measurement intervals of at least several days.[126] Rock (1996) has also showed that uninformed limit orders exacerbate the inventory problem of the market maker and thus delay the full adjustment to an inventory shock and that the book of limit orders can improve prices available to some traders by increasing the liquidity of the market.

Copeland and Galai (1983) introduced the idea of the *'free option' property of limit orders*, also tackled by Rock (1996) and Stoll (1990), based on the idea of Bagehot (1971),[127] by addressing the adverse selection problem from the part of a quote-setting dealer who by offering his bid and ask quotes is in effect writing a put and a call option (a 'free' straddle option) to an information motivated trader. This behaviour is equivalent to that of a limit order trader who, by submitting his orders, offers put or call options to the other traders in direct competition with the dealer.

F.3 How a trader chooses among the various types of orders available.

A number of papers have analyzed the decision to trade via a market or a limit order and the relative performance of the two types of order. This order-selection decision is associated with the strategic behaviour of traders. However, most of the early models of market microstructure[128] could not study the mix between market and

---

Hamao and Hasbrouck (1995) for Tokyo, de Jong, Nijman and Roell (1995) for Paris, Hollifield, Miller and Sandas (1996) for Stockholm and Sandas (1998) for Stockholm.

[126] Limit orders already in the book fail to reflect changes in information, a property they called "stickiness" of the limit order book. Stickiness causes prices to depart from the efficient markets model and this depends on factors such as the age of limit orders on the book, the number of orders on the book and their price distribution.

[127] They assumed that a dealer who wishes to maximize his profit adjusts the length of his bid-ask spread in such a way as to increase the difference between his losses to information-motivated traders and his gains from liquidity traders.

[128] Kumar and Seppi (1994) analyzed equilibrium in a static setting that allowed market and limit orders with discreetness. Seppi (1997) allowed only the specialist to act strategically whereas limit order submitters were assumed to be competitve. Parlour and Seppi (1997) compared the limit-order / specialist market structure in Seppi (1997) with the pure limit-order book in Glosten (1994). Angel(1990), Foucault (1993) and Harris (1994) modeled the order placement strategy of an informed investor for choosing between market and limit orders but the investor could select only one of these

limit orders because they did not allow agents to choose between the two types of order.[129] Almost all of the models reflect the essential trade off between market and limit orders, which is that between execution uncertainty and transaction costs.[130]

In his analysis of the limit order market Glosten (1994) allowed only for monopolistically competitive limit order submitters[131] and assumed that the trader will place a limit order only when his expected gain from transacting with liquidity traders exceeds his expected loss from informed traders but he did not explicitly model the investor's decision to trade via a market or a limit order.

Easley and O'Hara (1991) found that apart from the fact that different order forms lead to changes in the informativeness of the order flow, price-contingent orders lead to faster price adjustment and thus improve the efficiency of a market-maker system.

Chakravarty and Holden (1995) as well as Handa and Schwartz (1996) showed that placing a network of buy and sell limit orders around the current price is profitable and Handa and Schwartz (1996) who extended Glosten (1994) found evidence that returns to the limit order strategy given that the orders execute are greater than the unconditional returns to the market order whereas returns to the limit order strategy conditional on the orders not executing were lower than unconditional market order returns.

---

types of order. Hamon et al (1993), Handa and Schwartz (1996) and Harris and Hasbrouck (1996) focused on optimal order-submission strategies.

[129] Of these models Glosten and Milgrom (1985) focused explicitly on dealers; Biais, Foucault and Salanie (1998), Glosten (1994), Rock (1996), Seppi (1997) and Parlour and Seppi (1997) focused on limit order trading.

[130] Blume and Goldstein (1993) and Lee (1993) carried out comparative market studies based on price improvement defined as the difference between a transaction price and the prevailing quote.

[131] In his model a single risk-averse, utility-maximizing trader who may be informed or uninformed arrives and submits a market order which is transacted against the limit order book (he is not allowed to submit a limit order).

When limit order traders do not run the risk of being picked off[132] Parlour (1998) showed that a risk-neutral trader prefers to trade in a limit order market, compared to a hybrid or dealership market whereas when all dealers act strategically, as in the model of Viswanathan and Wang (1999)[133] and the customer's order is competed for and divided among a number of risk-averse market makers a risk-neutral customer prefers to trade in a limit-order market instead of in any hybrid market. Foucault (1999) found that higher volatility in the asset traded increases its risk and affects the mix between market and limit orders in favour of the latter. Since volatility decreases with firm size (Hasbrouck (1991b)), and since spreads are positively related to volatility, small firms should have a larger proportion of limit orders, lower fill rates and larger spreads than large firms in limit order markets.[134]

F.4 How the performance of limit orders can be compared to that of market orders

As a number of papers have concluded the NYSE generally offers investors better prices than those available on alternative exchanges[135] and therefore brokers should route market orders to the NYSE. However most such studies ignore limit orders despite the fact that they constitute a large portion of the order flow.

[132] Since the asset value fluctuates, the execution of limit orders is uncertain so they run the risk of becoming mispriced as new public information arrives. This creates a winner's curse problem since limit orders are more likely to be executed at a loss ('picked off') when they are mispriced than when they are not.

[133] Glosten (1994) has showed that the anonymity of markets and thus the ability to divide order flow creates a problem for dealership markets but not for the limit order book. Bernhardt and Hughson (1997) have analysed the equilibria that arise when order flow is divisible.

[134] The model suggested that the observation that spreads enlarge at the end of trading in limit order markets (as note, for example, by McInish and Wood (1992) for the NYSE) can be due to increased execution risk at the end of trading.

[135] Blume and Goldstein (1993), Lee (1993), Angel (1994), Petersen and Fialkowski (1994), Ready (1996), Ross, Shapiro and Smith (1996), Bessembinder and Kaufman (1997), SEC (1997), Battalio, Greene, and Jennings (1998).

In their previously discussed work Handa and Schwartz (1996) showed that when transaction prices change solely in response to information, trading via limit order is suboptimal for all traders because the advent of adverse news can trigger an undesired trade, while the advent of favourable news can result in the limit order not executing.[136] Harris and Hasbrouck (1996) found that limit orders placed at or better than the prevailing quote perform better compared to market orders even after imputing a penalty for unexecuted orders, and after taking into account market order price improvement and that in some cases the use of limit orders can reduce execution costs.

Measuring execution quality using non-price issues, as well as prices, Battalio, Greene, Hatch, and Jennings (1999) concluded that the routing of limit orders to regional exchanges frequently appears justified and fill rates and the economic performance of limit orders on regional exchanges may be higher compared with identical NYSE orders thus conflicting the evidence of other studies,[137] which showed that market order traders pay more for liquidity on regional exchanges than on the NYSE.

The SEC (1997) study[138] has shown that traders face similar adverse selection costs across exchanges.[139]

Considering several market structures when the order size is known[140] Biais, Foucault and Salanié (1998)[141] showed that the limit order market, where there is

---

[136] They referred to the cost of being bagged by informed traders which alludes to the situation where a limit order trader faces a negative expected gain from the execution of his order when this is not triggered by liquidity-driven changes in the transaction price.
[137] Blume and Goldstein (1993), Petersen and Fialkowski (1994), Bessembinder and Kaufman (1997), Battalio, Greene, and Jennings (1998)
[138] Lightfoot, Martin, Peterson, and Sirri (1999) and the SEC (1997) are the only two detailed limit-order execution quality studies but the former is not as detailed as the latter.
[139] Moreover, this study did not consider the possibility that orders routed to regional exchange might receive better execution quality if routed to the NYSE but it is possible that they are not routed there owing to conflicts of interest.

competitive pricing, efficient risk-sharing and no tacit collusion (in line with Wilson (1979) and the empirical findings of Christie and Schultz (1994a,b)[142] is the optimal trading mechanism under certain prerequisites.[143]

F.5 Components of the limit order spread and the limit order book

In their empirical study of the effect of limit orders on NYSE spreads Chung, Van Ness and Van Ness (1999) found that specialists quote more actively for low-volume stocks and during early hours of trading when there are fewer limit orders submitted which suggests that they intervene in order to provide liquidity when that supplied by limit order traders is low. The spread was shown to be widest when both bid and ask prices are quoted by specialists alone and narrowest when both sides of the quote originate from limit order traders which suggests that competition from limit order traders has a significant effect on spreads. They also showed that the well-documented U-shaped pattern of spreads reflects the intraday variation in spreads established by limit order traders which is associated with the intraday variation in competition among limit order traders.[144]

---

[140] Viswanathan and Wang (1999), Biais, Martimort, and Rochet (1997) and Roell (1998) considered market structures where the exact order size is unknown to the dealers when they submit their bidding schedules.

[141] Similar to Hedvall and Niemeyer (1996) they also focused on variations in the order flow due to transient changes in the state of the limit order book.

[142] Reiss and Werner (1995) found empirically that SEAQ dealers offer quantity discounts. Dutta and Madhavan (1997) analyzed how long term interaction can enable risk neutral (in contrast to this model's risk averse) dealers to collude on large spreads.

[143] Their analysis is related to the study of auctions for divisible goods and the papers of Wilson (1979), Berhheim and Whinston (1986), Back and Zender (1993), and Klemperer and Meyer (1989) even though their analysis is more complete and uses a different modeling framework.

[144] Their results were consistent with the information models of Madhavan (1992) and Foster and Viswanathan (1994), partially consistent with the market-power models of Stoll and Whaley (1990) and Brock and Kleidon (1992) since they did not find a rise in spreads at close and opposite to the inventory model of Amihud and Mendelson (1982) since they did not find a wider spread at or near the close.

In a study similar to that of Chung, Van Ness and Van Ness (1999), Kavajecz (1999) dealt with the management of the depth[145] and partitioned the specialist's quoted depth into that part contributed by the specialist and that contributed by the limit order book thus describing it as the difference between the quoted depth and that on the limit order book. Using the TORQ database he showed that the specialist's quotes reflect only the limit order book on the side of the market where he believes that there is a possibility of informed trading in order to avoid their own exposure to such trading. Changes in quoted depth were shown to be consistent with specialist's managing their inventory positions as well as having knowledge of the future value of the stock. He also showed that limit order traders and specialists reduce depth around information events in order to reduce their exposure to adverse selection costs.

F.6 Types of Limit Orders available and proposed

An important element in modeling limit orders is the time between the submission and the execution of limit orders and this is related to the time between transactions. Limit orders lack price risk but their execution is not guaranteed and their time-to-execution is a random function of many factors including limit price, number of shares, market conditions, private information etc.

A number of papers, like Angel (1994) and Holifield, Miller and Sandas (1996) have studied the probability of limit order execution both analytically and empirically, while another class[146] have been based on the idea that traders submit both types of

[145] According to rule 2092 of the NYSE, the specialist is obliged to give priority to limit orders at the quoted prices, contrary to his interest, unless he betters any limit order price before he can take the trade. The specialist contributes depth above that posted on the limit order book since he is responsible for maintaining a reasonable level of depth apart from a continuous price, a narrow spread and to stabilize trades. Moreover by offering depth he can signal his interest for trading to the trading crowd.

[146] Glosten (1989), Easley and O'Hara (1991), Glosten (1994), Parlour (1998), Chakravarty and Holden (1995), Harris and Hasbrouck (1996), Kavajecz (1999), Rock (1996), Seppi (1997).

order in their attempt to balance the opportunity cost of delaying execution against the risks of immediate execution.

Hausman, Lo and MacKinley (1992) found that time between trades is an important determinant of the variability in transaction-price changes whereas Lo, MacKinlay and Zhang (1997) indicated that execution times can be sensitive to certain explanatory variables like market depth, spread between limit price and quote midpoint and market volatility justifying the use of the types of order-placement strategies described in previous papers.[147] They also showed that hypothetical limit order executions are unreliable indicators of actual execution times and thus very poor substitutes for actual limit-order data.

Other researchers have proposed alternative types of limit orders. Black (1991, 1995) argued that in an unrestricted world equilibrium, if it could exist, uninformed traders would use 'unpriced' limit orders ('indexed limit orders,[148] orders with a limit price continually adjusted to market conditions) that specify the number of shares requested and a level of urgency and could not do better by using (1) conventional limit orders at a limit price, (2) specialized exchanges, (3) basket trading, or (4) sunshine trading[149] whereas informed traders would predominantly use market orders but might occasionally use unpriced limit orders.

Brown and Holden (1996) analyzed three alternative types of limit orders which they showed could offer benefits to different types of traders.

---

[147] For example Angel (1994), Foucault (1993), Harris (1994), Hollifield, Miller and Sandas (1996), Kumar and Seppi (1994) and Parlour (1998).
[148] Called Participating Limit Orders in Black (1971 a,b).
[149] An order to buy / sell a certain number of shares at a set future time at whatever price suffices to clear the market.

Appendix G

Tables of the Results and the Specification Tests of the Models

G.1 Summary Statistics of the Variables of the Quoted-Spread the Offer- and the Bid-Change Model

Tables AG.5.1 and AG.5.2

Summary Statistics of the Variables of the Quoted Spread, Offer-Change and Bid-Change Models, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.5.1 and AG.5.2 present Summary Statistics of the Variables of the Quoted Spread, Offer-Change and Bid-Change Models described by equations (5.11b), (5.9b) and (5.10b) respectively for the random Samples of NYSE (Table AG.5.1) and NASDAQ (Table AG.5.2) Common Stock data in October 1994.

Where

$\Delta \bar{S}_t^Q$: change in the quoted spread between times t and t-1 expressed as a fraction of the spread at time t-1.

S_t^p : spread at time t-1 expressed as a fraction of the spread at t-1 (proportional spread)

S_t^Q : quoted spread at time t-1

V_{t-1} : volume of shares traded at time t-1

$\Delta \bar{a}_t$: change in ask prices quoted between times t and t-1 expressed as a fraction of the spread at time t-1.

$\Delta \bar{b}_t$: change in bid prices quoted between times t and t-1 expressed as a fraction of the spread at time t-1.

$\Delta\tau_{t-1}$: difference (seconds) between the time elapsed between trades at t-1 and t-2, τ_{t-1} and the time elapsed between trades at t-2 and t-3, τ_{t-2}

ΔV_t^a : Change in the number of shares offered between times t and t-1

ΔV_t^b : Change in the number of shares bid between times t and t-1

$\bar{V}$: Average number of shares per transaction

q_{t-1} : Trade Indicator Variable : equal to +1 when the trade is buyer-initiated
equal to -1 when the trade is seller-initiated

ρ : First-order autocorrelation coefficient of the trade indicator variable, q_t

A total number of 200 common stocks for the NYSE and 600 common stocks for the NASDAQ sub-samples respectively are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. Trades for each decile are matched with preceding quotes and the above variables are calculated.

Table AG.5.1 - NYSE

PART A

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	20	20	20	20	20	20	20	20	20	20	200
Number of Matched trades	477	1,144	2,440	2,262	5,435	4,464	8,303	13,853	19,412	34,540	92,330
$\Delta \bar{S}_t^Q$											
Mean	0.241499	0.219045	0.193071	0.191445	0.242369	0.192452	0.219556	0.295965	0.27108	0.233225	0.229971
St. Deviation	1.075572	0.892186	0.895258	0.830109	0.969238	0.906065	0.908563	1.343147	1.647474	1.052191	1.05198
Maximum	8	7	10	9	9	10	12	32	83	79	83
Minimum	-0.875	-0.88889	-0.91667	-0.875	-0.9375	-0.875	-0.93443	-0.9697	-0.9881	-0.93939	-0.9881
S_t^p											
Mean	0.2415	0.2190	0.1931	0.1914	0.2424	0.1925	0.2196	0.2960	0.2711	0.2333	0.2300
St. Deviation	1.0767	0.8926	0.8954	0.8303	0.9693	0.9062	0.9086	1.3432	1.6475	1.0522	1.0522
Maximum	8.0000	7.0000	10.0000	9.0000	9.0000	10.0000	12.0000	32.0000	83.0000	79.0000	83.0000
Minimum	-0.8750	-0.8889	-0.9167	-0.8750	-0.9375	-0.8750	-0.9344	-0.9697	-0.9881	-0.9394	-0.9881
V_{t-1}											
Mean	949	1,076	1,132	1,787	1,234	1,835	1,802	1,769	2,276	2,688	1,655
St. Deviation	1,882	2,379	3,038	7,626	3,551	6,445	5,334	5,779	12,844	32,185	8,106
Maximum	25,000	42,500	91,500	298,300	80,000	200,000	194,100	200,900	1,000,000	5,650,000	5,650,000
Minimum	100	100	100	100	100	100	100	100	100	100	100
ΔV_t^a											
Mean	56	14	24	-69	10	173	1	0	-94	-473	-36
St. Deviation	1,640	2,771	1,495	8,815	2,746	8,172	4,225	5,793	7,320	13,878	5,686
Maximum	11,000	36,200	15,700	87,000	39,500	99,800	87,300	78,800	99,800	99,800	99,800
Minimum	-13,900	-29,900	-15,800	-87,000	-38,900	-99,800	-85,700	-78,800	-99,800	-99,800	-99,800

Table AG.5.1 - PART B

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
ΔV_t^b											
Mean	-41	-6	55	16	18	328	-18	87	-38	-381	2
St. Deviation	1,399	2,906	1,434	7,589	3,073	10,097	4,581	6,268	7,705	13,592	5,864
Maximum	9,900	33,200	12,900	99,400	49,500	99,800	49,900	96,000	99,800	99,800	99,800
Minimum	-13,400	-32,700	-11,900	-99,800	-49,900	-99,800	-59,800	-66,300	-99,800	-99,800	-99,800
$\Delta \tau_{t-1}$											
Mean	649	409	262	244	165	173	102	69	48	26	215
St. Deviation	4,357	3,245	2,427	2,385	1,521	1,543	831	542	394	216	1,746
Maximum	17,198	16,883	19,717	19,070	16,188	15,821	8,014	10,593	7,661	5,690	19,717
Minimum	-12,177	-14,893	-15,491	-14,434	-17,298	-16,374	-6,706	-6,537	-7,014	-4,763	-17,298
q_{t-1}											
Mean	0.0440	0.0787	0.1443	0.0866	0.1787	0.1281	0.1625	0.1150	0.1962	0.2589	0.1393
St. Deviation	1.0001	0.9973	0.9897	0.9965	0.9840	0.9919	0.9868	0.9934	0.9806	0.9659	0.9886
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000
$\Delta \bar{a}_t$											
Mean	0.126071	0.123891	0.104373	0.090757	0.104903	0.101169	0.102015	0.149249	0.119328	0.116226	0.113798
St. Deviation	0.583975	0.555767	0.589286	0.54016	0.5589	0.557931	0.533522	0.776982	0.575744	0.589371	0.586164
Maximum	4	5	8	4	6	9	9	15	6	23	23
Minimum	-1	-1	-1	-1.5	-1	-1	-1.5	-1	-1.5	-3	-3
$\Delta \bar{b}_t$											
Mean	-0.11543	-0.09515	-0.0887	-0.10069	-0.13747	-0.09128	-0.11754	-0.14672	-0.15175	-0.117	-0.11617
St. Deviation	0.700763	0.504889	0.521794	0.555302	0.64369	0.530884	0.582238	0.780911	1.455245	0.661489	0.693721
Maximum	1	1	1	2	1	1	1	1	2	2	2
Minimum	-5	-3	-4	-8	-8	-6	-9	-23	-82	-56	-82
S_{t-1}^Q											
Mean	0.4969	0.4926	0.4671	0.3800	0.3712	0.3431	0.3894	0.3883	0.3245	0.2871	0.3940
St. Deviation	0.2891	0.2838	0.2399	0.2277	0.2267	0.2088	0.2321	0.3683	0.3602	0.1740	0.2611
Maximum	1.8750	1.7500	1.7500	1.7500	2.0000	3.0000	7.6250	6.3750	10.6250	4.1250	10.6250
Minimum	0.1250	0.0625	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.0625
$\bar{V}$	948.64	1075.874	1132.09	1786.561	1233.542	1835.148	1801.951	1768.685	2275.628	2688.182	
ρ	-0.01884	-0.22691	-0.24679	-0.31285	-0.22041	-0.34697	-0.35576	-0.32281	-0.19808	-0.27152	

Table AG.5.2 - NASDAQ

PART A

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	60	60	60	60	60	60	60	60	60	60	600
Number of Matched trades	16	58	93	190	345	307	502	595	679	1,477	4,262
$\Delta\bar{S}_t^Q$											
Mean	0.256696	0.127997	0.114354	0.096836	0.097678	0.11275	0.141724	0.120949	0.154001	0.06798	0.129096
St. Deviation	0.583378	0.434871	0.589998	0.564847	0.53408	0.552435	0.587135	0.578437	0.640887	0.643433	0.57095
Maximum	1.5	1	3	2	2	2	3	2	3	4	4
Minimum	-0.66667	-0.5	-0.66667	-0.66667	-0.75	-0.66667	-0.66667	-0.66667	-0.66667	-2	-2
S_t^p											
Mean	0.2567	0.1280	0.1144	0.0955	0.0977	0.1128	0.1417	0.1201	0.1535	0.0669	0.1287
St. Deviation	0.6025	0.4387	0.5932	0.5639	0.5349	0.5533	0.5877	0.5795	0.6416	0.6445	0.5740
Maximum	1.5000	1.0000	3.0000	2.0000	2.0000	2.0000	3.0000	2.0000	3.0000	4.0000	4.0000
Minimum	-0.6667	-0.5000	-0.6667	-0.6667	-0.7500	-0.6667	-0.6667	-0.6667	-0.6667	-2.0000	-2.0000
V_{t-1}											
Mean	938	1,495	1,376	1,395	1,455	1,714	1,974	1,494	1,903	1,590	1,533
St. Deviation	894	1,998	1,763	1,553	1,636	2,503	6,141	1,849	3,416	3,319	2,507
Maximum	3,000	15,000	15,000	12,300	15,000	25,000	86,000	20,000	45,000	45,000	86,000
Minimum	100	100	100	100	100	100	100	100	100	100	100
ΔV_t^a											
Mean	-31	-9	0	16	-1	-15	-5	4	6	8	-3
St. Deviation	166	198	245	218	402	323	185	184	187	161	227
Maximum	300	500	500	500	4,500	500	800	800	2,000	1,500	4,500
Minimum	-500	-500	-500	-500	-4,500	-4,000	-800	-800	-800	-1,000	-4,500
ΔV_t^b											
Mean	0	26	11	-2	-3	25	-5	5	18	8	8
St. Deviation	258	237	208	184	205	218	172	197	216	157.4000	205
Maximum	500	500	500	500	500	800	800	800	2,000	1,500.0000	2,000
Minimum	-500	-500	-500	-800	-500	-500	-800	-800	-2,000	-1,000.0000	-2,000

Table AG.5.2 - PART B

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
$\Delta\tau_{t-1}$											
Mean	33	-321	-401	232	7	391	58	147	156	47	35
St. Deviation	7,140	4,280	3,195	3,059	2,609	2,966	2,368	1,632	1,430	605	2,929
Maximum	16,214	9,183	11,575	13,274	15,281	12,983	17,052	9,891	11,882	9,403	17,052
Minimum	-14,101	-14,239	-12,141	-10,428	-10,479	-13,289	-9,175	-10,004	-5,803	-8,199	-14,239
q_{t-1}											
Mean	-0.2500	-0.1379	0.0538	0.1368	0.1768	0.1270	0.0637	0.1462	0.1016	0.1903	0.0608
St. Deviation	1.0000	0.9991	1.0040	0.9932	0.9857	0.9935	0.9990	0.9901	0.9956	0.9821	0.9942
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000
$\Delta\bar{a}_t$											
Mean	0.096726	0.040189	0.028358	0.077563	0.062143	0.030293	0.053984	0.039172	0.052307	0.025396	0.050613
St. Deviation	0.46594	0.432109	0.463581	0.506143	0.452247	0.491491	0.511456	0.477928	0.519512	0.534725	0.485513
Maximum	1	1	2	3	2.5	3	4	2	3	4	4
Minimum	-0.66667	-1	-1	-1	-1.5	-2.5	-1	-2	-1	-3	-3
$\Delta\bar{b}_t$											
Mean	-0.15997	-0.08781	-0.086	-0.01927	-0.03553	-0.08246	-0.08774	-0.08178	-0.10169	-0.04258	-0.07848
St. Deviation	0.38695	0.31833	0.448077	0.414449	0.428334	0.525121	0.46211	0.441176	0.477692	0.489839	0.439208
Maximum	0.25	0.666667	0.5	1	1	2	1	1	1	3	3
Minimum	-1.5	-1	-3	-2	-2	-2	-2	-2	-2	-2	-3
S^Q_{t-1}											
Mean	0.6875	0.4289	0.4439	0.3796	0.4193	0.3153	0.3960	0.3693	0.3157	0.3308	0.4086
St. Deviation	0.4031	0.2397	0.2518	0.2190	0.3088	0.2274	0.2238	0.2076	0.2333	0.1912	0.2506
Maximum	1.7500	1.0000	1.0000	1.2500	3.0000	1.5000	1.2500	1.0000	1.0000	0.7500	3.0000
Minimum	0.1250	0.0625	0.0625	0.0000	0.0625	0.0313	0.0313	0.0000	0.0000	-0.2500	-0.2500
$\bar{V}$	937.50	1,494.83	1,376.34	1,394.74	1,454.78	1,714.01	1,974.30	1,494.29	1,903.24	1,590.12	
ρ	-0.33333	-0.23865	-0.71107	-0.16650	-0.13810	-0.09630	-0.22221	-0.18198	-0.22876	-0.10620	

G.2 Results from the GMM Estimations of the Quoted-Spread Model

Tables AG.5.3 and AG.5.4

Generalized Method of Moments Estimates of the Parameters of the Quoted Spread Model, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.5.3 and AG.5.4 present the Generalized Method of Moments estimates of the Quoted Spread Model described by equation (5.11b) for the NYSE and NASDAQ sub-samples respectively.

$$\Delta\bar{S}_t^{Q} = 2\delta_p q_{t-1} V_{t-1} + 2\beta_p \Delta\tau_{t-1} + \gamma_p\left(\Delta V_t^a + \Delta V_t^b\right) + \Delta\upsilon_t$$

where

$\Delta\bar{S}_t^{Q}$: change in the quoted spread at time t, expressed as a fraction of the spread at time t-1 (proportional change in the quoted spread).

V_{t-1} : volume of shares (in thousands) traded at time t-1

$\Delta\tau_{t-1}$: difference (seconds) between the time elapsed between trades at t-1 and t-2, τ_{t-1} and the time elapsed between trades at t-2 and t-3, τ_{t-2}

ΔV_t^a : Change in the number of shares (in thousands) offered between times t and t-1

ΔV_t^b : Change in the number of shares (in thousands) bid between times t and t-1

q_{t-1} : Trade Indicator Variable : equal to +1 when the trade is buyer-initiated
equal to -1 when the trade is seller-initiated

δ_p : Inventory-holding cost parameter expressed as a fraction of the spread at time t-1

β_p : Waiting time parameter expressed as a fraction of the spread at time t-1.

γ_p : Quoted-Volume parameter expressed as a fraction of the spread at time t-1

$\Delta\upsilon_t$: Error term at time t which is an i.i.d random variable

A total number of 200 common stocks for the NYSE and 600 common stocks for the NASDAQ sub-samples respectively are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. Trades for each decile are matched with preceding quotes and the above variables are calculated. The three parameters of the model, δ_p, β_p, and γ_p are estimated using the Generalized Method of Moments. Standard errors, followed by probability values appear in parentheses. The table also shows Hansen's J-Test statistic of Over-identifying Restrictions which is distributed as chi-square with degrees of freedom equal to the degree by which the number of parameters estimated is exceeded (over-identified) by the number of moments used in the estimation. For the Quoted-Spread Model this value is two. The confidence limits of this distribution are:

$$\Pr(\chi^2 > 4.60517) = 0.100$$
$$\Pr(\chi^2 > 5.99147) = 0.050$$
$$\Pr(\chi^2 > 9.21034) = 0.010$$

Table AG.5.3 - NYSE

Decile	1	2	3	4	5	6	7	8	9	The pan 10	Total Sample
Number of Stocks	20	20	20	20	20	20	20	20	20	20	200
Number of Matched trades	477	1,144	2,440	2,262	5,435	4,464	8,303	13,853	19,412	34,540	92,330
δ_p	-0.03579	-0.00396	-0.00065	-0.00024	-0.01667	-0.01188	-0.00937	-0.00114	-0.00060	-0.00003	-0.00019
	(0.02356)	(0.00930)	(0.00439)	(0.00135)	(0.00504)	(0.00362)	(0.00254)	(0.00304)	(0.00051)	(0.00007)	(0.00014)
	(0.12878)	(0.66958)	(0.88145)	(0.86186)	(0.00094)	(0.00104)	(0.00023)	(0.70732)	(0.24114)	(0.65573)	(0.15657)
β_p	-0.02451	0.01307	-0.00214	-0.02389	-0.01379	-0.00640	0.00340	0.08713	-0.02477	-0.07015	-0.00933
	(0.01698)	(0.01275)	(0.01173)	(0.01468)	(0.01360)	(0.01973)	(0.02437)	(0.03640)	(0.03803)	(0.03898)	(0.00584)
	(0.14894)	(0.30529)	(0.85526)	(0.10377)	(0.31084)	(0.74569)	(0.88897)	(0.01668)	(0.51486)	(0.07193)	(0.10997)
γ_p	-0.15277	-0.05153	-0.13764	-0.01523	-0.07010	-0.01539	-0.03783	-0.03307	-0.02325	-0.01132	-0.01499
	(0.03669)	(0.00865)	(0.01382)	(0.00261)	(0.00787)	(0.00170)	(0.00201)	(0.00172)	(0.00125)	(0.00031)	(0.00033)
	(0.00003)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
J-Statistic	1.4165	1.3697	0.72499	1.8111	5.6615	5.1248	2.9707	1.67	3.7552	61.032	59.709

Table AG.5.4 - NASDAQ

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	60	60	60	60	60	60	60	60	60	60	600
Number of Matched trades	16	58	93	190	345	307	502	595	679	1477	4,262
δ_p	0.08794	-0.00420	-0.00102	-0.00006	0.00140	0.00374	0.00250	-0.00301	-0.00660	-0.01932	-0.00664
	(0.03420)	(0.00921)	(0.01257)	(0.01396)	(0.00807)	(0.00713)	(0.00150)	(0.00793)	(0.00439)	(0.00504)	(0.00265)
	(0.01012)	(0.64843)	(0.93526)	(0.99628)	(0.86220)	(0.60009)	(0.09601)	(0.70439)	(0.13275)	(0.00013)	(0.01228)
β_p	0.02355	-0.01350	0.03581	-0.01307	0.02610	-0.00679	-0.01611	0.04149	0.00103	0.03874	0.00263
	(0.03934)	(0.02320)	(0.02006)	(0.02319)	(0.01482)	(0.01933)	(0.01631)	(0.03475)	(0.02836)	(0.03990)	(0.00794)
	(0.54942)	(0.56066)	(0.07418)	(0.57299)	(0.07811)	(0.72529)	(0.32328)	(0.23238)	(0.97109)	(0.33154)	(0.74066)
γ_p	1.23010	0.65127	0.62445	0.55034	0.23524	0.42105	0.47980	0.54583	0.52997	0.54951	0.48522
	(0.33838)	(0.14495)	(0.14787)	(0.14012)	(0.08726)	(0.08728)	(0.09473)	(0.07816)	(0.10064)	(0.06022)	(0.03912)
	(0.00028)	(0.00001)	(0.00002)	(0.00009)	(0.00702)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
J-Statistic	2.19350	1.61950	1.10490	1.45050	0.91845	2.76810	2.51090	0.31687	2.50130	0.39004	3.33320

G.3 Results from the GMM Estimations of the Offer- and Bid-Change Models

Tables AG.5.5 and AG.5.6

Generalized Method of Moments Estimates of the Parameters of the Offer-Change Model, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.5.5 and AG.5.6 present The Generalized Method of Moments estimates of the Offer-Change Model described by equation (5.9b) for the NYSE and NASDAQ sub-samples respectively.

$$\Delta \bar{a}_t = \theta_p q_{t-1} V_{t-1} - \theta_p \bar{V} \rho q_{t-1} + \delta_p q_{t-1} V_{t-1} + \beta_p \Delta\tau_{t-1} + \gamma_p \Delta V_t^a + \Delta\chi_t + \varpi_t$$

Where

$\Delta \bar{a}_t$: change in the ask prices quoted between times t and t-1, expressed as a fraction of the spread at time t-1 (proportional change in the quoted ask).

V_{t-1} : volume of shares (in thousands) traded at time t-1

$\Delta\tau_{t-1}$: difference (seconds) between the time elapsed between trades at t-1 and t-2, τ_{t-1} and the time elapsed between trades at t-2 and t-3, τ_{t-2}

V_t^a : Volume (in thousands) of shares offered between time t

$\bar{V}$: Average number of shares (in thousands)per transaction in a day

q_{t-1} : Trade Indicator Variable : equal to +1 when the trade is buyer-initiated
equal to -1 when the trade is seller-initiated

θ_p : Adverse-Selection cost parameter expressed as a fraction of the spread at time t-1

δ_p : Inventory-holding cost parameter expressed as a fraction of the spread at time t-1

β_p : Waiting time parameter expressed as a fraction of the spread at time t-1.

γ_p : Quoted-Volume parameter expressed as a fraction of the spread at time t-1

ρ : Coefficient of Correlation of the trade indicator variable

$\Delta\chi_t$: Random error term at time t assumed to be uniformly distributed with zero mean

ϖ_t : Random error term at time t assumed to be independently, identically distributed.

A total number of 200 common stocks for the NYSE and 600 common stocks for the NASDAQ sub-samples respectively are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. Trades for each decile are matched with preceding quotes and the above variables are calculated. The four parameters of the model, θ_p, δ_p, β_p, and γ_p are estimated using the Generalized Method of Moments and four moment conditions, therefore Hansen's J-Test of Over-identifying Restrictions is almost zero since the model is exactly identified. Standard errors, followed by probability values appear in parentheses.

Table AG.5.5 - NYSE

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	20	20	20	20	20	20	20	20	20	20	200
Number of Matched trades	477	1,144	2,440	2,262	5,435	4,464	8,303	13,853	19,412	34,540	92,330
θ_p	8.4647	0.62583	0.44173	0.23081	0.44645	0.17903	0.14812	0.02956	0.21251	0.097174	0.16666
	(7.14800)	(0.15656)	(0.04826)	(0.01947)	(0.15807)	(0.24291)	(0.01844)	(1.01590)	(0.00920)	(0.00437)	(0.00413)
	(0.23633)	(0.00006)	(0.00000)	(0.00000)	(0.00067)	(0.46110)	(0.00000)	(0.97679)	(0.00000)	(0.00000)	(0.00000)
δ_p	-8.4871	-0.65765	-0.4388	-0.22915	-0.46114	-0.19887	-0.14929	0.017288	-0.21278	-0.096721	-0.16627
	(7.27600)	(0.18699)	(0.05211)	(0.01961)	(0.18903)	(0.32665)	(0.02360)	(1.35330)	(0.00932)	(0.00438)	(0.00414)
	(0.24343)	(0.00044)	(0.00000)	(0.00000)	(0.00369)	(0.54265)	(0.00000)	(0.98981)	(0.00000)	(0.00000)	(0.00000)
β_p	-0.0078	0.024345	-0.00894	-0.04009	-0.01684	-0.00974	-0.01193	0.07942	0.038	-0.05165	0.00496
	(0.02012)	(0.01777)	(0.01609)	(0.02038)	(0.01590)	(0.02665)	(0.02607)	(0.11943)	(0.04182)	(0.04831)	(0.00844)
	(0.69743)	(0.17074)	(0.57825)	(0.04917)	(0.28962)	(0.71477)	(0.64736)	(0.50604)	(0.36358)	(0.28496)	(0.55678)
γ_p	-0.0891	-0.03096	-0.09375	-0.00731	-0.04195	-0.01060	-0.02529	-0.01892	-0.01508	-0.00730	-0.00999
	(0.02497)	(0.00902)	(0.01182)	(0.00127)	(0.00502)	(0.00191)	(0.00185)	(0.00345)	(0.00096)	(0.00028)	(0.00029)
	(0.00036)	(0.0006)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
J-Statistic	0	0	0	0	0	0	0	0	0	0	0

Table AG.5.6 - NASDAQ

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	60	60	60	60	60	60	60	60	60	60	600
Number of Matched trades	16	58	93	190	345	307	502	595	679	1477	4,262
θ_p	0.45205	0.21791	0.19505	0.72415	0.54849	0.35968	0.28268	0.10727	0.17611	0.30047	0.46441
	(1.33490)	(0.29941)	(0.08340)	(0.83395)	(0.32944)	(0.29291)	(0.04897)	(0.66854)	(0.09362)	(0.35046)	(0.12887)
	(0.73487)	(0.46673)	(0.01936)	(0.38521)	(0.09593)	(0.21946)	(0.00000)	(0.87253)	(0.05997)	(0.39124)	(0.00031)
δ_p	0.78009	-0.18547	-0.21997	-0.77249	-0.53714	-0.32954	-0.28355	-0.05502	-0.15885	-0.29026	-0.48064
	(1.69010)	(0.33894)	(0.12328)	(0.96663)	(0.36630)	(0.31196)	(0.04957)	(0.78703)	(0.10992)	(0.38401)	(0.15019)
	(0.64448)	(0.58424)	(0.07436)	(0.42420)	(0.14255)	(0.29082)	(0.00000)	(0.94427)	(0.14843)	(0.44973)	(0.00137)
β_p	-0.00489	0.01593	0.07047	-0.03648	0.01805	-0.00676	0.00398	0.04514	0.00755	0.01386	-0.00706
	(0.06356)	(0.03808)	(0.03783)	(0.03708)	(0.02799)	(0.02876)	(0.02550)	(0.05472)	(0.03594)	(0.05480)	(0.01691)
	(0.93865)	(0.67577)	(0.06247)	(0.32516)	(0.51895)	(0.81409)	(0.87602)	(0.40948)	(0.83357)	(0.80039)	(0.67621)
γ_p	1.32980	0.63009	0.62807	0.33898	0.17639	0.32715	0.59772	0.54893	0.50989	0.49449	0.38464
	(0.97434)	(0.18664)	(0.17946)	(0.26527)	(0.09570)	(0.09024)	(0.12717)	(0.11875)	(0.13702)	(0.10616)	(0.08930)
	(0.17230)	(0.00074)	(0.00047)	(0.20130)	(0.06530)	(0.00029)	(0.00000)	(0.00000)	(0.00020)	(0.00000)	(0.00002)
J-Statistic	0	0	0	0	0	0	0	0	0	0	0

Tables AG.5.7 and AG.5.8

Generalized Method of Moments Estimates of the Parameters of the Bid-Change Model, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.5.7 and AG.5.8 present The Generalized Method of Moments estimates of the Bid-Change Model described by equation (5.10b) for the NYSE and NASDAQ sub-samples respectively.

$$\Delta \bar{b}_t = \theta_p q_{t-1} V_{t-1} - \theta_p \bar{V} \rho q_{t-1} - \delta_p q_{t-1} V_{t-1} - \beta_p \Delta\tau_{t-1} - \gamma_p \Delta V_t^b + \Delta\lambda_t + \pi_t$$

where

$\Delta \bar{b}_t$: change in the bid prices quoted between times t and t-1, expressed as a fraction of the spread at time t-1 (proportional change in the quoted bid).

V_{t-1} : volume of shares (in thousands) traded at time t-1

$\Delta\tau_{t-1}$: difference (seconds) between the time elapsed between trades at t-1 and t-2, τ_{t-1} and the time elapsed between trades at t-2 and t-3, τ_{t-2}

ΔV_t^b : Change in the volume (in thousands) of shares bid between times t and t-1

$\bar{V}$: Average number of shares (in thousands) per transaction in a day

q_{t-1} : Trade Indicator Variable : equal to +1 when the trade is buyer-initiated
equal to -1 when the trade is seller-initiated

θ_p : Adverse-Selection cost parameter expressed as a fraction of the spread at time t-1

δ_p : Inventory-holding cost parameter expressed as a fraction of the spread at time t-1

β_p : Waiting time parameter expressed as a fraction of the spread at time t-1.

γ_p : Quoted-Volume parameter expressed as a fraction of the spread at time t-1

ρ : Coefficient of Correlation of the trade indicator variable

$\Delta\lambda_t$: Random error term at time t assumed to be uniformly distributed with zero mean

π_t : Random error term at time t assumed to be independently, identically distributed

A total number of 200 common stocks for the NYSE and 600 common stocks for the NASDAQ sub-samples respectively are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October 1994. Trades for each decile are matched with preceding quotes and the above variables are calculated. The four parameters of the model, θ_p, δ_p, β_p, and γ_p are estimated using the Generalized Method of Moments and four moment conditions, therefore Hansen's J-Test of Over-identifying Restrictions is almost zero since the model is exactly identified. Standard errors, followed by probability values appear in parentheses.

Table AG.5.7 - NYSE

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	20	20	20	20	20	20	20	20	20	20	200
Number of Matched trades	477	1,144	2,440	2,262	5,435	4,464	8,303	13,853	19,412	34,540	92,330
θ_p	7.34460	0.35823	0.48607	0.21297	0.59797	0.17703	0.18139	0.071948	0.23806	0.16240	0.23990
	(14.89900)	(0.13722)	(0.04265)	(36.64500)	(0.11777)	(0.10797)	(0.03014)	(0.84838)	(47.40300)	(46.93300)	(0.00795)
	(0.62203)	(0.00904)	(0.00000)	(0.99510)	(0.00000)	(0.10108)	(0.00000)	(0.93242)	(0.99599)	(0.99724)	(0.00000)
δ_p	7.2612	0.33766	0.48618	0.21027	0.60775	0.16876	0.17384	0.033872	0.23675	0.16195	0.23934
	(15.17200)	(0.16319)	(0.04517)	(46.38000)	(0.14075)	(0.14498)	(0.03965)	(1.12140)	(57.15900)	(60.34100)	(0.00795)
	(0.63224)	(0.03853)	(0.00000)	(0.99638)	(0.00002)	(0.24441)	(0.00001)	(0.97590)	(0.99670)	(0.99786)	(0.00000)
β_p	-0.0464	-0.00098	0.00002	-0.01448	-0.01640	-0.00559	0.02255	0.06221	-0.09868	-0.10054	-0.00039
	(0.02518)	(0.01414)	(0.01353)	(0.01866)	(0.01925)	(0.02201)	(0.03530)	(0.10123)	(0.05570)	(0.05515)	(0.00953)
	(0.06514)	(0.94471)	(0.99872)	(0.43770)	(0.39434)	(0.79944)	(0.52282)	(0.53889)	(0.07648)	(0.06829)	(0.96770)
γ_p	-0.0816	-0.02293	-0.09232	-0.00889	-0.04215	-0.00578	-0.02336	-0.01791	-0.01452	-0.00672	-0.00969
	(0.02357)	(0.00813)	(0.01186)	(0.00200)	(0.00593)	(0.00116)	(0.00195)	(0.00146)	(0.00122)	(0.00028)	(0.00033)
	(0.00053)	(0.00479)	(0.00000)	(0.00001)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
J-Statistic	0	0	0	0	0	0	0	0	0	0	0

Table AG.5.8 - NASDAQ

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	60	60	60	60	60	60	60	60	60	60	600
Number of Matched trades	16	58	93	190	345	307	502	595	679	1,477	4,262
θ_p	1.20150	0.18710	0.10803	0.14618	0.80523	0.75622	0.21629	0.09849	0.21892	0.45208	0.64536
	(3.13270)	(0.14265)	(0.06147)	(0.58528)	(0.37724)	(0.22364)	(0.04830)	(0.45572)	(0.09065)	(0.39374)	(0.21166)
	(0.70133)	(0.18967)	(0.07883)	(0.80278)	(0.03280)	(0.00072)	(0.00001)	(0.82889)	(0.01573)	(0.25090)	(0.00230)
δ_p	1.56430	0.13794	0.09884	0.10775	0.82934	0.75936	0.22139	0.03987	0.19266	0.41994	0.67816
	(4.03760)	(0.15520)	(0.09400)	(0.68108)	(0.42491)	(0.23483)	(0.04892)	(0.53529)	(0.10699)	(0.43341)	(0.24714)
	(0.69844)	(0.37412)	(0.29304)	(0.87430)	(0.05096)	(0.00122)	(0.00001)	(0.94062)	(0.07173)	(0.33259)	(0.00607)
β_p	-0.05444	-0.05176	0.02542	0.00341	0.03461	-0.01910	-0.03692	-0.04563	-0.02655	0.05832	0.00877
	(0.05481)	(0.03475)	(0.03782)	(0.02756)	(0.02259)	(0.02578)	(0.03566)	(0.08022)	(0.04429)	(0.05983)	(0.01671)
	(0.32063)	(0.13632)	(0.50147)	(0.90150)	(0.12546)	(0.45874)	(0.30050)	(0.56949)	(0.54884)	(0.32970)	(0.59953)
γ_p	0.33172	0.17952	0.36420	0.18501	0.32970	0.68467	0.25145	0.48802	0.32215	0.60387	0.43681
	(1.27210)	(0.11050)	(0.07535)	(0.14830)	(0.09101)	(0.12196)	(0.10252)	(0.08798)	(0.11863)	(0.08586)	(0.04166)
	(0.79427)	(0.10424)	(0.00000)	(0.21219)	(0.00029)	(0.00000)	(0.01419)	(0.00000)	(0.00662)	(0.00000)	(0.00000)
J-Statistic	0	0	0	0	0	0	0	0	0	0	0

G.4 Results from the Specification Tests of the Quoted-Spread, the Ask- and the Bid-Change Models

Tables AG.5.9 and AG.5.10

Eichenbaum, Hansen and Singleton (1988) C_T Statistic for testing the validity of moment conditions in the Quoted-Spread Model Equation, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.5.9 and AG.5.10 present the Eichenbaum, Hansen and Singleton (1988) C_T Statistic for the samples of NYSE and NASDAQ stocks of October 1994 respectively. The C_T Statistic tests the null hypothesis that fractions of the set of moment conditions used to estimate the Quoted Spread Model, described by equation (5.11b), by Generalized Method of Moments are true. For each exogenous variable omitted, which corresponds to a particular moment condition, entries in the table represent in the first line the J-statistic of over-identifying restrictions of Hansen, from the GMM estimation when the particular variable is excluded and in the second line the value of the C_T statistic, based on the same estimation. The C_T statistic is given by

$$C_T = T\left\{Q_T\left(\hat{\theta}_T\right) - Q_{IT}\left(\tilde{\theta}_{IT}\right)\right\}$$

Where

Q_T : is the value of the objective function minimized during GMM based on the full set of moment conditions and a sample of size T

Q_{IT} : is the value of the objective function minimized during GMM based on the set of moment conditions which are held to be true and on a sample of size T

$\hat{\theta}_T$: is the full parameter vector which minimizes the objective function through GMM estimation using the full set of moment conditions and a sample of size T

$\tilde{\theta}_{IT}$: is the part of the parameter vector which minimizes the objective function through GMM estimation using only that set of moment conditions which are held to be true based on a sample of size T

The statistic is distributed as chi-square with ν_1 degrees of freedom which is equal to the degree by which the number of parameters estimated is exceeded (over-identified) by the number of moments used in the estimation. For the Quoted-Spread Model when one of the moments is omitted this value is one. The confidence limits of this distribution are:

$$\Pr(\chi^2 > 2.70554) = 0.100$$
$$\Pr(\chi^2 > 3.84146) = 0.050$$
$$\Pr(\chi^2 > 6.63490) = 0.010$$

The J-statistic shown on the fourth line of the table is Hansen's J-Statistic for Over-identifying restrictions estimated using the full set of moment conditions, also presented in table AG.5.3 together with the definitions of the pertinent variables in the table.

Table AG.5.9 - NYSE

Decile	1	2	3	4	5	6	7	8	9	10
Number of Stocks	20	20	20	20	20	20	20	20	20	20
Number of Matched trades	477	1,144	2,440	2,262	5,435	4,464	8,303	13,853	19,412	34,540
J-Statistic	1.416500	1.369700	0.724990	1.811100	5.661500	5.124800	2.970700	1.670000	3.755200	61.032000
V_{t-1}										
J	1.410500	0.076558	0.072681	0.745240	5.154600	5.109600	2.154600	1.038700	3.329900	2.472800
C_T	0.006000	1.293142	0.652309	1.065860	0.506900	0.015200	0.816100	0.631300	0.425300	58.559200
ΔV_t^a										
J	0.001354	1.350100	0.561880	0.505330	0.662330	0.000257	0.774290	0.602270	1.165900	60.617000
C_T	1.415147	0.019600	0.163110	1.305770	4.999170	5.124543	2.196410	1.067730	2.589300	0.415000
ΔV_t^b										
J	0.000365	1.252000	0.534300	0.607840	0.528920	0.000267	0.959370	0.737800	1.218700	60.558000
C_T	1.416135	0.117700	0.190690	1.203260	5.132580	5.124533	2.011330	0.932200	2.536500	0.474000
$\Delta \tau_{t-1}$										
J	0.054994	0.032462	0.705661	1.204600	5.644100	0.521420	0.603630	0.054487	1.885100	1.191400
C_T	1.361506	1.337238	0.019329	0.606500	0.017400	4.603380	2.367070	1.615513	1.870100	59.840600
q_{t-1}										
J	0.554140	0.045850	0.119310	0.785840	5.341200	4.763800	1.375000	1.115200	3.332900	3.834100
C_T	0.862360	1.323850	0.605680	1.025260	0.320300	0.361000	1.595700	0.554800	0.422300	57.197900

Table AG.5.10 - NASDAQ

Decile	1	2	3	4	5	6	7	8	9	10
Number of Stocks	60	60	60	60	60	60	60	60	60	60
Number of Matched trades	16	58	93	190	345	307	502	595	679	1,477
J-Statistic	2.19350	1.61950	1.10490	1.45050	0.91845	2.76810	2.51090	0.31687	2.50130	0.39004
V_{t-1}										
J	2.180300	1.553300	0.343950	0.813840	0.524730	0.940820	1.091800	0.313750	2.311900	0.276200
C_T	0.013200	0.066200	0.760950	0.636660	0.393720	1.827280	1.419100	0.003120	0.189400	0.113840
ΔV_t^a										
J	2.044700	0.062690	1.022100	0.453270	0.431900	2.267300	1.278300	0.000478	0.160370	0.125990
C_T	0.148800	1.556810	0.082800	0.997230	0.486550	0.500800	1.232600	0.316392	2.340930	0.264050
ΔV_t^b										
J	1.796300	0.115500	1.086700	0.282290	0.300080	2.325400	1.252900	0.000755	0.100990	0.130650
C_T	0.397200	1.504000	0.018200	1.168210	0.618370	0.442700	1.258000	0.316115	2.400310	0.259390
$\Delta \tau_{t-1}$										
J	1.731900	1.044000	0.183690	0.311200	0.654440	0.838380	0.028846	0.165860	0.375090	0.276100
C_T	0.461600	0.575500	0.921210	1.139300	0.264010	1.929720	2.482054	0.151010	2.126210	0.113940
q_{t-1}										
J	1.544000	1.588300	0.259000	0.007133	0.698960	1.108300	1.139700	0.267830	1.827600	0.247940
C_T	0.649500	0.031200	0.845900	1.443367	0.219490	1.659800	1.371200	0.049040	0.673700	0.142100

Tables AG.5.11 and AG.5.12

GMM-BIC MSC Statistic for selecting moment conditions in the Quoted-Spread Model Equation, Decile 10, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.5.11 and AG.5.12 present the GMM-BIC (Bayesian Information Criterion) Moment Selection Criterion Statistic (MSC) reported in Andrews (1999) when one of the moment conditions used in the initial Generalized Method of Moments estimation is excluded from the set used to estimate the Quoted Spread Model described by equation (5.11b). For each exogenous variable omitted, which corresponds to a particular moment condition, entries in the table represent in the first line the J-statistic of over-identifying restrictions of Hansen, from the GMM estimation when the particular variable is excluded and in the second line the value of the MSC_{BIC} statistic, based on the same estimation, of the sample of NYSE and NASDAQ stocks respectively of October 1994. The MSC_{BIC} statistic is given by

$$MSC_{BIC,n}(c) = J_n(c) - (|c| - p)\log n$$

Where

$J_n(c)$: is the value of the Hansen J-Test statistic of over-identifying restrictions estimated using the moment vector c and a sample of size n.

$|c|$: is the number of moments used in the moment-selection vector c.

c : moment-selection vector (selects only certain moments)

p : is the dimension (number) of the parameter vector, θ, to be estimated by GMM

n : is number of data in the sample

The $J_n(c)$ statistic is distributed as chi-square with $[|c| - \min(p, |c|)]$ degrees of freedom under the null hypothesis that all moment conditions in c are correct. Since four moments are used to estimate three parameters there is one degree of freedom. Thus the confidence limits of this distribution are:

$$\Pr(\chi^2 > 2.70554) = 0.100$$
$$\Pr(\chi^2 > 3.84146) = 0.050$$
$$\Pr(\chi^2 > 6.63490) = 0.010$$

The definitions of the variables in this table are given in tables AG.5.3. and AG.5.4

Tables AG.5.11 - NYSE

Decile	1	2	3	4	5	6	7	8	9	10
Moment Omitted										
V_{t-1}										
J	1.410500	0.076558	0.072681	0.745240	5.154600	5.109600	2.154600	1.038700	3.329900	2.472800
MSC	*-1.268018*	*-2.981868*	*-3.314709*	*-2.609253*	*1.419400*	*1.459876*	*-1.764635*	*-3.102844*	*-0.958170*	*-2.065522*
ΔV_t^a										
J	0.001354	1.350100	0.561880	0.505330	0.662330	0.000257	0.774290	0.602270	1.165900	60.617000
MSC	*-2.677165*	*-1.708326*	*-2.825510*	*-2.849163*	*-3.072870*	*-3.649467*	*-3.144945*	*-3.539274*	*-3.122170*	*56.078678*
ΔV_t^b										
J	0.000365	1.252000	0.534300	0.607840	0.528920	0.000267	0.959370	0.737800	1.218700	60.558000
MSC	*-2.678153*	*-1.806426*	*-2.853090*	*-2.746653*	*-3.206280*	*-3.649457*	*-2.959865*	*-3.403744*	*-3.069370*	*56.019678*
$\Delta \tau_{t-1}$										
J	0.054994	0.032462	0.705661	1.204600	5.644100	0.521420	0.603630	0.054487	1.885100	1.191400
MSC	*-2.623524*	*-3.025964*	*-2.681729*	*-2.149893*	*1.908900*	*-3.128304*	*-3.315605*	*-4.087057*	*-2.402970*	*-3.346922*
q_{t-1}										
J	0.554140	0.045850	0.119310	0.785840	5.341200	4.763800	1.375000	1.115200	3.332900	3.834100
MSC	*-2.124378*	*-3.012576*	*-3.268080*	*-2.568653*	*1.606000*	*1.114076*	*-2.544235*	*-3.026344*	*-0.955170*	*-0.704222*

Table AG.5.12 - NASDAQ

Moment Omitted / Decile	1	2	3	4	5	6	7	8	9	10
V_{t-1}										
J	2.180300	1.553300	0.343950	0.813840	0.524730	0.940820	1.091800	0.313750	2.311900	0.276200
MSC	*0.976180*	*-0.210128*	*-1.624533*	*-1.464914*	*-2.013089*	*-1.546318*	*-1.608904*	*-2.460767*	*-0.519970*	*-2.893180*
ΔV_t^a										
J	2.044700	0.062690	1.022100	0.453270	0.431900	2.267300	1.278300	0.000478	0.160370	0.125990
MSC	*0.840580*	*-1.700738*	*-0.946383*	*-1.825484*	*-2.105919*	*-0.219838*	*-1.422404*	*-2.774039*	*-2.671500*	*-3.043390*
ΔV_t^b										
J	2.49 two am due to 1600	0.115500	1.086700	0.282290	0.300080	2.325400	1.252900	0.000755	0.100990	0.130650
MSC	*1.287480*	*-1.647928*	*-0.881783*	*-1.996464*	*-2.237739*	*-0.161738*	*-1.447804*	*-2.773762*	*-2.730880*	*-3.038730*
$\Delta \tau_{t-1}$										
J	1.731900	1.044000	0.183690	0.311200	0.654440	0.838380	0.028846	0.165860	0.375090	0.276100
MSC	*0.527780*	*-0.719428*	*-1.784793*	*-1.967554*	*-1.883379*	*-1.648758*	*-2.671858*	*-2.608657*	*-2.456780*	*-2.893280*
q_{t-1}										
J	1.544000	1.588300	0.259000	0.007133	0.698960	1.108300	1.139700	0.267830	1.827600	0.475230
MSC	*0.339880*	*-0.175128*	*-1.709483*	*-2.271621*	*-1.838859*	*-1.378838*	*-1.561004*	*-2.506687*	*-1.004270*	*-2.694150*

Tables AG.5.13 and AG.5.14

Likelihood Ratio (LR) Statistic for the parameter vector of the Quoted-Spread Model Equation, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.5.13 and AG.5.14 present the Likelihood Ratio (LR) Statistic for the joint hypothesis that all parameters in the Quoted Spread Model described by equation (5.11b) are equal to zero for the NYSE and NASDAQ sub-samples of common stock data of October 1994 respectively. The LR statistic is given by

$$LR_T = T\left[Q_T\left(\bar{\theta}_T\right) - Q_T\left(\hat{\theta}_T\right)\right]$$

where

Q_T : is the value of the objective function

$\hat{\theta}_T$: is the unrestricted estimator of the parameter vector

$\bar{\theta}_T$: is the restricted estimator of the parameter vector, and

T : is the number of observations

Under the null hypothesis that all restrictions are valid the LR statistic is distributed as a chi-square with degrees of freedom equal to the number of restrictions imposed. For the Quoted Spread Model which involves three variables there are three degrees of freedom when the null hypothesis assumes that all parameters are equal to zero. Thus the confidence limits of this distribution are:

$$\Pr(\chi^2 > 6.25139) = 0.100$$
$$\Pr(\chi^2 > 7.81473) = 0.050$$
$$\Pr(\chi^2 > 11.3449) = 0.010$$

Table AG.5.13 - NYSE

J-Statistic / Decile	1	2	3	4	5	6	7	8	9	10
Unrestricted Model	1,4165	1,3697	0,72499	1,8111	5,6615	5,1248	2,9707	1,67	3,7552	61,032
Restricted Model	26,652	86,052	150,06	145,21	436,1	261,26	652,86	774,75	853,45	2185,8
LR Statistic	25,2355	84,6823	149,33501	143,3989	430,4385	256,1352	649,8893	773,08	849,6948	2124,768

Table AG.5.14 - NASDAQ

J-Statistic / Decile	1	2	3	4	5	6	7	8	9	10
Unrestricted Model	2,1935	1,6195	1,1049	1,4505	0,9185	2,7681	2,5109	0,3169	2,5013	0,3900
Restricted Model	5,2274	6,6290	8,4269	14,1940	28,7400	27,8970	34,6280	43,4600	67,2930	71,8440
LR Statistic	3,03390	5,00950	7,32200	12,74350	27,82155	25,12890	32,11710	43,14313	64,79170	71,45396

Tables AG.5.15 and AG.5.16

Likelihood Ratio (LR) Statistic for the parameter vector of the Offer-Change and Bid-Change Model Equations, NYSE and NASDAQ TAQ data: October 1994.

Table AG.5.15 and AG.5.16 present the Likelihood Ratio (LR) Statistic for the joint hypothesis that all parameters in the Offer-Change and Bid-Change Models (described by equations (5.9b) and (5.10b) respectively) are equal to zero for the NYSE and NASDAQ sub-samples of common stock data for October 1994 respectively. The LR statistic is given by

$$LR_T = T\left[Q_T\left(\bar{\theta}_T\right) - Q_T\left(\hat{\theta}_T\right)\right]$$

where

Q_T : is the value of the objective function

$\hat{\theta}_T$: is the unrestricted estimator of the parameter vector

$\bar{\theta}_T$: is the restricted estimator of the parameter vector, and

T : is the number of observations

Under the null hypothesis that all restrictions are valid the LR statistic is distributed as a chi-square with degrees of freedom equal to the number of restrictions imposed. For the Quoted Spread Model which involves three variables there are three degrees of freedom when the null hypothesis assumes that all parameters are equal to zero. Thus the confidence limits of this distribution are:

$\Pr(\chi^2 > 6.25139) = 0.100$
$\Pr(\chi^2 > 7.81473) = 0.050$
$\Pr(\chi^2 > 11.3449) = 0.010$

Table AG.5.15 - NYSE

Offer-Change Model

Decile	1	2	3	4	5	6	7	8	9	10
J-Statistic										
Unrestricted Model	0	0	0	0	0	0	0	0	0	0
Restricted Model	40,123	88,152	161,8	165,67	366,86	236,64	565,49	728,67	1209,7	1987,9
LR Statistic	40,123	88,152	161,8	165,67	366,86	236,64	565,49	728,67	1209,7	1987,9

Bid-Change Model

Decile	1	2	3	4	5	6	7	8	9	10
J-Statistic										
Unrestricted Model	0	0	0	0	0	0	0	0	0	0
Restricted Model	39,519	70,922	167,4	141,54	388,15	275,26	551,08	642,19	548,43	1615
LR Statistic	39,519	70,922	167,4	141,54	388,15	275,26	551,08	642,19	548,43	1615

Table AG.5.16 - NASDAQ

Offer-Change Model

Decile	1	2	3	4	5	6	7	8	9	10
J-Statistic										
Unrestricted Model	0	0	0	0	0	0	0	0	0	0
Restricted Model	4,5819	7,4906	12,8680	11,3760	41,0400	21,8640	32,4930	42,0590	43,7540	42,6960
LR Statistic	4,5819	7,4906	12,8680	11,3760	41,0400	21,8640	32,4930	42,0590	43,7540	42,6960

Bid-Change Model

Decile	1	2	3	4	5	6	7	8	9	10
J-Statistic										
Unrestricted Model	0	0	0	0	0	0	0	0	0	0
Restricted Model	3,5111	10,4540	13,4960	7,8931	34,4520	24,5800	30,8000	47,3850	65,5480	106,4700
LR Statistic	3,5111	10,4540	13,4960	7,8931	34,4520	24,5800	30,8000	47,3850	65,5480	106,4700

G.5 Summary Statistics of the Variables of the Price-Change Model

Tables AG.6.1 and AG.6.2

Summary Statistics of the Variables of the Price-Change Model, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.6.1 and AG.6.2 present Summary Statistics of the Variables of the Price-Change Model described by equations (6.5b) for the random samples of NYSE and NASDAQ Common Stock data respectively in October 1994.

Where

$\Delta \bar{P}_t$	: Change in the transaction prices between times t and t-1, expressed as a fraction of the quoted spread at time t-1.
S_t^Q	: Spread quoted at time t
S_R	: Average Realized spread for a decile expressed as a percentage of the average quoted spread for the same decile.
V_{t-1}	: Volume of shares traded at time t-1
τ_{t-1}	: Time (seconds) elapsed between trades at times t-1 and t-2
V_t^a	: Volume (number) of shares offered at time t
V_t^b	: Volume (number) of shares bid at time t
$\bar{V}$	: Average number of shares per transaction in a day
q_t	: Trade Indicator Variable at time t : equal to +1 when the trade is buyer-initiated equal to -1 when the trade is seller-initiated
I_t^a, I_t^b	: Indicator Variables : equal to +1(0) when the trade at time t is buyer-initiated and 0(1) otherwise
ρ	: First-order autocorrelation Coefficient of the trade indicator variable, q_t

Table AG.6.1 - NYSE

PART A

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	20	20	20	20	20	20	20	20	20	20	200
Number of Matched trades	477	1,144	2,440	2,262	5,435	4,464	8,303	13,853	19,412	34,540	92,330
$\Delta \bar{P}_t$											
Mean	0.006838	0.003236	-0.00146	-0.01704	-0.01081	0.006176	-0.00433	-0.00364	0.002721	0.001067	-0.00172
St. Deviation	0.395746	0.31369	0.330626	0.402997	0.383707	0.389308	0.322863	0.379111	0.393971	0.406056	0.371807
Maximum	2	1	2	3	2.666667	2	2	2	3	4	4
Minimum	-2	-2	-3	-4.5	-2	-2.66667	-2.33333	-10	-3.5	-8	-10
V_{t-1}											
Mean	949	1,076	1,132	1,787	1,234	1,835	1,802	1,769	2,276	2,688	1,655
St. Deviation	1,882	2,379	3,038	7,626	3,551	6,445	5,334	5,779	12,844	32,185	8,106
Maximum	25,000	42,500	91,500	298,300	80,000	200,000	194,100	200,900	1,000,000	5,650,000	5,650,000
Minimum	100	100	100	100	100	100	100	100	100	100	100
τ_{t-1}											
Mean	2,886	1,972	1,425	1,364	810	826	472	314	202	91	1,036
St. Deviation	3,278	2,475	1,936	1,880	1,233	1,268	688	457	332	210	1,376
Maximum	18,766	16,949	20,378	19,085	17,650	16,791	8,103	15,489	7,366	5,784	20,378
Minimum	3	1	1	1	1	1	1	1	1	1	1
τ_{t-2}											
Mean	2,237	1,564	1,163	1,120	645	654	370	245	154	65	822
St. Deviation	2,854	2,139	1,709	1,691	1,097	1,098	606	395	290	173	1,205
Maximum	14,458	15,719	15,507	15,123	17,650	16,390	7,493	6,912	7,366	5,784	17,650
Minimum	1	1	1	2	1	1	1	3	1	1	1

Table AG.6.1 - PART B

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
V_t^a											
Mean	693	919	574	3,400	1,204	3,462	1,559	2,544	2,637	4,774	2,176
St. Deviation	1,417	2,747	1,647	10,422	2,752	8,900	4,499	6,367	6,729	11,623	5,710
Maximum	11,100	37,200	16,000	87,500	40,000	99,900	87,400	80,000	99,900	99,900	99,900
Minimum	100	100	100	100	100	100	100	100	100	100	100
V_t^b											
Mean	571	1,107	542	2,323	1,337	3,889	1,594	2,980	2,505	4,631	2,148
St. Deviation	1,219	3,353	1,406	7,011	3,854	10,908	4,538	8,356	7,407	12,079	6,013
Maximum	10,000	36,100	13,000	99,900	50,000	99,900	50,000	98,000	99,900	99,900	99,900
Minimum	100	100	100	100	100	100	100	100	100	100	100
V_{t-1}^a											
Mean	638	905	550	3,470	1,193	3,288	1,557	2,545	2,731	5,247	2,212
St. Deviation	1,433	2,587	1,626	10,858	2,675	8,512	4,407	6,342	6,863	12,170	5,747
Maximum	14,000	30,000	16,000	87,500	39,000	99,900	87,400	80,000	99,900	99,900	99,900
Minimum	100	100	100	100	100	100	100	100	100	100	100
V_{t-1}^b											
Mean	613	1,112	487	2,307	1,319	3,562	1,612	2,893	2,543	5,013	2,146
St. Deviation	1,329	3,319	1,261	7,153	3,740	10,045	4,536	8,033	7,123	12,389	5,893
Maximum	13,500	36,600	12,100	99,900	50,000	99,900	60,000	83,700	99,900	99,900	99,900
Minimum	100	100	100	100	100	100	100	100	100	100	100
q_t											
Mean	0.0901	0.1014	0.1615	0.0681	0.1930	0.1465	0.1642	0.1157	0.1974	0.2454	0.1483
St. Deviation	0.9970	0.9953	0.9871	0.9979	0.9813	0.9893	0.9865	0.9933	0.9804	0.9694	0.9877
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000

Table AG.6.1 - PART C

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
q_{t-1}											
Mean	0.0440	0.0787	0.1443	0.0866	0.1787	0.1281	0.1625	0.1150	0.1962	0.2589	0.1393
St. Deviation	1.0001	0.9973	0.9897	0.9965	0.9840	0.9919	0.9868	0.9934	0.9806	0.9659	0.9886
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000
q_{t-2}											
Mean	0.2663	0.2780	0.2754	0.2644	0.3152	0.3069	0.3511	0.3031	0.3833	0.4537	0.3197
St. Deviation	0.9649	0.9610	0.9615	0.9646	0.9491	0.9519	0.9364	0.9530	0.9237	0.8912	0.9457
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000
I_t^a											
Mean	0.5451	0.5507	0.5807	0.5340	0.5965	0.5733	0.5821	0.5579	0.5987	0.6227	0.5742
St. Deviation	0.4985	0.4976	0.4935	0.4990	0.4906	0.4947	0.4933	0.4967	0.4902	0.4847	0.4939
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
I_{t-1}^a											
Mean	0.5220	0.5393	0.5721	0.5433	0.5893	0.5641	0.5812	0.5575	0.5981	0.6294	0.5696
St. Deviation	0.5000	0.4987	0.4949	0.4982	0.4920	0.4959	0.4934	0.4967	0.4903	0.4830	0.4943
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table AG.6.1 - PART D

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
I_t^b											
Mean	0.4549	0.4493	0.4193	0.4660	0.4035	0.4268	0.4179	0.4421	0.4013	0.3773	0.4258
St. Deviation	0.4985	0.4976	0.4935	0.4990	0.4906	0.4947	0.4933	0.4967	0.4902	0.4847	0.4939
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
I_{t-1}^b											
Mean	0.4780	0.4607	0.4279	0.4567	0.4107	0.4359	0.4188	0.4425	0.4019	0.3706	0.4304
St. Deviation	0.5000	0.4987	0.4949	0.4982	0.4920	0.4959	0.4934	0.4967	0.4903	0.4830	0.4943
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
S_{t-1}^Q											
Mean	0.4969	0.4926	0.4671	0.3800	0.3712	0.3431	0.3894	0.3883	0.3245	0.2871	0.3940
St. Deviation	0.2891	0.2838	0.2399	0.2277	0.2267	0.2088	0.2321	0.3683	0.3602	0.1740	0.2611
Maximum	1.8750	1.7500	1.7500	1.7500	2.0000	3.0000	7.6250	6.3750	10.6250	4.1250	10.6250
Minimum	0.1250	0.0625	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.1250	0.0625
S_R	37.37	32.19	35.08	41.27	42.63	41.87	36.51	37.14	41.97	42.22	
$\bar{V}$	948.64	1075.874	1132.09	1786.561	1233.542	1835.148	1801.951	1768.685	2275.628	2688.182	
ρ	-0.01884	-0.22691	-0.24679	-0.31285	-0.22041	-0.34697	-0.35576	-0.32281	-0.19808	-0.27152	

Table AG.6.2 - NASDAQ

PART A

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	60	60	60	60	60	60	60	60	60	60	600
Number of Matched trades	16	58	93	190	345	307	502	595	679	1477	4,262
$\Delta \bar{P}_t$											
Mean	-0.18992	0.068432	0.006395	0.004298	0.043749	-0.04806	-0.02049	-0.01088	-0.0059	-0.02534	-0.01777
St. Deviation	0.551453	0.411186	0.462185	0.542679	0.550299	0.574167	0.537494	0.533463	0.547224	0.602517	0.531267
Maximum	0.875	1	1	3	2	2	4	3	3	3	4
Minimum	-1.5	-0.5	-1.5	-1.5	-2	-2.5	-2	-3	-2	-3	-3
V_{t-1}											
Mean	938	1,495	1,376	1,395	1,455	1,714	1,974	1,494	1,903	1,590	1,533
St. Deviation	894	1,998	1,763	1,553	1,636	2,503	6,141	1,849	3,416	3,319	2,507
Maximum	3,000	15,000	15,000	12,300	15,000	25,000	86,000	20,000	45,000	45,000	86,000
Minimum	100	100	100	100	100	100	100	100	100	100	100
τ_{t-1}											
Mean	3,097	1,885	1,034	1,451	1,018	1,415	881	624	509	163	1,208
St. Deviation	4,326	3,232	1,848	2,595	2,173	2,489	1,841	1,345	1,332	567	2,175
Maximum	16,255	13,175	11,671	13,306	15,330	13,226	17,177	10,017	14,947	9,765	17,177
Minimum	103	3	4	1	1	1	1	1	2	1	1
τ_{t-2}											
Mean	3,064	2,206	1,435	1,219	1,011	1,025	823	478	353	116	1,173
St. Deviation	4,449	3,364	2,596	2,129	1,983	1,989	1,763	1,155	828	437	2,069
Maximum	14,408	14,269	12,245	10,877	10,685	14,583	13,089	10,015	5,863	8,214	14,583
Minimum	3	0	2	1	1	1	1	1	1	1	1

Table AG.6.2 - PART B

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
V_t^a											
Mean	794	845	866	795	832	830	885	934	945	969	869
St. Deviation	284	233	223	247	330	239	220	179	205	147	231
Maximum	1,000	1,000	1,000	1,000	5,000	1,000	1,000	1,000	2,500	2,000	5,000
Minimum	200	500	500	500	200	200	200	200	200	200	200
V_t^b											
Mean	688	836	898	830	825	843	892	928	968	973	868
St. Deviation	250	237	203	240	239	234	211	180	175	135	210
Maximum	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	2,500	2,000	2,500
Minimum	500	500	500	200	500	200	200	200	500	200	200
V_{t-1}^a											
Mean	825	853	866	779	834	845	890	930	939	962	872
St. Deviation	277	230	223	249	330	338	212	184	187	160	239
Maximum	1,000	1,000	1,000	1,000	5,000	5,000	1,000	1,000	2,500	2,000	5,000
Minimum	200	500	500	500	200	200	200	200	200	200	200
V_{t-1}^b											
Mean	688	810	887	832	828	817	897	923	949	964	860
St. Deviation	250	245	210	237	238	252	207	188	201	161	219
Maximum	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	2,500	2,000	2,500
Minimum	500	500	500	500	500	200	200	200	500	200	200
q_t											
Mean	-0.5000	0.1379	0.1613	0.0421	0.1594	0.0489	0.0478	0.1563	0.2224	0.1605	0.0637
St. Deviation	0.8944	0.9991	0.9923	1.0018	0.9886	1.0004	0.9999	0.9885	0.9757	0.9874	0.9828
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000

Table AG.6.2 - PART C

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
q_{t-1}											
Mean	-0.2500	-0.1379	0.0538	0.1368	0.1768	0.1270	0.0637	0.1462	0.1016	0.1903	0.0608
St. Deviation	1.0000	0.9991	1.0040	0.9932	0.9857	0.9935	0.9990	0.9901	0.9956	0.9821	0.9942
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000
q_{t-2}											
Mean	0.2500	0.3448	0.5054	0.3790	0.3913	0.4007	0.4422	0.4555	0.5346	0.5464	0.4250
St. Deviation	1.0000	0.9469	0.8676	0.9279	0.9216	0.9177	0.8978	0.8910	0.8457	0.8378	0.9054
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000
I_t^a											
Mean	0.2500	0.5690	0.5807	0.5211	0.5797	0.5244	0.5239	0.5782	0.6112	0.5802	0.5318
St. Deviation	0.4472	0.4996	0.4961	0.5009	0.4943	0.5002	0.4999	0.4943	0.4878	0.4937	0.4914
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
I_{t-1}^a											
Mean	0.3750	0.4310	0.5269	0.5684	0.5884	0.5635	0.5319	0.5731	0.5508	0.5951	0.5304
St. Deviation	0.5000	0.4996	0.5020	0.4966	0.4928	0.4968	0.4995	0.4950	0.4978	0.4910	0.4971
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table AG.6.2 - PART D

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
I_t^b											
Mean	0.7500	0.4310	0.4194	0.4790	0.4203	0.4756	0.4761	0.4219	0.3888	0.4198	0.4682
St. Deviation	0.4472	0.4996	0.4961	0.5009	0.4943	0.5002	0.4999	0.4943	0.4878	0.4937	0.4914
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
I_{t-1}^b											
Mean	0.6250	0.5690	0.4731	0.4316	0.4116	0.4365	0.4681	0.4269	0.4492	0.4049	0.4696
St. Deviation	0.5000	0.4996	0.5020	0.4966	0.4928	0.4968	0.4995	0.4950	0.4978	0.4910	0.4971
Maximum	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimum	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
S_{t-1}^Q											
Mean	0.6875	0.4289	0.4439	0.3796	0.4193	0.3153	0.3960	0.3693	0.3157	0.3308	0.4086
St. Deviation	0.4031	0.2397	0.2518	0.2190	0.3088	0.2274	0.2238	0.2076	0.2333	0.1912	0.2506
Maximum	1.7500	1.0000	1.0000	1.2500	3.0000	1.5000	1.2500	1.0000	1.0000	0.7500	3.0000
Minimum	0.1250	0.0625	0.0625	0.0000	0.0625	0.0313	0.0313	0.0000	0.0000	-0.2500	-0.2500
S_R	54.55	49.58	63.09	71.68	63.41	61.92	59.11	56.75	57.09	69.32	
$\bar{V}$	937.50	1,494.83	1,376.34	1,394.74	1,454.78	1,714.01	1,974.30	1,494.29	1,903.24	1,590.12	
ρ	-0.33333	-0.23865	-0.71107	-0.16650	-0.13810	-0.09630	-0.22221	-0.18198	-0.22876	-0.10620	

G.6 Results from the GMM Estimations of the Price-Change Model

Tables AG.6.3 and AG.6.4

Generalized Method of Moments Estimates of the Parameters of the Price Change Model, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.6.3 and AG.6.4 present the Generalized Method of Moments estimates of the Price Change Model described by equation (6.5b) for the NYSE and NASDAQ samples of data respectively in October 1994.

$$\Delta \bar{P}_t = Bq_t - Cq_{t-1} - \theta\rho \bar{V} q_{t-2} + \kappa_t$$

where

$$B = \alpha_p + \beta_p \tau_{t-1} + \frac{\gamma_p}{2}\left(V_t^a + V_t^b + q_t\left(V_t^a - V_t^b\right) + \theta_p V_t\right)$$

$$C = \alpha_p + \beta_p \tau_{t-2} + \frac{\gamma_p}{2}\left(V_{t-1}^a + V_{t-1}^b + q_{t-1}\left(V_{t-1}^a - V_{t-1}^b\right)\right) + \rho\theta_p \bar{V} - \delta_p V_{t-1}$$

$\Delta \bar{P}_t$: change in the transaction price between times t and t-1, expressed as a fraction of the spread at time t-1 (proportional change in price).

V_{t-1} : volume of shares (in thousands) traded at time t-1

τ_{t-1} : time (hours) elapsed between trades at t-1 and t-2

V_t^a : Volume (thousands) of shares offered at time t

V_t^b : Volume (thousands) of shares bid at time t

$\bar{V}$: Average number (in thousands) of shares per transaction in a day

q_{t-1} : Trade Indicator Variable : equal to +1 when the trade is buyer-initiated
equal to -1 when the trade is seller-initiated

α_p : Fixed Order Processing Cost parameter for a single transaction expressed as a fraction of the spread at time t-1

θ_p : Adverse-Selection cost parameter expressed as a fraction of the spread at time t-1

δ_p : Inventory-holding cost parameter expressed as a fraction of the spread at time t-1

β_p : Waiting time parameter expressed as a fraction of the spread at time t-1.

γ_p : Quoted-Volume parameter expressed as a fraction of the spread at time t-1

ρ : Coefficient of Correlation of the trade indicator variable

κ_t : Random error term at time t assumed to be independently, identically distributed.

A total number of 200 common stocks for the NYSE and 600 common stocks for the NASDAQ sub-samples respectively are randomly selected from the TAQ database and grouped under deciles of trading activity based on the total number of shares traded in October. Trades for each decile are matched with preceding quotes and the above variables are calculated. The five parameters of the model, α_p, θ_p, δ_p, β_p, and γ_p are estimated using the Generalized Method of Moments. Standard errors, followed by probability values appear in parentheses. The table also shows Hansen's J-Test of Over-identifying Restrictions which is distributed as chi-square with degrees of freedom equal to the degree by which the number of parameters estimated is exceeded (over-identified) by the number of moments used in the estimation. For the Price-Change Model this value is five. The confidence limits of this distribution are:

$$\Pr(\chi^2 > 9.23635) = 0.100$$
$$\Pr(\chi^2 > 11.0705) = 0.050$$
$$\Pr(\chi^2 > 15.0863) = 0.010$$

Table AG.6.3 - NYSE

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	20	20	20	20	20	20	20	20	20	20	200
Number of Matched trades	477	1,144	2,440	2,262	5,435	4,464	8,303	13,853	19,412	34,540	92,330
α_p	-0.35299	0.10187	0.12234	0.21134	0.16222	0.18582	0.14358	0.16429	0.04001	0.18371	0.03919
	(0.39896)	(0.05516)	(0.02278)	(0.08134)	(0.02820)	(0.03010)	(0.02123)	(0.01541)	(0.01335)	(1.08360)	(0.00648)
	(0.37629)	(0.06475)	(0.00000)	(0.00937)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00272)	(0.86537)	(0.00000)
θ_p	0.11695	0.02892	0.06352	0.00162	0.05841	0.00947	0.02636	0.00813	0.00295	0.00815	0.00136
	(0.11883)	(0.02104)	(0.00890)	(0.00204)	(0.01470)	(0.00873)	(0.00593)	(0.00889)	(0.00082)	(0.32868)	(0.00016)
	(0.32504)	(0.16925)	(0.00000)	(0.42639)	(0.00007)	(0.27785)	(0.00001)	(0.36036)	(0.00031)	(0.98022)	(0.00000)
δ_p	-0.08323	0.03112	0.02001	0.00471	-0.00548	0.02023	0.00356	0.02265	0.00199	0.00844	0.00058
	(0.15679)	(0.01399)	(0.00785)	(0.00265)	(0.01617)	(0.01234)	(0.00637)	(0.01076)	(0.00066)	(0.36795)	(0.00020)
	(0.59552)	(0.02611)	(0.01084)	(0.07549)	(0.73465)	(0.10097)	(0.57620)	(0.03533)	(0.00260)	(0.98170)	(0.00297)
β_p	0.65221	0.03884	-0.01234	0.99304	-0.15354	-0.07107	-0.20435	-0.10719	2.63860	-0.14867	2.22070
	(0.48225)	(0.13594)	(0.06433)	(0.21390)	(0.11037)	(0.08294)	(0.18200)	(0.12639)	(0.24607)	(37.17800)	(0.06703)
	(0.17624)	(0.77508)	(0.83693)	(0.00000)	(0.16418)	(0.39153)	(0.26152)	(0.39639)	(0.00000)	(0.99689)	(0.00000)
γ_p	0.00837	0.01881	0.00534	0.01725	0.01403	0.01250	0.00239	0.01053	0.00611	0.00433	0.00385
	(0.00641)	(0.00770)	(0.01004)	(0.00834)	(0.00508)	(0.00302)	(0.00226)	(0.00182)	(0.00139)	(0.06182)	(0.00040)
	(0.19198)	(0.01462)	(0.59489)	(0.03861)	(0.00574)	(0.00004)	(0.29013)	(0.00000)	(0.00001)	(0.94421)	(0.00000)
J-Statistic	1.31240	6.91880	19.63900	13.34500	6.85470	11.41200	19.99800	7.03070	43.86700	26.29900	209.4900

Table AG.6.4 - NASDAQ

The

Decile	1	2	3	4	5	6	7	8	9	10	Total Sample
Number of Stocks	60	60	60	60	60	60	60	60	60	60	600
Number of Matched trades	16	58	93	190	345	307	502	595	679	1,477	4,262
α_p	0.60140	-0.18502	0.10551	0.25541	0.06827	0.28177	0.25187	0.24761	0.11020	0.07201	0.10164
	(0.31151)	(0.18129)	(0.14554)	(0.16167)	(0.19808)	(0.09954)	(0.10411)	(0.09902)	(0.12594)	(0.12354)	(0.06006)
	(0.05353)	(0.30744)	(0.46848)	(0.11415)	(0.73034)	(0.00464)	(0.01555)	(0.01240)	(0.38156)	(0.55996)	0.09056)
θ_p	-0.01787	0.05370	0.00271	0.00014	0.08814	0.04592	0.03815	0.00357	0.03552	0.01942	0.093292
	(0.01846)	(0.04876)	(0.00728)	(0.00177)	(0.04828)	(0.01767)	(0.00858)	(0.00422)	(0.04105)	(0.02164)	(0.01642)
	(0.33320)	(0.27074)	(0.71001)	(0.93542)	(0.06792)	(0.00936)	(0.00001)	(0.39747)	(0.38687)	(0.36957)	(0.00000)
δ_p	0.03255	0.05313	0.06517	0.07087	0.02642	0.01106	-0.00191	0.08606	0.01725	0.08069	-0.02400
	(0.03022)	(0.01736)	(0.02969)	(0.02170)	(0.04728)	(0.01652)	(0.00240)	(0.01682)	(0.02501)	(0.02694)	(0.01978)
	(0.28140)	(0.00222)	(0.02817)	(0.00109)	(0.57638)	(0.50309)	(0.42553)	(0.00000)	(0.49022)	(0.00274)	(0.22496)
β_p	-0.01614	-0.20578	-0.11632	0.04486	0.18285	-0.02965	-0.01165	0.02881	0.95675	-0.30044	-0.03619
	(0.07948)	(0.06960)	(0.09994)	(0.10306)	(0.50768)	(0.08956)	(0.08594)	(0.09450)	(1.09280)	(0.17509)	(0.12671)
	(0.83909)	(0.00311)	(0.24448)	(0.66338)	(0.71872)	(0.74058)	(0.89214)	(0.76046)	(0.38132)	(0.08617)	(0.77516)
γ_p	-0.24727	0.61603	0.24663	0.03517	0.11622	-0.04018	-0.03588	0.06020	-0.01483	0.26125	0.06662
	(0.41711)	(0.21838)	(0.16481)	(0.18205)	(0.07354)	(0.11093)	(0.10619)	(0.10486)	(0.08455)	(0.12071)	(0.05210)
	(0.55330)	(0.00479)	(0.13454)	(0.84682)	(0.11405)	(0.71722)	(0.73543)	(0.56589)	(0.86081)	(0.03044)	(0.20096)
J-Statistic	4.38080	2.97200	4.50750	2.74670	10.77100	9.47570	37.06700	4.68200	0.49471	9.75950	0.96602

G.7 Results from the Specification Tests of the Price-Change Models

Tables AG.6.5 and AG.6.6

Eichenbaum, Hansen and Singleton (1988) C_T Statistic for testing the validity of moment conditions in the Price-Change Model Equation, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.6.5 and AG.6.6 present the Eichenbaum, Hansen and Singleton (1988) C_T Statistic which tests the null hypothesis that fractions of the set of moment conditions used to estimate the Price-Change Model described by equation (6.5b) and estimated by the Generalized Method of Moments are true. Table AG.6.5 is for the NYSE and table AG.6.6 for the NASDAQ sub-sample. For each exogenous variable omitted, which corresponds to a particular moment condition, entries in the table represent in the first line the J-statistic of over-identifying restrictions of Hansen, from the GMM estimation when the particular variable is excluded and in the second line the value of the C_T statistic, based on the same estimation, for the samples of NYSE and NASDAQ stocks of October 1994 respectively. The C_T statistic is given by

$$C_T = T\left\{Q_T\left(\hat{\theta}_T\right) - Q_{IT}\left(\tilde{\theta}_{IT}\right)\right\}$$

Where

Q_T : is the value of the objective function minimized during GMM based on the full set of moment conditions and a sample of size T

Q_{IT} : is the value of the objective function minimized during GMM based on the set of moment conditions which are held to be true and on a sample of size T

$\hat{\theta}_T$: is the full parameter vector which minimizes the objective function through GMM estimation using the full set of moment conditions and a sample of size T

$\tilde{\theta}_{IT}$: is the part of the parameter vector which minimizes the objective function through GMM estimation using only that set of moment conditions which are held to be true based on a sample of size T

The statistic is distributed as chi-square with ν_1 degrees of freedom which is equal to the degree by which the number of parameters estimated is exceeded (over-identified) by the number of moments omitted. For the Price – Change Model when one of the moments is omitted this value is four. The confidence limits of this distribution are:

$$\Pr(\chi^2 > 7.77944) = 0.100$$
$$\Pr(\chi^2 > 9.48773) = 0.050$$
$$\Pr(\chi^2 > 13.2767) = 0.010$$

The J-statistic shown on the table is Hansen's J-Statistic for Over-identifying restrictions estimated using the full set of moment conditions, also shown in tables AG.6.3 and AG.6.4 together with the definitions of the pertinent variables in the tables.

Table AG.6.5 - NYSE

PART A

Decile	1	2	3	4	5	6	7	8	9	10
Number of Stocks	20	20	20	20	20	20	20	20	20	20
Number of Matched trades	477	1,144	2,440	2,262	5,435	4,464	8,303	13,853	19,412	34,540
J-Statistic	*1.312400*	*6.918800*	*19.639000*	*13.345000*	*6.656000*	*11.412000*	*19.998000*	*7.030700*	*43.867000*	*26.299000*
V_t										
J	0.613800	4.638100	16.542000	3.064700	5.453100	9.834400	17.532000	6.243800	37.714300	19.422000
C_T	0.698600	2.280700	3.097000	10.280300	1.202900	1.577600	2.466000	0.786900	6.152700	6.877000
V_{t-1}										
J	1.129600	1.541100	3.853200	13.188000	4.403300	10.453000	10.849000	6.957100	18.582000	13.738000
C_T	0.182800	5.377700	15.785800	0.157000	2.252700	0.959000	9.149000	0.073600	25.285000	12.561000
τ_{t-1}										
J	1.300100	5.858700	6.491400	2.930300	4.034200	6.501300	10.084000	3.398200	6.784300	25.484000
C_T	0.012300	1.060100	13.147600	10.414700	2.621800	4.910700	9.914000	3.632500	37.082700	0.815000
τ_{t-2}										
J	0.316480	5.893800	18.963000	12.466000	5.083000	7.756400	12.498000	5.312700	26.950000	13.497000
C_T	0.995920	1.025000	0.676000	0.879000	1.573000	3.655600	7.500000	1.718000	16.917000	12.802000
ΔV_t^a										
J	0.495690	6.572200	15.911000	6.691100	6.245300	21.537000	11.959000	6.712500	42.952000	23.484000
C_T	0.816710	0.346600	3.728000	6.653900	0.410700	-10.125000	8.039000	0.318200	0.915000	2.815000

Table AG.6.5 - PART B

Decile		1	2	3	4	5	6	7	8	9	10
ΔV_t^b											
	J	0.762480	4.971400	18.383000	8.508000	4.542700	8.615200	17.375000	3.072000	43.800000	23.307000
	C_T	0.549920	1.947400	1.256000	4.837000	2.113300	2.796800	2.623000	3.958700	0.067000	2.992000
ΔV_{t-1}^a											
	J	1.024000	6.695500	18.188000	9.670500	3.816400	11.055000	12.716000	5.981800	41.049000	26.048000
	C_T	0.288400	0.223300	1.451000	3.674500	2.839600	0.357000	7.282000	1.048900	2.818000	0.251000
ΔV_{t-1}^b											
	J	1.180700	6.602300	19.116000	9.484400	6.433700	7.681900	19.972000	5.412700	35.678000	6.509300
	C_T	0.131700	0.316500	0.523000	3.860600	0.222300	3.730100	0.026000	1.618000	8.189000	19.789700
q_t											
	J	1.260400	6.462300	2.285900	12.221000	6.443000	8.596700	9.007200	5.982800	15.116000	24.455000
	C_T	0.052000	0.456400	17.353100	1.124000	0.213000	2.815300	10.990800	1.047900	28.751000	1.844000
q_{t-1}											
	J	1.073900	5.121470	4.522400	13.134000	6.098000	10.450000	11.521000	5.158500	28.088000	22.921000
	C_T	0.238500	1.797330	15.116600	0.211000	0.558000	0.962000	8.477000	1.872200	15.779000	3.378000

Table AG.6.6 - NASDAQ

PART A

Decile	1	2	3	4	5	6	7	8	9	10
Number of Stocks	60	60	60	60	60	60	60	60	60	60
Number of Matched trades	16	58	93	190	345	307	502	595	679	1,477
J-Statistic	*4.380800*	*2.972000*	*4.507500*	*2.746700*	*10.771000*	*9.475700*	*37.067000*	*4.682000*	*0.494710*	*9.759500*
V_t										
J	4.2238	1.5337	0.43733	2.7252	9.7608	8.4034	22.513	0.48468	0.3918	8.7822
C_T	0.157000	1.438300	4.070170	0.021500	1.010200	1.072300	14.554000	4.197320	0.102910	0.977300
V_{t-1}										
J	3.913500	2.495900	1.902800	2.238100	2.759500	6.299900	7.287300	4.001200	0.376300	7.157500
C_T	0.467300	0.476100	2.604700	0.508600	8.011500	3.175800	29.779700	0.680800	0.118410	2.602000
τ_{t-1}										
J	3.680500	2.624600	4.164700	2.729700	2.878700	9.398800	36.010000	1.859800	0.134920	9.554000
C_T	0.700300	0.347400	0.342800	0.017000	7.892300	0.076900	1.057000	2.822200	0.359790	0.205500
τ_{t-2}										
J	2.628100	2.863600	3.690600	2.532600	10.555000	4.237800	36.417000	3.713100	0.418810	6.172700
C_T	1.752700	0.108400	0.816900	0.214100	0.216000	5.237900	0.650000	0.968900	0.075900	3.586800
ΔV_t^a										
J	3.961300	2.547000	4.500000	2.250500	8.990600	8.318600	32.745000	4.623800	0.295000	6.300800
C_T	0.419500	0.425000	0.007500	0.496200	1.780400	1.157100	4.322000	0.058200	0.199710	3.458700

Table AG.6.6 - PART B

Decile		1	2	3	4	5	6	7	8	9	10
ΔV_t^b											
	J	3.075400	0.602720	3.106100	1.633800	7.376400	6.616900	36.390000	4.455600	0.260400	8.964900
	C_T	1.305400	2.369280	1.401400	1.112900	3.394600	2.858800	0.677000	0.226400	0.234310	0.794600
ΔV_{t-1}^a											
	J	3.210300	0.849640	3.323700	2.027300	1.717300	8.693400	33.617000	4.311600	0.488940	9.189700
	C_T	1.170500	2.122360	1.183800	0.719400	9.053700	0.782300	3.450000	0.370400	0.005770	0.569800
ΔV_{t-1}^b											
	J	4.049300	2.501800	4.303400	2.482900	1.528900	4.794900	33.008000	4.622900	0.404520	9.634400
	C_T	0.331500	0.470200	0.204100	0.263800	9.242100	4.680800	4.059000	0.059100	0.090190	0.125100
q_t											
	J	1.363000	0.578000	1.331000	0.270310	1.030500	2.781700	7.021700	3.346100	0.225000	6.909700
	C_T	3.017800	2.394000	3.176500	2.476390	9.740500	6.694000	30.045300	1.335900	0.269710	2.849800
q_{t-1}											
	J	3.471800	2.902600	3.338000	1.063500	8.800700	8.770000	4.093900	2.239300	0.215800	8.639300
	C_T	0.909000	0.069400	1.169500	1.683200	1.970300	0.705700	32.973100	2.442700	0.278910	1.120200

Tables AG.6.7 and AG.6.8

GMM-BIC MSC Statistic for selecting moment conditions in the Price-Change Model Equation, Decile 10, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.6.7 and AG.6.8 present the GMM-BIC (Bayesian Information Criterion) Moment Selection Criterion Statistic, reported in Andrews (1999), when two of the moment conditions used in the initial Generalized Method of Moments estimation are excluded from the set used to estimate the Price-Change Model described by equation (6.5b). Entries in tables AG.6.7 and AG.6.8 represent the MSC_{BIC} statistic when the two particular moment conditions are omitted for decile 10 of the samples of NYSE and NASDAQ stocks of October 1994 respectively. The MSC_{BIC} statistic is given by

$$MSC_{BIC,n}(c) = J_n(c) - (|c| - p)\log n$$

Where

$J_n(c)$: is the value of the Hansen J-Test statistic of over-identifying restrictions estimated using the moment vector c and a sample of size n.

$|c|$: is the number of moments used in the moment vector c.

c : moment-selection vector (selects only certain moments)

p : is the dimension (number) of the parameter vector, θ, to be estimated by GMM

n : is number of data in the sample

The $J_n(c)$ statistic is distributed as chi-square with $[|c| - \min(p, |c|)]$ degrees of freedom under the null hypothesis that all moment conditions in c are correct. Since eight moments are used (two omitted from the initial vector) to estimate five parameters there are three degrees of freedom. Thus the confidence limits of this distribution are:

$$\Pr(\chi^2 > 6.25139) = 0.100$$
$$\Pr(\chi^2 > 7.81473) = 0.050$$
$$\Pr(\chi^2 > 11.3449) = 0.010$$

The definitions of the variables in this table are shown in tables AG.6.3 and AG.6.4.

Table AG.6.7 - NYSE

2nd Moment Omitted \ 1st Moment Omitted	V_{t-1}	τ_{t-1}	τ_{t-2}	ΔV^a_t	ΔV^b_t	ΔV^a_{t-1}	ΔV^b_{t-1}	q_t	q_{t-1}
V_t									
J	14.466000	13.562000	13.852000	13.019200	6.244800	11.344500	7.433100	13.822000	8.254700
MSC	*0.851033*	*-0.052967*	*0.237033*	*-0.595767*	*-7.37017*	*-2.270467*	*-6.181867*	*0.207033*	*-5.360267*
V_{t-1}									
J		13.528000	13.719000	13.952000	4.950000	12.651000	7.763200	14.044000	8.947100
MSC		*-0.08697*	*0.104033*	*0.337033*	*-8.66497*	*-0.963967*	*-5.851767*	*0.429033*	*-4.667867*
τ_{t-1}									
J			13.301000	22.708000	22.304000	26.230000	6.329000	25.730000	34.308000
MSC			*-0.313967*	*9.093033*	*8.689033*	*12.615033*	*-7.285967*	*12.115033*	*20.693033*
τ_{t-2}									
J				11.968000	11.102000	13.746000	2.589900	30.168000	32.404000
MSC				*-1.646967*	*-2.51297*	*0.131033*	*-11.02507*	*16.553033*	*18.789033*
ΔV^a_t									
J					21.215000	1.760900	1.795000	28.075000	36.833000
MSC					*7.600033*	*-11.85407*	*-11.81997*	*14.460033*	*23.218033*
ΔV^b_t									
J						23.002000	5.996400	19.862000	32.782000
MSC						*9.387033*	*-7.618567*	*6.247033*	*19.167033*
ΔV^a_{t-1}									
J							1.826900	24.302000	34.290000
MSC							*-11.78807*	*10.687033*	*20.675033*
ΔV^b_{t-1}									
J								7.090000	7.055100
MSC								*-6.524967*	*-6.559867*
q_t									
J									3,354900
MSC									*-10,26007*

Table AG.6.8 - NASDAQ

2nd Moment Omitted \ 1st Moment Omitted	V_{t-1}	τ_{t-1}	τ_{t-2}	ΔV_t^a	ΔV_t^b	ΔV_{t-1}^a	ΔV_{t-1}^b	q_t	q_{t-1}
V_t									
J	4.863100	9.350700	6.508300	5.175000	5.557800	8.783900	8.457200	5.593500	7.573700
MSC	*-4.64504*	*-0.15744*	*-2.99984*	*-4.33314*	*-3.95034*	*-0.72424*	*-1.05094*	*-3.91464*	*-1.93444*
V_{t-1}									
J		6.324200	4.078400	4.910900	6.062200	7.247100	7.100500	7.062200	8.306900
MSC		*-3.18394*	*-5.42974*	*-4.59724*	*-3.44594*	*-2.26104*	*-2.40764*	*-2.44594*	*-1.20124*
τ_{t-1}									
J			1.663600	3.683400	4.344800	8.895700	9.656400	4.116400	5.590100
MSC			*-7.84454*	*-5.82474*	*-5.16334*	*-0.61244*	*0.148259*	*-5.39174*	*-3.91804*
τ_{t-2}									
J				2.909500	3.833900	5.628100	6.220300	2.933000	3.898300
MSC				*-6.59864*	*-5.67424*	*-3.88004*	*-3.28784*	*-6.57514*	*-5.60984*
ΔV_t^a									
J					6.495400	5.974600	4.207100	5.372000	5.990200
MSC					*-3.01274*	*-3.53354*	*-5.30104*	*-4.13614*	*-3.51794*
ΔV_t^b									
J						6.867000	8.142600	6.584500	7.760300
MSC						*-2.64114*	*-1.36554*	*-2.92364*	*-1.74784*
ΔV_{t-1}^a									
J							9.235000	6.317500	8.294900
MSC							*-0.27314*	*-3.19064*	*-1.21324*
ΔV_{t-1}^b									
J								6.375100	8.166900
MSC								*-3.13304*	*-1.34124*
q_t									
J									6.909700
MSC									*-2.59844*

Tables AG.6.9 and AG.6.10

Likelihood Ratio (LR) Statistic for the parameter vector of the Price-Change Model Equation, NYSE and NASDAQ TAQ data: October 1994.

Tables AG.6.9 and AG.6.10 present the Lagrange Multiplier (LR) Statistic for the joint hypothesis that all parameters in the Price-Change Model described by equation (6.5b) are equal to zero for the NYSE and NASDAQ samples of common stock data respectively. The LR statistic is given by

$$LR_T = T\left[Q_T\left(\bar{\theta}_T\right) - Q_T\left(\hat{\theta}_T\right)\right]$$

where

Q_T : is the value of the objective function

$\hat{\theta}_T$: is the unrestricted estimator of the parameter vector

$\bar{\theta}_T$: is the restricted estimator of the parameter vector , and

T : is the number of observations

Under the null hypothesis that all restrictions are valid the LR statistic is distributed as a chi-square with degrees of freedom equal to the number of restrictions imposed. For the Price-Change Model which involves five parameters there are five degrees of freedom when the null hypothesis assumes that all parameters are equal to zero. Thus the confidence limits of this distribution are:

$$\Pr(\chi^2 > 9.23635) = 0.100$$
$$\Pr(\chi^2 > 11.0705) = 0.050$$
$$\Pr(\chi^2 > 15.0863) = 0.010$$

Table AG.6.9 - NYSE

Decile	1	2	3	4	5	6	7	8	9	10
J-Statistic										
Unrestricted Model	1,3124	6,9188	19,6390	13,3450	6,8547	11,4120	19,9980	7,0307	43,8670	26,2990
Restricted Model	91,0910	178,6600	433,0100	336,8000	812,3600	683,3800	1142,3000	1908,9000	2767,3000	4781,4000
LR statistic	89,7786	171,7412	413,3710	323,4550	805,5053	671,9680	1122,3020	1901,8693	2723,4330	4755,1010

Table AG.6.10 - NASDAQ

Decile	1	2	3	4	5	6	7	8	9	10
J-Statistic										
Unrestricted Model	4,38080	2,97200	4,50750	2,74670	10,77100	9,47570	37,06700	4,68200	0,49471	9,75950
Restricted Model	4,83880	15,18600	18,98800	41,01200	72,05100	65,23600	98,10100	110,30000	125,54000	283,08000
LR statistic	0,45800	12,21400	14,48050	38,26530	61,28000	55,76030	61,03400	105,61800	125,04529	273,32050

Nomenclature

English Symbols

A : Event where investor buys at the ask price

A_t : Adverse-selection cost component at time t

A_a : Adverse-selection cost component in the ask price (buyer-initiated trade)

A_b : Adverse-selection cost component in the bid price (seller-initiated trade)

A_0 : Intercept of the linear adverse selection spread component equation in the Glosten and Harris (1988) model, equation (3.12)

A_1 : Slope of the linear adverse selection spread component equation in in the Glosten and Harris (1988) model, equation (3.12)

a_t : Ask price quoted at time t

Δa_t : Change in ask prices quoted between times t and t-1

$\Delta \bar{a}_t$: Change in the ask prices quoted between times t and t-1, expressed as a fraction of the spread at time t-1 (proportional change in the quoted ask).

B : Event where investor sells at the bid price

b_t : Bid price quoted at time t

Δb_t : Change in bid prices quoted between times t and t-1

$\Delta \bar{b}_t$: Change in the bid prices quoted between times t and t-1, expressed as a fraction of the spread at time t-1 (proportional change in the quoted bid).

C_i : Cost of trading amount Q_i (Stoll (1978) model).

C_t : Gross-profit component at time t

C_a : Gross-profit component at the ask price (buyer-initiated trade)

C_b : Gross-profit component at the bid price (seller-initiated trade)

C_0 : Intercept of the linear gross-profit (or transitory) spread component in the Glosten and Harris (1988) model, equation (3.12)

C_1 : Slope of the linear gross-profit (or transitory) spread component in the Glosten and Harris (1988) model, equation (3.12)

$|c|$: Number of moments selected in the moment- selection vector c in the GMM-BIC statistic of Andrews(1999), section 5.6.2

c : Moment-selection vector c in the GMM-BIC statistic of Andrews (1999), section 5.6.2

c_i : Percentage cost that the dealer requires in order that he takes position Q_i in stock i, (Stoll (1978) model)

c_c : Current consumption of dealer (Glosten and Milgrom (1985) model)

cov : Covariance

$D_T(\theta)$: Matrix of partial derivatives in the first-order condition in the minimization of the objective function (GMM estimation)

e_t : Error term in the vector function of population moments in the GMM estimation of the Quoted-Spread Model Equation (5.11b)

e_t^a : Error term in the vector function of population moments in the GMM estimation of the Offer-Change Model Equation (5.9b)

e_t^b : Error term in the vector function of population moments in the GMM estimation of the Bid-Change Model Equation (5.10b)

e_t^p : Error term in the vector function of population moments in the GMM estimation of the Price-Change Model Equation (6.5b)

F_t : Informed trader's total information set (Glosten and Milgrom (1985) model)

$g(\theta, y_T)$: Vector (r x 1) valued function denoting the sample average of $h(\theta, w_t)$ (GMM estimation)

$h(\theta,w_t)$: Vector (r x 1) valued function of instruments selected for GMM estimation

H_t : Public information set (Glosten and Milgrom (1985) model)

IV_t : Inventory-holding cost component at time t

I_t^a, I_t^b : Indicator Variables : equal to +1(0) when the trade at time t is buyer-initiated and 0(1) otherwise

$I_{\{Z_t \rangle A\}}$: Indicator variable on whether event $Z \rangle A$ has occurred (Glosten and Milgrom (1985) model).

$J_n(c)$: Hansen J-Test statistic of over-identifying restrictions estimated using the moment vector c and a sample of size n.

J_t : Informed trader's private information set (Glosten and Milgrom (1985) model)

M_t : Mid-point between the bid and ask prices at time t

m_t : True (common information) price of a stock at time t

n : Number of data in the sample

N_f : Number of orthogonality conditions used in GMM estimation

N_β : Number of parameters to be estimated by GMM estimation

O_t^a : Order-processing cost component at the ask

O_t^b : Order-processing cost component at the bid

P : Common (public information) value of a stock

p : Number of parameters estimated by GMM (GMM-BIC statistic of Andrews (1999), section 5.6.2)

P^* : Full-information value of a stock

P_i^* : True price of stock i before the transaction (Stoll (1978) model)

P_t : Observed transaction price at time t

P_a : Ask price

P_b : Bid Price

ΔP_t : Change in the observed transaction prices between times t and t-1

$\Delta \bar{P}_t$: Change in the observed transaction prices between times t and t-1, expressed as a fraction of the spread at time t-1

q_{t-1} : Trade Indicator Variable : equal to +1 when the trade is buyer-initiated; equal to -1 when the trade is seller-initiated

Q_i : True value of trade on asset i (Stoll (1978) model)

Q_i^a : True value of a sale of asset i to the dealer (Stoll (1978) model)

Q_i^a : True value of a purchase of asset i from the dealer (Stoll (1978) model)

Q_T : The value of the objective function minimized during GMM based on the full set of moment conditions and a sample of size T (Eichenbaum, Hansen and Singleton (1988) statistic).

Q_{IT} : The value of the objective function minimized during GMM based on the set of moment conditions which are held to be true and on a sample of size T (Eichenbaum, Hansen and Singleton (1988) statistic).

$Q(\theta; y_T)$: The value of the objective function to be minimized in GMM estimation based on parameter θ and vector of T observations of observed variables (GMM estimation)

r_t : Round-off error (drift between observed discrete price and unobserved price when no discrete prices are used), assumed to be a zero mean, asymptotically uniformly distributed random variable (Glosten and Harris (1988) model)

$\tilde{R}^*$: Rate of return on the dealer's initial portfolio (Stoll (1978) model)

$\tilde{R}_i$: Rate of return on stock i (Stoll (1978) model)

R_f : Risk-free rate for borrowing and financing (Stoll (1978) model)

S : Realized spread (Roll (1984) Model)

S_R : Average realized spread as a percentage of quoted spread

s_p : proportional spread in the Glosten (1987) model

S^Q : Quoted spread assumed to be constant (Stoll (1979) model)

S_t^Q : Quoted spread at time t

ΔS_t^Q : Change in the quoted spread between times t and t-1

$\Delta \bar{S}_t^Q$: Change in the quoted spread at time t, expressed as a fraction of the spread at time t-1

S_t : Specialist's information set (Glosten and Milgrom (1985) model)

T : Number of observations (GMM estimation)

T_t : Time elapsed between trades at times t and t-1 (Glosten and Harris (1988) model)

$\bar{V}$: Average number of shares per transaction

V_t^a : Volume (number) of shares offered prior to the trade at time t

V_t^b : Volume (number) of shares bid prior to the trade at time t

V_t : Volume of shares traded at time t

ΔV_t^a : Change in the number of shares offered between times t and t-1

ΔV_t^b : Change in the number of shares bid between times t and t-1

U_t : Indicator variable signifying information event (Glosten and Milgrom (1985) model).

$\tilde{W}$: Terminal wealth of the dealer (Stoll (1987) model).

W_0 : Initial wealth of the dealer (Stoll (1987) model).

W_T : Weighting matrix in the objective function (GMM estimation)

w_t : Vector (h x 1) of variables observed at time t (GMM estimation)

$\{W_T\}_{T=1}^{\infty}$: Sequence of (r x r) positive definite weighting matrices in the minimization of the objective function (GMM estimation)

V^* : Variance – covariance matrix of the parameters in (GMM estimation)

V : Random value of asset (Glosten and Milgrom (1985) model)

x : Number of shares of stock (Glosten and Milgrom (1985) model)

z : Dealer's coefficient of relative risk aversion (Stoll (1978) model)

z_t : Effective spread for trade at time t (Lin, Sanger and Booth (1995) model).

Z_t : Decision rule of investor (Glosten and Milgrom (1985) model)

Greek Symbols

α : Fixed Order-Processing Cost parameter for a single transaction

α_p : Fixed Order-Processing Cost parameter for a single transaction, expressed as a fraction of the spread at time t-1

$1-\alpha$: Order-processing component in the George, Kaul and Nimalendran (1991) model

β : Waiting-time-between-trades parameter

β_p : Waiting-time-between-trades parameter expressed as a fraction of the spread at time t-1

γ : Quoted-Volume parameter

γ_e : Order-processing cost component of the effective spread (Lin, Sanger and Booth (1995) model).

γ_p : Quoted-Volume parameter expressed as a fraction of the spread at time t-1

δ : Inventory-holding cost parameter

δ_c : The magnitude of a price continuation as a proportion of the quoted spread (Stoll (1979) model)

δ_p : Inventory-holding cost parameter expressed as a fraction of the spread at time t-1

$\Delta\zeta_t$: Random error term, asymptotically, uniformly distributed with zero mean

$\Delta\eta_t$: Random error term, asymptotically, uniformly distributed with zero mean

$\Delta\nu_t$: Random error term, asymptotically, uniformly distributed with zero mean

$\Delta\xi_t$: Random error term, asymptotically, uniformly distributed with zero mean

ε_t : Public information innovation (represents the innovation in beliefs regarding the true price of the stock) between trades at times t and t-1, assumed to be an independently, identically distributed, zero mean normal random variable.

ζ_t : Random error term, asymptotically, uniformly distributed with zero mean

η_t : Random error term, asymptotically, uniformly distributed with zero mean

θ : Adverse-Selection parameter

θ_p : Adverse-Selection parameter expressed as a fraction of the spread at time t-1

ϑ_t : Random error term, independently, identically distributed

$\hat{\theta}_T$: The full parameter vector which minimizes the objective function through GMM estimation using the full set of moment conditions and a sample of size T (Eichenbaum, Hansen and Singleton (1988) statistic); unrestricted estimator of the parameter vector (Lagrange Multiplier Statistic section 5.6)

$\tilde{\theta}_{IT}$: The part of the parameter vector which minimizes the objective function through GMM estimation using only that set of moment conditions which are held to be true based on a sample of size T (Eichenbaum, Hansen and Singleton (1988) statistic).

$\bar{\theta}_T$: Restricted estimator of the parameter vector (Lagrange Multiplier Statistic section 5.6)

κ_t : Random error term, independently, identically distributed

λ_e : Adverse-selection cost component of the effective spread (Lin, Sanger and Booth (1995) model).

$\Delta\lambda_t$: Error term at time t, uniformly distributed zero mean random variable

ν_t : Random error term, asymptotically, uniformly distributed with zero mean

ξ_t : Independently, identically distributed random variable with zero mean models the effect of rounding errors due to price discreteness or time-varying returns (Madhavan, Richardson and Roomans (1997) model)

π_t : Error term at time t, independently, identically distributed random variable

π_g : The proportion of the spread due to factors other than adverse selection (gross profit, Glosten (1987) model)

π : The probability of a price reversal (Stoll (1979) model)

ρ : The first-order autocorrelation coefficient of the trade indicator variable

ρ_l : Liquidity parameter (Glosten and Milgrom (1985) model)

σ_{ip} : Correlation between the rate of return of stock i and that of the optimal efficient portfolio (Stoll (1978) model)

σ_i^2 : Variance of the return of stock i (Stoll (1978) model)

τ_{t-1} : Time (seconds) elapsed between trades at t-1 and t-2

$\Delta\tau_{t-1}$: Difference in times (seconds) elapsed between trades at t-1 and t-2 and between times t-2 and t-3 (τ_{t-1} - τ_{t-2})

$\Delta\upsilon_t$: Error term at time t, independently, identically distributed random variable

φ : Transitory effect per share which covers the dealer's compensation for providing liquidity, the inventory holding cost and risk bearing (Madhavan, Richardson and Roomans (1997) model)

$\Delta\chi_t$: Error term at time t, uniformly distributed, zero mean random variable

ϖ_t : Error term at time t, independently, identically distributed random variable

Ω : Common information set relating to the stock

Bibliography

Abhyankar, A., D. Ghosh, E. Levin, and R. J. Limmack, 1997, *Bid-Ask Spreads, Trading Volume and Volatility : Intra-Day Evidence From the London Stock Exchange,* Journal of Business Finance and Accounting, 24, 3 and 4 (April), 343-362.

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

Admati, A., and P. Pfleiderer, 1989, *Divide and Conquer: A Theory of Day-of-the-Week Mean Effects, The* Review of Financial Studies, 2, 2, 189-223.

Affleck-Graves, J., S. P. Hedge, and R. E. Miller, 1994, *Trading Mechanisms and the Components of the Bid-Ask Spread*, The Journal of Finance, 49,4 (September 1994), pp. 1471-1488.

Ahn, S. C., 1995, *Model Specification Testing Based on Root-T Consistent Estimators,* Discussion Paper, Department of Economics, Arizona State University, Tempe, AZ.

Aitken, M., and A. Frino, 1996, *"The Accuracy of the Tick Test: Evidence from the Australian Stock Exchange"*, Journal of Banking and Finance, 20, 1715-1729.

Alexander, S. S., 1961, *'Price Movements in Speculative Markets: Trends or Random Walks, Industrial Management Review'*, 2, 7-26.

Amihud, Y., T. S. Y. Ho, and R. A. Schwartz, 1985, *Overview of the Changing Securities Markets,* in Y. Amihud, T. S. Y. Ho, and R. A. Schwartz, eds.: Market Making and the Changing Structure of Securities Industries (Lexington Books, Massachusetts), pp. 1-15.

Amihud, Y., and H. Mendelson, 1980, *Dealership Market: Market-Making with Inventory*, The Journal of Financial Economics, 8, pp. 31-53.

Amihud, Y., and H. Mendelson, 1982, *Asset Pricing Behaviour in a Dealership Market,* Financial Analysts Journal, 38 (May/June), pp. 50-59.

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

Amihud, Y., and H. Mendelson, 1987, *Trading Mechanisms and the Stock Returns: An Empirical Investigation*, The Journal of Finance, 42, July 1987, pp. 533-555.

Amihud, Y., and H. Mendelson, 1989, *The Effects of Beta, Bid-Ask Spread, Residual Risk, and Size on Stock Returns,* Journal of Finance, 44, pp. 479-786.

Amihud, Y., and H. Mendelson, 1991, *'Volatility, Efficiency and Trading: Evidence from the Japanese Stock Market'*, Journal of Finance, 46, 1765-1790.

Amihud, Y., H. Mendelson, and M. Murgia, 1990, Stock '*Market Microstructure and return volatility: Evidence from Italy*', Journal of Banking and Finance, 14, 423-440.

Andrews, D. W. K., 1991, *Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation*, Econometrica, 59, 817-858.

Andrews, D. W. K., 1993, *Tests for Parameter Instability and Structural Change with Unknown Change Point*, Econometrica, 61, 821-856.

Andrews, D. W. K., 1999, *Consistent Moment Selection Procedures for Generalized Method of Moments Estimation*, Econometrica, 67, 3 (May), 543-564.

Andrews, D. W. K., and R. Fair, 1988, *Inference in Econometric Models with Structural Change,* Review of Economic Studies, 55, 615-640.

Andrews, D. W. K., and J. C. Monahan, 1992, *An Improved Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimator*, Econometrica, 60, 953-966.

Andrews, D. W. K., and W. Ploberger, 1994, *Optimal Tests when a Nuisance Parameter is Presented Only Under the Alternative,* Econometrica, 62, 1383-1414.

Angel, J., 1990, *Order Placement Strategy of Informed Investors: Market versus Limit Orders,* Working Paper, University of California at Berkeley.

Angel, J., 1994, *Who Gets Price Improvement on the NYSE?*, Working Paper, Georgetown University.

Angel, J., 1995, *Market Versus Limit Orders*, Working Paper, Georgetown University.

Angel, J., 1998, *Nonstandard Settlement Transactions*, Financial Management, 27, 31-46.

Bachelier, L., *Theory of Speculation (translation of the 1900 French Edition).* In the Random Character of Stock Prices, P. H. Cootner, ed. Cambridge, MA : MIT Press (1964), 17-78.

Back, K., and J. Zender, 1993, *Auctions of Divisible Goods: On the Rationale for the Treasury Experiment*, Review of Financial Studies, 6, 733-764.

Baesel, J., G. Shows, and E. Thorp, 1983, *The Cost of Liquidity Services in Listed Options: A Note*, Journal of Finance, 38 (June), pp. 989-995.

Bagehot, W. (Jack Treynor), 1971, *The Only Game in Town*, Financial Analysts Journal, 22, pp. 12-14.

Bagehot, W., 1971, *The Only Game in Town,* Financial Analysts Journal, 8, 31-53.

Bamber, L. S., O. E. Barron, and T. L. Stober, 1999, *Differential Interpretations and Trading Volume,* Journal of Financial and Quantitative Analysis, 34, 3 (September), 369-386.

Barclay, M., W. Christie, J. Harris, E. Kandel and P. Schultz, 1999, *The Effects of Market reform on the Trading Costs and Depths of Nasdaq Stocks*, Journal of Finance, 54, 1-34.

Barnea, A., 1974, *Performance Evaluation of New York Stock-Exchange Specialists,* Journal of Financial and Quantitative Analysis, (September), 511-535.

Barnea, A., and D. E. Logue, 1975, *The Effect of Risk on the Market Maker's Spread,* Financial Analysts' Journal, Nov.-Dec., 45-49.

Battalio, R. H., 1997, *Third Market Broker-Dealers: Cost Competitors or Cream Skimmers ?*, The Journal of Finance, 52, 1 (March 1997), pp. 341-352.

Battalio, R., J. Greene, B. Hatch, and R. Jennings, 1999, *Does the Limit Order Routing Decision Matter?*, Working Paper, Indiana University Business School, January, 1-41.

Battalio, R., J. Greene, and R. Jennings, 1997, *Do Competing Specialists and Preferencing Dealers Affect Market Quality?,* Review of Financial Studies, 10, 969-994.

Battalio, R., J. Greene, and R. Jennings, 1998, *Order Flow Distribution, Bid-Ask Spreads, and Liquidity Costs : Merrill Lynch's Decision to Cease Routinely Routing* Orders to Regional Stock Exchanges, Journal of Financial Intermediation, 7, 338-358.

Baumol, W. J., 1957, *Speculation, Profitability and Stability,* Review of Economics and Statistics, 39, August, 263-271.

Baumol, W. J., 1965, *The Stock Market and Economic Efficiency,* The Millan Lectures, No. 6 New York: Fordham University Press.

Beaver, W. H., 1968, *The Information Content of Annual Earnings Announcements*, Empirical Research in Accounting: Selected Studies. Supplement to Journal of Accounting Research 6: 67-92.

Beebower, G., 1989, *Evaluating Transaction Cost*, In the Complete Guide to Securities Transactions, Wayne Wagner, ed. New York, NY: John Wiley & Sons, 137-150.

Beja, A., and M. B. Goldman, 1980, *On the Dynamic Behaviour of Prices in Disequilibrium,* The Journal of Finance, 35, 2 (May), 235-248.

Beja, A., and M. B. Goldman, 1978, *Market Prices vs. Equilibrium Prices: Returns Variance, Serial Correlation, and the Role of the Specialist*, W.P. no. 78-51, N.Y. University, New York.

Beja, A., and N. A. Hakansson, 1977, *Some Alternative Market Mechanisms*, , W.P. no. 104, N.Y. University, New York.

Beja, A., and N. A. Hakansson, 1979, On the *Feasibility of Using a Mechanical Specialist in Securities Trading*, Paper presented at the XxiV Meeting of the Institute of Management Science.

Beja, A., and N. A. Hakansson, 1979, From *Orders to Trades: Some Alternative Market Mechanisms*, In E. Block and R. A. Schwartz, eds, Impeding Changes for Securities Markets: What Role for the Exchange?, Greenwich, Conn.: JAF Press, pp. 144-161.

Benston, G. J., and R. L. Hagerman, 1974, *Determinants of Bid-Ask Spreads in the Over-the-Counter Market,* Journal of Financial Economics, 1 (December), pp. 353-364.

Bernhardt, D., and E. Hughson, 1997, *Splitting Orders*, Review of Financial Studies, 10, 69-102.

Bernhardt, D., V. Dvoracek, E. Hughson, and I. Werner, 1999, *"Why do Large Orders Receive Discounts on the London Stock Exchange?"*, Working Paper #99-0114, University of Illinois.

Bernheim, B. D., and M. D. Whinston, 1986, *Menu Auctions, Resource Allocation and Economic Influence*, Quarterly Journal of Economics, 101, 1-31.

Bessembinder, H., 1999, *Trade Execution Costs on NASDAQ and the NYSE : A Post-Reform Comparison*, Journal of Financial and Quantitative Analysis, 34, 3(Sept.), 387-407.

Bessembinder, H., 2002, *Issues in Assessing Trade Execution Costs*, Forthcoming in Journal of Financial markets.

Bessembinder, H., and H. Kaufman, 1997, *A Cross-Exchange Comparison of Execution Costs and Information Flows for NYSE-listed Stocks*, Journal of Financial Economics, 46, 293-319.

Bessembinder, H., and H. Kaufman, 1997, *A Comparison of Trade Execution Costs for NASDAQ and NYSE-listed Stocks*, Journal of Financial and Quantitative Analysis, 32, 287-310.

Bhattacharya, M., 1986, *Direct Tests of Bid/Ask Spread Models,* University of Michigan, Graduate School of Business, Working Paper.

Biais, Bruno, 1993, *Price Formation and Equilibrium Liquidity in Fragmented and Centralized Markets*, The Journal of Finance, 48, 1 (March), pp. 157-185.

Biais, B., and P. Hillion, 1990, *Option Prices, Insider Trading, and Interdealer Competition*, Working Paper, INSEAD.

Biais, B., F., Foucault, and F. Salanie, 1998, *Floors, Dealer Markets and Limit Order Markets, Journal of Financial Markets*, 1, 3-4 (September), 253-284.

Biais, B., D. Martimort, and J. C. Rochet, 1997, *Competing Mechanisms in a Common Value Environment*, Working Paper, Toulouse University.

Biais, B., P. Hillion, and C. Spatt, 1995, *An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse,* The Journal of Finance, 50, 5 (December), 1655-1689.

Black, F., 1971a, *Toward a Fully Automated Exchange, Part A*, Financial Analysts Journal, July-August, 28-44.

Black, F., 1971b, *Toward a Fully Automated Exchange, Part B*, Financial Analysts Journal, November-December, 24-87.

Black, F., 1991, Trading in *Equilibrium with Bluffing, Credits and Debits,* Working Paper, Goldman Sachs, New York.

Black, F., 1995, *Equilibrium Exchanges*, Financial Analysts Journal, May-June, 23-29.

Black, F., and M. Scholes, 1974, *From Theory to a New Financial Product*, Journal of Finance, 29 (May), 399-412.

Blattberg, R. C., and N. J. Gonedes, 1974, *A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices,* Journal of Business, 47, (April), 244-280.

Bleakley, F., 1971, *Is NASDAQ Really the Answer?*, Institutional Investor, 5 (July), 21.

Blume, L., D. Easley, and M. O'Hara, 1994, *Market Statistics and Technical Analysis: The Role of Volume*, The Journal of Finance, 49 (March), 153-181.

Blume, M. E., and M. A. Goldstein, 1992, *Differences in Execution Prices among the NYSE, the Regionals and the NASD,* Working Paper, Philadelphia, University of Pennsylvania.

Blume, M. E., and M. A. Goldstein, 1993, *Differences in Execution Prices among the NYSE, the Regionals and the NASD*, Working Paper, Rodney L. White Center for Financial Research, Wharton School.

Blume, M. E., and M. A. Goldstein, 1997, *Quotes, Order Flow, and Price Discovery*, The Journal of Finance, 52, 1, March 1997, pp.221-244.

Blume, M. E., and R. F. Stambaugh, 1983, *Biases in Computed Returns,* Journal of Financial Economics, 12 (November), pp. 387-404.

Board, J., and C. Sutcliffe, 1995, *The Effects of Trade Transparency in the London Stock Exchange: A Summary*, W.P., London School of Economics.

Board, J., A. F. Vila, and C. Sutcliffe, 1996, *Market Maker Heterogeneity and Order Preferencing: Evidence from the London Stock Exchange*, W.P., London School of Economics.

Bollen, N. P. B., T. Smith, and R. E. Whaley, 2002, *Modelling the Bid-Ask Spread: Measuring the Inventory-Holding Premium*, Forthcoming in the Journal of Financial Economics.

Bradfield, J., 1979, *A Formal Dynamic Model of Market Making,* Journal of Financial and Quantitative Analysis, 14, 2 (June 1979), 275-291.

Bradfield, J., 1982, *Optimal Dynamic Behaviour of a Market-Maker in Choosing Bid and Asked Prices,* Journal of Economics and Business, 34, 303-316.

Bradfield, J., and E. Zabel, 1979, *Price Adjustment in a Competitive Market and the Securities Exchange Specialist,* in *General Equilibrium Growth and Trade*, ed. By J. Green and J. A. Scheinkman, New York: Academic Press.

Branch, B., and W. Freed, 1977, *Bid-Ask Spreads on the AMEX and the Big Board,* The Journal of Finance, 32, 1 (March), pp. 159-163.

Brock, W. A., and A. W. Kleidon, 1992, *Periodic Market Closure and Trading Volume: A Model of Intraday Bids and Asks*, Journal of Economic Dynamics and Control, 16, pp. 451-489.

Brock, W., J. Lakonishok, and B. LeBaron, 1992, *Simple Technical Trading Rules and the Stochastic Properties of Stock Returns,* Journal of Finance, 47, 1731-1764.

Bronfman, C., 1991, *Limit Order Placement and Execution Risk: An Empirical Analysis*, Working Paper, Commodity Futures Trading Commission.

Bronfman, C., and R. A. Schwartz, 1991, *Implications of Risk Aversion for the microstructure of a dealers market*, W.P. New York University.

Brooks, R., and J. Masson, 1996, *Performance of Stoll's Spread Component Estimator : Evidence from Simulations, Time-series and Cross-Sectional Data*, The Journal of Financial Research, 14, 459-476.

Broyles, J. E., 1971, *A Stochastic Model of New York Stock Exchange Transactions*, Ph.D. Thesis, University of London.

Brown, P., G. Clinch, and G. Foster, 1991, *Market Microstructure and Capital Market Information Content Research,* American Accounting Association Monograph.

Brown, D. P., and C. W. Holden, 1996, *The Design of Limit Orders Under a Hybrid Mechanism with Endogenous Depth*, Working Paper, Indiana University Business School, May, 1-28.

Bunting, J. P., *Moving from Today's to Tomorrow's Trading System*, in "The Challenge of Information Technology for the Securities Market", (H. C. Lucas, Jr., and R. A. Schwartz, Eds.). Dow Jones – Irwin, Homewood.

Burdett, K., and M. O'Hara, 1983, *Building Blocks: An Introduction to Block Trading*, W.P., Cornell Univ. (October).

Campbell, J. Y., A.W. Lo, and A.C. MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey, U.S.A.

Chakravarty, S., F. H. deB. Harris, and R. A. Wood, 2001, *Do Bid-Ask Spreads Or Bid and Ask Depths Convey New Information First?*, Working Paper, Purdue University.

Chakravarty, S., and C. W. Holden, 1995, *An Integrated Model of Market and Limit Orders*, Journal of Financial Intermediation, 4, 3, 213-241.

Chan, K., Y. P. Chung, and H. Johnson, 1995, *The Intraday Behaviour of Bid-Ask Spreads for NYSE Stocks and CBOE Options*, Journal of Financial and Quantitative Analysis, 30 (September), pp. 329-346.

Chan, K. C., W. G. Christie, and P. H. Schultz, 1995, *Market Structure and the Intraday Pattern of Bid-Ask Spreads for NASDAQ Securities*, Journal of Business, 68, 1, pp. 35-60.

Chen, D-H., and L. P. Blenman, 2003, *An Extended Model of Serial Covariance Bid-Ask Spreads, International Journal of Business and Economics*, 2, 1, 75-83.

Chiang, R., and P. C. Venkatesh, 1988, *Insider Holdings and Perceptions of Information Asymmetry: A Note,* Journal of Finance, 43, 4 (September), 1041-1048.

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

Chordia, T., and A. Subrahmanyam, 1992, *Off-Floor Market Making, Payment for Order Flow, and the Tick Size*, Working Paper, University of California at Los Angeles.

Chotikapanich, D. and W.E. Griffiths, 1993, *Learning SHAZAM: A Computer Handbook for Econometrics*, John Wiley and Sons, Inc., New York.

Christie, A., 1982, *The Stochastic Behaviour of Common Stock Variances*, Journal of Financial Economics, 10, 407-432.

Christie, W., J. Harris and P. Schultz, 1994, *Market Structure and the Intraday Pattern of Bid – Ask Spreads for NASDAQ Securities,* Journal of Business 68, 35-60.

Christie, W., and R. Huang, 1994, *Market Structures and Liquidity: A Transactions Data study of Exchange Listings,* Journal of Financial Intermediation, 3 (June), pp. 300-326.

Christie, W. G., and P. H. Schultz, 1994a, *Why Do NASDAQ Market Makers Avoid Odd-Eighth Quotes?*, The Journal of Finance, 49, 5 (December), 1813-1840.

Christie, W. G., and P. H. Schultz, 1994b, *Why Did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes?*, The Journal of Finance, 49, 5 (December), 1841-1860.

Chung, K. H., B. F. Van Ness, R. A. Van Ness, 1999, *Limit Orders and the Bid-Ask Spread,* Journal of Financial Economics, 53, 255-287.

Clark, P. K., 1973, *A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices,* Econometrica, 41, 1 (January), 135-159.

Clark, R. A., J. J. McConnell, and M. Singh, 1992, *Seasonalities in NYSE Bid-Ask Spreads and Stock Returns in January,* Journal of Finance, 47 (December), pp. 1999-2014.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1977, *The Impact of Designated Market Makers on Security Prices: II, Policy Proposals*, Journal of Banking and Finance, 1, 236-247.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1978a, *The Returns Generation Process, Returns Variance, and the Effect of Thinness in Securities Markets*, The Journal of Finance, 33, 1 (March), 149-167.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1978b, *Limit Orders, Market Structure, and the Returns Generation Process*, The Journal of Finance, 33, 3 (June), 723-736.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1979a, *Market Makers and the Market Spread: A Review of Recent Literature,* Journal of Financial and Quantitative Analysis, 14, 4 (November), 813-835.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1979b, *On the existence of Serial Correlation in an Efficient Securities Market,* in E. J. Elton and M. J. Gruber (eds.), Portfolio Theory, 25 Years After: Essays in Honor of Harry Markowitz, New York:Elsevier North-Holland, pp. 151-168.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1981, *Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread*, Journal of Political Economy, 89, 2, 287-305.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1983, *A Simulation Model of Stock Exchange Trading*, Simulation, 50, 181-191.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1986, *The Microstructure of Securities Markets,* Prentice-Hall, New Jersey.

Cohen, K. J., S. F. Maier, W. L. Ness, H. Okuda, R. A. Schwartz, and D. K. Whitcomb, 1977, *The Impact of Designated Market Makers on Security Prices: I, Empirical Evidence*, Journal of Banking and Finance, 1, 219-235.

Comiskey, E. E., R. A. Walking, and M. A. Weeks, 1984, *Dispersion of Expectations and Trading Volume*, Working Paper, GA Inst. of Tech., (October).

Conrad, J. S., A. Hameed, and C. Niden, 1994, *Volume and Autocovariances in Short-Horizon Individual Security Returns*, The Journal of Finance, 49 (September), 1305-1329.

Cornell, B., 1981, *The Relationship between Volume and Price Variability in Futures Markets*, The Journal of Futures Markets, 1 (Fall), 303-316.

Copeland, T. E., 1976, *A Model of Asset Trading under the Assumption of Sequential Information Arrival,* The Journal of Finance, 31, 4 (September), 1149-1168.

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

Copeland, T., and H. R. Stoll, 1990, *Trading Markets*, in Dennis E. Logue (ed.), *Handbook of Modern Finance*, 2nd ed., Warren, Gorham and Lamont, Boston.

Cramer, H., 1946, *Mathematical Methods of Statistics, Princeton*, N.J.: Princeton University Press.

Crouch, R. L., 1970, *A Non-Linear Test of the Random Walk Hypothesis*, American Economic Review, 60 (March), 199-202.

Crouch, R. L., 1970, *The Volume of Transactions and Price Changes on the New York Stock Exchange*, Financial Analysts Journal, 26 (July – Aug.), 104-109.

Cuneo, L., and W. Wagner, *Reducing the Cost of Stock Trading*, Financial Analysts Journal, 31 (November-December 1975), 35-44.

de Jong, F., T. Nijman and A. Roell, 1995, *A Comparison of the Cost of Trading French Shares on the paris Bourse and on SEAQ International*, European Economic Review, 39, 1277-1301.

Demsetz, H., 1968, *The Cost of Transacting*, Quarterly Journal of Economics, 82, February, pp. 33-53.

Demsetz, Harold, 1997, *Limit Orders and the Alleged Nasdaq Collusion*, Journal of Financial Economics, 45, 91-95.

Domowitz, I., 1990, *The Mechanics of Automated Trade Execution*, Journal of Financial Intermediation, 1, 167-194.

Dubofsky and Groth, 1987, *An Examination of Asked-Bid Spreads for two Over-the-Counter Market Segments*, Financial Review, 21 (May), pp.289-298.

Dupont, D. Y., 1995, *Market Making, Prices and Quantity Limits,* Working Paper, Board of Governors of the Federal Reserve System.

Dutta, P., and A. Madhavan, 1997, *Competition and Collusion in Dealer Markets,* The Journal of Finance, 52, 1(March 1997), pp. 245-275.

Easley, D., and M. O'Hara, 1991, *Order Form and Information in Securities Markets,* Journal of Finance, 46, 3 (July), 905-927.

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

Easley, D. and M. O'Hara, 1987b, *Large Trade Volume and Market Efficiency,* Working Paper, Cornell University.

Easley, D. and M. O'Hara, 1992, *Time and the Process of Security Price Adjustment,* The Journal of Finance, 47, 2(June 1992), pp. 577-605.

Efron, B., and G. Gong, 1983, *A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation*, American Statistician, 37, 36-48.

Eichenbaum, M., L. P. Hansen, and K. S. Singleton, 1988, *A Time Series Analysis of Representative Agent Models of Consumption and Leisure Under Uncertainty,* Quarterly Journal of Economics, 103, 51-78.

Ellis, K., R. Michaely, and M. O'Hara, 2000, *"The Accuracy of Trade Classification Rules: Evidence from Nasdaq"*, Journal of financial and Quantitative Analysis, 35, 529-551.

Engle, R., and J. Russell, 1995, *Forecasting Transaction Rates: The Autoregressive Conditional Duration Model,* UCSD Discussion Paper.

Epps, T. W., 1975, Security *Price Changes and Transaction Volumes: Theory and Evidence*, American Economic Review, 65 (Sept.), 586-597.

Epps, T. W., 1976, *The Demand for Brokers' Services: The Relation Between Security Trading Volume and Transaction Cost,* Bell Journal of Economics, 7 (Spring).

Epps, T. W., 1977, *Security Price Changes and Transaction Volumes: Some Additional Evidence*, Journal of Financial and Quantitative Analysis, 12 (March), 141-146.

Epps, T. W., and M. L. Epps, 1976, *The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis*, Econometrica, 44 (March), 305-321.

Fama, Eugene F., 1970, *Efficient Capital Markets : A Review of Theory and Empirical Work,* Journal of Finance, 25, 383-417.

Fama, Eugene F., 1991, *Efficient Capital Markets : II,* Journal of Finance, 46 (December), 1575-1617.

Farrell, M. J., 1966, *Profitable Speculation,* Economica, 33, May, 183-193.

Fedenia, M., and T. Grammatikos, 1992, *Options Trading and the Bid-Ask Spread of the Underlying Stocks*, The Journal of Business, 65, 3, 335-35.
Fedenia, M., and T. Grammatikos, 1992, *Options Trading and the Bid-Ask Spread of the Underlying Stocks*, The Journal of Business, 65, 3, 335-35.
Ferguson, T. S., 1958, *A Method of Generating Best Asymptotically Normal Estimates with Application to the Estimation of Bacterial Densities*, Annals of Mathematical Statistics, 29, 1046-1062.

Ferson, W. E., and S.R. Foerster, 1994, *Finite Sample Properties of the Generalized Method of Moments in Tests of Conditional Asset Pricing Models,* Journal of Financial Economics, 36, 29-55.

Finucane, T. J., 2000, *"A Direct Test of Methods for Inferring Trade Direction from Intra-Day Data"*, Journal of Financial and Quantitative Analysis, 35, 553-576.

Flood, M. D., R. Huisman, K. G. Koedijk, and R. K. Lyons, 1998, *Search Costs: The Neglected Spread Component*, Fortcoming in the Journal of Financial Intermediation.

Foster, F., and S. Viswanathan, 1989, *Variations in Volumes, Spreads, and Variances*, Research Paper, 88-108, Fuqua School of Business, Duke University, Durham, NC.

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

Foster, F., and S. Viswanathan, 1993, *Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models,* Journal of Finance, 48, (March), pp. 187-211.

Foster, F., and S. Viswanathan, 1994, *Strategic Trading with Asymmetrically Informed Traders and Long-Lived Information,* Journal of Financial and Quantitative Analysis, 29, pp. 499-518.

Foucault, F., 1993, *Price Formation in a Dynamic Limit Order Market,* Working Paper, Groupe Hec, Financial Economics Department.

Foucault, F., 1999, *Order Flow Composition and Trading Costs in a Dynamic Limit Order Market*, Journal of Financial Markets, 2, 2, 99-134.

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

Frino, A., and M. McCorry, 1995, *Why Are Spreads Tighter on the Australian Stock Exchange than on the NYSE? An Electronic Limit order Book versus the Specialist Structure,* Working Paper, University of Sydney.

Gallant, A. R., 1987, *Nonlinear Statistical Models*, Wiley, NewYork.

Gallant, A. R., P. E. Rossi, and G. Tauchen, 1992, *Stock Prices and Volume*, The Review of Financial Studies, 5 (2), 199-242.

Garbade, K., 1978, *Electronic Quotation Systems and the Market for Government Securities*, Federal Reserve Bank of New York Quarterly Review, 3 (Summer), 13-30.

Garbade, K., 1982, *Securities Markets*, McGraw-Hill Inc., New York.

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

Garbade, K., and W. Silber, 1979a, *Structural Organization of Secondary Markets: Clearing Frequency, Dealer Activityand Liquidity Risk*, Journal of Finance, 36, pp. 577-593.

Garbade, K., and W. Silber, 1979b, *Dominant and Satellite Markets: A Study of Dually-Traded Securities,* Review of Economics and Statistics, 61, pp. 455-460.

Gardner, G., A. C. Harvey, and G. D. Phillips, 1980, *An Algorithm for the Exact Maximum Likelihood Estimations of Autoregressive Moving Average Models by means of Kalman Filtering*, Applied Statistics, 29, 311-322.

Garman, M. B., 1976, *Market Microstructure,* Journal of Financial Economics, 3, 257-275.

Gemmill, G., 1996, *Transparency and Liquidity: A Study of Block Trading on the London Stock Exchange under Different Trading Rules*, Journal of Finance, 51, 1765-1790.

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

George, T. J., and F. A. Longstaff, 1993, *Bid-Ask Spreads and Trading Activity in the S&P 100 Index Options Market* , The Journal of Financial and Quantitative Analysis, 28, 3 (September), pp. 381-397.

Geweke, J., R. Meese, and W. Dent, 1983, *Comparing Alternative Tests of Causality in Temporal Systems: Analytic Results and Experimental Evidence*, Journal of Econometrics, 21, 161-194.

Ghysels, E., and A. R. Hall, 1990b, *A Test for Structural Stability of Euler Condition Parameters Estimated via the Generalized Method of Moments,* International Economic Review, 31, 355-364.

Ghysels, E., A. Guay, and A. R. Hall, 1997, *Predictive Test for Structural Change with Unknown Breakpoint,* Journal of econometrics, 82, 209-233.

Glosten, L. R., 1984, *Estimating and Adjusting for the Bid Ask Spread*, University of Chicago Centre for Research in Security Prices, W.P. # 137. November.

Glosten, L.R., 1987, *Components of the Bid - Ask Spread and the Statistical Properties of Transaction Prices*, The Journal of Finance, 42, 5 (December 1987), pp. 1293-1307.

Glosten, L.R., 1989, *Insider Trading, Liquidity, and the Role of the Monopolist Specialist,* Journal of Business, 62 (April), pp. 211-236.

Glosten, L., 1994, *Is the Electronic Limit Order Book Inevitable?,* Journal of Finance, 49, 1127-1161.

Glosten, L. R., and P. R. Milgrom, 1985, *Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders*, Journal of Financial Economics, 14, pp. 71-100.

Glosten, L. R., and L. E. Harris, 1988, *Estimating the Components of the Bid / Ask Spread,* Journal of Financial Economics, 21, pp. 123-142.

Gode, D. K., and S. Sunder, 1990, *Efficiency of Double Auctions with Weakly Rational Traders: Market as a Partial Substitute for Individual Rationality,* unpublished manuscript, GSIA, Carnegie-Mellon University.

Godfrey,M.D., C. W. J. Granger, and O. Morgenstern, 1964, *The Random Walk Hypothesis of Stock Market Behaviour,* Kyklos, 17 (Fasc. 1), 1-30.

Goldman, M. B., and A. Beja, 1979, *Market Prices and Equilibrium Prices: Returns Variance, Serial Correlation, and the Role of the Specialists*, Journal of Finance, 34, 3 (June), pp. 595-607.

Goldman, M. B., and H. B. Sosin, 1979, *Information Dissemination, Market Efficiency and the Frequency of Transactions,* Journal of Financial Economics, 7, 1, 29-61.

Goldstein, M., 1993, *Specialist versus Dealer Markets: A Comparison of Displayed Bid-Ask Spreads on NASDAQ and the NYSE*, Working Paper, Philadelphia, University of Pennsylvania.

Grammatikos, T., and A. Saunders, 1986, *Futures Price Variability: A Test of Maturity and Volume Effects*, Journal of Business, 59 (April), 319-330.

Granger, C. W. J., and O. Morgenstern, 1963, *Spectral Analysis of New York Stock Market Prices*, Kyklos, 16 (Fasc. 1), 1-27.

Greene, J., 1997, *The Impact of Limit Order Executions on Trading Costs in NYSE Stocks*, Working Paper, Georgia State University, Atlanta.

Grossman, S., and M. Miller, 1988, *Liquidity and Market Structure*, Journal of Finance, 43, pp. 617-637.

Grossman, S., and J. Stiglitz, 1976, *Information and Competitive Price Systems*, American Economic Review Proceedings, May, 66, 246-53.

Hagerty, K., 1991, *Equilibrium Bid-Ask Spreads in Markets with Multiple Assets*, The Review of Economic Studies, 58, pp. 237-257.

Hall, A. R., and A. Sen, 1996a, *Structural Stability Testing in Models Estimated by the Generalized Method of Moments,* Discussion Paper, Department of Economics, North Carolina State University, Raleigh, NC.

Hall, A. R., and A. Sen, 1996b, *On the Interpretation of Optimality in Certain Tests of Structural Stability,* Discussion Paper, Department of Economics, North Carolina State University, Raleigh, NC.

Hamao, Y., and J. Hasbrouck, 1995, *Securities Trading in the Absence of Dealers: Trades and Quotes on the Tokyo Stock Exchange,* Review of Financial Studies, 8, 849-878.

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

Hamilton, J., 1978, *Marketplace Organization and Marketability: NASDAQ, the Stock Exchange and the National Market System*, The Journal of Finance, 33 (March), 487-503.

Hamilton, J. L., 1979, *Market Place Fragmentation, Competition, and the Efficiency of the Stock Exchange,* Journal of Finance, 34, pp. 171-187.

Hamilton, J. L., 1987, *Off-board Trading of NYSE-listed Stocks:The Effects of Deregulation and the National Market System,* Journal of Finance, 42, pp. 1331-1446.

Hamilton, J. D., 1994, *Time Series Analysis*, Princeton University Press, Princeton, New Jersey, U.S.A.

Hamon, J., B. Jacquillat, P. Handa, and R. Schwartz, 1993, *The Profitability of Limit Order Trading on the Paris Stock Exchange*, Working Paper, CEREG, Universite Paris-Dauphine.

Handa, P., and R. A. Schwartz, 1996, *Limit Order Trading*, Journal of Finance, 50 (1), 5 (December), 1835-1861.

Hanna, M., 1978, *Security Price Changes and Transaction Volumes: AdditionalEvidence.*" American Economic Review, 68 (Sept.), 692-695.

Hansch, O., N. Naik, and S. Viswanathan, 1998a, *Do Inventories Matter in Dealership Markets? Evidence from the London Stock Exchange*, Journal of Finance, 53, 5 (October), 1623-1656.

Hansch, O., N. Y. Naik, and S. Viswanathan, 1998b, *Preferencing,Internalization, Best Execution and Dealer Profits*, Journal of Finance, forthcoming.

Hansch, O., and A. Neuberger, 1996, *Strategic Trading by Market Makers on the London Stock Exchange*, W.P. no. 224-1996, London Business School.

Hansen, L. P., 1982, *Large Sample Properties of the Generalized Method of Moments Estimators,* Econometrica, 50, 1029-1054.

Hansen, B. E., 1990, *Lagrange Multiplier Tests for Parameter Instability in Non-Linear Models,* Discussion Paper, Department of Economics, University of Rochester, Rochester, NY.

Harris, L., 1984a, *Transactions Data Tests of the Mixture of Distributions Hypothesis*, Working Paper, Univ. of Southern CA, (April).

Harris, L., 1984b, *The Joint Distribution of Speculative Prices and of Daily Trading Volume*, Working Paper, Univ. of Southern CA, (April).

Harris, L., 1986, *Cross-Security Tests of the Mixture of Distributions Hypothesis*, Journal of Financial and Quantitative Analysis, 21 (March), 39-46.

Harris, L., 1986, *Estimation of True Price Variances and Bid-Ask Spreads from Discrete Observations*, Working Paper, University of Southern California, Los Angeles, CA.

Harris, L., 1990, *Liquidity, Trading Rules, and Electronic Trading systems,* New York University Salomon Center Monograph Series in Finance, Monograph 1990-4.

Harris, L., 1994, *Optimal Dynamic Order Submission Strategies In Some Stylized Trading Problems*, Financial Markets, Institutions and Instruments, 7, April, 1-76.

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

Harris, M., and A. Raviv, 1981, *Allocation Mechanisms and the Design of Auctions,* Econometrica, 49, 6 (November), 1477-1499.

Harris, L., and J. Hasbrouck, 1996, *Market vs. Limit Orders: The SuperDot Evidence on Order Submission Strategy*, Journal of Financial and Quantitative Analysis, 31, 2 (June), 213-231.

Hasbrouck, Joel, 1988, *Trades, Quotes, Inventories, and Information*, The Journal of Financial Economics, 22, pp. 229-252.

Hasbrouck, Joel, 1991a, *Measuring the Information Content of Stock Trades*, The Journal of Finance, 46, 1 (March), 179-206.

Hasbrouck, J., 1991b, *The Summary Informativeness of Stock trades: an Econometric Analysis*, Review of Financial Studies, 4 (January), 571-595.

Hasbrouck, J., 1992, *Using the TORQ Database*, Working Paper, NYSE, Research and Planning Section.

Hasbrouk, Joel, 1993, *Assessing the Quality of a Security Market: A New Approach to Transaction Cost Measurement*, Review of Financial Studies, 6, 191-212.

Hasbrouk, Joel, 1996, *Order Characteristics and Stock Price Evolution : An Application to Program Trading*, Journal of Financial Economics, 46, 129-149.

Hasbrouck, J., and R. A. Schwartz, 1986, *The Liquidity of Alternative Market Centers: A Comparison of the New York Stock Exchange, the American Stock Exchange, and the NASDAQ National Market System*, American Stock Exchange, New York.

Hasbrouck, J., and R. A. Schwartz, 1988, *An Assessment of Stock Exchange and Over the Counter Markets*, Journal of Portfolio Management, 14, 10-16.

Hasbrouck, J. and T. S. Y. Ho, 1987, *Order Arrival, Quote Behaviour, and the Return - Generating Process*, The Journal of Finance, 42, 4 (September 1987), pp. 1035-1048.

Hasbrouck, J., and G. Sofianos, 1993, *The Trades of Market Makers: An Empirical Analysis of NYSE Specialists*, The Journal of Finance, 48, 5 (Dec.), 1565-1593.

Hasbrouck, J., G. Sofianos and Sosebee, 1993a, *New York Stock Exchange Systems and Trading Procedures*, Working Paper #93-01, New York Stock Exchange.

Hasbrouck, J., G. Sofianos, and D. Sosebee, 1993b, *Orders, Trades, Reports and Quotes at the New York Stock Exchange*, Working Paper, NYSE, Research and Planning Section.

Hausman, J., 1978, *Specification Tests in Econometrics*, Econometrica, 46, 1251-1271.

Hausman, J., A. Lo, and A. C. MacKinlay, 1992, *An Ordered Probit Analysis of Transaction Stock Prices*, Journal of Financial Economics, 31, 319-379.

Hayashi, Fumio, 2000, *Econometrics*, Princeton University Press, Princeton, New Jersey.

He, Hua, and J. Wang, 1995, *Differential Information and Dynamic Behaviour of Stock Trading Volume,* The Review of Financial Studies, 8, 4 (Winter), 919-972.

Hedvall, K., 1994, *Essays on the Market Microstructure of the Helsinki Stock Exchange,* Unpublished Dissertation, Swedish School of Economics and Business Administration.

Hedvall, K., and J. Niemeyer, 1996, *Order Flow Dynamics: Evidence from the Helsinki Stock Exchange,* Working Paper, StockholmSchool of Economics.

Ho, T., 1984, *Dealer Market Structure: A Dynamic Competitive Equilibrium model,* W.P. New York university.

Ho, T. S. Y., and R. G. Macris, 1984, *Dealer Bid-Ask Quotes and Transaction Prices: An Empirical Study of Some AMEX Options*, The Journal of Finance, 39, 1 (March 1984), pp. 23-45.

Ho, T. S. Y., and R. G. Macris, 1985, *Dealer Market Structure and Performance,* in Y. Amihud, T. Ho, and R. A. Schwartz, eds.: Market Making and the Changing Structure of the Securities Industries, Lexington Books, Lexington, Mass.

Ho, T. S. Y., and H. Stoll, 1980, *On Dealer Markets under Competition,* Journal of Finance, 35, 2 (May), pp. 259-267.

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

Ho, T. S. Y., and H. Stoll, 1983, *The Dynamics of Dealer Markets under Competition,* Journal of Finance, 38, 4 (September), pp. 1053-1074.

Hollifield, B., R. Miller and P. Sandas, 1996, *An Empirical Analysis of a Pure Limit Order Market*, Working Paper, University of British Columbia.

Holthausen, R., R. Leftwich, and D. Mayers, 1987, *The Effect of Large Block Transactions on Security Prices: A Cross-Sectional Analysis,* Journal of Financial Economics, 19, pp. 237-267.

Huang, R., and H. Stoll, 1994, *Market Microstructure and Stock Return Predictions,* Review of Financial Studies, 7, 179-213.

Huang, R., and H. Stoll, 1996, *Dealer versus Auction Markets: A paired Comparison of Execution Costs on Nasdaq and the NYSE*, Journal of Financial Economics, 41, 313-357.

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

Huffman, G. W., 1987, *A Dynamic Equilibrium Model of Asset Prices and Transactions Volume*, Journal of Political Economy, 95, 138-159.

Huffman, G. W., 1988, *A Representative Agent Model of Asset Pricing and Transaction Volume*, Working Paper, Federal Reserve Bank of Minneapolis.

Hughson, E., *Intraday Trade in Dealership Markets*, W.P. Carnegie Mellon University, Pittsburgh, PA., 1989.

Hutchinson, J. M., A. W. Lo, and T. Poggio, 1994, *A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks*, The Journal of Finance, 49 (July), 851-889.

Institutional Investor Study Report of the Securities and Exchange Commission, 1971, 92nd Congress, 1st Session House Document No. 92-64, March 12, Washington: G.P.O., Part 4, Ch. 12.

Jaffe, J. F., and S. Patel, (undated), *A Direct Investigation of the Specialist's Spread and Intraday Price Behaviour,* Unpublished Paper, (The Wharton School, University of Pennsylvania, Philadelphia, PA.)

Jaffe, J., and R. Winkler, 1976, *Optimal Speculation Against an Efficient Market*, Journal of Finance, 31 (March).

Jain, P., and G. Joh, 1988, *The Dependence between Hourly Prices and Trading Volume,* Journal of Financial and Quantitative Analysis, 23, pp. 269-284.

James, C., and R. O. Edmister, 1983, *The Relation between Common Stock Returns Trading Activity and Market Value*, Journal of Finance, 38 (Sept.), 1075-1086.

Jameson, M., and W. Wilhelm, *Market Making in the Options Markets and the Costs of Discrete Hedge Rebalancing*, Journal of Finance, 47, pp. 765-779.

John, K., A. Koticha, and M. Subrahmanyam, 1991, *The Microstructure of Options Markets: Informed Trading, Liquidity, Volatility and Efficiency*, Working Paper, New York University.

Jones, C. M., G. Kaul, and M. L. Lipson, 1994, *Transactions, Volume, and Volatility*, The Review of Financial Studies, 7 (Winter), 631-651.

Kamara, A., 1988, *Market Trading Structures and Asset Pricing: Evidence from the Treasury Bill Markets*, Review of Financial Studies, 1 (Winter), pp. 357-376.

Kandel, E., and N. Pearson, 1995, *Differential Interpretation of Public Signals and Trade in Speculative Markets,* Journal of Political Economy, 103 (August), 831-872.

Karpoff, J. M., 1985, *Costly Short Sales and the Correlation of Returns with Volume,* Working Paper, University of Washington, October.

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

Kavajecz, K. A., 1998, *The Specialist's Quoted Depth as a Strategic Choice Variable*, Working Paper, The Warton School, University of Pennsylvania.

Kavajecz, K. A., 1999, *A Specialist's Quoted Depth and the Limit Order Book*, Journal of Finance, 54, 2 (April), 747-771.

Keim, D.B., 1983, *Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence,* Journal of Financial Economics, 12, pp. 13-32.

Keim, D.B., 1989, *Trading Patterns, Bid-Ask Spreads, and Esimated Security Returns*, Journal of Financial Economics, 25, pp. 75-97.

Ketterer, J. A., and A. Marcet, 1989, *Introduction of Derivative Securities: A General Equilibrium Approach*, Working Paper, Carnegie Mellon University.

Kim, S., and J. P. Ogden, 1996, *Determinants of the Components of Bid-Ask Spreads ofon Stocks, European Financial Management*, 1, 127-145.

Kleidon, A., and I. Werner, 1993, *Round-the-Clock Trading: Evidence from U.K. Cross-listed Securities,* Working Paper, Cambridge, Mass.: National Bureau of Economic research.

Klemperer, P. D., and M. A. Meyer, 1989, *Supply Function Equilibria in Oligopoly under Uncertainty*, Econometrica, 57, 1243-1277.

Knez, P. J., and M. J. Ready, 1996, *"Estimating the Profits from Trading Strategies"*, Review of Financial Studies, 9, 1121-1163.

Koski, J. L., and J. T. Scruggs, 1998, *Who Trades Around the Ex-dividend Day? Evidence from NYSE Audit File Data*, Financial Management, 27, 58-72.

Kraus, A., and H. Stoll, *Price Impacts of Block trading on the NYSE*, Journal of Finance, 27, 569-588.

Kumar, P., and D. Seppi, 1994, *Limit and Market Orders with Optimizing Traders*, Working Paper, Carnegie Mellon University.

Kyle, A., 1989, *Informed Speculation with Imperfect Competition*, Review of Economic Studies, 56, 317-356.

Kyle, A. S., 1985, *Continuous Auctions and Insider Trading,* Econometrica, 53, 6 (November), pp. 1315-1335.

Lai, M., 1996, *Market Transparency and Intra-day Trade Behaviour in the London Stock Exchange*, Ph.D. dissertation, London Business School.

Laszlo Matyas (ed.), 1999, *Generalized Method of Moments Estimation*, Cambridge University Press, Cambridge, U.K.

Leach, C., and A. Madhavan, 1992, *Intertemporal Price Discovery by Market Makers: Active Versus Passive Learning*, Journal of Financial Intermediation, 2, 207-235.

Leach, C., and A. Madhavan, 1993, *Price Experimentation and Security Market Structure,* Review of Financial Studies, 6, 375-404.

Lee, C., 1993, *Market Integration and Price Execution for NYSE-listed Securities,* Journal of Finance, 48 (July), pp. 1009-38.

Lee, C. M. C., and M. J. Ready, 1991, *Inferring Trade Direction from Intraday Data,* The Journal of Finance, 46, 2 (June 1991), pp. 733-746.

Lee, C. M. C., B. Mucklow, and M. J. Ready, 1993, *Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis*, The Review of Financial Studies, 6, 2, pp. 345-374.

Lee, C. M. C., and R. Radhakrishna, 1999, *"Inferring Investor Behaviour: Evidence from TORQ Data"*, Working Paper, Cornell University.

Lee Ruben, 1998, *What is an Exchange?: The Automation, Management, and Regulation of Financial Markets*, Oxford University Press, Oxford, U.K.

Lehmann, B. N., and D. M. Modest, 1994, *Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View,* Journal of Finance, 49, 951-984.

Leroy, S., 1973, *Risk Aversion and the Martingale Property of Stock Returns,* International Economic Review, 14, 436-446.

Lightfoot, L. E., P. G. Martin, M. A. Peterson, and E. R. Sirri, 1999, *"Order Preferencing and Market Quality on United States Equity Exchanges"*, Working Paper, U.S. Securities and Exchange Commission.

Lin, Ji-Chai, G. C. Sanger, and G. G. Booth, 1995, *Trade Size and Components of the Bid-Ask Spread*, The Review of Financial Studies, 8, 4, pp. 1153-1183.

Lo, A. W., and A. C. MacKinlay, 1988, *Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test*, Review of Financial Studies, 1, 41-66.

Lo, A. W., A. C. MacKinlay, and J. Zhang, 1997, *Econometric Models of Limit Order Executions*, Working Paper 6257, National Bureau of Economic Research Working Paper Series, Cambridge, MA.

Logue, D. E., 1975, *Market-Making and the Assessment of Market Efficiency*, The Journal of Finance, 30, 1 (March), 115-123.

Loomis, P., and E. Rotberg, 1964, *Over the Counter Market Quotations*, Michigan Law Review, 62, 589-606.

Lucas, R., Jr., 1978, *Asset Prices in an Exchange Economy,* Econometrica, 46, 1429-1446.

Lucas, H. C., Jr., and R. A. Schwartz, 1989, *The Challenge of Information Technology for the Securities Markets: Liquidity, Volatility, and Global Trading,* Dow Jones-Irwin, Homewood.

Lyons, R. K., 1995, *Tests of Microstructure Hypotheses in the Foreign Exchange Market,* Journal of Financial Economics, 39, 321-351.

Lyons, R., 1996, *Optimal Transparency in a Dealership Market with an application to Foreign Exchange,* Journal of Financial Intermediation, 5, 225-256.

Macey, J., and M. O'Hara, 1997, *The Law and economics of Best Execution*, Journal of Financial Intermediation, 6, 188-223.

Madhavan, A., 2000, *Market Microstructure: A survey*, Journal of Financial Markets, 3, pp. 205-258.

Madhavan, A., 1992, *Trading Mechanisms in Security Markets*, Journal of Finance, 47, 2 (July), pp. 607-641.

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

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

Madhavan, A., and S. Smidt, 1993, *An Analysis of Changes in Specialist Inventories and Quotations*, The Journal of Finance, 48, 5 (December), pp. 1595-1628.

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

Malinvaud, E., 1970, *Statistical Methods of Econometrics*, Amsterdam, North Holland.

Manaster, S., and S. C. Mann, 1996, *Life in the Pits: Competitive Market Making and Inventory Control,* The Review of Financial Studies, 9, 3, 953-975.

Mandelbrot, B., (1971), *When Can Price be Arbitraged Efficiently?,* Review of Economics and Statistics (August), pp. 225-236.

Marsh, T., and K. Rock, 1986a, *Exchange Listing and Liquidity: A Comparison of the American Stock Exchange and the NASDAQ National Market System*, American Stock Exchange, New York.

Marsh, T. A., and K. Rock, 1986b, *The Transactions Process and Rational Stock Price Dynamics,* Unpublished W.P. (Graduate School of Business Administration, University of California, Berkeley, Berkeley, CA).

Martin, W., 1969, chairman, *Joint Treasury-Federal Reserve Study of the U.S. Government Securities Market* (Washington, D.C.: Board of Governors of the Federal Reserve System, 1969).

Mayhew, S., 1993, *Intraday Patterns in Option Market Spreads*, Working Paper, Univ. of California, Berkeley.

McInish, T. H., and R. A. Wood, 1987, *Stock Return Variances: Trading versus Non-Trading Periods*, Unpublished Paper, Dept. of Finance, University of Texas at Arlington, Texas.

McInish, T., and R. Wood, 1992, *An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks*, Journal of Finance, 47 (June), pp. 753-764.

McInish, T. H., and R. A. Wood, 1995, *Hidden Limit Orders on the NYSE*, Journal of Portfolio Management, 21, 19-26.

McInish, T., and B. Van Ness, 2002, *An Intraday Examination of the Components of the Bid-Ask Spread*, The Financial Review, 37, 4, 507-524.

McInish, T., B. Van Ness, and R. Van Ness, 1998, *The Effect of the SEC's Order-Handling Rules on NASDAQ*, Journal of Financial Research, 21, 3, 247-254.

Melamed, L., 1977, *The Mechanics of a Commodity Futures Exchange: A Critique of Automation of the Transaction Process*, Hofstra Law Rev. 6, 149-172.

Melton, W., *Corporate Equities and the National Market System*, Federal Reserve Bank of New York Quarterly Review, 2 (Winter 1978-79), 13.

Meltzer, A., and G. von der Linde, 1960, *A Study of the Dealer Market for Federal Government Securities*, Joint Economic Committee Print, 86th Cong., 2nd Sess. (United States Government Printing Office).

Mendelson, H., 1982, *Market Behaviour in a Clearing House*, Econometrica, 50, 6 (November), 1505-1524.

Menyah, K., and K. Paudyal, 1996, *The Determinants and Dynamics of Bid-Ask Spreads on the London Stock Exchang*, Journal of Financial Research, 19, 377-394.

Menyah, K., and K. Paudyal, 2000, *The Components of Bid-Ask Spreads on the London Stock Exchange*, Journal of Banking and Finance, 24, 1767-1785.

Merton, R. C., 1971, *Optimum Consumption and Portfolio Rules in a Continuous-time Model*, Journal of Economic Theory, 3, 373-413.

Merton, R. C., 1973, *An Intertemporal Capital Asset Pricing Model*, Econometrica, 41 (September), 867-87.

Mildenstein, E., and H. Schleef, 1983, *The Optimal Pricing policy of a Monopolistic Market Maker in the Equity Market,* The Journal of Finance, 38, 1 (March), 218-231.

Miller, M. H., 1991, *Financial Innovations and Market Volatility*, Blackwell Publishers, Cambridge, MA.

Monroe, M., 1988, *Indeterminacy of Price and Quantity in Futures Markets*, J. Futures Markets, 8, 575-588.

Morgan, I. G., 1976, *Stock Prices and Heteroscedasticity*, Journal of Business, 49 (Oct.), 496-508.

Myerson, R., 1981, *Optimal Auction Design*, Mathematics of Operations Research, 6, 58-73.

Naik, N. Y., A. Neuberger, and S. Viswanathan, 1996, *Disclosure Regulation in Competitive Dealership Markets,* W.P., London Business School.

Naik, N. Y., and P. Yadav, 1996a, *Risk Sharing Among Dealers: Evidence of Interdealer Trading on the London Stock Exchange*, W.P., London Business School.

Naik, N. Y., and P. Yadav, 1996b, *Execution Costs and Order Flow Characteristics in Dealership Markets: Evidence from the London Stock Exchange*, W.P., London Business School.

Neal, R., 1992, *A Comparison of Transaction Costs between Competitive Market Maker and Specialist Market Structures*, The Journal of Business, 65, 3, pp. 317-334.

Neftci, S. N., 1991, *Naive Trading Rules in Financial Markets and Weiner – Kolmogorov Prediction Theory : A Study of Technical Analysis,* Journal of Business, 64, 549-571.

Neuberger, A., 1992, *An Empirical Examination of Market Maker Profits on the London Stock Exchange*, Journal of Financial Services Research, 6, 343-372.

Neuberger, A., and R. Schwartz, 1990, *London Equity Market*, Financial Markets and Portfolio Management, 4, 281-301.

Newey, Whitney K., 1985a, *Generalized Method of Moments Specification Testing,* Journal of Econometrics, 29, 229-256.

Newey, W. K., 1985b, *Maximum Likelihood Specification Testing and Instrumented Score Tests,* Econometrica, 53, 1047-1070.

Newey, W. K., and K. D. West, 1987, *Hypothesis Testing with Efficient Method of Moments Estimation*, International Economic Review, 28, 3, 777-787.

Newton, W., and R. E. Quandt, 1979, An *Empirical Study of Spreads*, W.P. Princeton Univ. Dept. Econ.

New York Stock Exchange (NYSE), 1980, *The Quality of the NYSE marketplace in 1979*, Report on the Quality of Markets Committee of the Board of Directors of the NYSE, March 6.

Niederhoffer, V., and M. F. M. Osborne, 1966, *Market Making and Reversals on the Stock Exchange*, Journal of the American Statistical Association, 61, 316 (December), pp. 897-916.

Niemeyer, J., 1994, *Market Microstructure – Empirical Evidence from some Nordic Exchanges*, Unpublished Dissertation, Stockholm School of Economics.

Niemeyer, J., and P. Sandas, 1993, *An Empirical Analysis of the Trading Structure at the Stockholm Stock Exchange*, Journal of Multinational Financial Management, 3, 63-101.

Odders-White, E. R., 2000, *"On the Occurrences and Consequences of Inaccurate Trade Classification"*, Journal of Financial Markets, 3, 259-286.

O'Hara, Maureen, 1997, *Market Microstructure Theory*, Blackwell Publishers Ltd, MA, U.S.A.

O'Hara, M., and G. S. Oldfield, 1986, *The Microeconomics of Market Making*, Journal of Financial and Quantitative Analysis, 21, 4 (December), 361-376.

Oldfield, G., R. Rogalski, and R. Jarrow, 1977, *An Autoregressive Jump Process for Common Stock Returns*, Journal of Financial Economics, 5, 3, 389-418.

Pagano, M., 1989, *Trading Volume and Asset Liquidity*, The Quarterly Journal of Economics, May, 255-274.

Parlour, C. A., 1998, *Price Dynamics in Limit Order Markets*, Review of Financial Studies, 11, 4 (Winter), 789-816.

Parlour, C. A., and D. Seppi, 1997, *Liquidity Based Competition for Order Flow*, Working Paper, Carnegie Mellon University.

Perold, A. F., 1988, *The Implementation Shortfall: Paper Versus Reality*, Journal of Portfolio Management, 14, 4-9.

Petersen, M. A., and D. Fialkowski, 1992, *Price Improvement: Stocks on Sale*, Working Paper, Chicago: University of Chicago, Center for Research in Security Prices.

Petersen, M. A., and D. Fialkowski, 1994, *Posted versus Effective Spreads: Good Prices or Bad Quotes ?*, Journal of Financial Economics, 35, pp. 267-292.

Petersen, M., and S. Umlauf, 1990, *An Empirical Examination of the Intraday Behaviour of the NYSE Specialist*, CRSP W.P. (University of Chicago, Chicago, IL).

Phillips, S., and C. Smith, 1980, *Trading Costs for Listed Options : The Implications for Market Efficiency,* Journal of Financial Economics, 8, 179-201.

Porter, D.C., 1988, *Bid-Ask Spreads: An Examination of Systematic Behaviour Using Intra-Day Data on Canadian and U.S. Exchanges,* Doctoral Dissertation (Faculty of Graduate Studies, University of Western Ontario, London, Ont.).

Porter, D.C., 1992, *The Probability of a Trade at the Ask: An Examination of Interday and Intraday Behaviour*, Journal of Financial and Quantitative Analysis, 27, 2 (June 1992), pp. 209-227.

Porterba, J., and L. Summers, 1988, *Mean Reversion in Stock Returns: Evidence and Implications*, Journal of Financial Economics, 22, 27-60.

Pratt, J. W., 1964, *Risk Aversion in the Small and the Large*, Econometrica, 32, 122-36.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1996, *Numerical Recipes in Pascal: The Art of Scientific Computing*, Cambridge University Press, Cambridge, U.K.

Ready, M., 1996, *The Specialist's Discretion: Stopped Orders and Price Improvement*, Working Paper, University of Wiscocin-Madison.

Ready, M., 1999, *The Specialist's Discretion : Stopped Orders and Price Improvement*, The Review of Financial Studies, 12, 5(Winter), 1075-1112.

Reiss, P., and I. Werner, 1995, *Transaction Costs in Dealer Markets: Evidence from the London Stock Exchange.* In: Lo, A. (Ed.), The Industrial Organization and Regulation of the Securities Industry, University of Chicago Press, 125-176.

Reiss, P., and I. Werner, 1998, *Does Risk Sharing Motivate Interdealer Trading?,* Journal of Finance 53, 1657 – 1703.

Richardson, G., S.E. Sefcik, and R. Thompson, 1986, *A Test of Divident Irrelevance Using Volume Reaction to a Change in Divident Policy*, Journal of Financial Economics, 17 (Dec.), 313-333.

Rock, K., 1996, *The Specialist's Order Book and Price Anomalies*, Review of Financial Studies, forthcoming.

Roell, A., 1998, *Liquidity in Limit Order Book Markets and Single Price Auctions with Imperfect Competition*, Working Paper, Princeton University.

Rogalski, R. J., 1978, *The Dependence of Prices and Volume*, The Review of Economics and Statistics, 36 (May), 268-274.

Roll, Richard, 1984, *A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market*, The Journal of Finance, 39, 4(September 1984), pp. 1127-1139.

Rothenberg, T. J., 1973, *Efficient Estimation with A Priori Information*, New Haven, Conn., Yale University Press.

Riley, J. G., and W. F. Samuelson, 1981, *Optimal Auctions,* American Economic Review, 71, 381-392.

Ross, K., J. Shapiro, and K. Smith, 1996, *Price Improvement of SuperDOT Market Orders on the NYSE,* NYSE Working Paper 96-02.

Ruder, D., and A. Adkins, 1990, *Automation of Information Dissemination and Trading in the U.S. Securities Markets*, paper presented at the Annenberg Washington Program Forum on Technology and Financial Markets, in "Innovation and Technology in the Markets: A Reordering of the World'' Capital Market System"" (D. Siegel, Ed.)., Probus, Chicago.

Rutledge, D. J. S., 1984, *Trading Volume and Price Variability: New Evidence on the Price Effects of Speculation*, In Selected Writings on Futures Markets: Research Directions in Commodity Markets, 1970-1980, A. E. Peck, ed. Chicago: Chicago Board of Trade, 237-251.

Samuelson, P., 1965, *A Proof that Properly Anticipated Prices Fluctuate Randomly*, Industrial Management Review, 6, Spring, 41-49.

Sandas, P., 1998, *Adverse Selection and Competitive Market Making: Empirical Evidence from a Pure Limit Order Market*, Working Paper, The Wharton School, University of Pennsylvania.

Schleef, H. J., and E. Mildenstein, 1979, *A Dynamic Model of the Security Dealer's Bid and Ask Prices*, Paper presented at meetings of the Western Economic Association, Las Vegas.

Schmidt, H., 1977, *Advantages and Disadvantages of an Integrated Market Compared with a Fragmented Market*, (Brussels: Commission of the European Communities).

Scholes, M., and J. Williams, 1977, *Estimating Betas from Nonsynchronous Data*, The Journal of Financial Economics, December, 359-380.

Schultz, P., 1983, *Transaction Costs and the Small Firm Effect*, Journal of Financial Economics, 12, pp. 81-88.

Schwartz, R. A., 1988, *Equity Markets: Structure, Trading and Performance*, Harper and Row, New York.

Schwartz, R. A., 1991, *Reshaping the Equity Markets*, New York, NY: Harper Collins.

Schwartz, R., and D. Whitcomb, 1977, *The Time-Variance Relationship: Evidence on Autocorrelation in Common Stock Returns*, Journal of Finance, 32, 1 (March), 41-55.

Schwartz, R., and D. Whitcomb, 1988, *Transaction Costs and Institutional Investors Trading Strategies*, Monograph Series in Finance and Economics, Salomon Brothers Center for the Study of Financial Institutions, N.Y.U.

Securities and Exchange Commission: *Report of the Special Study of the Securities Markets of the Securities and Exchange Commission*, Part 2, 89th Congress, 1st Session, House Document No. 95, part 2. Washington, D.C.: U.S. Government Printing Office, 1963, pp. 55-171.

Seppi, D., 1990, *Equilibrium Block Trading and Asymmetric Information*, Journal of Finance, 45, 73-94.

Seppi, D., 1997, *Liquidity Provision with Limit Orders and a Strategic Specialist*, Review of Financial Studies, 10, 103-150.

Shapiro, J. E., 1993, *U.S. Equity Markets: A View of Recent Developments*, Working Paper 93-02, New York Stock Exchange.

Sias, R. W., and L. T. Starks, (1997), *Institutions and Individuals at the Turn-of-the-Year*, Journal of Finance, 52, 1543-1562.

Silber, W. L., 1975, *Thinness in Capital Markets: the Case of the Tel Aviv Stock Exchange*, Journal of Financial and Quantitative Analysis, 10, 1, March, 129-142.

Simmons, D., 1971, *Common-Stock Transaction Sequences and the Random Walk Model*, Operations Research, 19, July-August, 845-861.

Simmons, G., 1994, *The Bid-Ask Spread with Shared Trades*, Working Paper, York University.

Simon, M., and R. Colby, 1986, *The National Market System for Over-the-Counter Stocks*, George Washington Law Rev., 17, 34-38.

Smidt, S., 1965, *A Test of the Serial Independence of Price Changes in Soybean Futures*, Food Research Institute Studies, 5, 2, 117-136.

Smidt, S., 1968, *A New Look at the Random Walk Hypothesis*, Journal of Financial and Quantitative Analysis, September, 235-262.

Smidt, S., 1971, *Which Road to an Efficient Stock Market?*, Financial Analysts' Journal, sep. – Oct., 18-20; 64-69.

Smidt, S., 1979, *Continuous versus Intermittent Trading on Auction Markets*, Journal of Financial and Quantitative Analysis, 14, 4 (November), 837-872.

Smirlock, M., and L Starks, 1985, *A Further Examination of Stock Price Changes and Transactions Volume*, Journal of Financial Research, 8 (Fall), 217-225.

Smith, V., *Bidding Theory and the Treasury Bill Auction: Does Price Discrimination Increase Bid Prices?,* Review of Economics and Statistics, 48 (May 1966), 141-146.

Snell, A., and I. Tonks, 1995, *Determinants of Price-quote Revisions on the London Stock Exchange,* The Economic Journal, 105, 77-94.

Snell, A., and I. Tonks, 1998, *Testing for Asymmetric Information and Inventory Control Effects in Market Maker Behaviour on the London Stock Exchange,* Journal of Empirical Finance, 5, 1-25.

Son, G., *Dealer Inventory Position and Intraday Patterns of Price Volatility, Bid/Ask Spreads and Trading Volume,* W.P., Department of Finance, Kellogg School, Northwestern University.

Sowell, Fallaw, 1993, *A New Approach to Testing Parameter Instability in the Generalized Method of Moments Framework,* Working Paper, Carnegie Mellon University, Graduate School of Industrial Administration.

Sowell, Fallaw, 1994, *An Approach to Specification Tests in the Generalized Method of Moments Framework,* Working Paper, Carnegie Mellon University, Graduate School of Industrial Administration.

Sowell, Fallaw, 1996a, *Optimal Tests for Parameter Instability in the Generalized Method of Moments Framework,* Working Paper, Carnegie Mellon University, Graduate School of Industrial Administration.

Sowell, Fallaw, 1996b, *Tests for Violations of Moment Conditions,* Working Paper, Carnegie Mellon University, Graduate School of Industrial Administration.

Stigler, G., 1961, *The Economics of Information,* Journal of Political Economy, June, 69, 213-25.

Stigler, G., 1964, *Public Regulation of the Securities Markets,* The Journal of Business, 37, 117-133.

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

Stoll, H. R., 1978a, *The Supply of Dealer Services in Securities Markets,* The Journal of Finance, 33, 4 (September), 1133-1151.

Stoll, H. R., 1978b, *The Pricing of Security Dealer Services : An Empirical Study of the NASDAQ Stocks,* The Journal of Finance, 33, 4 (September), 1153-1172.

Stoll, H. R., 1979, *Regulation of Securities Markets: An Examination of the Effects of Increased Competition,* Monograph Series in Finance and Economics, 1979-2, 83 pages.

Stoll, H. R., 1985, *The Stock Exchange Specialist System: An Economic Analysis,* In Monograph Series in Finance and Economics, Monograph 1985-2. New York Univesity, Salomon Brothers Center.

Stoll, H. R., 1989, *Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests*, The Journal of Finance, 44, 1 (March 1989), pp. 115-134.

Stoll, H., 1990, *Principles of Trading Market Structure*, W.P. 90-31, Vanderbilt University.

Stoll, H., and R. E. Whaley, 1983, *Transaction Costs and the Small Firm Effect,* Journal of Financial Economics, 12, pp. 57-79.

Stoll, H., and R. E. Whaley, 1990, *Stock Market Structure and Volatility,* Review of Financial Studies, 3, pp. 37-71.

Subrahmanyam, A., 1990, *A Theory of Trading in Stock Index Futures*, Review of Financial Studies, 4, pp. 17-51.

Tauchen, G., and M. Pitts, 1983, *The Price Variability-Volume Relationship on Speculative Markets*, Econometrica, 51 (March), 485-505.

Telser, L. G., 1959, *A Theory of Speculation Relating Profitability and Stability*, Review of Economics and Statistics, 41, August, 295-301.

Teweles, R. J., and E. S. Bradley, 1998, *The Stock Market*, 7th edn, John Wiley and Sons, Inc., New York.

The Downstairs Insider: The Specialist and Rule 10b-5., New York University Law Review, 42 (1967), 695-715.

The TAQ Database, New York Stock Exchange, Inc., version 3.1, January 1995.

Tinic, S. M., 1972, *The Economics of Liquidity Services,* Quarterly Journal of Economics, 86, (February), 79-93.

Tinic, S. M., and R. R. West, 1972, *Competition and the Pricing of Dealer Services in the Over-the -Counter Market,* Journal of Financial and Quantitative Analysis, 7 (June)....

Tinic, S., and R. West, 1974, *Marketability of Common Stock s in Canada and the U.S.A.: A Comparison of Agent versus Dealer Dominated Markets*, Journal of Finance, 29 (June), 729-746.

Tinic, S., and R. West, 1980, *The Securities Industry under Negotiated Brokerage Commissions: Changes in the Structure and Performance of the NYSE Member Firms*, Bell Journal of Economics, 11 (Spring), 29-41.

Tobin, James, 1984, *On the Efficiency of the Financial System,* in 'The Market on Trial, Lloyds Bank Annual Review', Pinter Publishers Ltd, London, 1989.

Van Ness, B., R. Van Ness, and R. Warr, 2002, *Is the Adverse Selection Component Really Higher on the NYSE/AMEX than on the NASDAQ?*, The Journal of Business Finance and Accounting, 29, 5&6, 807-824.

U. S. Securities and Exchange Commission (SEC), 1971, *Institutional Investor Study Report.* 92d Cong., 1st sess. House Document no. 92-64.

U. S. Securities and Exchange Commission (SEC), 1963, *Report of the Special Study of Securities Markets.* 88thCongress, 1st session, Washington D.C.: U.S. Government Printing Office, Chap. 5. 1963

U.S. Securities and Exchange Commission, 1997, *Report on the Practice of Preferencing.*

Vickrey, W., 1961, *Counterspeculation, Auctions, and Competitive Sealed Tenders,* Journal of Finance, 16 (March), 8-37.

Vijh, A., 1990, *Liquidity of CBOE Equity Options*, Journal of Finance, 45 (September), pp. 1157-1180.

Visuanathan, S., and J. J. D. Wang, 1996, *Auctions with When-Issued Trading: A Model of the U.S. Treasury Markets,* Working Paper, Duke University.

Visuanathan, S., and J. J. D. Wang, 1998, *Why is Inter-Dealer Trading so Pervasive in Financial Markets?,* Working Paper, Duke University.

Visuanathan, S., and J. J. D. Wang, 1999, *Market Architecture: Limit Order Books Versus Dealership Markets*, Working Paper, Fuqua School of Business, Duke University, March, 1-50.

Wahal, S., 1996, *Entry, Exit, Market Makers and the Bid-Ask Spread*, Review of Financial Studies, forthcoming.

Wang, Jiang, 1994, *A Model of Competitive Stock Trading Volume,* Journal of Political Economy, 102, 1, 127-168.

Wang, J., and J. Zender, 1996, *Auctioning Divisible Goods*, Working Paper, Duke University.

Werner, I., and A. Kleidon, 1996, *U.K. and U.S. Trading of British Cross-Listed Stocks : An Intraday Analysis of Market Integration*, Review of Financial Studies, 9, 619-654.

West, R. R., and S. M. Tinic, 1971, *The Economics of the Stock Market,* New York: Praeger Publishers.

Westerfield, R., 1977, *The Distribution of Common Stock Price Changes: An Application of Transactions Time and Subordinated Stochastic Models,* Journal of Financial and Quantitative Analysis, 12, (December), 743-765.

White, K. J., 1997, *SHAZAM: The Econometrics Computer Program, Version 8.0, Users's Reference Manual*, Irwin / McGraw-Hill.

Wilson, R., 1979, *Auctions of Shares*, Quarterly Journal of Economics, 93, 675-698.

Wolfson, N., and T. Russo, 1970, *The Stock Exchange Specialist: An Economic and Legal Analysis*, Duke Law Journal.

Wolinski, A., 1990, *Information revelation in a Market with Pairwise Meetings*, Econometrica, 58, 1-25.

Wood, R., T. McInish, and J. Ord, 1985, *An Investigation of Transactions Data for NYSE Stocks*, Journal of Finance, 40, pp. 723-739.

Working, H., 1954, *Price Effects of Scalping and Day Trading*, Proceedings of the 7th Annual Symposium Commodity Markets and the Public Interest.

Working, H., 1958, *A Theory of Anticipatory Prices*, American Economic Review, 188-199, May.

Ying, C. C., 1966, *Stock Market Prices and Volumes of Sales*, Econometrica, 34 (July), 676-686.

Zabel, E., 1981, *Competitive Price Adjustment without Market Clearing*, Econometrica, 49, 5 (September), 1201-1221.