AN ANALYSIS OF BID - ASK SPREADS CONSIDERING ASPECTS OF RISK INSURANCE, DEGREE OF COMPETITION, AND MARKET LIQUIDITY

Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

by

Silvia A. Gerber - Helbling

UNIVERSITY OF YORK

Department of Economics and Related Studies

MAY 1994

An meine Eltern und

an Ruedi

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am grateful to my supervisor Professor Peter Simmons whose patience and constructive counsel have well exceeded the obligations of a thesis supervisor.

My sincere thanks are due to the members of my supervision panel, Professor David Mayston and Peter Smith for their helpful comments and also thanks are due to the external examiner Dr. Ailsa Roell from the London School of Economics for the interesting discussion. I am also indebted to Ester Arisi who generously provided a detailed data set.

Financial support from the ESRC in project grant is gratefully acknowledged which enabled me to undertake this research.

A special mention is due to all my collegues and fellow research students who have given me their support in many invaluable conversations during the preparation of this thesis. Especially, I thank Gaia Garino for her kind hospitality and patience during the final preparation of the thesis.

Above all, I would like to thank my husband who has been a constant source of support and encouragement to me throughout my time at University.

DECLARATION

Chapter 2 and chapter 6 are based on works undertaken jointly with Professor Peter Simmons.

An early version of that work was presented at the European Meeting of Econometric Society in Brussels in August 1992.

In those joint works the effort and contribution may be attributed to both the authors in equal parts.

Section 4.3. in chapter 4 is also collaborative work with Professor Peter Simmons.

ABSTRACT

Financial market trading is investigated with respect to profit margins of the market makers. We analyse the bid-ask spread of market makers in a centralised market and in a fragmented market structure in respect of risk insurance and degree of competition which in turn influences market liquidity. Risk insurance can be obtained by sharing of the market order between risk averse dealers or through diversification of the portfolio which enables the market maker to hedge some of the risk. Market makers can reduce their risk exposure by trading in various assets or by being active in more than one market at the same time. Thus under the assumption of decreasing returns to scale risk averse market makers are prepared to share a market order. We also investigate the influence of futures trading on the spot market bid-ask spread.

In part one, the bid-ask spread is analysed in respect of divisibility of the market order and diversification possibilities into different correlated markets such as the spot and the futures market. We show that a market where market makers can split the order is Pareto superior to a market where the order is indivisible. In addition, our finding is that trading in futures contracts has various impacts on the spot market bid-ask spread depending on the trading information available in the spot market. Our analysis of the bid-ask spread follows the inventory control argument and does not investigate any influence of asymmetry of information.

Part two provides empirical evidence of some of the theoretical issues. Based on daily data of the Italian secondary market for government bonds we obtain supportive evidence of the inventory control argument and the next best dealer aspect based on our theoretical models. An alternative bid-ask spread analysis partly confirms these findings.

LIST OF TABLES

LIST OF ILLUSTRATIONS

(Figures)

GLOSSARY

Chapter 4:

$c_0 = c_t$	cash holding at time t
c_{t+1}	cash holding at time t+1
cov(.)	covariance between spot and futures prices
cov(a), cov(b), cov(NT)	covariance of spot and futures prices of a seller, a buyer, and an inactive dealer in the spot market respectively
E(.)	expectation taken over a random variable
$E(U)_0$	expected utility of "no trade"
$E(U)_1$, $E(U)_2$	expected utility at the end of period 1 and 2
$I_0 = I_t$	inventory position in a risky asset at time t
I_{t+1}	inventory position in a risky asset at time t+1
k	coefficient of absolute risk aversion
λ	probability that an order arrives in the market
M = T	total number of dealers in the futures market
μ_s, μ_f	mean of the expected spot and futures prices
$N = N_1$	net futures position at time t+1
N_2	net futures position at time t+2
N_s, N_B, N_{NT}, N_{SP}	net futures positions of the seller, the buyer, and the inactive dealer in the spot market, and the speculator respectively
pa, pb	spot market bid and ask prices
p_t, p_{t+1}	prices at time t and t+1
p^s, p^f	spot market price and futures market price
$p^f_t = p^f_0$	futures price at time t
r	risk free interest rate

ρ	correlation coefficient $[cov(.)/\sigma^2_{p_s}\sigma^2_{p_f}]$
$\sigma^2_{p_s}, \sigma^2_{p_f}$	price variance in the spot and the futures market (with subscript 1, 2 = in period 1, 2)
U(.)	von-Neuman Morgenstern utility function
V_t, V_{t+1}	value of the portfolio at time t and t+1
VAR	variance
V	number of buyers in the spot market
W	number of inactive dealers in the spot market
W_1, W_2	terminal wealth at the end of period 1 and 2
W_0	initial wealth at time t
X	size of purchase (and sale) order
(X-Y)	difference between purchases and sales
Y	number of sellers in the spot market
Z	number of speculators in the futures market

INTRODUCTION

The theme of this thesis is that of profit margins of market makers in dealership markets.

Market makers are dealers [1] who are obliged to quote their bid and ask prices (buying and selling prices respectively) for a particular asset and to meet incoming orders from the public at these quoted prices within a certain time period. The ever-changing environment leads to the requirements of continuous presence and high flexibility of market makers in the market. These requirements are not without cost and the market maker gets her return of market making by the bid-ask spread which has been recognized in the literature by the 'transaction cost approach' studied among others by Demsetz (1968).

This theory about the determinants of the bid-ask spread which has grown rapidly over recent years is called the theory of the microstructure of market making.

The bid-ask spread is defined as the return to the market makers for standing ready to buy or sell an asset at their bid and ask price quotes. The dealer or market maker faces uncertainty by having to quote prices without knowing about the nature and the size of the incoming public order. In order to be able to fulfill the order the dealer has to carry inventory of the risky asset. The inventory carrying aspect has been developed by Garman (1976), Amihud and Mendelson (1980), Ho and Stoll (1980, 1981, 1983), and by O'Hara and Oldfield (1986), to mention just the most prominent studies.

---

[1] We use the expression market maker, dealer, and trader interchangeably with the same meaning unless it is otherwise explicitly mentioned. The public is referred to as the private investor, the private trader, or the customer.

Most of these theoretical models, except the Ho and Stoll (1981, 1983), models [2], analyse the bid-ask spread of a monopoly dealer and fail to account for competition between market makers. The results of Ho and Stoll show that the equilibrium bid and ask prices are determined by the next best dealer's price quotes which deviates from the monopoly case.

Another source of uncertainty in the market is the presence of informed investors who possess superior information which results in a loss for an uninformed dealer who trades with such an informed investor. The analysis of Bagehot (1974) is the first study which considers the asymmetry of information in the market. Other subsequent and more elaborate studies are Jaffe and Winkler (1976) and Copeland and Galai (1983) which are followed by a number of other investigations. Thus, to stay in the market, the market makers set their bid and ask prices in a way that the resulting bid-ask spread covers the cost of a dealer coming from the risk inherent in such uncertainty.

However, there is a problem in respect of the size of the spread. As the the bid-ask spread becomes larger the less likely it is that there are some incoming orders of the public, because the trading in the market is too expensive. In turn, the high cost encountered by the market makers in such a thin market will not attract other market makers as there are no profit opportunities. Hence, the market becomes less liquid. The final consequence is that the market becomes illiquid and breaks down. The problem of market thinness and market liquidity has been examined by Garbade and Silber (1979), Grossman and Miller (1988) and Pagano (1989).

One of the main concerns of our research is the problem of risk insurance, particularly in respect of the inventory position of the market maker.

---

[2]The bid and ask prices are actually analysed by using a duopoly model.

Hence we will focus on the issues of inventory carrying costs and leave aside the asymmetry of information problem.

By analysing the determinants of the bid-ask spread we will try to find ways which reduce costs of market making and which result in a smaller bid-ask spread and a more liquid market.

There are several ways of risk insurance for market makers. Risk averse market makers may reduce their risk exposure by trading in smaller quantities. Another source of risk reduction is inter-dealer trading. Due to differences between dealers positions, i.e. degrees of risk aversion, inventory positions and future price expectation, inter-dealer trading may be profitable for one or the other dealer. It will also narrow the difference between the dealers' positions. The risk inherent in trading can also be reduced through diversification. Market makers can choose whether they want to diversify into various assets or into different markets.

The first approach to such risk reduction is the investigation of risk averse dealers with decreasing returns to scale in their cost structure which results in a convex cost curve. In order to meet the demand and supply of the public, dealers have to hold a stock of a risky security with unknown future price. In a competitive market dealers must quote the best price, in case of selling it is the lowest ask price and in case of buying it is the highest bid price, in order to get any trade.

If we think of a trading environment where market makers quote their prices for a fixed order quantity which is known to them, then the risk can be reduced by allowing the splitting of the incoming order between the best quoting dealers. Thus, the market makers are able to trade smaller quantities which reduces their cost of inventory carrying and hence with reduced costs, the bid-ask spread is smaller.

Until now, we argued about the cost of market making based on the individual dealer's costs of uncertainty which has its roots in the unknown future price of the asset, the unknown time period of carrying the inventory, and the cost arising from risk aversion. However, if we analyse the market maker's pricing strategy we find that the market structure, or more general the trading environment, has an important influence on the market bid-ask spread.

The transparency of markets is crucial for the trading procedure. Pagano and Roëll (1990) investigate trading procedures of various regulated stock exchanges and over-the-counter markets (OTC) and find that the pricing strategies of market makers are influenced by the market structure. The factors which determine the trading procedure in the market are the knowledge of the dealers about the incoming order, the knowledge of the reservation prices of each other, and the knowledge about the trading history, especially about the last trade, before they have to quote their prices. We can define the type of market structure by these factors.

If, for instance, market makers know each others' reservation prices and the last trade is made public immediately after it was executed and market makers know the order flow, then we speak of a centralised market structure. In contrast, if market makers only know their own reservation price and the last trade is not immediately made public then we call this structure a fragmented market.

Biais (1993) compares the bid-ask spread of a centralised market and a fragmented market and finds that the bid-ask spread is the same for both markets. This result is obtained under quite restrictive assumptions. Biais assumes that the average of the expected prices in a rational expectations equilibrium is equal to the average of prices which can be observed in a centralised market. We attempt to show that the bid-ask spread is not the same in a centralised and in a fragmented market structure. Again, we investigate the risk reduction under the assumption of decreasing returns to scale.

Another possibility of keeping the cost of trading low is diversification. With correlated asset returns market makers can reduce the price risk by trading in different assets. Ho and Stoll (1983) analysed the effects of a diversified portfolio on the bid-ask spread in a model with two risky assets. The impact on the spread comes from the risk incurred by the deviation from the optimal (or preferred) inventory level after a transaction has been executed for one asset. As dealers are assumed to balance their inventory at the end of the period, the spread is independent of the inventory level. Therefore, their finding is that the bid-ask spread is not affected by the diversification into two assets. However, they make the crucial assumption that the prices of the assets are correlated but not the transactions of the assets. If we change this assumption the result will change.

On the other hand, market makers have the possibility to be active in various markets at the same time which may give them the opportunity to reduce the risk if markets are correlated. One such possibility is that dealers may be in the position to reduce their risk of carrying inventory of the risky asset by hedging the risk through trading in futures contracts and so hedge the price risk.

The fact that market makers are active in more than one market at the same time calls for an investigation of the interaction between such correlated markets. We can find a variety of studies which analyse the effects of futures trading on the spot market prices. The well known theories regarding the interaction of spot and futures markets are the traditional theory of storage (Keynes (1930) and Hicks (1939) amongst others), the theory of risk premium (Dusak 1973) and Breeden (1980)), and the forecast power of futures prices (Grossman (1976) and Kyle (1985)).

In addition the work of Anderson and Danthine (1983) investigates the effects of futures trading on the spot prices within the framework of the microstructure of market making.

However, none of these analyses investigates the influence of futures trading on the spot market bid-ask spread which may show that market makers can obtain risk insurance by trading in futures contracts which enables them to narrow the bid-ask spread in the spot market.

On the empirical side, we find a variety of studies examining the determinants or components of the bid-ask spread for a centralised and a fragmented market structure. There is evidence of the inventory control aspect and a component which explains the asymmetry of information between market makers, as in Hasbrouck (1988) and Stoll (1989). However the next best dealer argument has not been empirically investigated so far.

This thesis aims to contribute to the theory of the microstructure of market making by examining the aspect of risk insurance and degree of competition which influences market liquidity.

There is an ongoing change in the design of financial markets such as spot or futures and options markets. The type of trading procedure is subject to the particular characteristics of the market. These characteristics are the number and types of market participants which changed from relatively small investors to institutional traders who encourage or even call for large block trading. This is a challenge for the market makers who have to be able to absorb such a demand in trading. These market makers for instance in a dealership market are more professional than a few years earlier. These market makers are mostly international banks or large broker companies who know the market very well.

Our analyses are intended to give some support for the decision makers in designing the respective market structure which ensures an efficient trading procedure and which is Pareto optimal for a particular trading environment.

The thesis is divided into two parts of which the first part contains theoretical work and the second part presents the empirical analyses.

The first part contains chapters one to four.

Chapter one gives an overview of the most relevant theoretical research in the area of the microstructure of market making. The literature survey gives the reasoning for the existence of a bid-ask spread and the role of the spread in respect of market liquidity. Furthermore, theoretical models are presented which explain the determinants of the bid-ask spread including empirical studies which investigate the components of the bid-ask spread.

In chapter two we investigate how the bid-ask spread is affected by assuming risk averse dealers and decreasing returns to scale of the dealers' reservation price function. We allow for splitting of the public order which means that the dealer faces lower costs by trading a smaller quantity and therefore she can reduce the bid-ask spread. Such a model is set in a competitive market where we can have the situation that the number of active dealers is different on the buying and on the selling side. Such a framework has not been investigated until today. In addition, we present such a model for a centralised market structure and also for a fragmented market where market makers do not know each others reservation prices.

Chapter three investigates the influence of futures trading on the spot market prices. The interaction of the spot and futures market is presented in this survey chapter. We analyse and discuss the theory of storage, the concept of risk premium, the forecast power of futures prices and the term structure of interest rates. All these models explain the bias between the spot price at time T and the price of a maturing futures contract at time T. The lack of the investigation of futures trading on the spot market bid-ask spread leads us to the next chapter.

In chapter four we present a bid-ask spread model which accounts for trading in futures contracts. We carry out our analysis for two different trading situations. On one side, we assume that market makers know the order flow. On the other side, we assume that the market makers do not know the order flow in the spot market which means that they face two types of uncertainty which are price and quantity uncertainty.

We expect that with trading in futures the market makers are able to hedge some of the price risk of their inventory position and thus they reduce the spot market bid-ask spread. Our findings show that under the assumption of symmetry of trading on the selling and on the buying side there is no influence of trading in futures on the spot market bid-ask spread in the case where the market makers do not know the order flow. Therefore, we extend our analysis and let the amounts of selling and buying differ. We also analyse how our results change if we consider two periods instead of one period only. The results of this rigorous analysis give interesting insights regarding risk insurance for market makers and the interaction of markets.

Part Two of the thesis includes chapters five to seven which are all empirical studies.

This second part is intended to provide some empirical evidence for the theoretical issues discussed in part one. The empirical studies are based on data of the Italian Secondary market of government bonds. This market was reorganised in May 1988 with the creation of the secondary market in which primary dealers are obliged to quote their bid and ask prices for at least five assets for a given period. These price quotes are binding for a quantity up to a fixed amount and the prices are displayed on a computerised information system. The actual trade with the public is done on the telephone. The traded deals are reported to a central unit and the aggregate volume but not its division is public information.

The data obtained are daily time-series of bid and ask quotes which have been taken from the information system between 12.00 a.m. and 1.00 p.m. which represents the most active trading time of the day. [3]

Chapter five investigates the pricing strategies of the primary dealers in the market. The daily data exhibit a distinct pattern of quoting frequency of the various dealers. One can ask whether some dealers may take advantage in quoting more frequently in one asset or another.

We argued in part one that there is no asymmetry of information in the dealership market which can be explained by the professionalism of market trading with sophisticated information systems which allow that information is quickly and evenly spread among dealers. To test whether this is the case in the Italian secondary market, we analyse the quoting behaviour of the market makers.

We assume that dealers who quote very actively in a particular asset can gain better information about the asset, especially in respect of the future price. If this is the case we have asymmetry of information among dealers. Under the assumption that the other dealers recognize that the "specialised" dealers have superior information, we expect that the bid-ask spread in such an asset is larger compared to the other assets due to the

---

[3]We are very grateful to Ester Arisi for providing such an extensive data set. The data set used for the research is available from the author.

asymmetry of information. We assume that all the assets in this market have the same systematic risk which implies that the difference in the bid-ask spread (or returns) of various assets comes from differences in information. We employ different methodologies and compare the outcomes. In particular, we analyse the level of activity of the dealers by means of a cluster analysis. We then compare the findings of grouping together the various dealers with the results of an ordinary least squares analysis on the returns based on the price quotes. If the OLS result shows differences in returns for some dealers or some assets we may say that the market is segmented. If this is so, dealers have arbitrage opportunities which indicates that the market is inefficient and not Pareto optimal.

Chapter six is closely linked to chapter two in which we develop bid-ask spread models for the centralised and the fragmented market structure. These models assume that market makers are allowed to share the market order. The analyses are based on the inventory control argument. Thus we investigate the determinants of the bid-ask spread within a similar setting to the models of chapter two. Furthermore, we test the hypothesis of the next best dealer's price quotes.

We investigate the price quotes in respect of the inventory control argument for the centralised market structure in the Italian secondary market. In addition, we analyse whether equilibrium prices are in fact determined by the second best dealer, i.e. we try to find evidence of the next best dealer argument.

Our study includes two different analyses. One of them is an ordinary least square analysis which examines the determinants of the price quotes for each dealer separately for the bid and the ask side of the market. In contrast to the existing studies our models assume risk averse dealers who have reservation price functions with decreasing returns to scale.

In addition we investigate which dealer is likely to quote the best price due to her individual parameters which determine the reservation price. Such an analysis is based on a probit estimation. According to the theoretical model in chapter two we expect that a dealer who is not competitive which implies that her reservation buying (selling) price is below (above) the market price does quote her reservation price or does not quote at all. If the dealer's reservation price is the same as the market price the dealer quotes her reservation price and will share the market order. If the actual reservation buying (selling) price of the dealer is above (below) the market price then the dealer quotes just below the market price and gets the whole order. The results show how well our hypotheses predict the pricing behaviour of the market makers.

Chapter seven contains an empirical bid-ask spread analysis based on the model of Stoll (1989). The serial covariance of price changes is compared with the respective bid-ask spread. The serial covariance is explained by the inventory control effect. This analysis is an alternative measure of the components of the bid-ask spread. The underlying assumption is that, in an efficient market with a constant bid-ask spread over time, any change in the price can only be due to the spread or better the cost of trading (as discussed by Demsetz (1968)). This measure based on the inventory control argument means that a dealer who holds inventory intends to remain on this inventory level. If, for instance, the dealer sells a certain quantity then she deviates from this level. In order to induce trade which enables her to get back to the initial level, she increases her bid and ask prices which makes it more likely that the next transaction will be a purchase. The spread is then determined as a function of the probability of a price reversal and the magnitude of an adverse price change.

We analyse empirically the relationship between the serial covariance of returns calculated from daily price quotes and the square of quotes spreads where the empirical results of Stoll do not give any conclusive evidence. We extend Stoll's model by adding a variance component analysis which helps us to identify whether there are market inefficiencies and whether there are differences in the covariance between the bonds.

The empirical investigation is carried out with daily data. The result indicates a positive serial correlation instead of a negative which is what we expect according to Stoll's inventory control theory. As a consequence, we also estimate the model on the basis of weekly data which slightly changes the findings.

The final chapter contains our concluding remarks in which we summarise and discuss the various results of our theoretical and empirical analyses. In addition, we present an outlook for further research to be undertaken in this area.

CHAPTER ONE

THE MICROSTRUCTURE OF MARKET MAKING AND THE BID-ASK SPREAD

1.1. Introduction

Technological and informational developments, especially in the financial markets indicate that the traditional economic models of the financial market no longer describe the situation in the real world.

Traditional theories focus on more static analysis and consider the trading activity in the financial market as a one shot process. This kind of process gradually changed to a continuous trading procedure. In addition, the agents of the market place are confronted with a random demand and supply function which implies uncertainty about the flow of orders both on the demand as well as on the supply side.

The earlier studies of bid-ask spread analysis examine a single market maker and try to determine the cost of trading for such a dealer. For determining the bid-ask spread all the costs of a dealer in the market have to be considered. The main components of such cost are firstly the inventory carrying cost, i.e. the opportunity cost of financing inventory.

Second, there is the cost of immediacy, i.e. the cost of providing immediate service in the market by matching buy or sell orders at any point in time or even continuously, and thereby carrying the risk of uncertainty of future order arrivals and price changes. This uncertainty in turn influences the cost of holding inventory and is therefore a major determinant of the cost of a dealer.

Third, there is a risk bearing cost added, because the dealer faces uncertainty not only from uncertainty about order arrivals in the market; but also from uncertainty about the future price of the asset.

An additional complication which increases the cost to the dealer is asymmetry of information in the market, i.e. there are some dealers who possess superior information than others and it arises an adverse selection problem for the uninformed dealer. If we focus on the cost of immediacy or the cost of transaction we have to consider the problem of uncertainty of the future price of the asset. An examination of the price volatility of an asset leads to an analysis of the underlying market structure, i.e. the market depth and the market liquidity.

By assuming that the risk increases with the time horizon, it is evident that the rate of arrival of the market orders are crucial for the dealer's profit function as the longer she has to carry a position the greater the risk taken and the greater the inventory carrying cost involved.

More recently, researchers noted that the market structure and the organisation of the market plays an important part in determining the cost of transaction. This may be due to the development of information technology and the change to continuous trading which even led to "international" trading by which we mean that stocks can be listed on more than one exchange at the same time and therefore stocks are traded simultaneously at different exchanges. These exchanges differ in their market organisation, i.e. they employ different trading systems, which may bring an advantage to one exchange or another. A further aspect of trading is that competition, amongst market makers on the one side and between exchanges on the other side, is more pronounced than before.

This chapter gives a critical introduction to the field of the analysis of the bid-ask spread and then leads to the unanswered questions which will be dealt with in the subsequent chapters of this thesis.

Section two starts with a basic discussion about the cost of transaction.

It is a good introduction and presents the earliest works in this area. We also include the analysis of market liquidity which in turn is one of the determinants of the bid-ask spread in equilibrium. If the market is thin (which means that trading can not take place due to no (or insufficient) demand or supply), the market makers encounter uncertainty in the form of long intervals of trading. This in turn implies that they take on risky stock which they have to carry for a long time before they can sell it again. Such additional cost increases the bid-ask spread. This circumstance may lead to the market eventually collapsing as the investors (or customers) are not willing to trade at such high costs.

Section three deals with various models which explain the determinants of the bid-ask spread in detail. The analysis is divided into three parts.

Part one examines the problem of a so called "preferred inventory position". A dealer decides, based on her price expectation, how much of the risky stock she wants to keep in her position to be able to meet the demand.

By trading in the market the dealer deviates from this "optimal" position, which means that the dealer faces increased costs due to the larger or smaller inventory. All of these models analyse the market situation with a single dealer as supplier in the market. However, as we already mentioned, the competition among dealers seems to be more and more the case in today's trading environment. Hence, although these studies give a valuable insight they are not accurate anymore.

Part two presents more viable models in so far as several dealers are considered. Market makers are risk averse and due to their differences in price expectation, risk aversion, and inventory positions they compete in prices for the order demand. Such inventory control models are the basis of further research which also captures the importance of the underlying market structure which has not been exploited so far.

Part three contains studies on the asymmetry of information between market makers. Some dealers may have superior information about a particular asset which gives them a profit opportunity. Market makers who cannot distinguish whether they trade with an informed or an uninformed dealer may make a loss due to trading with an informed dealer. Thus, by taking into account this adverse selection problem they increase their bid-ask spread to compensate for an eventual loss. However, if we consider today's markets in which a computerized information system is present such superior information may become less important than other determinants. Asymmetry of information is also influenced by the information aggregation through the trading procedure.

Although we focus on the inventory control aspect in our thesis, we still include this line of argument to have a complete overview of the literature.

The final part of section three gives the results of empirical studies which evaluate the components of the bid-ask spread for various markets.

The principal factors by which market structures differ are given in section four. The respective questions for determining a particular market structure (or system) are the following:

Firstly, does trading take place in discrete intervals or continuously? Secondly, do the market makers know the order demand before they have to quote their prices? This question determines the nature of the trading procedure to be either an auction or a pure dealership market.

Thirdly, do market makers know the reservation prices of each other which means that a market is either centralised or fragmented.

Based on such an analysis, the differences between an auction market and a dealership market are analysed. The final section of this chapter naturally contains the conclusions and the outlook for the subsequent chapters in this thesis.

1.2. Cost of Transacting

Market makers, acting as specialists in the market, quote their prices which are fixed for a given period of time and for a particular asset. They undertake to buy an asset at the quoted bid price and similarly they undertake to sell an asset at the quoted ask price. It is common practice that the price is quoted for a standard volume of the respective asset.

This first change in the financial market concerns the evolution from the 'call markets', which means trading synchronously at pre-established discrete times, to 'continuous markets', implying asynchronous trading during continuous intervals of time.

The market makers (specialists) earn their living on the bid-ask spread which is their return of offering their services of continuous trading.

In order to ensure such a service they often hold their own portfolio which seems to be important considering uncertainty about the order flow and the future market price.

The specialist hopes, of course, to realize a profit on inventory turnover. She would like to acquire inventory at low prices and resell at high prices and to do so very rapidly.

One of the earliest analysis of the bid-ask spread was carried out by Demsetz (1968). In his general approach he analyzes the cost of transaction at the New York Stock Exchange (NYSE) and he examines the determinants of the bid-ask spread in a dealership market.

The major aim of his paper is to investigate the extent to which transaction costs are affected by the scale of trading. He argues that the inclusion of the bid-ask spread in transaction costs can be understood best by considering the neglected problem of 'immediacy' in supply and demand analysis. On the grounds that waiting costs are important cost for trading in organized markets, it is obvious that they dominate the determination of the spread. In addition, it seems reasonable to assume that waiting costs will be reduced most rapidly when the transaction rate is small and increasing.

The bid-ask spread is then the markup that is paid for predictable immediacy of exchange in organized markets.

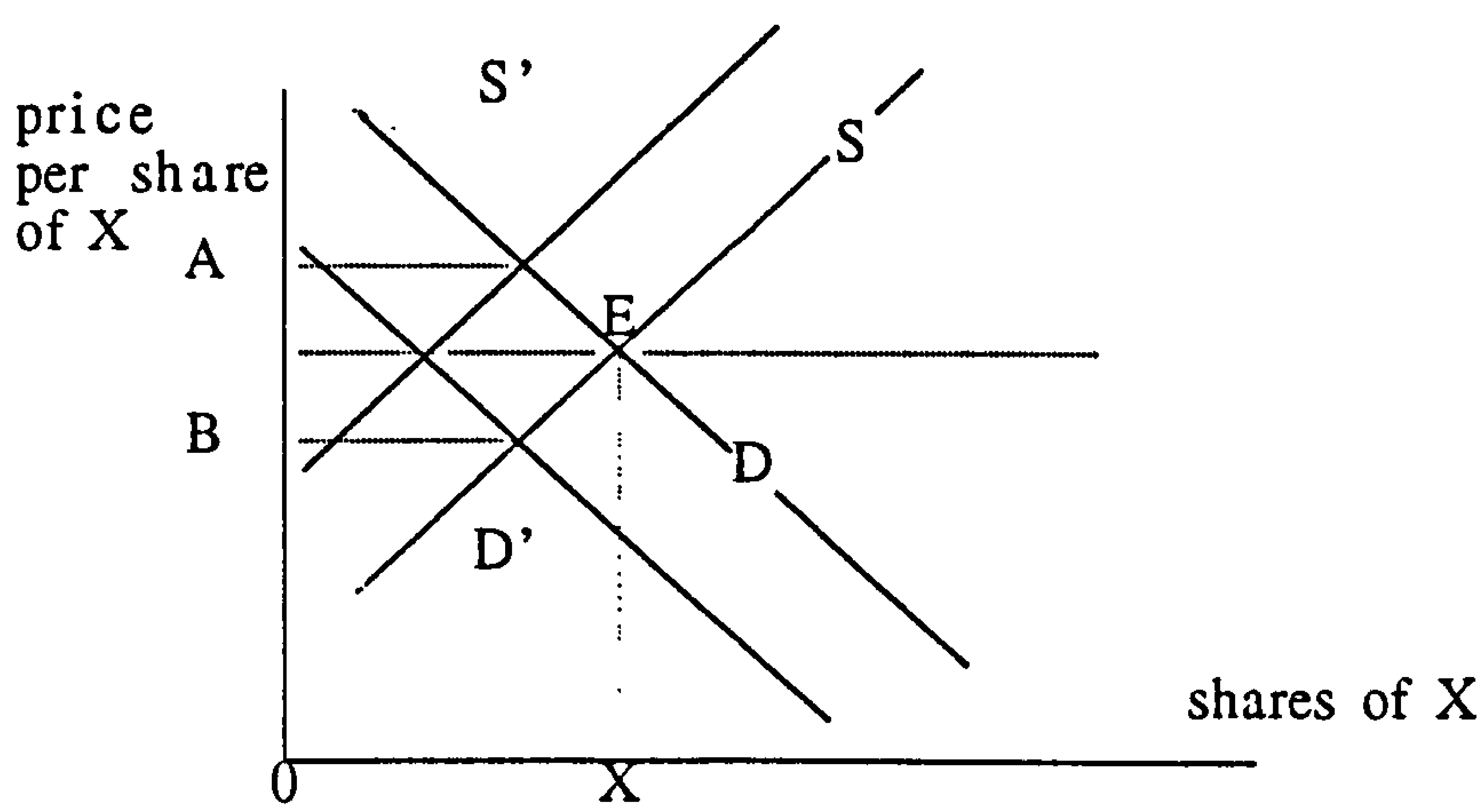

Figure 1.1.: Cost of immediacy

Figure 1.1. represents the price formation in an asset market. If a buy and a sell order arrive at the same time, with dealers having a demand (buying) function of D and a supply (selling) function of S, then the equilibrium quantity is X. The average of the bid and the ask price is E.

However, most of the time orders do not coincide in time. Therefore, dealers are prepared to offer a service of immediacy, but at new demand and supply curves of D' S'. By trading a quantity X the equilibrium prices are at A for selling and B for buying. The difference between A and B is the bid-ask spread.

Regarding the spread, Demsetz argues that, even though scale economies are present in the specialist's trading activities, there is little likelihood of her maintaining spread much above the cost of waiting. Competition of several types will keep the observed spread close to cost. Furthermore he defines that the main types of competition emanate from 1) rivalry for the specialist's job, 2) competing markets, 3) outsiders, who submit limit orders rather than market orders, 4) floor traders who may bypass the specialist by crossing buy and sell orders themselves, and 5) other specialists.

He predicts that the cost of exchanging a security declines as trading activity in that security increases which is based on the assumption that the market is in a competitive situation.

Garbade and Silber (1976) enlarge this approach by arguing that the nature of the exchange process has been ignored. For many goods we can observe competing inventory specialists who stand ready to buy and/or sell on demand at prices they have posted. Such quotes will be dispersed over some range, giving an incentive for search by public transactors. This phenomena of price dispersion search, and bid-ask spreads are alien to a Walrasian world but appear pervasive in the real world.

Garbade and Silber are examining the dispersion of quotations made by the dealers. They point out that it is of interest to note that the presence of a dealer market is itself an efficient response to the greater price dispersion which would exist in its absence.

For a security traded in the market with competing specialists, the expected round trip cost is the expected transaction spread $(pa_{t+1}-pb_t)$, with pa_{t+1} being the ask price at time t+1 and pb_t being the bid price at time t, plus a term which reflects the cost and extent of searching for favorable quotations. If the volume of the transaction is sufficiently large, the investor may choose to contact every dealer, so that if there are n market makers the cost of liquidity is min (p_a) - max (p_b) + 2nC/V whereby C is the cost of search and V is the volume of the transaction. In the limit as V increases, the cost of liquidity services converges to the spread between the best quotes on either side of the market. On the other hand, for small-volume transactions or for investors entering the market only infrequently the expected cost of liquidity may be substantially greater. Garbade and Silber comment that this suggests that those investors who are concerned with the cost of liquidity services, will, ceteris paribus, restrict their investment to issues which trade on narrow spreads and which are characterized by compact dispersions. The larger volume of trading in these issues will tend to further reduce both the spread and the dispersion as well. This implies that there is simultaneity between trading volume and dispersion as well as between trading volume and spreads.

If we turn now to the market makers behaviour in respect of price dispersion we can say that there are five major reasons for the difference in prices: different inventory policies, heterogeneous expectations of future security prices, instability in supply-demand conditions, different cost functions, and ignorance of other dealer quotes.

Garbade and Silber extended the concept of the cost of transacting to include search costs in a dealer market. Dealers quote different prices because they are ignorant of the quotations of other dealers. This could lead to the extreme case, that one dealer may be bidding on an issue above another dealer's asked price. However, the more trading, and hence search, the greater the probability that some investor will uncover the arbitrage. Such arbitrage limits the range of dispersion of quotations.

In addition, the dispersion of quotations in a dealer market leads to transactional inefficiencies as well as the imposition of search costs. Furthermore, interdealer transactions allow dealers to adjust their inventory positions efficiently and thereby limit the dispersion of their quotations.

In a subsequent paper, an interesting aspect of risk in the financial market has been taken up by Garbade and Silber (1979). Their key variable is the liquidity of the financial market which they link to the clearing frequency in the market. The longer an asset has to be carried in the inventory position the bigger the risk about the future price of this asset taken by the market maker. The measure of risk is defined as the variance of the difference between the equilibrium value of an asset and its value at the time a market participant decides to trade and the time when the trade is reversed.

The price variance can be divided into two parts. The first part includes the risk run by the investor that the equilibrium price may change from the moment the investor decides to trade until the time the trade is completed. The second part of the liquidity risk is the variance of the difference between contemporaneous transactions prices and equilibrium values. Hence, the clearing prices will usually differ from the equilibrium price derived from a Walrasian auction. It follows that the longer the time between clearings, the greater the number of participants in the clearing.

As a consequence of above results we can derive the optimal clearing frequency which is the time interval that minimizes the liquidity risk.

Furthermore, Garbade and Silber show that dealer participation reduces the liquidity risk born by the public transactor.

An empirical investigation of Tinic (1972) in the market of the New York Stock Exchange (NYSE) shows that

1. The price of liquidity service increases as a direct function of the price of the asset and the level of trading concentration.

2. Liquidity costs are lower for issues that experience continuous and heavy trading activity.

3. Dealers can make better markets in which there are greater opportunities for self-equating block transactions. Therefore, assets with a larger number of institutional investors possess better marketability than others in which only a few investors hold very large blocks.

4. Sample findings indicate that units registered in more numerous securities charge higher prices, on the average, for their liquidity services.

5. Prices for liquidity services are more stable for stocks that experience continuous trading activity, a larger number of transactions, and lower prices.

In another paper issued by Tinic and West (1972) we can find the examination of the influence of competition among dealers on the bid-ask spread. Their basic hypothesis is that the spread behaviour is a function of 1) a stock's trading volume, 2) its price level, 3) a measure of its price volatility, and 4) the extent of competition among dealers. Based on their results they conclude that the explanatory variables such as price, trading activity, and the intensity of competition are probably the basic determinants of the size of bid-ask spreads.

Their principal conclusion is that increases in the amount of interdealer competition in this market tends to reduce the price of dealer services (reduce spreads) and thus, tends to increase the marketability of issues. This conclusion suggests that dealership activities in the OTC stocks do not entail economies of scale as significant as those that have been reported by Demsetz and the NYSE for the exchanges.

In the same line of argumentation are Cohen, Maier, Schwartz and Whitcomb (1981). Based on their empirical investigation they conclude that thinner securities will, ceteris paribus, have larger equilibrium market spreads. They come to this result by carrying out the following analysis.

They have established that with transaction costs the probability of a limit order executing does not rise to unity as the price at which the order is placed gets infinitesimally close to a counterpart market quote. This can be explained by examining the investor's behaviour. We assume that an investor places a limit order to buy with a price below the market order price. If these two prices move closer together then, at a certain point, the investor has to consider whether a small increase in the price is more desirable which means to trade a market order instead of waiting until the limit order is executed. Hence, the closer the prices of limit and market orders [1] the more likely it is that the investor trades a market order instead of a limit order. This situation is referred to as the "gravitational pull effect".

This means, essentially, in the neighborhood of the current market bid and ask quotations, what would have otherwise been limit orders, are instead submitted as market orders (at slightly less desirable prices) so as to

[1]With a limit order the dealer places an order at a certain (limit) price in the order book. The execution of such an order is not certain. On the other hand, the market order clears the limit order at the market price and thus the execution of a market order is certain.

achieve certainty of execution.

Such market orders trigger trades which clear limit orders off the book, widening the market spread. The gravitational pull effect explains why market spreads may be substantial even in markets composed of many traders. Thus they have shown that the market bid-ask spread (equilibrium spread) is positively related to a security's thinness (measured inversely by the order arrival rate).

Their policy recommendations are to expand the extent and frequency with which investors interact with the market by minimizing various transaction costs. For example, decreasing variable transaction cost will decrease individual spreads and generate a greater order flow.

Overall, these models show that a bid-ask spread exists because of transaction costs in the market. We observe that the size of the spread is linked to the market liquidity. However, none of the studies establishes the level of liquidity in equilibrium. This is analysed in the next section.

1.2.1. Market Liquidity

Grossman and Miller (1988) examined the liquidity and the market structure by formulating a simple model which captures the essence of market liquidity. Exogenous liquidity events coupled with the risk of delayed trades create demand for immediacy. However, in the long run the number of market makers adjusts to equate the demand and supply for immediacy. This determines the equilibrium level of liquidity in the market. They argue that the lower is the autocorrelation in rate of returns, the higher is the equilibrium level of liquidity.

The basic feature of their model is discussed in order to describe their predictions. There are two groups of market participants, market makers and outside customers, with identical risk tolerance. The model incorporates three dates (1, 2, and 3) which is illustrated in figure 1.2. below.

At date 1, a liquidity event occurs which creates a temporary order imbalance of size i. Market makers offset this temporary imbalance by taking trading positions which they hold until the next date 2.

At date 2, the market makers offset their positions as other outside customers arrive to offset the imbalance.

Date 3 is introduced only as a terminal condition for valuing the securities as of date 2.

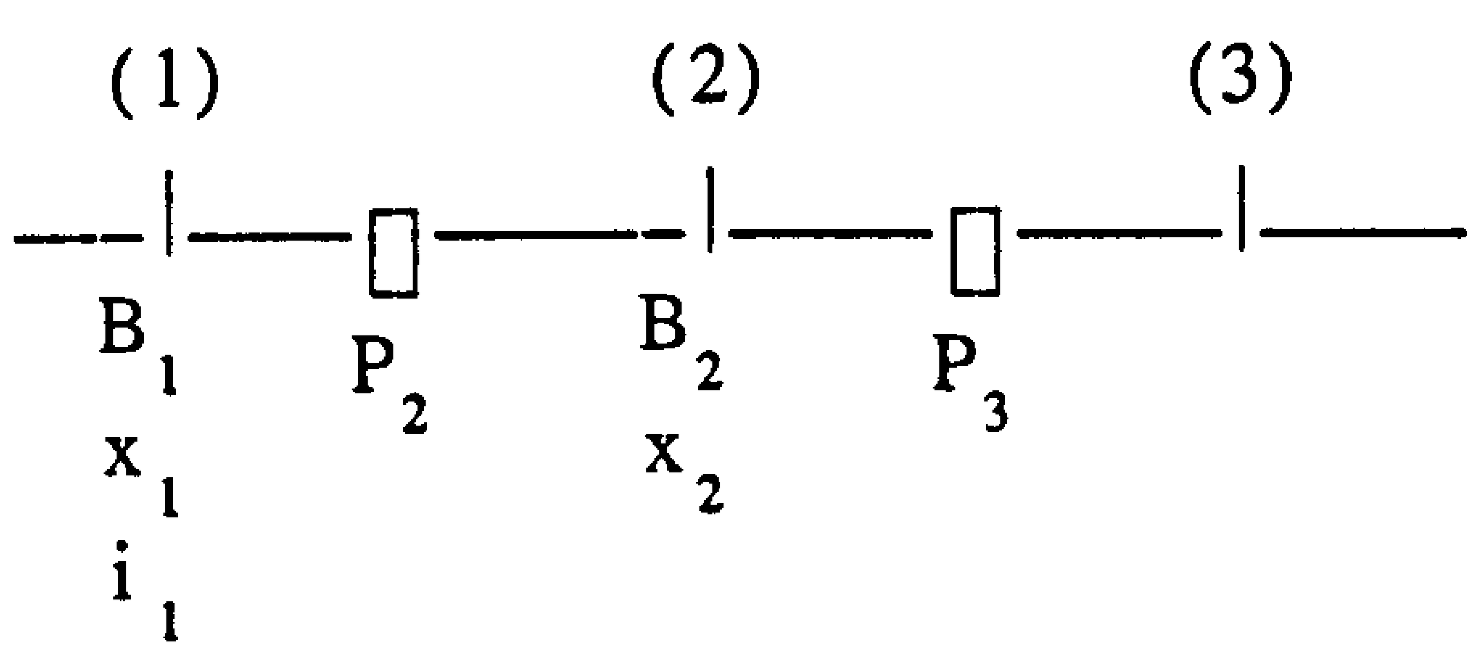

Figure 1.2. : Time sequence of events

Furthermore, two assets are considered, a risk free asset (cash with zero rate of return) and a risky asset with liquidation value P_3. It is assumed, that at times t=1,2 the customer chooses asset holdings $\bar{x}_t$ and a risk free asset position B_t to maximize the expected utility of terminal wealth (i.e. at date 3)

$EU(W_3)$

subject to the constraints

$W_3 = B_2 + \bar{x}_3\tilde{P}_3$ $(= B_2 + \tilde{P}_3\bar{x}_2 + i\tilde{P}_3)$

$\tilde{P}_2\bar{x}_2 + B_2 = W_2 = B_1 + \tilde{P}_2\bar{x}_1$

$P_1\bar{x}_1 + B_1 = W_1 = P_1 i_1 + W_0$

where i_1 represents the excess holding of the asset on top of the initial endowment of the asset [2] and W_0 represents other wealth with $i_1 = i$ and $\bar{x}_3 = \bar{x}_2$. These constraints say that the only gain in wealth comes from trading in this asset.

By elimination of B_1 and B_2 we get

$$W_3 = W_0 + (\tilde{P}_2 - \tilde{P}_1)(\bar{x}_1 - i_1) + (\tilde{P}_3 - \tilde{P}_2)(\bar{x}_2 - i_1) + \tilde{P}_3 i$$

where $\bar{x}_t - i_1$ is the excess demand for the asset with $t = 1,2$.

Dealers are assumed to maximise

$$EU(W_3) = EU(W_0 + (\tilde{P}_2 - P_1)x_1 + (\tilde{P}_3 - \tilde{P}_2)x_2 + \tilde{P}_3 i)$$

The utility function is defined as $U(W) = -e^{-aW}$ (a = constant absolute risk aversion coefficient).

In addition, it is assumed that $\tilde{P}_t$ is normally distributed and that the dealers have mean-variance utilities. The optimisation problem is solved by backward induction.

We define x^{cd}_2 to be the optimal value of x_2 (chosen at date 2). Hence, the maximisation of E(U) over x_2 is

$\max E_2U(W_2 - P_2 i_1 + (\tilde{P}_3 - \tilde{P}_2)x_2 + \tilde{P}_3 i)$ and by solving it we get

$$x^{cd}_2 = \left[(E_2\tilde{P}_3 - P_2)/a\ var_2(\tilde{P}_3)\right] - i$$

Under the assumption that all the customers are identical except in respect of i, x^{cd}_2 represents the aggregate demand of the customers.

In addition, M is defined as the number of market makers who do not have any endowment of the risky asset at time 1, i = 0.

Hence, the excess demand per market maker is x^{md}_2 and the total excess demand of all the market makers in period 2 is:

---

[2] i_1 is regarded as an excess holding of the asset in respect of the customer's preference of an optimal inventory level.

$$M\, x^{md}_{2} = M \left[(E_2\tilde{P}_3 - P_2)/a\ var_2(\tilde{P}_3)\right]$$

The service of immediacy is required by the asynchronisation of trades. The imbalance of period 1 is offset in period 2. Only due to this asynchronisation of trades market makers enter the market otherwise there would not be any trade.

Thus, above excess demand in period 2 is counterbalanced in period 1 by an "excess demand" of opposite sign of new customers which is

$$\left[(E_2\tilde{P}_3 - P_2)/a\ var_2\tilde{P}_3\right] + i.$$

The market clearing condition is that the various demands should sum to zero which is:

$$\left[(E_2\tilde{P}_3 - P_2)/a\ var_2\tilde{P}_3\right] - i + M \left[(E_2\tilde{P}_3 - P_2)/a\ var_2\tilde{P}_3\right]$$
$$+ \left[(E_2\tilde{P}_3 - P_2)/a\ var_2\tilde{P}_3\right] + i = 0$$

As period 3 is regarded as only a terminal condition it follows that

$$(E_2\tilde{P}_3 - P_2) = 0.$$

Thus, the equilibrium excess demand at date 2 of the customer arriving at the market at date 1 is:

$$x_2^{cd} = -i\ .$$

For the market makers, clearing at date 1 requires $Mx_1^{m} + x_1^{cd} = 0$.

The date 1 demand for the customer can be derived from the maximisation over x_1

$$\max E_1 U(W_0 + x_1(\tilde{P}_3 - P_1) + i\ E_2\tilde{P}_3)$$

The respective excess demands of the customers and the market makers become

$$x^{cd}_{1} = \left[(E_1\tilde{P}_3 - P_1)/a\ var_1(E_2\tilde{P}_3)\right] - i \text{ and } x^{md}_{1} = \left[(E_1\tilde{P}_3 - P_1)/a\ var_1(E_2\tilde{P}_3)\right]$$

The market clearing condition at date 1 is: $M\ x^{md}_{1} + x^{cd}_{1} = 0$. By substitution the clearing condition can be expressed as

$\left[(E_1\tilde{P}_3 - P_1)/a\ var_1(E_2\tilde{P}_3)\right] = [i/(1 + M)].$

Let $\tilde{r} = \tilde{P}_2/(p_1 - 1)$ be the excess return earned by the market makers, then

$E_1\tilde{r} = P_1 i/[(1+M)aVar_1(\tilde{r})]$

$E_1\tilde{r}$ deviates from zero due to the asynchronization of the order flow and the finite risk bearing capacity of market makers. In determining the number of market makers, we can say that the gain from being in the market is the ability to trade at price P_1. Thus free entry of market makers will occur until

$EU(W_0 - c + (\tilde{P}_2 - P_1)x_1^m) = EU(W_0)$

with c being the dealer's opportunity cost of being in the market.

The results in equilibrium show that the lower the cost of maintaining a market presence, the greater the number of market makers in equilibrium. That number would also be larger, the smaller the risk aversion coefficient "a" for the market makers.

Hence, the opportunity cost of maintaining a presence in the market is very important in determining the supply for immediacy and the services for market making. The contribution of the market makers can be found in the correlation between successive price changes. Grossman and Miller prove that the correlation between successive price changes is negative and is determined by the cost of being in the market. Therefore, the demand for immediacy depends on the volatility of the underlying price and the diversifiability of the risk of an adverse price move.

Finally, we can say that the greater the demand for immediacy and the lower the cost to market makers of maintaining a continuous position in the market, the larger the proportion of the transactions between ultimate customer effected initially through market makers, and hence, the more liquid the market. In such a liquid market, the spread is expected to be small.

The amount of immediacy provided in equilibrium can be measured by the amount of customer trade, since the total size of the trade desired is -i, the fraction completed in period 1 is determined by M. When M is very large the transaction is executed immediately and the market is said to be liquid.

We have discussed the model of Grossman and Miller in the context of the spot market, but it is equally applicable for an analysis of trading in futures. However, there is little attention given to their model in respect of the interaction of the spot and futures market although they show the influence of trading in futures on the spot prices.

Another approach has been taken by Pagano (1989) for examining market thinness and and stock price volatility. Generally, thin markets are characterized by small numbers of transactors per unit time, and subsequently their prices are more sensitive to the impact of individual trader's demand shocks. This leads to the observed relationship between market size and price volatility, by taking market size as the exogenous factor. The market size is measured by the amount of orders and the ability to absorb, for instant, large bulk orders without an increase in the price volatility.

However, in this study, Pagano argues that the volatility of a speculative market may feed back on its size, in the sense that the high liquidation risk implied by very volatile prices can induce potential entrants to keep out of the market. Thus, thinness and the related price volatility may become joint self-perpetuating features of a market, irrespective of the volatility of the asset fundamentals.

The paper shows that in a stock market with transaction costs, this interaction between thinness and volatility can produce multiple steady-state equilibria, some characterized by low trade and high volatility, and others by high trade and low volatility. If expectations are formed on the basis of the previous history of the market, its thinness or depth will become a self-perpetuating feature.

An important extension of the model is the introduction of imperfect competition. It stresses that there are two distinct ways by which entry of additional traders can be said to make a market more liquid. This can be done either by reducing the price volatility due to uncorrelated demand shocks or by decreasing the adverse price response to the order flow.

Until now, we discussed the issue of transaction costs in the market. We analysed models which explain the existence of a bid-ask spread which is regarded as a return to market makers who provide a service of immediacy of trading a risky asset. In turn, the supply of immediacy is dependent on the market activity such as the trading volume which again feeds back to market liquidity. Thus, the bid-ask spread is a crucial factor in sustaining a proper market functioning.

If the bid-ask spread is large less customers are attracted and trading is not very active. This thinness of the market increases the risk for market makers to supply immediacy and also the price volatility.

Hence, the next issue which we examine is the analysis of the determinants of the bid-ask spread.

1.3. Determinants of the Bid-Ask Spread

The literature about the bid ask spread can mainly be divided into three groups. The first group contains models concerned with the market maker's pricing strategies and the 'optimal' inventory level. These models assume that market makers maximize their expected profits which consists of the gains from trading and the profits from their portfolio. The second group examines the effects of uncertainty, and risk on the bid-ask spread. The third group considers the influence of asymmetry of information on the bid-ask spread.

1.3.1. Pricing Strategies and 'Optimal' Inventory Level

A first fully developed issue of explaining the pricing behaviour of traders in an auction and in a dealership market has been presented by Garman (1976) [3].

His objective is to describe the 'temporal microstructure' of one shot trading activities in asset markets. He departs from the usual approach of the theory of exchange by (1) making the assumption of asynchronous, temporally discrete market activities on the part of market agents and (2) adopting a viewpoint which treats the temporal microstructure, i.e. moment to moment aggregate exchange behavior, as an important descriptive aspect of such markets. Garman's definition of demand and supply functions is set in a stochastic framework which gives rise to the concept of temporal imperfections.

A stochastic process is defined $\{N(t), t \in [0, \infty)\}$ with $N_i(t) \in \{0,1,...\}$ being the cumulative number of discrete points in time where the good has been demanded up to time t. Furthermore, $Y_{in}(t_{in})$ is the amount demanded by

---

[3]We restrict our discussion to the case of the dealership market which includes all the relevant issues on the determinants of the bid-ask spread.

customer i at the n'th point in time, given that it occurs at time t_{in}. The total amount demanded by the i'th customer in the interval [0, t] is

$$X_i(t) = \sum_{n=1}^{N_i(t)} Y_{in}(t_{in}).$$

On the assumptions:

1. that there are a large number of market agents, 2. that agents act independently in selecting the timing of their orders, 3. that no small subset of agents dominates overall order generation, 4. that no agent can generate an infinite number of orders in a finite period of time. Garman defines the superposition of the individual demand processes as $X(t) = \sum_{i=1}^{M} \sum_{n=1}^{N_i(t)} Y_{in}(t_{in})$ which converges to a Poisson process as M (the number of individual market agents) becomes very large. [4]

In addition, the mean-value function of X(t) is $E[X(t)] = \lambda_B(t) = \sum_i \lambda_{Bi}(t)$ with λ being the Poisson rate.

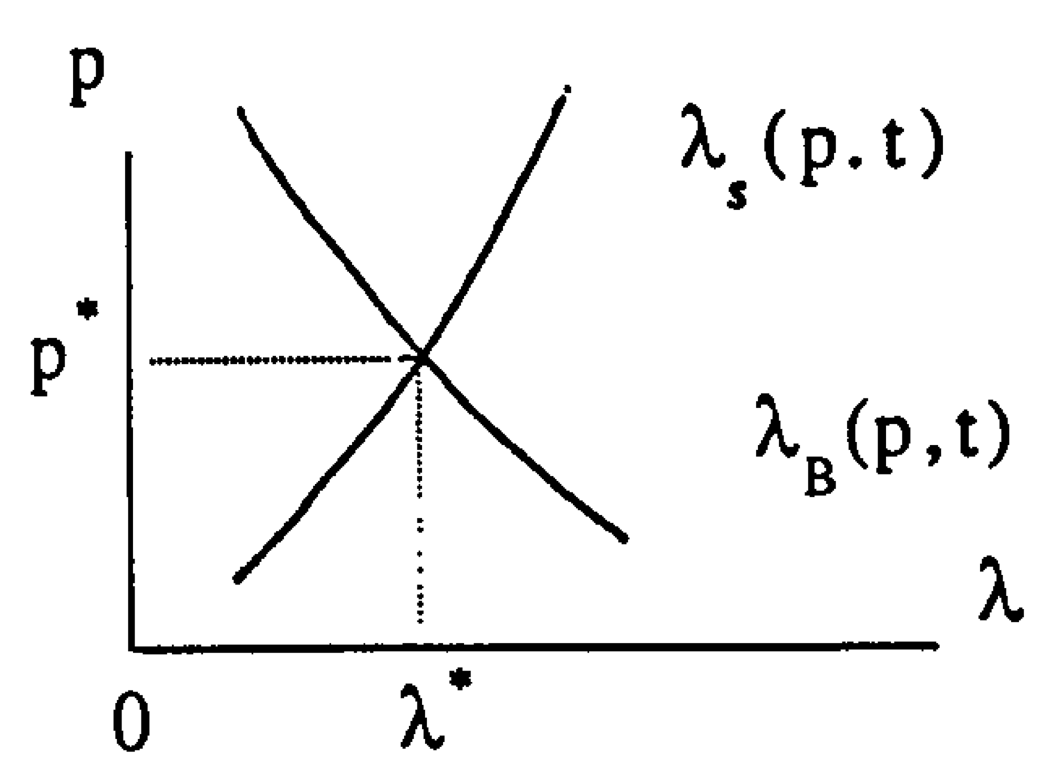

Figure 1.3. : Type of equilibrium

The Poisson rates $\lambda_s(p,t)$ and $\lambda_B(p,t)$ for the supply and the demand side respectively (given in figure 1.2.) are instantaneous rates. These rates are mean-value functions which depend on the price p and t.

---

[4]What has been assumed so far is equally valid for the supply side.

Garman describes the rates as follows: [5]

...these rates are 'on the average' amounts and do not necessarily have specific physical realizations in the marketplace; we acknowledge that there will be sampling fluctuations of actual demand and supply within a continuous market. This leads in turn to a different interpretation of the market 'equilibrium' point (p^*, λ^*) as it now represents a 'stochastic' equilibrium in which actual prices and quantities may fluctuate randomly, even under conditions of stationarity in the stochastic order process.

Furthermore, the stochastic demand can be expressed by aggregate price probability functions. This concept is similar to traditional theory where it is assumed that there are several "latent" demands, given that the market price is at an arbitrary price p at time t which means that the demand rate will be $\lambda(p,t)$. Usually the price p is the equilibrium price at time t. At this point in time t, only this demand rate is active and the remainder of the demand rate function is not coming into force. The actions of the customers are influenced by a range of latent demand and supply functions.

In the stochastic case, we do not need this scenario. Instead, we define aggregate price probability functions. So, the probability that an incoming order may be traded at price p at time t is defined as the price probability function.

For a dealer dominated market (monopoly situation), Garman makes the following additional assumptions: [6]

1. Arrivals of buy and sell orders to the market are Poisson distributed in time, with stationary rate functions $\lambda_B(p)$ and $\lambda_S(p)$; q (order quantity) is assumed equal to 1.
2. All exchanges are made through a single 'central market maker', who possesses a monopoly on all trading. No direct exchanges between buyers and

[5]Garman pp. 260/261

[6]Garman pp. 263

sellers are permitted.

3. The market maker is a price setter, in the sense that he may control the price probability functions for aggregate demand and supply. Specifically it is assumed that he sets a price p_B at which he will fill buy orders and correspondingly a price p_S for sell orders, yielding the resultant order rates $\lambda_B(p_B)$ and $\lambda_S(p_S)$, respectively.

4. At time zero, the central market maker has cash and stock inventories of $I_c(0)$ and $I_s(0)$, respectively. Subsequent negative inventories imply the market maker's failure, i.e. inability to continue in his role.

5. The market maker seeks to maximize expected profit per unit time, subject to the avoidance of certain ultimate failure.

6. There are no transaction costs for the market maker.

In such a setting, all the trade has to be executed through the market maker who has the opportunity to control the price probability functions.

Garman then describes the actual exchange process within such a framework for a dealership market and an auction market. We restrict our analysis to the dealership market.

In such a market, the assumption is that there is a centralised market maker who dominates the trading. The aggregate demand and supply, which we can also regard as orders, is exogenous to the market maker who only reacts to the incoming orders.

The problem to solve is to find the ultimate failure probabilities. We do that by formulating the inventories of cash and the asset to be:

$I_c(t) = I_c(0) + psN_B(t) - pbN_s(t)$ and

$I_s(t) = I_s(0) + N_s(t) - N_b(t)$

where $N_B(t)$ and $N_s(t)$ are the cumulative numbers of bids and offers which have been executed by time t.

However, calculating an exact solution for the ultimate failure probabilities turns out to be complicated due to the fact that there are two interrelated state variables which requires the solution of polynomial order p_B+p_S.

As an alternative, Garman derived the ultimate failure probability as a function of the market maker's pricing strategy as an approximation to that problem. By deriving approximate ultimate failure probabilities, it is shown that, in order to avoid any failure, the monopolistic market maker must set p_B and p_S in such a way that the following simultaneous conditions hold:

$p_B\lambda_B(p_B) > p_S\lambda_S(p_S)$ and

$\lambda_S(p_S) > \lambda_B(p_B)$.

It is evident that the prices set by the market maker need not necessarily straddle the equilibrium price p^* where the condition must hold that the expected sell order equals the expected buy order which is $\lambda_B(p^*)=\lambda_S(p^*)$.

The market maker may be prepared to increase the inventory position and thus she increases both the bid and ask price which may be above p^* as illustrated in figure 1.4..

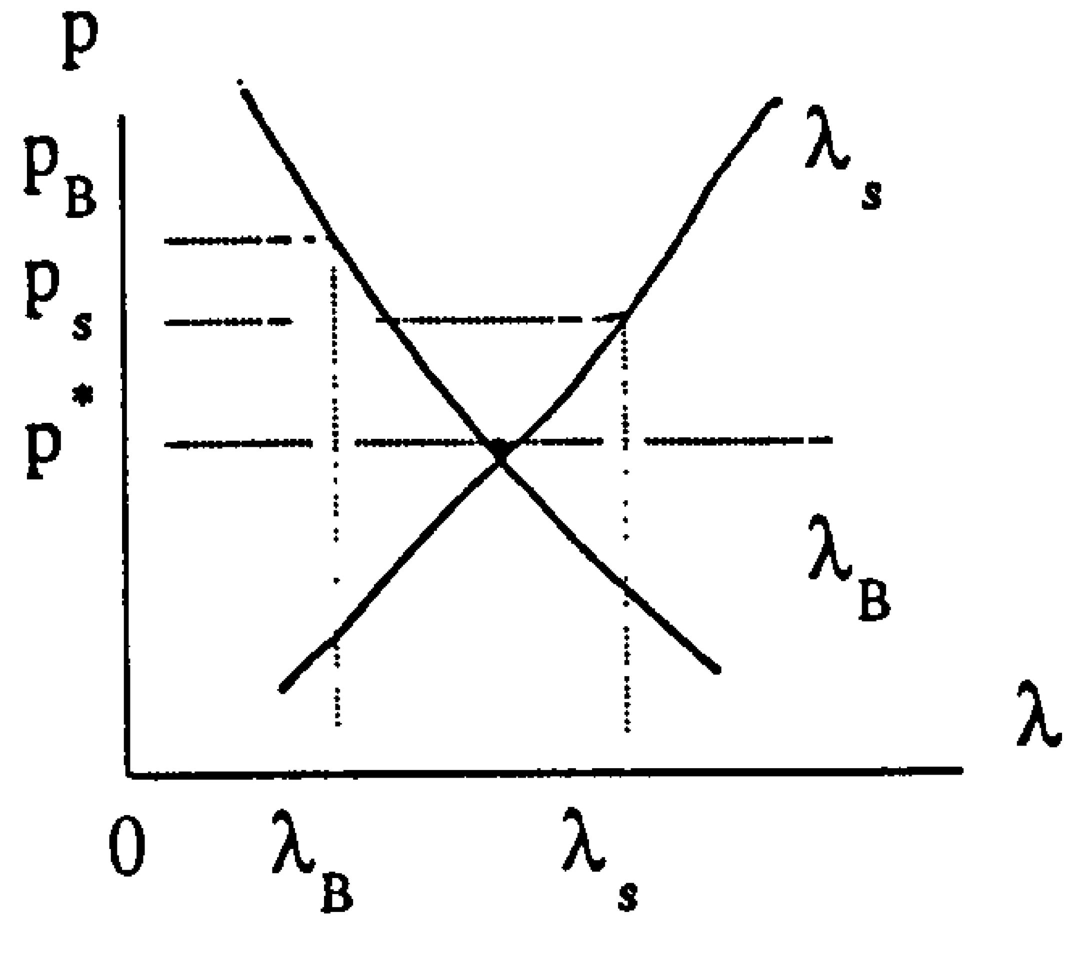

Figure 1.4. : Equilibrium prices

If we change assumption 4 to allow the market maker's inventories to be essentially infinite and interpret assumption 5 in the sense that the market maker takes profits in cash by permitting no upward drift in his stock inventory, i.e. she will keep a "preferred" inventory position, and put these altered assumptions together with the limiting conditions derived above, then it turns out that the market maker will set the prices p_B and p_S which will equate the rates of her incoming buy and sell orders at some value $\lambda' = \lambda_B = \lambda_S$.

Thus, the market maker's profit rate is shown in figure 1.5..

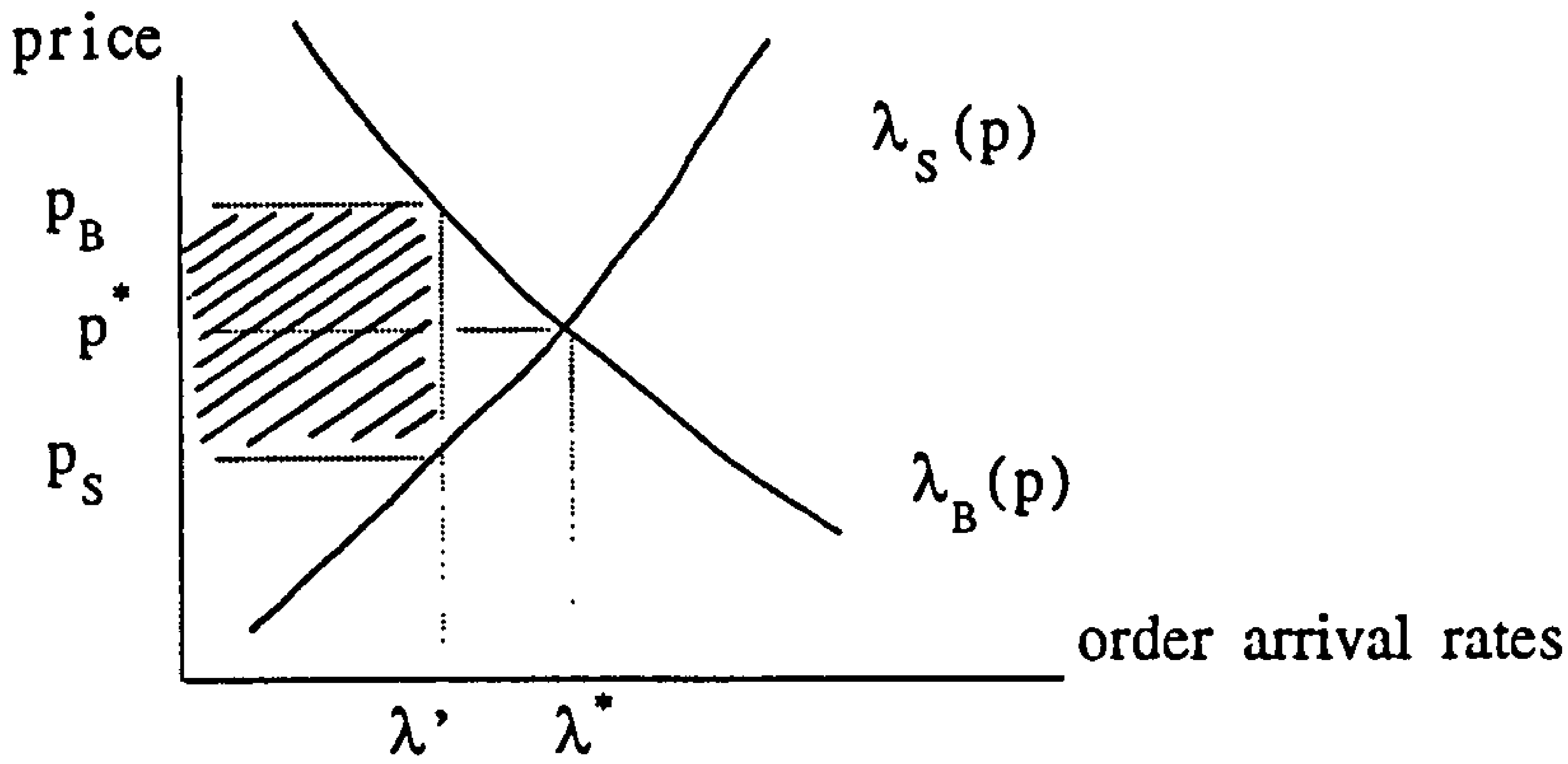

Figure 1.5. : Order arrival rate and equilibrium

The shaded area in figure 1.5. represents the profit rate per unit of the market maker.

In respect of implications on the inventory, by assuming that $p_B = p_S = p^*$ and by canceling assumption 5, Garman shows that the expected time to failure is maximized when the market maker divides his wealth equally between stock and cash. This proves that the market maker has to take into account her inventory position by setting her prices in order to avoid any ultimate failure.

Unlike Garman, Amihud and Mendelson (1980) consider a more dynamic approach, in so far that they allow the market maker to make price adjustments over time. They derive an optimal pricing policy of the market maker in a similar dealership market. They assume a monopoly, but the market maker is subject to constraints on short and long stock inventory positions. The inventory is assumed to have an upper bound by some constant

L and a lower bound of -K. The stock inventory levels are {-K, -K+1, -K+2,..., L-2, L-1, L} which are renumbered as {0, 1, 2,..., M-1, M} with M = L + K and $M \geq 3$.

They slightly change the third assumption of Garman's model and formulate it as for a given pair of prices, P_a and P_b, the next incoming order will be a buy order with probability $D(p_a)/(D(p_a)+S(P_b))$, or a sell order with probability. $S(P_b)/(D(P_a)+S(p_b))$. The time until the next arriving order has an exponential distribution with mean $1/(D(P_a)+S(P_b))$.

The dynamic process is characterized by the order arrival rates which are Poisson rates and the inventory development process which is a birth and death process with λ_k being the birth rate in state k and μ_k being the death rate in state k. Since $\lambda_k=S(P_{bk})$ is a monotone increasing function of P_{bk}, there is a one-to-one correspondence between λ_k and P_{bk} and as μ_k is a monotone decreasing function of P_{ak} there is also a one-to-one correspondence between μ_k and P_{ak}. Hence λ_k and μ_k are used as decision variables in state k. The characteristics of the Poisson process are the independent exponentially distributed interarrival times whose mean is τ_a = 1/D(pb) and τ_b = 1/S(pa). The market maker's revenue from sales and cost from purchases are given by

$R(\mu) = \mu P_a(\mu) = \mu D^{-1}(\mu)$ and

$C(\lambda) = \lambda P_b(\lambda) = \lambda S^{-1}(\lambda)$

whereby the regularity assumption are

1. R(.) is strictly concave ($R''(\mu) < 0$)
2. C(.) is strictly convex ($C''(\lambda) > 0$)
3. $R'(0) > C'(0)$, $R'(\infty) < C'(\infty)$.

Furthermore $\mu_0 = \lambda_M = 0$. q_k is the earning rate $(R(\mu_k)-C(\lambda_k))$ in state k.

Their model stipulates the objective of the market maker as the maximization of his expected average profit per unit time.

Profit is defined as net cash inflow. The objective function can be expressed as the expected return i.e. $g(\underline{\lambda},\underline{\mu})=\sum_{k=0}^{M} \phi_k q_k$ where ϕ_k is the limiting probability of state k and $\underline{\lambda}=(\lambda_0, ...,\lambda_{M-1})$ and $\underline{\mu}=(\mu_1, ...,\mu_M)$.

When the dealer sets her price in state k which results in an arrival rate of λ_k, at the same time she affects the arrival rate on the other side of the market of state k+1 which we can formulate as $\lambda_k\phi_k = \mu_{k+1}\phi_{k+1}$. This yields $\phi_k = \phi_0(\Pi\lambda/\Pi\mu)$.

The optimal market maker's behaviour can be derived from the objective function and the necessary conditions are:

$$\lambda_k : \sum_{j=k+1}^{M} \phi_j [R(\mu_j)-C(\lambda_j)]-\lambda_k\phi_k C'(\lambda_k)=g(\underline{\lambda},\underline{\mu}) \sum_{j=k+1}^{M} \phi_j \text{ and}$$

$$\mu_k : \sum_{j=k}^{M} \phi_j[R(\mu_j)-C(\lambda_j)]-\mu_k\phi_k R'(\mu_k)=g(\underline{\lambda},\underline{\mu}) \sum_{j=k}^{M} \phi_j.$$

If we subtract the (k+1)st equation of the FOC of λ_k from the kth equation of the FOC for μ_k and using the condition $\lambda_k\phi_k = \mu_{k+1}\phi_{k+1}$ we get $R'(\mu_{k+1}) = C'(\lambda_k)$ for k=0,1,..,M-1. Since $pa(\mu_{k+1}) > R'(\mu_{k+1}) = C'(\lambda_k) > pb(\lambda_k)$ it follows from the optimality condition that a loop of transitions starting from any state k, traversing other states and returning to state k yields a positive profit with probability one.

Consequently, as long as the market maker's resources exceed $\sum_{k=0}^{M-1} pb_k$, the probability of cash failure is zero, even in the worst possible case.

Considering that the market maker tries to avoid a drift in his inventory position he will set prices in order to equate the rates of buy and sell orders, i.e. $\mu=\lambda$, this regardless of his inventory position. Furthermore, it is shown that the optimal relation (λ_k,μ_k) and hence, the optimal relation between the bid and the ask prices, are aligned along the curve defined as $[R(\mu) - \mu R'(\mu)] - [C(\lambda) - \lambda C'(\lambda)] = g$ which is a downward sloping function.

The analysis on the bid-ask spread shows that the optimal bid and ask prices are monotone decreasing functions of the stock in hand and in addition, that the bid-ask spread is always positive.

Amihud and Mendelson argue that the profits of this monopolist are lower than the profits of a market maker who does not restrict to $\mu=\lambda$. Yet the market maker has constraints on the long and short positions which he can take, whereas Garman's monopolist has no such constraints.

In addition, the profit maximizing market-maker will never choose to refrain from making buy and sell transactions. This implies that if the constraints are relaxed by expanding the allowed short or long positions the market maker's profit would be increased. Thus the existence of positive costs of providing dealership services leads to a positive spread which straddles the expected price p^e.

To compare the studies of Garman and Amihud/Mendelson we can look at how their results differ. The difference occurs through the choice of different birth and death rates to obtain the optimal inventory level. This can be illustrated as follows:

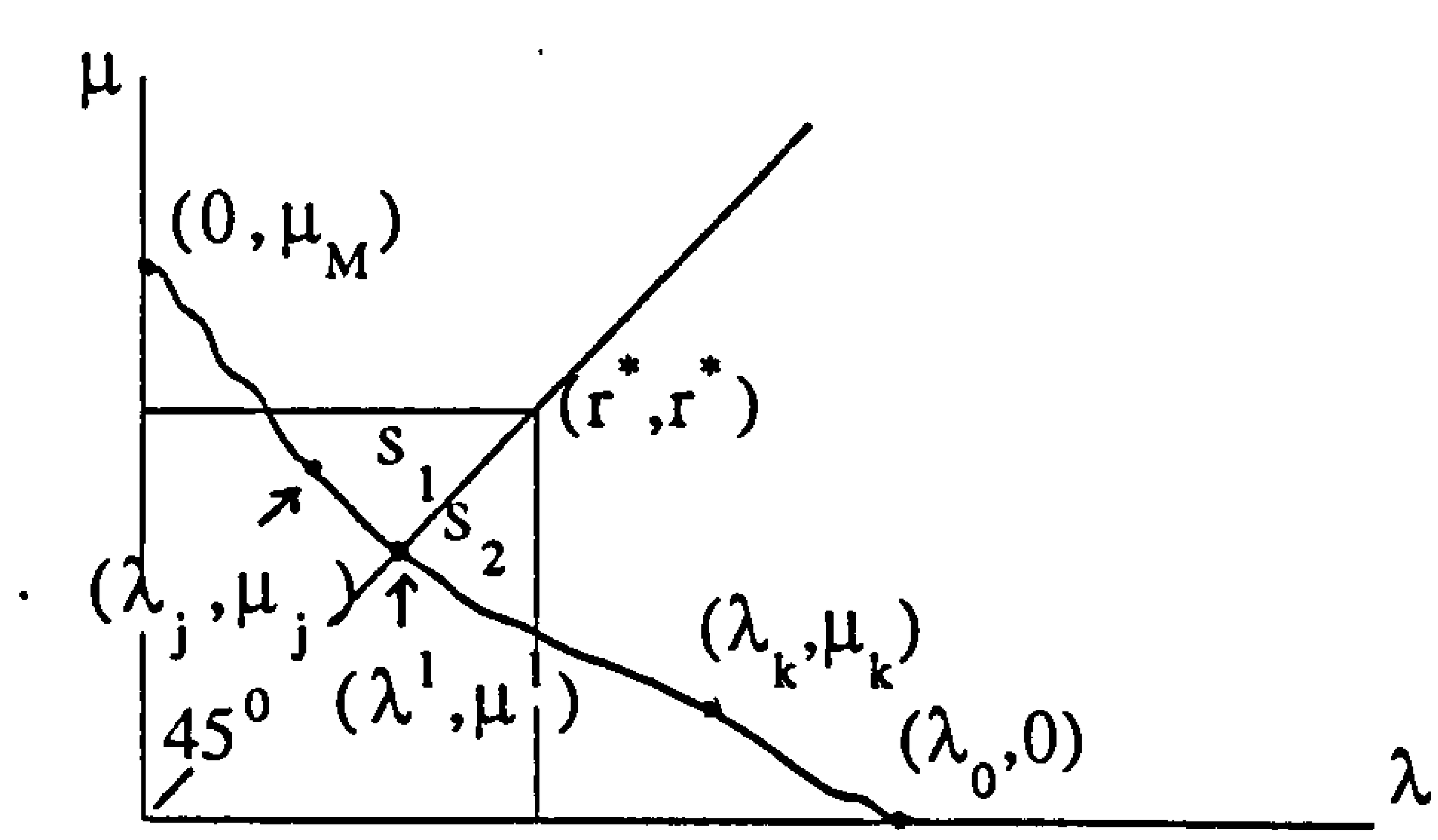

Figure 1.6. : "Optimal" pricing policies

By examining the optimal pricing policy, it is shown that such a policy produces a "preferred" inventory position $J(\lambda_j,\mu_j)$ which is located away from the limiting positions O and M. At such a position J, λ and μ are approximately equal. They are exactly equal at (λ^1,μ^1) where the curve intersects with the 45^0 line. For comparison purposes, r^* is the optimal rate of Garman. The preferred rates (λ_j,μ_j), derived by Amihud and Mendelson, are both less than or equal to Garman's rates (r^*,r^*), thus they are contained in the segment s_1, s_2.

A disadvantage of the studies of Garman as well as of Amihud and Mendelson is that, although they recognize the existence of competition among market makers, they base their models on a monopoly market maker. In such a setting, the monopoly market maker may face capacity constraints.

However, with several competing dealers, this assumption seems no longer of importance. Next, we discuss some models which include a dealer's risk aversion as an important factor of the bid-ask spread determination. The following section supplements the previously discussed models by analysing the pricing behaviour of risk averse market makers in a competitive market.

1.3.2. Bid-Ask Spread, Risk Considerations, and Uncertainty

Unlike the studies discussed, Ho and Stoll (1980) consider a bid-ask spread model taking into account the risk taken by a dealer. The model formulates trading under competition with more than one dealer. This model is an extension to a previously developed model (by Stoll [1979]) with one market maker only who is trading in one asset. A new aspect, going along with the extension to multi dealers trading, is the inter-dealer trading which is examined in this work.

The model defines the situation in which each dealer has his own strategy that maximizes his expected utility of terminal wealth taking into account all the future actions of his competitors as well.

The solution, obtained from a dynamic programming problem, indicates the optimal reservation selling fee a and an optimal reservation buying fee b. The reservation fee represents the minimum fee for the dealer with which she is willing to trade without lowering her expected utility of terminal wealth.

In other words, the reservation fee represents the cost to the dealer which she faces if she enters a transaction that changes her optimal ("preferred") inventory position in a way that is non-optimal and includes higher risk for the dealer.

Furthermore, if we define p to be the true price of the stock in the opinion of the dealer which is common to all dealers, she then would earn the fee by buying at $p-b=p_b$, the bid price, and selling at $p+a=p_a$, the ask price. The quantity of the order is fixed and is defined to be Q.

The expected utility of a dealer is: [7]

$EU = (p + a)Q + p(I-Q) -(R/2)\sigma^2_p(I-Q)^2$ for selling and

$EU = (p - b)Q + p(I+Q) -(R/2)\sigma^2_p(I+Q)^2$ for buying. The respective fees a and b are derived under the assumption that the expected utility of a dealer should be at least equal to her expected utility without any trade which is $EU = p + pI - (R/2)\sigma^2_p I^2$.

Thus, the fees are:

$$a = R\sigma^2((Q/2) - I) \quad (1)$$

$$b = R\sigma^2((Q/2) + I) \quad (2)$$

---

[7]Expected utility is derived by assuming mean-variance preferences.

Which dealer makes the next transaction and what amount she will charge over the reservation fee depends on the relative position of the two competing dealers.

In the one period model and for two dealers the reservation fees are:

Dealer A: $a=R\sigma^2_I((1/2)Q-I)$ $b=R\sigma^2_I((1/2)Q+I)$

Dealer B: $a^0=R^0\sigma^2_I((1/2)Q-I^0)$ $b^0=R^0\sigma^2_I((1/2)Q+I^0)$

where R is the coefficient of absolute risk aversion; σ^2_I is the per period variance of the stock's return; Q is the fixed transaction size; I is the dealer's inventory holding of stock. Dealer B's parameters are indicated by the superscript ^{0}. It is assumed that a transaction can occur in the next instance. A dealer purchase or sale can occur with equal probability λ; or there may be no transaction with probability $1-2\lambda$. Furthermore, investors are assumed to interrogate dealers to elaborate the maximum buying price (p-b) and minimum selling price (p+a) that dealers are willing to bid. This is like a Bertrand price competition which eventually drives any monopoly profits to zero.

There could be the situation that a dealer may earn a profit over her reservation fee -a producer surplus- because the dealer is in an advantageous inventory position with respect to her competitor and therefore, she can slightly outbid her and still earn a profit.

In respect of trading patterns and pricing behavior and assuming no inter-dealer trading, we can argue that if $I=I^0$ and if a transaction occurs, A will trade if $R<R^0$. In the other case B will trade.

In other words, if the inventory positions and price expectations are identical the less risk averse dealer can offer the lower buying or selling fee. However, this does not mean that the less risk averse dealer does have a natural monopoly as, in addition, she must have a sufficiently large inventory position or is allowed to go short in inventory. On the other hand, the dealer with the lower reservation fee has no incentive to quote that fee. It is more advantageous for her to quote the reservation fee of her competitor less a small amount, of course.

The conclusion is that the reservation spread can be negative which depends on the inventory position of the dealer and the respective size of the order. If ((Q/2) - I) is negative the respective fee is negative and the resulting reservation spread is negative as well. This situation can be interpreted that the dealer is willing to pay a fee in order to trade and thus to reduce the risk exposure coming from holding the inventory position. However, the market spread is always positive. The reason is that the lower bound on the market spread is the reservation spread of the "worst" dealer. The worst dealer is the one with the greatest risk aversion.

Under the assumption of inter-dealer trading each dealer must calculate the utility of trading with the other dealer at the quoted price compared with the utility of trading with the next market order with probability λ.

Dealers are assumed to be identical except in their inventory positions.

In the one period case, assuming that dealer A is holding the larger inventory than dealer B (other things identical), A has two options, either to sell to a market order with probability λ to earn a fee of a^0 or to sell to dealer B paying to dealer B a fee of π.

Hence, A's expected utility under option one can be expressed as

$$EU = U(W) + U'\Delta W + 1/2U''\sigma^2(W) + R(I-I^0)\lambda\sigma^2_I U'Q \qquad (3)$$

where the first three terms on the right hand side of (3) represent the expected utility of total of the end-of-period wealth in the absence of any transactions and given the underlying return dynamics that make uncertain the future wealth. U(W) is the dealer's basic utility function which has first and second derivatives of U' and U'' and W defines the dealer's total wealth. The last term of (3) represents the expected utility of the profit from a sale transaction. It is obvious that if A has to sell to the market (which means that $(I-I^0)=0$) there is no additional profit to the expected utility as A has to sell at her reservation price.

The expected utility of option two is derived by changing the inventory of A and B to become (I-Q) and (I^0+Q) respectively. The return of A is only $Q(1-\Pi)$. Therefore, A will only trade if the following condition is met

$$\Pi < \frac{Z}{\bar{W}} \sigma_I^2 [I - Q(\frac{1}{2} + \lambda)] \quad (4)$$

with Z being a random variable such that $Z \sim N(0,\sigma^2)$ and whereby $\frac{Z}{\bar{W}} \sigma_I^2$ represents the return uncertainty.

If inter-dealer trading is allowed at market quotes only, then Π is the buying fee which is set by A (Π=b) and the buying fee is given by (2). If we substitute these two values into (4) we see that this condition can never be met. In this case, no inter-dealer trading would occur, in a two dealer scenario as, even if the probability λ is zero, A will not sell to B *and* pay the market buying fee to reduce inventory.

However, if $\lambda>0$, then A has the additional possibility of selling to a market order and *earning* a fee. There is a negotiated fee at which A would sell to B rather to take a chance on a market order. In this case $I > Q(1/2 + \lambda)$. If we relax the assumption of two dealers only and consider more than two dealers, inter-dealer trading becomes possible. The reason is that the market buying fee is not determined by the dealer who intends to sell at the market price.

There is a so called "gravitational pull effect" which means that inter-dealer trading will take place which can be described best by an example as given by Ho and Stoll: [8]

'Suppose that there are three dealers with identical R; and suppose dealers B and C also have identical inventory of $I^0<I$, A's inventory. Therefore, B and C set the market bid and ask price. Because they have the smaller inventory, B and C are in the better position to buy. Because they compete, they will be forced to offer to buy at their reservation bid price. A has the larger inventory and is in the better position to sell. Since there is no competition on the sell side, A is able to quote the higher reservation ask price of B and C rather than his own reservation ask price'. Condition (4) can now be met at the market bid since the market bid is not A's reservation bid. Indeed, in this example Π is given by (2) with a subscript "0" which refers to the dealer setting the market bid price (B or C). Thus, (4) becomes $R^0\sigma_I^2((Q/2)+I^0)<R\sigma_I^2[I-Q(1/2 + \lambda)]$ and because $R^0=R$ in this example this becomes $I^0 + Q < \lambda Q$. If the market inventory (i.e. of B or C) after an inter-dealer transaction, $Q+I^0$, is less than A's expected inventory without an inter-dealer transaction, $I-\lambda Q$, an inter-dealer trade will occur.'

A further development of the model under competition to a model of equilibrium has been carried out by Ho and Stoll (1983). This time, they are concerned with behaviour and interaction of individual competing dealers and with the determination of the market bid-ask spread.

The model of equilibrium examines markets with several dealers and several assets within several periods.

The formulation of the model restricts to two dealers A and B who are active in two stocks. The two dealers have homogeneous expectation about the "true" future price (p) of the stocks.

The expected utility of terminal wealth is defined as $U(W_0)$ with

$W_0 = F_0 + Y_0 + M_0 + N_0$

whereby W_0 is the terminal wealth, F_0 is the initial cash position, Y_0 is the base wealth, and M_0 and N_0 are the inventories of the stocks.

The first step in the model is to examine the quotes under a one period

[8] Ho and Stoll pp. 264

horizon as only under this assumption does the bid and the ask quote not depend on the inventory position. The dynamics of the dealer are given by:

1. Inventory: The value of the inventories of the two stocks

$M_{t-1}=(1+r_M)[M_t+q_M(Q,-Q)]+[M_t+q_M(Q,-Q)]Z_M$ and

$N_{t-1}=(1+r_N)[N_t+q_N(Q,-Q)]+[N_t+q_N(Q,-Q)]Z_N$

where M,N are the dollar values of the dealer's inventory of stock M or N, t is the subscript which gives the number of periods remaining to the horizon date and r_i, i=M,N is the dealer's expected per period rate of return in stock i in the absence of a bid or ask fee.

$Z\sim N(0, \sigma_I^2)$, i=M,N is the stochastic component in the return in stock i.

Q is the dollar transaction size in each stock

$$q_i(Q,-Q) \begin{cases} \lambda_i - Q & \text{if } b_i < b_i^0, 0 \text{ otherwise} \\ \lambda_i - -Q & \text{if } a_i < a_i^0, 0 \text{ otherwise} \\ 1-2\lambda_i - 0 \end{cases}$$

λ_i,i=M,N is the probability of a public sale (dealer purchase) of Q dollars or of a public purchase (dealer sale) of -Q dollars in each period where the Bertrand price competition condition is included, i.e. that only the dealer with the lowest reservation fee will get the market orders.

a_i, b_i, i=M,N is the dealer's proportional reservation selling fee and proportional reservation buying fee, respectively.

The superscript "0" means the variable of dealer B.

2. Cash position

$F_{t-1}=(1+r)[F_t+q_M(-Q+b_MQ,Q+a_MQ)+q_N(-Q+b_NQ,Q+a_NQ)]$

3. Base wealth

$Y_{t-1}=(1+r_Y)Y_t+Y_tZ_Y$

where r_Y is the expected return on base wealth and $Z_Y\sim N(0,\sigma_Y^2)$

The respective objective function for the dealer is defined as

$$J(t,M,N,Y,F^0,M^0,N^0,Y^0)= \max_{\substack{a_M, b_M \\ a_N, b_N}} Eu(W_0 | t,F,M,N,Y^0,F^0,M^0,N^0,Y)$$

Based on above formulation, the optimal spread is shown not to depend on the inventory level within a one period framework. The reservation buying and selling fee are given by $b_M=(1/2)\sigma_M^2R(Q + 2I_M)$ and $a_M=(1/2)\sigma_M^2R(Q - 2I_M)$ where $I_M = M + \beta_{NM}N$ and $\beta_{NM}=\sigma_{NM}/\sigma_M^2$ and σ_M^2 is the variance of the return of the stock M and σ_{NM} is the covariance of return between stock M and N. R is a discounted coefficient of absolute risk aversion defined as

$$R = \frac{-U''(W)}{(1+r)U'(W)}$$

The market bid-ask spread with several dealers is derived by examining which dealer trades the next transaction.

The question of which trader makes the next transaction and what market fee above the reservation fee can be charged depends on the relative positions of the dealers.

Ho and Stoll show that the dealer with the lowest reservation fee does not quote this fee, but instead the fee of the second best dealer plus a small amount. This means that the next-best dealer sets the market spread.

Furthermore, the equilibrium market spread is limited when dealers have identical coefficients of absolute risk aversion and identical opinions of the true price of the stock.

It is shown that under homogeneous preferences and opinions, the equilibrium market bid-ask spread satisfies the following conditions:

Two dealers: $s \geq R\sigma^2Q$

Three dealers: $s = R\sigma^2Q$

More than three dealers: $0 \leq s \leq R\sigma^2Q$

If the assumption of heterogeneous opinions is considered then the market bid-ask spread, will still be independent of inventory, as long as the risk behaviour of the dealers is the same.

This finding is somehow obvious as it is assumed that the market order is executed by the same dealer on the selling side as well as on the buying side.

Ho and Stoll argue that since the inventory obviously does not matter in the one period horizon, it follows that the degree of diversification of the dealer's inventory has no effect on the dealer's reservation spread. This is not true for the market spread which is determined by two different dealers on each side of the market which means that their inventories are not the same. In such a case with two or several dealers in the market, the market bid-ask spread depends on the inventories of the market makers.

To show that we assume that the inventories of the two dealers are denoted by I and I^0 with $I < I^0$. If both dealers have identical risk aversion and price expectation then the dealer with I will buy the order of size Q (assuming that only one order arrives within the period considered). In the next period, the inventories are (I+Q) and I^0. Only if $(I+Q) > I^0$ the same dealer executes the next order which we assume is a sale. As a consequence, the bid-ask spread does not depend on the inventory level of this dealer. However, if the other dealer executes the sale order then the spread depends on the difference of the inventory levels.

In addition, as Ho and Stoll rightly point out, if the transactions in different stocks are dependent, the reservation fees and the spread are affected by the degree of which the transactions in the dealer's stocks are correlated.

If the model is examined in the context of two periods, assuming one asset only (M) and that the dealers have identical absolute risk aversion and identical initial inventory positions, the bid-ask spread at t=2 is derived as

$$a_2 = \frac{1}{2(1+r)} \sigma_M^2 R Q(1 - 2 \mu)$$

where μ is the market's conditional probability, given a purchase at t=1, of a sale by the dealer at t=2.

The consequences are that the greater the probability of a reverse transaction in the following period of trade, the lower the reservation fee in period t=1.

In addition, since actively traded stocks have a larger μ, it follows that these stocks have a lower spread than stocks traded not so frequently.

This study by Ho and Stoll can be regarded as a valuable contribution to determining the bid-ask spread in equilibrium. However, their analysis of the market under competition is somewhat limited as they assume that the same and only dealer executes the market order on both sides of the market.

How does the bid-ask spread change if different market makers, who have heterogeneous price expectations, different degrees of risk aversion and differences in their inventory positions, trade on either side of the market? We do not get any answer to that problem from their study.

Furthermore, the examination of the diversification problem does not show any influence on the bid-ask spread. This finding may well change if we assume that not only the asset prices are correlated, but also the transactions of the assets which we actually observe in today's markets. We will come back to these issues in the subsequent chapters.

Another examination of the bid-ask spread in a multi-period framework has been done by O'Hara and Oldfield (1986).

They also look at the influence of risk aversion on the bid-ask spread of an asset. However, unlike the previous analyses, they do not specify a particular order flow process nor an intrinsic "known" price of the asset.

In contrast to the results of Ho and Stoll, they show that within a one period horizon, the bid-ask spread is dependent on the inventory. The assumption made is that the market maker has constant absolute risk aversion and that the market maker maximizes the expected utility of trading profits over an infinite horizon of trading days, j=1,2 ... .

The market maker's order flow includes both limit orders and market orders. We can decompose the market order flow into a price dependent component and a liquidity induced component for each side, i.e. the ask and the bid side separately which is

$\tilde{A}^m_t = \alpha^m - a_t\gamma^m + \tilde{w}_t$ for the ask side and

$\tilde{B}^m_t = \beta^m + b_t\phi^m + \tilde{\varepsilon}_t$ for the bid side

where "m" denotes the market order and with $\alpha^m - a_t\gamma^m$ and $\beta^m + b_t\phi^m$ being the price dependent component and $\tilde{w}_t$ and $\tilde{\varepsilon}_t$ being the liquidity induced component which are random variables.

The limit orders in the market maker's order book are described by the linear cumulative order functions. These are given by the integrals of the incremental orders:

$\alpha^L - \gamma^L a_t = \int_a^{\bar{a}} q_a(a)da$ which is the limit buy function

$\beta^L + \phi^L b_t = \int_{\underline{b}}^{b} q_b(b)db$ which is the limit sell function

with α^L, β^L, γ^L and ϕ^L being parameters of the cumulative order flow functions. a_t and b_t are the ask and bid prices. $\bar{a}$ is the highest buying reservation price and $\underline{b}$ is the lowest selling reservation price. Furthermore, $q_a(a)$ and $q_b(b)$ are the incremental quantities for buying and selling at the ask or bid prices. The solution of above integral gives the incremental orders $q_a(a) = \gamma^L$ and $q_b(b) = \phi^L$.

A period's total order flows for the ask and the bid side are $\tilde{A}_t$ and $\tilde{B}_t$ which are

$\tilde{A}_t = \alpha - a_t\gamma + \tilde{w}_t$ if $\alpha^L - a_t\gamma^L \geq 0$

$\alpha^m - a_t\gamma^m + \tilde{w}_t$ otherwise

$\tilde{B}_t = \beta + b_t\phi + \tilde{\varepsilon}_t$ if $\beta^L + b_t\phi^L \geq 0$

$\beta^m + b_t\phi^m + \tilde{\varepsilon}_t$ otherwise

where $\alpha = a^L + \alpha^m$; $\gamma = \gamma^L + \gamma^m$; $\beta = \beta^L + \beta^m$; and $\phi = \phi^L + \phi^m$

In addition the following constraints are to be imposed:

$\alpha^L - a_t\gamma^L \geq 0$, $\beta^L + b_t\phi^L \geq 0$ which means that the market maker does not accept limit orders for negative quantities. Furthermore, the market makers are not allowed to buy at the sell price or to sell at the buy price.

The dealers optimization problem is defined as

$$\max E\,[\sum_{j=0}^{\infty} \alpha^j\, U(\sum_{t=1}^{n} (\tilde{\pi}_{jt}))],$$

where U represents a von-Neumann Morgenstern utility function which is increasing concave, bounded and twice differentiable (U"<0); α is a discount rate $0 \leq \alpha < 1$ and $\tilde{\pi}$ is the trading profit in period t of day j. Based on above assumptions the market maker may end up with a positive or negative inventory from trading.

In addition it is assumed that each day has n trading periods. Since the current inventory is the basis for trading in the next period, inventory represents the state variable of this dynamic system.

Thus, the market makers infinite horizon problem as stated above, can be written as

$$\max E(U(\sum_{t=1}^{n}(\tilde{\pi}_t)) + V(\tilde{I}_n))$$

whereby I_n is the market maker's inventory position at the end of the day and V is the market maker's derived value for inventory. This value function represents the market maker's utility of the inventory position taken into the next trading day. The overall utility of the market maker is the cash position plus such an inventory value.

Unlike the previously discussed studies, the expectation of the future asset price is included in the value function and thus the future value of the inventory will be determined endogenously.

At the end of the day, the trader has to determine which bid price b_n and which ask price a_n he will set which affect the volume of trades in period n. Under above assumptions, the market maker faces a constrained maximization defined as

$$\max_{\{a_n, b_n\}} E\,[U(\Sigma_{t=1}^{n-1}\pi_t+a_n(\alpha-a_n\gamma+\tilde{w}_n)-b_n(\beta+b_n\phi+\tilde{\varepsilon}_n) + r\tilde{p}(I_{n-1}+\beta+b_n\phi+\varepsilon_n-\alpha+a_n\gamma-\tilde{w}))$$

$$+V(I_{n-1}+\beta+b_n\phi+\tilde{\varepsilon}-\alpha+a_n\gamma-w_n)];$$

subject to: $\alpha^L-a_n\gamma^L \geq 0$; $\beta^L+b_n\phi^L \geq 0$.

The optimal bid and ask prices can be derived and the interpretation of the terms gives the following evidence:

$$a_n = \alpha/2\gamma + E(U'\tilde{w}_n)/E(U')2\gamma+rE(U'\tilde{p})/2E(U')+E(V')/2E(U')$$

$$b_n = -\beta/2\phi-E(U'\tilde{\varepsilon}_n)/E(U')2\phi+rE(U'\tilde{p})/2E(U')+E(V')/2E(U')$$

The first expression is determined from the limit orders and the expected market orders. The second term indicates the adjustment due to the uncertainty of the market orders. The last two terms represent the adjustment of the price level induced by inventory changes. In addition, the last term indicates the market maker's expectation of the future price which affects both a_n and b_n equally. Derived from above, the optimal bid-ask spread is

$$a_n - b_n = (\alpha\phi + \beta\gamma)/2\phi\gamma + (\phi E(\tilde{w}_n) + \gamma E(\tilde{\varepsilon}_n)/2\gamma\phi$$
$$+ (\phi\ cov(U', \tilde{w}_n) + \gamma\ cov(U', \tilde{\varepsilon}_n))/2\phi\gamma E(U').$$

About the determination of the bid-ask spread we can say that the first term shows the market maker's total expected supply and demand during trading interval n and represents a risk neutral market maker's charge. However the market makers are faced with uncertainty which is expressed in the third term. The covariance terms show the adjustment to the spread depending on the degree of risk aversion of the market maker. The final term shows how far the overnight inventory affect the bid-ask spread.

Summarizing the findings of O'Hara and Oldfield, it is evident that inventory has a pervasive influence on the bid-ask spread as well as on the bid and ask prices themselves. Furthermore, it is shown that risk aversion influences both the spread and the bid and ask prices. In addition it is found that if prices from risk neutral traders prevail, a stable trading situation implies that the expected overnight price lies between the quoted bid and ask.

In the next section, particular attention is drawn on asymmetric information, i.e. it is assumed that the market does not work efficiently and that some participants in the market have superior information available.

1.3.3. Dealer Quotes and Asymmetry of Information

Walter Bagehot (1974) was the first who examined the bid-ask spread with respect to asymmetry of information in the dealership market. Under his "B-T theory" he stipulates that the dealer gains from liquidity transactors and loses from inside information transactors.

The fact that a better informed trader has an advantage in the market, irrespective of how small such an informational advantage is, has been shown by Jaffe and Winkler (1976). They show that a market maker can always expect to lose when trading with a rational individual, even if the market maker is more knowledgeable and even after the bid-ask spread is included. They argue that under the specialist system, speculators have profitable opportunities which would not arise in a more perfect market. In their model, they define the situation involving one asset, a risk-neutral investor called A, and a market maker called B who sets the market price p_B+T at which A can buy the asset and a price p_B-T at which A can sell the asset. The spread is 2T.

Furthermore, p is defined as the value of the asset. Hence the estimates of the investor and the market maker can be expressed as $p_A=p+u_A$ and $p_B=p+u_B$, where p is the true value of the asset and u_A and u_B are error terms which means that the investor and the market maker have diverse expectations of the true price with $u = (u_A, u_B)$ having a bivariate normal distribution. It is assumed that $E(u_A)=E(u_B)=0$ and $\sigma^2_A=\sigma^2_B$; in addition $cov(u_A,u_B)=0$ and it is assumed that this distribution is known to A.

If A did not know p_B, A's prior distribution for p is a normal distribution with mean p_A and variance σ^2_A.

After having learnt the value of p_B, A's distribution (posterior) for p is a normal distribution with mean p^*_A and variance σ^{*2}_A, where

$$p^*_A = \frac{k^2 p_A + p_B}{k^2+1}$$

$$\sigma^{*2}_A = \frac{k^2 \sigma^2_A}{k^2+1}$$

with $k=\sigma_B/\sigma_A$.

The expected return to A from buying the asset is $p^*_A-(p_B+T)$, so A will buy if $p_A-p_B > \beta T$, with $p^*_A - (p_B+T) > 0$ or if

$$\beta = 1 + \frac{1}{k^2}.$$

As a result we can say that the expected return to A from buying the asset is positive when p_A-p_B is greater than T by at least T/k^2.

It follows that the smaller k is, the larger the price difference that is required before it is advantageous for A to buy the asset. If A is selling the asset the situation is analogous and the condition is $p_B-p_A > \beta T$ which means that considering both cases A will trade whenever

$$|p_A - p_B| > \beta T.$$

These findings imply that any investor who has superior information compared to the market maker can make a profit by trading with the market maker. However, the probability of trading is inversely related to the precision of the forecast of the investor. Hence, the better informed investor is more likely to trade and make a profit than an investor with only a small amount of information.

Furthermore, Jaffe and Winkler show that as A's error variance decreases relative to the market maker's error variance, A is more likely to trade. Based on the results we can conclude that any investor with some information can profit by employing a decision rule such as described by their model. The volume to be traded is determined by the estimation ability of the investor relative to the one of the market maker.

In such a situation, the market maker would always make a loss if he is bound to trade with informed traders only.

In case there are not any liquidity traders the dealer would then set T to infinity in order to reduce any trading loss. However, this would result in a non-trading situation. This means for liquidity traders that they, in a way, subsidize the market.

A different approach has been chosen by Copeland and Galai (1983). They formulate their model in the way that the market maker sets his bid-ask spread in order to optimize his position. In this kind of model the objective is to maximize expected profits. Hence, the bid-ask spread is said to be optimal if it maximizes the difference between expected revenues received from liquidity-motivated traders and expected losses to information-motivated traders. Informed traders are in possession of superior information compared with the information of the public and this allows them to have a better estimate of the future price of the asset than a dealer or a liquidity trader.

The cost function of the market maker is

$$p_I\{ \int_{K_A}^{\infty} (S-K_A)f(S)dS + \int_0^{K_B} (K_B-S)f(S)dS\}$$

where the true underlying future price of the asset is S.

The expected loss of a trader depends on the probability that the next trader will be an informed one p_I, the dealer's knowledge of the stochastic process governing price changes f(S) and on the dealer's ask (K_A) and bid (K_B) prices set. The true underlying future price perceived by the dealer is denoted by S_0. The assumptions of the model are:

- there are no taxes
- short selling is not constrained
- the instantaneous risk-free borrowing and lending rate is constant $r_f \geq 0$
- the true underlying asset value S, follows a stochastic process f(S) which is known (ex ante) to all the market participants.
- all traders arrive at the market trading post according to a stationary stochastic process $g(\tau)$, which is known to all participants and which has

calendar time arrival $\tau>0$ and finite mean $E(\tau)$ which is the mean arrival rate

- the probability that the next trader is an informed trader is $p_I (0 < p_I < 1)$, determined exogenously, and the probability that the next trader is a liquidity trader is $p_L = 1-p_I$.
- the dealer's quote is limited for a fixed quantity n which will be given to the first trader only.
- the dealer is assumed to be risk neutral and maximizes expected profits
- the occurrence of trades is a function of the bid-ask spread which means that both liquidity and informed trader have price-elastic demand.

The revenue function shows the revenue coming from the liquidity motivated traders who are willing to pay premia of K_A-S_0 or receive S_0-K_B as a price for immediacy as they do not know the true price S.

Examining the relation between the bid-ask spread and the probability that a liquidity motivated trader will trade, the probability can be defined as $p_{BL}+p_{SL}=p_{TL}$(whereby p_{BL} is the probability of a purchase by a liquidity trader and p_{SL} is the probability of a sale by a liquidity trader). Furthermore, p_{TL} and p_{NL} are the probabilities of trading and non-trading. We assume that p_{BL} is a decreasing function of K_A-S_0 and p_{SL} is an increasing function of S_0-K_B.

Therefore, the trader's expected revenue per trade with a liquidity trader is $(1-p_I)\{p_{BL}(K_A-S_0)+p_{SL}(S_0-K_B)+p_{NL}0\}$.

With an elastic demand (as assumed above) it is less likely that a trade with a liquidity trader occurs the greater the bid-ask spread is.

We can even argue that if the spread is too wide due to a large number of informed traders the market will be shut down unless more information will be made available.

The optimisation problem of a risk-neutral trader is given by

$$\max_{K_A, K_B} \{(1-p_I)[p_{BL}(K_A-S_0)+p_{SL}(S_0-K_B)] - p_I[\int_{K_A}^{\infty} (S-K_A)f(S)dS+\int_0^{K_B}(K_B-S)f(S)dS]\} \geq 0$$

where the first term on the LHS is the revenue received from trading with a liquidity trader and the second expression on the LHS represents the cost from trading with a dealer with superior information.

If the trader is in a monopolistic situation, he will maximize the difference between the expected revenue and cost function by setting the ask price K_a^{**}. Under the assumption of free entry, the competitive situation leads to an ask price of K_a^* where expected costs and revenues are equal and the long run profit is zero. Hence, the competitive ask spread $(K_a-S_0)^*$ occurs where expected revenue equals expected cost and the monopoly ask spread $(K_a-S_0)^{**}$ occurs where expected profits are maximized.

This whole analysis is equally applicable on the bid side.

The situation is shown in figure 1.7.below.

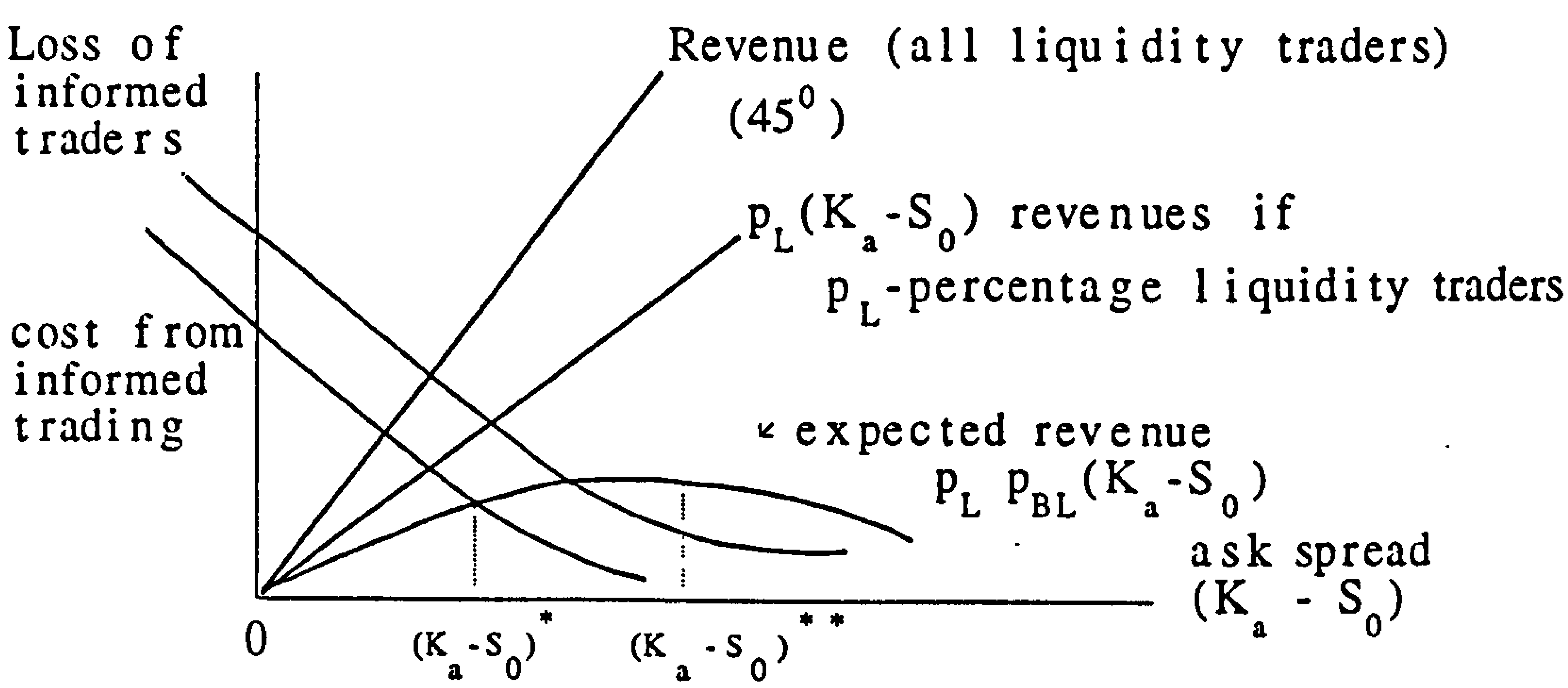

Figure 1.7. : Dealer's optimisation problem

Only the ask side is illustrated, but the same analysis is valid for the bid side. Based on above model three conclusions about the bid-ask spread can be drawn. First, with the movement from the monopoly to the competition, the bid-ask spread decreases. Second, with an increase of the percentage of informed traders the difference between the monopoly and the competitive spread expressed in percentage of the asset price $(K_A^{**}-K_A^{*})/S_0$ decreases. Third, if there is a decrease in the elasticity of demand for liquidity trading, other things equal, then the market maker's revenue curve will shift to the left which leads to a decrease of the ask price. In addition, one prediction of the model is that if p_I increases then also the bid-ask spread will increase.

In respect of the bid-ask spread and the trading volume it can be said that first, if the asset is not very actively traded, p_I may be higher and as a consequence there is a negative correlation between the bid-ask spread and the volume of trading. Second, it could be possible that p_I may increase if more information is associated with the size of transaction.

Under this assumption there will be a positive relationship between the bid-ask spread and the volume of transaction.

Glosten and Milgrom (1985) present a model which is similar to the model of Copeland and Galai (1983). The main difference between the models is that Glosten and Milgrom investigate the dynamic properties of the spread and transaction prices. In particular, they examine how markets process privately available information. In contrast, Copeland and Galai assume that private information is revealed immediately after each trade.

The findings of Glosten and Milgrom are that adverse selection itself (coming from asymmetry of information in the market) can be the reason of the existence of a bid-ask spread. The average spread depends amongst others on the exogenous arrival patterns of insiders and liquidity traders, and the quality of the information held by insiders. In their model, transaction prices are informative, and hence spreads tend to decline with trade.

Glosten (1989) examines the effects of the market structure, i.e. monopoly or competition, in respect of the spread. Especially, he is looking at whether inefficiencies have been created by trading on insider information and the institutional reaction to such inefficiencies.

The result is that market makers reduce the liquidity of the market in response to traders with private information. Such change reduces the amount of trade and the amount of risk sharing.

As argued by Ho and Stoll (1983), competitive market makers will tend to create a more liquid market, because the bid-ask spread will be smaller with competing market makers. On the other hand, monopolistic market makers are in the position to average profits over time. One can think of a situation where competitive market makers are not quoting at all due to the adverse selection problem as irrespective of which price they quote they could not break even. The result is that the market shuts down.

However, the monopolist may have an advantage to keep the market open as he still can get some information of the informed traders. Thus, Glosten argues that both liquidity traders and informed traders are made better off relative to the competing market maker system.

The major results of this study is that informed trading leads to a welfare loss in that it reduces the liquidity of the market. A market maker in a monopolistic situation may provide a more liquid market and hence, the welfare loss occurred from informed trading may be negated.

All these models which we presented in this section describe the problem of asymmetry of information in the market which is a totally new issue compared to the inventory control model of the previous section. Until now, there does not exist a model in the literature which can explain both problems in a setting which determines the bid-ask spread.

The controversy about both issues may be a reason why such a model has not yet been developed. We also think that in a dealership market the inventory control aspect is more important than the asymmetry of information problem. Our opinion is based on the fact that we find professional market makers in a dealership market who very often are linked to a computerised information system which provides instantaneous information. We hope that the next section will shed some light on this controversy by giving some empirical evidence on the determinants of the bid-ask spread.

1.3.4. Empirical Evidence on the Components of the Bid-Ask Spread

We can find extensive empirical work about the determinants of the bid-ask spread in the literature until today. In order to provide an overview the most relevant results shall be presented.
Together with his approach of "cost of transacting", Demsetz (1968) has carried out an empirical investigation based on data of the New York Stock Exchange (NYSE). He found evidence that the spread per share increases in proportion to an increase in the price per share. This can be explained that the cost of transacting per dollar exchanged will be equalized.
Furthermore, he proved that the cost of exchanging a security declines as trading activity in that security increases.
Another aspect examined by Demsetz is the effect of competition in the market where he concludes that there is no significant evidence that

competition will reduce the spread.

In contradiction to the findings of Demsetz, Tinic and West (1972) show that competition tends to reduce the spread. However, it has to be pointed out that Tinic and West based their observations on data of the Over the Counter (OTC) market. Thus, it is difficult to judge about this inconsistency. However, a reason could be the difference in methodology.

With respect to spreads and the number of dealers, they argue that the relationship depends in large part on the extent of economies of scale associated in the dealership function. Furthermore, Tinic and West find that the spread varies inversely with the trading volume.

There has been carried out an analysis by Benston and Hagerman (1974) in which security markets were examined. Referring to the theory of Bagehot, they take risk considerations into account. The argument is that insider trading increases the risk which a trader faces. They also examine whether dealers are natural monopolists. Both systematic risk and unsystematic risk have been examined in respect to the bid-ask spread. The respective risk measurements have been derived from the capital asset pricing model. They show that there is a positive relation between unsystematic risk and the spread which implies that insufficient diversification and insider trading increases the risk taken by the dealer.

Furthermore there is a negative relationship between economies of scale and the spread which is consistent with the result of Demsetz, and Tinic and West. Another result is that there is a negative relationship between the number of dealers and the spread.

A combination of inventory control and asymmetry of information effect has been examined by Hasbrouck (1988). His results show that there is no conclusive evidence for the inventory control effect. Trades for low-volume stocks show negative autocorrelation; whereby for high-volume stocks, there is no such evidence.

On the other hand, the impact of trades on the quote revisions turns out to be significantly positive which supports the informational effects of volume. Other interesting findings are that effects of dealer inventory control behavior on quotes are not significant. However, for high-volume stocks, the order size influences the quote revisions which means that large orders convey more information.

Madhavan and Smidt (1991) examine the effects of both trading volume and unanticipated information. Their model analyses intraday security price movements in a specialist market (like NYSE). They also find strong evidence of information asymmetry, but the inventory control effect appears to be weak. Furthermore, they show that the degree of information asymmetry depends on the specialist participation rate.

The estimation of a simple model of asymmetric information, done by Glosten and Harris (1988), has brought the following insights to the problem of measuring the bid-ask spread. Their model breaks the bid-ask spread into a transitory component and an adverse-selection component. The results are twofold. First, the time-series analysis confirms that the adverse-selection component is positive. Second, the cross-section analysis brings evidence on the asymmetry of information influence. Furthermore, they show that the spread is a function of trade size.

Finally, Stoll (1989) examined the components of the bid-ask spread and the results of the analysis confirm that the spread can be broken down into adverse information cost, inventory holding cost, and order processing cost.

Another group of empirical investigation about the bid-ask spread is concerned with the influence of asymmetric information on the spread with special consideration of the efficiency of the market. Such analyses have been carried out by Goldman and Beja (1979), Roll (1984), French and Roll (1986), Choi, Salandro, and Shastri (1988).

Especially, Roll has defined a measure of the effective bid-ask spread which is "spread=2√-cov" (where "cov" is the first order serial covariance of price changes). He proved that the serial covariances of returns are negatively associated with the square of quoted spreads.

1.4. Market Structure and Market Organisation

We have already discussed the relevance of the market structure throughout this chapter. By trying to determine the bid-ask spread we note that the market liquidity is crucial in finding the equilibrium prices. Market liquidity in turn depends on the underlying market structure. Hence, a thin market in which trading intervals are long, tends to have larger spreads than in more active markets.

The problems of designing an appropriate market structure (or trading mechanism) are stated by Pagano and Röell (1990) as follows:[9]

Upon designing the mechanisms of a stock exchange, policy makers confront four key choices. First, is the exchange to work as an auction market or as a dealership market? Second, if one opts for the auction market, should the auction to be structured as a sequence of discrete batches, possibly at daily intervals, or as a continuous auction? (With a continuous auction, the market clears every time at least two orders can be executed against each other. A dealership market is inherently continuous.) Third, should one provide incentives or impose rules to favour the concentration of trade on a single market, or rather allow off-exchange dealing? Finally, again in the context of an auction market, should one allow exchange members to act as brokers and to deal on own account?

[9]pp. 83

The latter point is especially important in markets where mostly banks are the market makers and trade also for their own account. There could be a conflict of interest by executing orders from the customers and trading for their own account.[10]

It is interesting to see how the design of the market structure can influence the volatility and volume traded of a stock at the exchange. In a comparison between the Paris Bourse, which recently underwent a reform to a continuous system, and the London exchange, Pagano and Röell (1993a) analysed the effects of the reform. Their results are, that for some shares, there is a significant decrease in volatility after the reform of the Paris exchange. The trading volume is also lower with continuous trading, holding other factors constant. There is some evidence of intense competition between the two exchanges. If we analyse the spread of cross-listed stocks, their analysis shows that the spread in the London market is smaller during the trading hours of the Paris exchange.

One of the most important points in designing a market structure is the decision whether the market should operate as an auction market or a dealership market. This problem has been discussed by Pagano and Röell (1992b). They make a distinction between market structures and they divide them into three main categories: Batch auctions, continuous auctions, and pure dealership markets.

In a batch market, dealers submit their bids to a central auctioneer or auction mechanism. The trades are then executed according to a particular auction rule at an equilibrium price where all the orders are cleared. Such auctions take place in a regular interval, mostly daily. Examples of such

[10]For a detailed analysis see Pagano and Röell (1993b).

auctions are, among others, the ones in Paris, New York and Tokyo.

The continuous auction works through a computerized system. Dealers submit their bids to this system where the best limit orders are displayed and the incoming orders are then automatically executed against these limit orders.

Both these structures allow the submission and advertisement of public limit orders. After the execution under the auction procedure the actual trade (price and quantities) are displayed so that the dealers observe the history of the order flow.

In an dealership market, market makers quote their prices at which they are willing to trade. These prices are set for trade volumes up to a specific amount and are displayed on a screen. The actual trade is done over the telephone. In contrast to the auction procedure, matching orders of the public can not be traded directly, but all the orders have to be executed through the designated market makers. The main difference to the auction procedure is that the individual trade information is only known to the market maker who executed the order unless the last trade information is made public immediately on the screen. It is evident that the information dissemination is different for the various market structures. The auction procedures (both the batch and the continuous auction) ensure that information is made public immediately after a trade. However, in a dealership market, market makers may have to quote their prices before learning about the past trade. This means that there is greater uncertainty for market makers in a dealership market and the spread is expected to be wider than in an auction market. We can illustrate the problem by ranking the trading systems according to transparency: [11]

---

[11]Pagano and Röell (1992a) pp.5

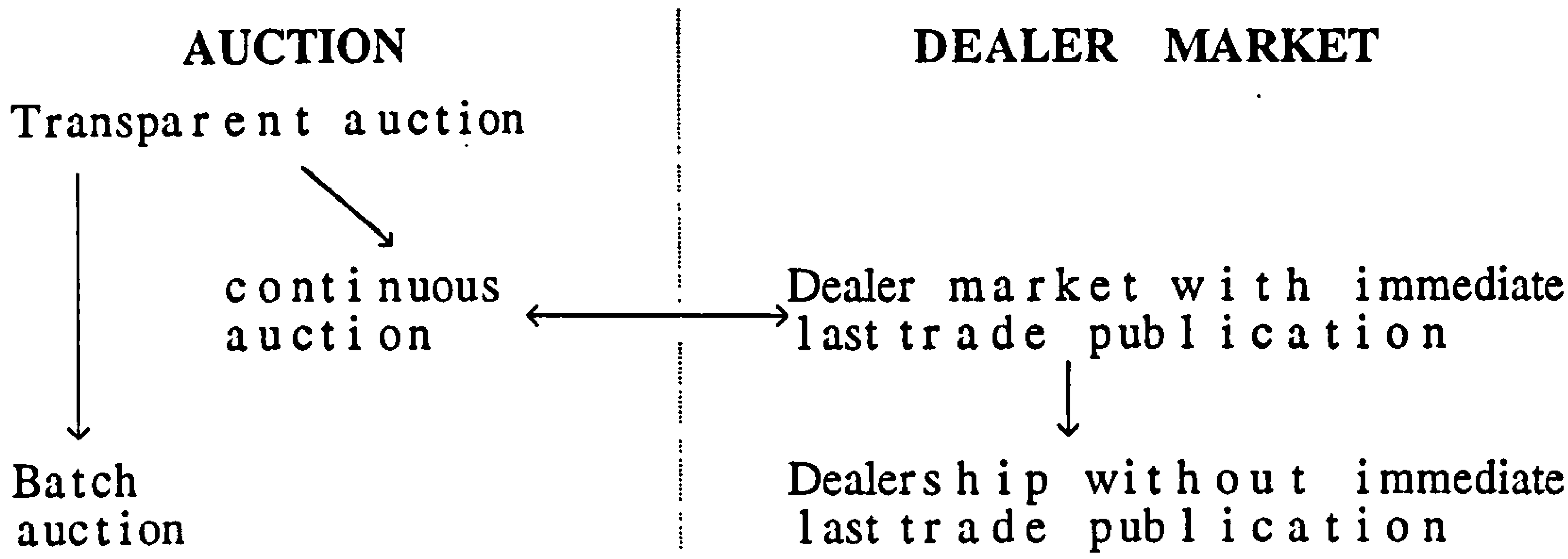

Figure 1.8.: Trading systems ranked by decreasing transparency

Another important distinction between an auction market and a dealership market is the execution risk. In a dealership market the market makers stand ready to trade any incoming order instantaneously at the quoted prices whereas in an auction market the participants have to bear the risk that no matching order arrives.

Pagano and Röell (1993c) investigated that problem and they find, based on a theoretical model, that the dealership market is superior to the auction market, i.e. the market makers are in the position to give this insurance against execution risk, if the market makers are risk neutral or sufficiently less risk averse than their customers.

1.5. Conclusions

In this chapter we made an attempt to capture all the relevant issues of the literature about financial market structures and financial market trading which influence the cost of market making, i.e. the bid-ask spread. The earlier works examine the individual dealer's cost within a

monopolistic market setting. However, soon it has been discovered that other market forces, such as competition between dealers in a market and more recently, competition between exchanges play an important part in determining the costs of transacting in a market. The structure of the bid-ask spread models changed in order to take into account the nature of competitive markets and the difference of market structures.
The splitting of the order may be often the case in a dealership market where one dealer is not in the position to fulfill the whole market order. If dealing costs rise with quantity as in inventory control models, there is a natural incentive to share the order.

Most of the existing bid-ask spread research examine the bid and ask prices assuming an indivisible order. There are models in which orders are split, but these models are based in a Walrasian world. Examples of such analyses are Kyle (1989) and Madhavan (1992).

The analysis of the different market structures includes the theory about auctions. There exist various types of auction procedures with different outcomes. However, until today, most of the studies consider an auction where bidders compete for an indivisible good. In the same light, all the existing models about the bid-ask spread are based on the fact that the most competitive market maker gets all the trade and the others get nothing.
Another point which has not been investigated so far is the determination of the bid-ask spread in a pure dealership market where dealers do not know about the market order size at the time when they quote their prices. Such a situation may occur when dealers do not only trade in one market at the

same time. The reason for trading in different markets at the same time is that the market makers aim for a risk reduction through diversification of the portfolio.

For instance, they may trade in the spot market as well as in the futures market. This calls for an investigation of the bid-ask spread by examining the interaction of these markets.

Hence, further research is still required in these areas which are investigated in the subsequent chapters of this thesis.

In chapter two, we discuss a theoretical study which considers trading in an indivisible good in a fragmented market. We then present our model where market makers compete for an incoming order of a divisible good which is applied to two different market structures, a centralised and a fragmented market.

Another shortcoming of the existing literature is that the interaction between different markets, i.e. spot and futures market has not been studied in respect of the effect on the bid ask spread. As we have pointed out earlier, Ho and Stoll (1983) consider a two asset framework, but they fail to fully capture the influence of diversification on the bid-ask spread as their spread does not depend on the inventory level.

Chapter three gives an introduction and overview of the existing literature about futures markets. This is followed by the analysis of the bid-ask spread in the spot market by simultaneously trading in futures which forms chapter four. The empirical investigation of the Italian secondary market for government bonds includes an examination of the market characteristics, i.e. the analysis of the quoting behaviour of the primary dealers and the returns of the various bonds which is presented in chapter five. In addition, we investigate the inventory control argument and the next best dealer aspect which are the main findings as the determinants of the bid-ask spread by Ho and Stoll (1980, 1983). This is analysed in chapter six. An investigation of the time series properties and the relation to the bid-ask spread is given in chapter seven as another empirical study examining the nature of the Italian secondary market.

CHAPTER TWO

AUCTION MECHANISMS AND DEALERSHIP MARKETS: A THEORETICAL ANALYSIS OF THE BID-ASK SPREAD

2.1. Introduction

The so called theory of the microstructure of market making produced several interesting models which analyse the existence of a bid-ask spread most of them or almost all of them are set in a centralised auction market. Most of the analyses are carried out by considering a monopolistic market maker and so do not model appropriately the competition among dealers. An exception is the model of Ho and Stoll (1983) [1]. They analyse a market with two competing dealers in a centralized market which is described as an auction procedure which gives an outcome of second best prices which is equivalent to a Vickrey auction (1961). Although they argue that their model can be equally applied to a dealership market the application is not so straightforward.

The emphasis of their study is to evaluate the determinants of the bid-ask spread where market makers have full information which means that they know the incoming order and they have knowledge of each others reservation prices. The order will be executed by the market maker with the best price (i.e. the lowest quoted ask price or the highest quoted bid price).

More recent studies (Pagano and Röell [1990, 1992, 1993]) [2] recognize the fact that the market structure is important in evaluating the respective cost in the market.

There are factors such as the market structure, which influence the bid-ask spread. These factors of the market structure may be whether the market is centralized or fragmented, the transparency of the market, i.e. whether

[1]A detailed discussion of their models can be found in the previous chapter.

[2]Details are presented in chapter one.

market makers have information of each others reservation prices and the order flow, and the organisation of the market, i.e. whether the quoted prices are transaction prices or the trade is done over the telephone, which influence the size of the bid-ask spread.

These circumstances differ in the fact that the less transparent the market and the more fragmented the market the more likely it is that the public or private investor is given more power in trading with the market makers. As an example we can think of a pure dealership market where dealers quote their prices without knowing the order flow and without information about the reservation prices of their competitors. Hence, the private investor can exploit the situation by choosing the dealers she wants to trade with and also the quantity to be traded.

In his paper, Biais (1993) compares the theories of the "benchmark" market structures of an open auction procedure (transparent auction) with full information and a dealership market where dealers have only private information about their reservation prices and the individual deal traded. He analyses the bid-ask spread in a centralised market, where dealers know each other's reservation prices and the order flow, and in a fragmented market, where dealers know about the size of the incoming order, but they do not know the reservation prices of the competitors. Dealers quote their prices based on expectations of the reservation prices of their rivals. The trading strategies of the dealers can be compared to the bidding strategies of a high bid auction described by Riley and Samuelson (1981).

Biais also examines the volatility of the asset in both markets as well as the liquidity in equilibrium.

Like all the other bid-ask spread models he assumes that the whole incoming order will be executed by the best quoting dealer who is assumed to be the same on both sides of the market. The features of his model are based on Ho and Stoll's inventory control model. According to their order arrival process only one order arrives within the one period framework which can be a purchase or a sale transaction which is known to the dealer before she quotes her price. Under the assumption that dealers balance their inventory position at the end of the period, the same dealer is assumed to trade the transaction in the next period which is of opposite sign.

Our aim is to add a model which describes a situation where dealers trade in a divisible good. We analyse the market situation where risk averse dealers are allowed to share the market order, but we assume incomplete information in the sense that market makers do not know their rivals' reservation prices. We also assume that only one order arrives at a time, but the sharing of the order gives a different inventory dynamic in the sense that the fact of who executes the next transaction depends on the relative inventory positions of the dealers. The bid-ask spread is calculated from the bid and ask prices of different dealers. This situation is modeled by a discriminating auction procedure which might be socially superior to the Vickrey auction.

The design of markets is an important factor which influences the bid-ask spread and hence the cost of trading.

Risk averse dealers face decreasing returns to scale. Their risk exposure increases with the quantity traded. By allowing the splitting of the order dealers can quote their prices for a lower quantity or on an average of the expected quantity to trade and not on a large trade which is riskier. The result is that the bid-ask spread is expected to be smaller due to the reduced risk and the market is expected to be more liquid. We present two models of which one is set in a centralised market where all the dealers know each others reservation prices and a second one in which dealers only know their own reservation price, but they do know the incoming order.

In section 2.2. we analyse the difference of bid and ask prices and the spread between the Ho and Stoll model (abbreviated as HS) and our model applied to a centralised market structure. The distinction between the two models is the possibility of sharing the incoming order. HS assume trading in an indivisible good whereas our model allows the splitting of the incoming order.

In section 2.3., we present the model of Biais which is based on trading in an indivisible good and we compare this outcome within a fragmented market structure with the centralised market bid and ask spread. An interesting extension to Biais model is worked out by assuming heterogeneous price expectations among market makers.

Section 2.4. contains our bid-ask spread model which evaluates the determinants of the bid-ask spread in a fragmented market with trading in a divisible good which is compared to the corresponding model in the centralised market and to the model of Biais.

The final section gives the summary and the conclusions.

2.2. Trading in an Indivisible and a Divisible Good within a Centralised Market Structure

2.2.1. The Ho and Stoll Approach

The model of Ho and Stoll has been presented in the previous survey chapter and we only sketch the main features of their model in order to have a comparison with the following studies. Their model is a static, one period analysis. They assume a single monopolist market maker who deals on both sides of the market, i.e. is the best quoting dealer on the ask side and on the bid side. Furthermore, they assume a symmetric market. If an order arrives it is equally likely to be a buy or a sell transaction of the same quantity. It is assumed that only one order arrives which is observed by each dealer. This assumption allows us to deal with the optimization problem on each side separately. The dealer's wealth consists of cash and the inventory of a risky asset. The dealer is willing to be active in the market if the expected utility is not less or equal to the expected utility without any trade.

By assuming mean-variance preferences [3] the resulting reservation ask price is $pa = E(p) + \gamma\sigma^2[Q - 2I]$ and the bid reservation price is

$pb = E(p) - \gamma\sigma^2[Q + 2I]$ with

$E(p)$ = future price expectation of the risky asset, γ = coefficient of risk aversion, σ^2 = price variance of the risky asset, Q = order size, I = market maker's inventory of the risky asset

The spread is simply the difference between the ask and bid price which is $pa - pb = s = 2\gamma\sigma^2 Q$

The equilibrium price, considering a market under competition, is determined by the next best dealer argument. Under the assumption that dealers know each others reservation prices, the dealer with the best reservation price, i.e. the lowest ask and the highest bid price, will not quote her own reservation price, but the reservation price of the next best dealer plus (minus) a small margin. Thus, the next best dealer is not able to raise (lower) her own reservation price without incurring a loss. Hence, we have the outcome of a second price auction with certainty about rival bids or an English auction.

---

[3]We assume that the expected futures prices are normally distributed and that market makers have constant absolute risk aversions (CARA).

2.2.2. Trading in a Divisible Good

In this section, we examine the trading procedure of market makers in a competitive dealership market. The determination of the equilibrium bid-ask spread is based on an inventory control model similar to the one of HS, but allowing for splitting of the incoming order. The dealers know the size and the nature of the order before they have to quote their prices. We assume that only one order arrives within one period. The bid-ask spread is then composed of prices from different dealers as it is not given that the same dealer executes the transaction in the subsequent period. We allow for different possible trading situations in the market, for instance, one dealer sells all units of the asset and several other dealers buy the asset, or there is only one dealer on either side of the market buying or selling the asset.

The difference from the traditional model is that the risk averse dealer is confronted with decreasing returns to scale so allowing the splitting of the order. A smaller traded quantity means less risk involved for an active dealer and hence less costs which may result in a smaller bid-ask spread.

The approach that we take is to use the basic HS framework of Bertrand price competition between market makers but extended to allow for many active dealers.

All trade is at the best price but because of differences in the inventory, risk aversion or price expectations of different dealers, particular dealers can quote better prices than others.

Let μ_i be the ith dealers expected future price of the asset and σ^2_{pi} his view of the variance and γ_i be the degree of risk aversion of the dealer. C_i is defined as the cash holding of dealer i. With this notation, the expected utility of a dealer who does no buying or selling in interval t is

$$Eu_i^{NT} = (1 + r_t) C_{it} + \mu_i I_{it} - \gamma_i \sigma_{pi}^2 I_{it}^2 \quad (1)$$

The dealer quotes an ask price of p_{ai} and a bid price of p_{bi} and secures a share *k* of the market buy order if it arrives and a share *l* of the market sell order if it arrives.

They also know about the incoming order of size Q and whether this order is a purchase or a sale. Hence, as only one order arrives at the time, we can examine the bid and the ask prices separately.

Then expected utility is

$$Eu(0,l) = (1 + r_t)(C + p_a lQ) + \mu(I - lQ) - \beta(I - lQ)^2 \quad (2)$$

for a sale and the expected utility for a purchase is

$$Eu(m,0) = (1 + r_t)(C - p_b mQ) + \mu(I + mQ) - \beta(I + mQ)^2 \quad (3)$$

where $0 \leq m, l \leq 1$, $\beta = \gamma\sigma_p^2$ and we have dropped the subscript i to save notation.

We also assume that dealers know about the incoming order and the size Q of this order.

We can write $p_{bi} = \frac{1}{1+r_t}\mu_i - b_i$ and $p_{ai} = \frac{1}{1+r_t}\mu_i + a_i$

where a and b are the selling and buying fees earned by the dealer. If we equate the expected utility of trading with the expected utility of no trade (as in (1)) for each side of the market we get reservation fees of:

$$a = \beta(lQ - 2I) \quad (4)$$

$$b = \beta(mQ + 2I) \quad (5)$$

The determination of the optimal share now depends on the assumption that dealers try to maximize their expected utility of terminal wealth due to their possible advantageous position regarding the price expectation, the degree of risk aversion, and the inventory position.

2.2.3. Equilibrium Prices

We base the determination of the equilibrium prices on the process where market makers quote their prices simultaneously.

We restrict our model in the way so that l=Q/k and m=Q/k, hence l=m=Q/k.

We define the reservation selling price to be

$$pa^r_i = \frac{1}{1+r_t}\mu_i - 2\beta_i I_i + \beta_i lQ = \alpha_i + \beta_i(Q/k) \tag{6a}$$

and the reservation buying price

$$pb^r_i = \frac{1}{1+r_t}\mu_i - 2\beta_i I_i - \beta_i mQ = \alpha_i - \beta_i(Q/k) \tag{6b}$$

with $\alpha_i = \frac{1}{1+r_t}\mu_i - 2\beta_i I_i$.

The ask (bid) reservation price increases (decreases) with the share of any order that the dealer assumes; a smaller traded quantity reduces the dealers inventory risk exposure and so cet par the dealer is prepared to trade at closer prices. The amount of gain from trading increases with the share traded so long as the trade is at prices better than the reservation price.

In equilibrium, market makers set their prices for a division of Q so that none of them wishes to trade a lower or a higher quantity at the quoted price and so that the whole market order Q is satisfied which implies that we do not have any excess demand from the public.

For our analysis we assume that the order can be split into discrete bundle sizes. Market makers are assumed to quote their prices in a way that they either get the whole order or that they will share the order equally between them at an identical price or that they do not get anything. The best price, i.e. the lowest ask and the highest bid price, are obtained by a process of Bertrand price competition which means that dealers undercut each other's prices as long as they do not quote below their marginal costs. The equilibrium is a Nash type equilibrium, but as we show below it is typically not unique.

We can define the price quotes as $(P_1, P_2, ..., P_n)$ and we assume that the number of dealers who quote the best price p is k. Thus, each of the best quoting dealers gets a market share of Q/k and the other market makers with a higher (lower) ask (bid) price do not get anything.

If dealer i quotes $p_i<p$ on the ask side (or $p_i > p$ on the bid side) then she gets the whole order. Due to their differences in inventory, price expectations, and risk aversions, price quotes are given by

$p_1,..., p_{i-1}, p_{i+1},...,p_n$.

For simplicity reasons, we carry out the analysis for the selling side, but the respective argument is equally valid for the buying side.

Given k-1 best quotes, dealer i can, in principle, make three choices. If the market price is p dealer i can quote the same price $p_i = p$ which implies that she will share the market order with the other k-1 best quoting dealers. Another possibility is that she can quote marginally below the market price and attract the whole order $p_i=p-\varepsilon$ or she can quote above the market price $p_i > p$ which means that she does not get any trade at all. If the market maker decides to quote the same price as the market price then she will get a share of Q/k of the market order (with k best quoting dealers). Then, under the assumption of maximisation of expected utility, the expected gain of dealer i is given by the difference between the quoted price and her reservation price given in (6a) which is

$$Q/k[p - \alpha_i - \beta_i(Q/k)] \qquad (7)$$

If the dealer gets the whole order the expected gain is

$$Q[(p - \varepsilon) - \alpha_i - \beta_i Q \qquad (8)$$

In the last case where the dealer does not get any trade the expected gain naturally is zero. Hence, if $p < [\alpha_i + \beta_i(Q/k)]$ then it is best for dealer i to quote $p_i > p$ which means that she will not get any trade.

If $p > [\alpha_i + \beta_i(Q/k)]$ then the dealer has to decide whether she wants to trade the whole order or whether she prefers to share the order as under both options she will still make a profit.

In order to decide whether to trade Q or Q/k we have to find the critical price at a level where the dealer is just indifferent between the two choices.

Such a price level is illustrated as p^* in figure 2.1. below.

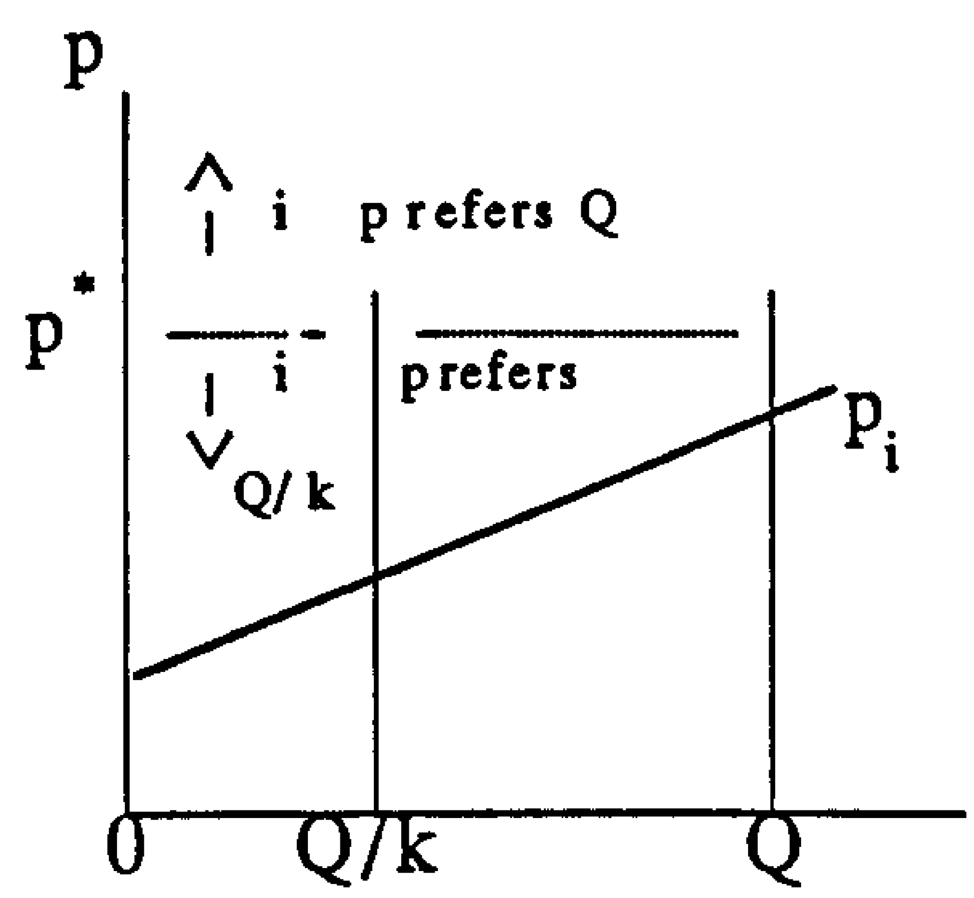

Figure 2.1. : Trading possibilities

The price p^* can be found by comparing the expected gains as in (7) and (8). The dealer is indifferent if she gets the same return for sharing or executing the whole order which is

$1/k[p^* - \alpha_i - \beta_i(Q/k)] = p^* - \varepsilon - \alpha_i - \beta_i Q$ which means that the return from getting a share of (1/k) at the reservation price of $(p-\alpha-\beta(Q/k))$ is equal to the return of getting the whole order at the price $(p-\varepsilon-\alpha-\beta Q)$.

If we do some multiplication we get:

$p^*(1-(1/k)) = \alpha_i(1-(1/k)) + \beta_i Q(1-(1/k^2)) + \varepsilon$ which we can simplify by dividing through $(1-(1/k))$ and finally we have

$$p^* = \alpha_i + [\varepsilon k/(k-1)] + \beta_i Q[(k+1)/k] \qquad (9)$$

At this price p^* and with $\varepsilon \to 0$, the dealer is indifferent between trading the whole order or sharing the order [4]. To summarise:

If $p < \alpha_i + \beta_i(Q/k)$ dealer i quotes $p_i > p$ and gets nothing.

If $\alpha_i + \beta_i(Q/k) < p < p^*_i$ dealer i quotes p and gets a share of Q/k.

If $p^*_i < p$ dealer i quotes p-ε and gets Q.

The equilibrium conditions to hold for all such quotes derived under Bertrand competition which we denote as $\bar{p}_1, ..., \bar{p}_n$ are:

a) that the number of quotes which are identical to p is k and that there does not exist a quote which is smaller than those

b) that for each of these quoters i the following condition holds:

$\alpha_i + \beta_i(Q/k) < p < \alpha_i + [\varepsilon k/(k-1)] + \beta_i Q[(k+1)/k]$ again with $\varepsilon \to 0$ and

c) that for all other quoters j it is true that $p < \alpha_j + \beta_j(Q/k)$

With n dealers we have various possible outcomes where these conditions hold, and as a consequence, such a Bertrand/Nash equilibrium exists, but it will not be unique.

To show that we find a non-unique equilibrium we take an example with three dealers.

For example in figure 2.2., we assume that three dealers are in the market. Then we may have the situation that on the ask side dealers 1 and 2 share the order and on the bid side dealer 3 executes the whole order which we denote as case one. Another possible equilibrium situation may occur where the market order is shared between the three dealers on the ask side as

---

[4]We do not have the problem of non-existence of a Bertrand/Nash equilibrium as we assume a discrete number of shares of the order instead of infinitely divisible shares. The issue of non-existence of a Bertrand equilibrium is discussed for instance in Tirole (1988).

well as on the bid side which is case two.

We can show these types of equilibrium graphically:

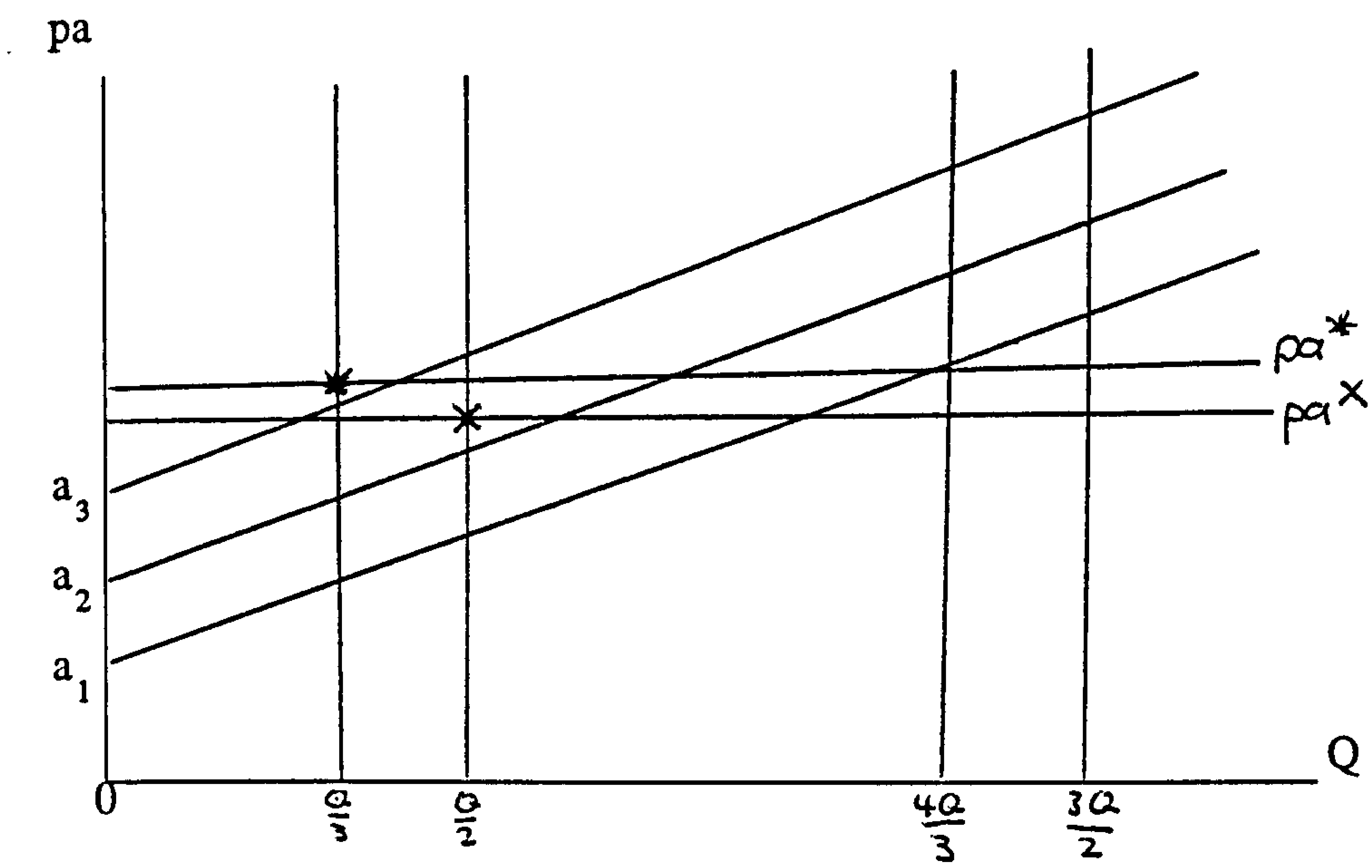

Figure 2.2. : Different types of equilibrium

with x being a possible equilibrium price in case one with the following condition to be satisfied:

$\alpha_i + \beta_i(Q/2) < p < \alpha_i + \beta_i Q(3/2)$

and with * being a possible equilibrium price in case two with the following condition to be satisfied:

$\alpha_i + \beta_i(Q/3) < p < \alpha_i + \beta_i Q(4/3)$.

To see that how this relates to the Ho-Stoll framework, we take the case of only two dealers. For each dealer i there are two critical values of the fees. One is the fee where she is indifferent between not trading and sharing an order at the reservation price for sharing which is $[\alpha_i + \beta_i(Q/k)]$ denoted as pa_{i1} and the critical price p^* at which she is indifferent between executing the whole order and sharing the order which is $p^* = \alpha_i + \beta_i Q[(k+1)/k]$ which we denote as pa_{i2}. The first subscript indicates the dealer and the second suscript refers to the type of critical price for the decision of not trading and sharing or executing the whole order and sharing the order.

Hence in the two dealer case, the respective prices are:

$$pa_{i1} = \frac{1}{1+r_t}\mu_i - 2\beta_i I_i + \beta_i Q(1/2) \text{ and } pa_{i2} = \frac{1}{1+r_t}\mu_i - 2\beta_i I_i + \beta_i Q(3/2)$$

$$pb_{i1} = \frac{1}{1+r_t}\mu_i - 2\beta_i I_i - \beta_i Q(1/2) \text{ and } pb_{i2} = \frac{1}{1+r_t}\mu_i - 2\beta_i I_i - \beta_i Q(3/2)$$

with i=1,2 (dealers).

The possible types of equilibrium are:

(a) Monopoly equilibrium: the same dealer (denoted as dealer 1) is active on both sides of the market and she does not share the market. The equilibrium prices are

$$pa_{11} = \mu_1 - 2\beta_1 I_1 + \beta_1 Q(1/2) \text{ and } pb_{11} = \mu_1 - 2\beta_1 I_1 - \beta_1 Q(1/2).$$

Only in this case, where the same dealer executes the order on both sides of the market, we have independence of the inventory levels and the spread.[5]

(b) Specialised trading equilibrium: one dealer sets the lowest ask price and does all the selling; the other dealer sets the highest bid price and does all the buying.

The prices must again be at levels which just prevent either dealer from entering the side of the market on which they are inactive.

If 1 is the seller and 2 is the buyer this gives equilibrium prices of

$$pa_{11} = \mu_1 - 2\beta_1 I_1 + \beta_1 Q(1/2) \text{ and } pb_{21} = \mu_2 - 2\beta_2 I_2 - \beta_2 Q(1/2)$$

and now, contrary to the Ho and Stoll case, the spread does depend on the inventories and other characteristics of the two dealers.

---

[5]If dealer 1 holds an inventory of I and dealer 2 holds an inventory of I^0 with $I < I^0$ then (other things equal) dealer 1 is the buyer and we assume that she executes the whole order. At the time of the next incoming order, the inventory position of dealer 1 is (I+Q) and I^0 for dealer 2. Only if $(I+Q) > I^0$ (other things equal) dealer 1 again will be active and will sell.

(c) Shared trading equilibrium: both traders set identical bid and ask prices; the prices must be set so as to give no incentive for either dealer to either meet all the orders on one side of the market or to drop out of either side of the market. Dealers gain from a wider spread so long as it does not disturb the equilibrium trading pattern; consequently the common market bid price will be set by the higher of pb_{11} and pb_{21}; any further reduction in the bid price would make one of the dealers wish to take on all the buying. The market ask price will be set by the lower of pa_{11} and pa_{21} for analogous reasons.

(d) Mixed sharing equilibrium: one dealer shares one side of the market but has all the dealing on the other side of the market; the second dealer is inactive on the second side of the market but shares trade on the first side. For example suppose that dealer 1 is active on both sides of the market but dealer 2 only sells. The bid price is set at pb_{21} while the ask price is set at the lower of pa_{21} and pa_{11}.

For a given dealer the reservation prices divide price space into 9 regions; as in figure 2.3.. However, only if the difference between the inventories is relatively large compared to the order size is situation 4 possible.

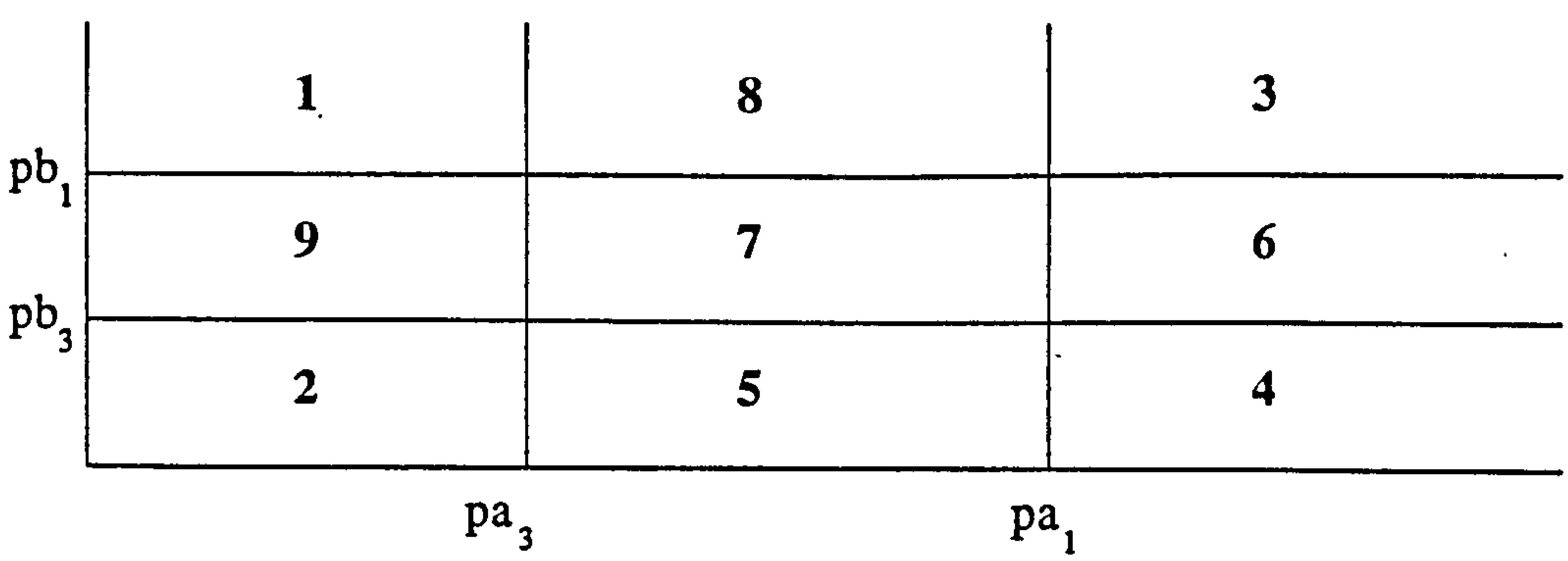

Figure 2.3. : Trading patterns

Table 2.1. gives the trading pattern for each region:

Table 2.1. : Trading combinations

	Sole Buyer	Share Buying	No Buying
Sole seller	4	6	3
Share selling	5	7	8
No Selling	2	9	1

Within such a setting we have shown that the equilibrium outcome is a Bertrand/Nash equilibrium. Such an equilibrium is based on the assumption that market makers differ in their inventory positions degree of risk aversion and price expectations. Under the assumption of decreasing returns to scale and risk averse dealers, we have shown that a dealer always prefers to share the order. This means that the dealer faces a reduced risk exposure by trading a smaller quantity and therefore she will quote a lower spread.

The sharing of the order gives more than one possible equilibrium outcome which can be obtained by a Bertrand price competition among market makers. The process of reaching such an equilibrium can be described by a first price auction. Under the assumption that the dealers have full information including the knowledge of each others' reservation prices there can be collusion among the best quoting dealers. The result is that all of them quote the same price which is just slightly below (above) the next best dealer's ask (bid) price.

Such a market structure can be implemented through appropriate trading rules. Such a rule is that the best quoting dealers get an equal share of the market order.

2.3. Trading within a Fragmented Market Structure

In this section, we analyse the pricing behaviour of dealers where the reservation prices are private information to the dealers. However, the dealers have knowledge about the incoming order which will be executed by the best quoting dealer which means by the dealer with the lowest ask price and the highest bid price. Also this time, only one order arrives at the time which is known to the dealers.

2.3.1. The Model of Biais with Trading in an Indivisible Good

Biais (1993) assumes a market for one risky security with two types of agents, the liquidity traders who are the public and who demand liquidity, and the risk averse agents who supply liquidity.

In a centralised market, these agents are limit order traders, also called market makers. In a fragmented market, these agents are the dealers. As discussed above, in the centralised market, market makers can observe price quotes and actual trades, whereas in fragmented markets, dealers do not have such information. They only know the distribution of the other dealers' inventories of the risky security, denoted as G(.) (the cumulative d.f.). It is assumed that all dealers inventories are identically distributed.

All agents are assumed to have constant coefficients of absolute risk aversion and homogeneous expectations about the final value of the security, i.e. E(P). The realisation of the final value P can be written as $P = (1+z)$ with $z \sim N(0,\sigma^2)$. So the problem is reduced to a mean-variance analysis.

The model is broken down into the following stages: [6]

"1.N out of M agents decide whether to become liquidity suppliers, at a given cost [F].

2. All M agents receive inventory positions $[I_i]$ in the risky security.[7]

3. With probability λ the liquidity shock on the risk exposure of the public occurs. In this case the public places one market order....[]. This is modeled as a random inventory position...[]. If the liquidity shock occurs, the risk averse outside investor is endowed with a long position +L, with probability 1/2, or with a short position -L, with probability 1/2.

4. The N suppliers of liquidity compete for the order flow from the public. The buy (sell) market order is executed at the best ask (bid) price, ...[] denoted by A_i and B_i...[] the agent i serves the market order to buy (sell), if his inventory is larger (lower) than those of his competitors (I_{-i}). The probability that this is the case is $P(I_i>I_{-i}) = \pi_{a,i}$ or $P(I_i<I_{-i}) = \pi_{b,i}$.

5. The final value of the security is realised. It is denoted by P. It can be thought of as the liquidation value of the asset. At that point in time, all uncertainty about the payoff of the asset is assumed to be resolved. "

The M agents have identical utility functions with constant absolute risk aversion parameter A: $U(X) = -e^{-Ax}$, $\forall$ x

With the expected final value to be E(P) we can write the ask and bid quotes as $A_i = E(P)(1+a_i)$ and $B_i = E(P)(1+b_i)$. In the following analysis E(P) is normalised to one. In such a setting, Biais assumes that the dealers only differ in their inventory positions which is quite crucial for the outcome of this analysis.

Hence, due to different initial inventory endowments, agents want to trade with different intensities which results in different inventory positions. Thus bidding strategies are assumed to be decreasing functions of the agents' inventories and we can write $a_i = a_i(I_i)$ and $b_i = b_i(I_i)$ which implies that we have a symmetric equilibrium with common bidding functions which are assumed to be monotone. Furthermore, we define $(I_i^*)_{i=1,...,N}$ to

[6]Biais (1993) pp.160/161

[7]They are also endowed with a cash position of C_i.

be the set of order statistics which are formed from the inventories $(I_i)_{i=1,..,N}$. (I_N^*) is the agent with the longest inventory, and (I_{N-1}^*) is the agent with the second longest inventory and so on.

2.3.2. Price Quotes in Equilibrium

The dealer, endowed with cash C_i and inventory I_i, pays a cost F to be in the market. Thus the final wealth of a dealer who is not trading is

$$W_i(0) = C_i - F + I_i(1+z) \tag{10}$$

where the risk free rate of interest is normalised to zero.

If the dealer sells a quantity Q at a price $1+a_i$, the final wealth is

$$W_i(a_i) = C_i - F + I_i(1+z) + (a_i-z)Q \tag{11}$$

for the selling side and

$$W_i(b_i) = C_i - F + I_i(1+z) + (b_i+z)Q \tag{12}$$

for the buying side.

Expected utility of trading has to be at least equal to expected utility of no trading otherwise the dealer does not want to trade.

$$E(U[W_i(0)]|I_i) \leq E_iU(W_i(a_i)) \ , \ E_iU(W_i(b_i))$$

which results in the following reservation fees, given the above assumptions:

$$a_{r,i} = a_r(I_i) = (A\sigma^2/2)(Q - 2I_i) \tag{13}$$

and $b_{r,i} = b_r(I_i) = (A\sigma^2/2)(Q + 2I_i)$ (14)

which are the same as the ones of Ho and Stoll.

Now, what are the optimal prices to quote for a dealer in such a fragmented market? We know that $\pi_{a,i} = P(I_i > \max(\tilde{I}_{-i}))$ [8] $= G(I_i)^{N-1}$ where -i denotes everybody but i, and we define $G(I_i)^{N-1} = H(I_i)$. The common bidding

---

[8]Generally a ~ means a random variable.

strategy of each dealer is $(P+a_i) = (P+a_i(I_i))$ and $(P+b_i) = (P+b_i(I_i))$.

As we assume a symmetric market where the size of the purchases is identical to the size of a sale and it is equally likely that a buy or a sell transaction occurs, only the ask side of the market is analysed.

When a dealer quotes his price he does not know about the competitors' inventories. The dealer forms expectations about it. He knows that he will execute the order from the liquidity traders if his ask price is lowest compared to the others. Thus, the dealer computes the probability of his ask price being lower than the ones of the competitors which we denote as $\pi_{a,i}$ [9]. The expected utility before trade but after knowing the private inventory shock is

$E\left[U(W_i(0)) + \pi_{a,i}[U(W_i(a_i)) - U(W_i(0))] \,|\, I_i\right]$ and he solves

$$\max_{a_i} \pi_{a,i} \left(E[U(W_i(a_i)) \,|\, I_i] - E[U(W_i(0)) \,|\, I_i]\right) \tag{15}$$

We can write:

$$E[U(W_i(a_i)) \,|\, I_i] = E[U(W_i(0))e^{-A(a_i-a_{r,i})Q} \,|\, I_i] \tag{16}$$

substituting (16) into (15) we get:

$$\max_{a} \pi_{a,i} \left(E[U(W_i(0))e^{-A(a_i-a_{r,i})Q} \,|\, I_i] - E[U(W_i(0)) \,|\, I_i]\right)$$

which is $\max_{a} \pi_{a,i} \; [E[U(W_i(0)) \,|\, I_i][e^{-A(a_i-a_{r,i})Q}-1]$

Thus, we can simplify by $E[U(W_i(0)) \,|\, I_i]$ which is a negative constant and we multiply by -1. The maximisation problem can then be expressed as maximising the expected surplus from trade which is:

$$\max_{a_i} \pi_{a,i}(1-e^{-A(a_i-a_{r,i})Q})$$

We can simplify this expression by using a Taylor approximation. By Taylor expansion and by neglecting terms of the magnitude $((a_i-a_{r,i})AQ)^2$ we get

---

[9]See also point 4 above.

$$\max_{a_i} \pi_{a,i}\left(A(a_i-a_{r,i})Q\right)$$

We solve the maximisation problem and get the first order condition of:

$$\delta\pi_{a,i}/\delta a_i(a_i-a_{r,i}) + \pi_{a,i} = 0$$

Then, $\delta\pi_{a,i}/\delta a_i = \left[\delta\pi_{a,i}/\delta I_i\right]\left[dI_i/da_i\right] = (\delta H(I_i)/\delta I_i)(1/a'_i)$

with a'_i being the derivative of $a(I_i)$ with respect to I_i.

We then can write the differential equation as

$$(\delta H(I_i)/\delta I_i)(a_i-a_{r,i})(1/a'_i) + H(I_i) = 0 \text{ or}$$

$$\delta\left[H(I_i)a_i\right]/\delta I_i = h(I_i)a_{r,i}$$

where h(.) is the derivative of H(.) with respect to I_i.

Now, we integrate H(x)a(x) over the interval between -R and I_i which gives

$$\left[H(x)a(x)\right]_{-R}^{I_i} = \int_{-R}^{I_i} h(x)a_r(x)dx + c$$

with c being a constant. The LHS of above equation is zero when we evaluate it at I_i = -R which results in the RHS being zero as well.

It follows that a is such that

$$a(I_i)H(I_i) = \int_{-R}^{I_i} h(x)a_r(x)dx$$

By substituting the reservation fees calculated above and integrating the RHS by parts we get optimal fees of [10]

$$a_i = a_{r,i} + A\sigma^2\left[\int_{-R}^{I_i} G(x)^{N-1}\, dx \;/\; G(I_i)^{N-1}\right] \tag{17}$$

and by applying the same procedure to the bid side we get

$$b_i = b_{r,i} + A\sigma^2\left[\int_{I_i}^{R} (1-G(x))^{N-1}\, dx \;/\; (1-G(I_i))^{N-1}\right] \tag{18}$$

---

[10]This result is similar to the optimal bidding strategy in a high bid auction of Riley and Samuelson (1981). They also showed that, in equilibrium, it is optimal for each dealer to adopt the same strategy. However, their bidders are assumed to be risk neutral which is not the case in Biais model which makes it more appropriate for a dealership market. However they all do not investigate the second order condition which is always assumed to be satisfied.

Under the assumption that the same dealer executes the order on both sides of the market, Biais defines the spread as:

$$S_i = a_i + b_i = S_{r,i} + A\sigma^2\left[\left[\int_{-R}^{I_i} G(x)^{N-1}dx/G(I_i)^{N-1}\right]+\left[\int_{I_i}^{R} (1-G(x))^{N-1}dx/(1-G(I_i))^{N-1}\right]\right] \quad (19)$$

However this result is not always a valid solution. The reason is simple and can be explained by the fact that market makers differ in their inventory only. Hence, it may not be possible that the same dealer, among several market makers, can bid the best bid price and get the order and, in the following period, she bids the best ask price which is, according to Biais assumption, a function of the dealer's inventory position.

For example, we denote the inventories of two dealers as I and I^0 and we assume that $I < I^0$ which is the only difference between the dealers who are assumed to have common price expectations and risk aversions. In this example the dealer with the inventory I is in the position to bid a better buying price and gets the order of size Q. Hence her inventory changes to (I + Q). Now, for the next period, it depends on the relative inventory positions whether $(I + Q) \gtrless I^0$ which of the dealers bids the best price for selling. If $(I + Q) < I^0$ then the other dealer does all the selling. On the other hand if $(I + Q) > I^0$ then the same dealer executes the sale. Thus, Ho and Stoll's and Biais' result gives only one part of the solution which is the case where the difference of the inventory positions of two dealers is less than the order size.

Hence, the spread equation may depend on the relative inventory levels of the market makers which we express in the case below .

If we assume an uniform distribution with G(.) uniform over [-R, R] the bidding fees and the spread become

$$a_i = a_{r,i} + A\sigma^2\left[(R + I)/N\right] \quad \text{and} \quad b_i = b_{r,i} + A\sigma^2\left[(R - I)/N\right]$$

$$S_i = a_i + b_i = S_{r,i} + 2R(A\sigma^2/N) = A\sigma^2(Q + (2R/N)) \quad (20)$$

which is in the more general case:

$S_i = a_{ri1} + A\sigma^2\left[(R + I_1)/N\right] + b_{ri2} + A\sigma^2\left[(R - I_2)/N\right]$

where dealer 1 is the seller and dealer 2 is the buyer.

If we compare the spread in the centralised market and in the fragmented market we have one important difference which is that in a fragmented market, the spread also depends on the number of dealers active in the market. We see that as the number of dealers N increases the surplus earned in the fragmented market gets smaller and smaller. This can be explained by the fact that the divergence of inventories is less with many active dealers than with only a small number of dealers which puts competitive pressure on the dealers and forces them to narrow the spread which at the limit is equal to the spread in the centralised market.

2.3.3. Modification of Heterogeneous Price Expectations

Biais assumes that dealers have common price expectations and the same degrees of risk aversion. The consequence is that the quoted price strategy is a function of the inventory only. If we change this assumption and let the quoted price be a function of the reservation price with the price expectation and the inventory position being random variables, we get a different result.

As a reminder, the reservation price is $pa_{r,i} = \tilde{\mu}_i + \gamma\sigma^2[Q - 2\tilde{I}_i]$.

It is obvious that the reservation price is decreasing in inventory, and increasing in future price expectation.

We can define $(pa^{R*}_i)_{i=1,..,N}$ to be the set of order statistics which is formed from the reservation prices $(pa^R_i)_{i=1,..,N}$. (pa^{R*}_N) is the lowest reservation price, and (pa^{R*}_{N-1}) is the second lowest reservation price and so on. We assume the sample of reservation prices is a random draw from a common cumulative distribution function G(.) and that the quote each dealer makes increases with the reservation price. Underlying this we could have a common distribution of inventories and price expectations. All dealers use the same bidding function $\bar{p}a(pa^R)$ where $\bar{p}a$ is the price quote which is increasing. [11]

We then can express the probability that dealer i wins the ask bid at $\bar{p}a_i$ as equivalent to the probability that dealer i's reservation price is lower than the second lowest reservation price of all the other N-1 dealers of which dealer i has some expectations. We can write such a probability as:

$Pr(\bar{p}a_i < \tilde{p}a_j \; i \neq j) = Pr(pa^R_i < \tilde{p}a^R_j \; i \neq j) = (1-G(pa^R_i))^{N-1}$.

$(1-G(pa^R_i))$ is the distribution function which describes the probability that dealer i wins the bid at pa^R_i which implies that $pa^R_i < \tilde{p}a^R_j$ with $i \neq j$. This in turn means that [12] at least one of the following conditions hold:

$I_i > I_j$ or $\mu_i < \mu_j$. [13]

The expected utility is composed of the probability of winning with $\pi(pa^R_i) = (1-G(pa^R_i))^{N-1}$ and the gain from trading which is:

$EU = \pi(pa^R_i)\,[U(Q,\bar{p}a_i) - U(0)]Q.$

The optimisation problem of dealer i is to maximize:

$$\max_{\bar{p}a_i} \pi(pa^R_i)\,[U(Q,\bar{p}a_i) - U(0)]Q$$

and we have $[U(Q,\bar{p}a_i) - U(0)] = (\bar{p}a_i - pa^R_i)Q$

By differentiation we get the first order condition which is [14]

---

[11]The proof can be found in appendix A.

[12]We assume that dealers have common degrees of risk aversions.

[13]For the bid side, the probability of winning is $Pr(\bar{p}b_i > \tilde{p}b_j) = Pr(pb^R_i > \tilde{p}b^R_j) = G(pb^R_i)^{N-1}$. It also means that at least one of the following conditions hold: $I_i < I_j$ or $\mu_i > \mu_j$.

[14] Mathematical methods in Chiang (1984)

$(\delta\pi(pa^R_i)/\delta pa^R_i)(\delta pa^R_i/\delta\bar{p}a_i)(\bar{p}a_i - pa^R_i) + \pi(pa^R_i) = 0$

The solution of this maximisation problem gives us reservation prices of [15]

$$\bar{p}a_i = pa^R_i + \left[\int_{pa^R_i}^{R} (1-G(X))^{N-1} dX/(1-G(pa^R_i))^{N-1}\right] \tag{21}$$

and

$$\bar{p}b_i = pb^R_i - \left[\int_0^{pb^R_i} G(X)^{N-1} dX/G(pb^R_i)^{N-1}\right] \tag{22}$$

In contrast to the analysis in the previous section, we have two different distribution functions. In order to calculate the bid-ask spread we first calculate the extreme values of the distribution functions of the inventories and the price expectations.

We can write the reservation prices with respect to these two parameters, i.e. inventory levels and price expectations, as $p^R_b = \tilde{\mu} - \beta\tilde{I} - \beta(Q/2)$ and $p^R_a = \tilde{\mu} - \beta\tilde{I} + \beta(Q/2)$, but the last term in both equations is common to all dealers. To find the limits of the distributions we have to examine the maximum and the minimum of the possible values which the variables can take on. The respective parameter ranges are illustrated in figure 2.4. below.

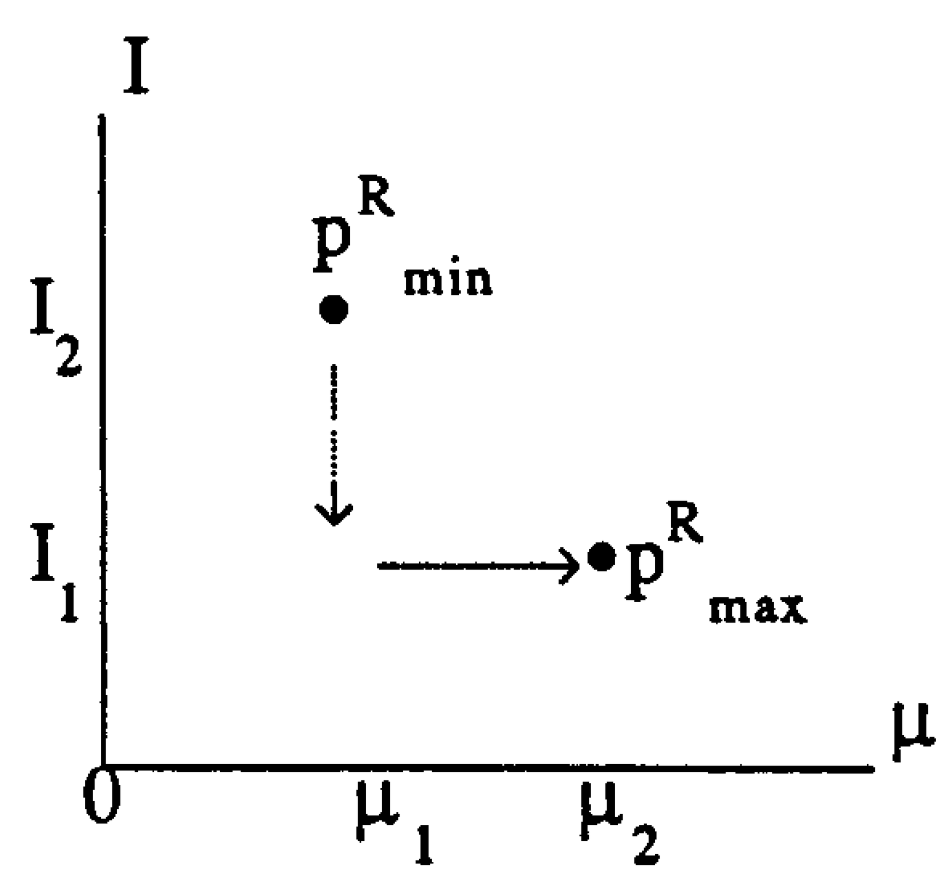

Figure 2.4. : Parameter ranges

Since $(\delta\bar{p}a/\delta pa^R) > 0$ the highest ask quote is at the highest reservation ask price which is $pa^R_i = \mu_2 - \beta I_1 + \beta(Q/2)$ and for the bid side we get the

---

[15]The detailed calculation is presented in appendix A.

lowest bid price at the lowest bid reservation price which is $pb^R_i = \mu_1 - \beta I_2 - \beta(Q/2)$. These price quotes are the extreme values which give the "worst" or largest spread that is possible which is not necessarily the market or transaction spread.

In order to get the mean reservation spread we have to take expectations over the random variables in the price quote equations above.[1] Our result shows that in a rational expectations equilibrium the mean reservation quotes in a fragmented market are the same as the average of the prices we can expect in the centralised market which is identical to the findings of Biais.

In order to compare our result with the result of Biais, we also assume a uniform distribution of pa^R_i and pb^R_i over the interval [-R, R].[2]

Now, we have $\int_{-R}^{pb^R} G(X)^{N-1}dX = \int_{-R}^{pb^R} [(X+R)/2R]^{N-1}dX = (1/2R)^{N-1} \int_{-R}^{pb^R} (X+R)^{N-1}dX$

we have $(1/2R)^{N-1} = (1/2R)^N 2R$ and $\int_{-R}^{pb^R} (X+R)^{N-1} = (1/N)(X+R)^N \Big|_{-R}^{pb^R_i}$ which gives

$$(1/2R)^N(2R/N)(X + R)^N \Big|_{-R}^{pb^R_i} = (2R/N)\ (1/2R)^N[pb^r_i + R)]^N$$
$$= (2R/N)[(pb^r_i + R)/2R]^N$$

Our expression from the price quote in (21) and (22) is $\int G(X)^{N-1}/G(pb^R)$ which gives us $(1/N)(pb^r_i + R)^N$. Thus,

$\bar{pb}_i = pb^R_i - (1/N)(pb^r_i + R)$.

If we apply the same procedure to the ask side we get

---

[1]The proof is given in appendix A.

[2]$G(X)^{N-1} = [(X-(-R))/(R-(-R))]^{N-1} = [(X + R)/2R]^{N-1}$.

$\bar{p}a_i = pa^R_i + (1/N)(R - pa^R_i)$.

The bid-ask spread at the extremes i.e. at the highest bid and the lowest ask price is

$$\bar{p}a_i - \bar{p}b_i = (\mu_1 - \mu_2) + 2\gamma\sigma^2(Q + (I_1 - I_2)$$
$$+ (1/N)\Big[2R + (\mu_2 - \mu_1) - 2\gamma\sigma^2(Q + (I_2-I_1))\Big] \qquad (23)$$

If we examine the result in respect of the number of dealers, the same is true as in Biais' model that in the limiting case, i.e. when N goes to infinity, the expression

$(pa_i-pa^R_i) = \Big[\int^R_{pa^R_i} (1-G(X))^{N-1}dpa^R_i/(1-G(pa^R_i))^{N-1}\Big]$ and also $(pb_i - pb^R_i)$

approaches zero. Thus, in a competitive market with a large number of dealers, the differences among dealers get smaller and smaller and the spread becomes the same as in the centralised market.

We extend our analysis to take into account the fact that the underlying random variable ($\tilde{p}a^R$, $\tilde{p}b^R$ respectively) of the distribution function is actually composed of two different independent random variables which are the price expectations $\tilde{\mu}$ and the inventory positions $\tilde{I}$. In order to calculate the joint density of the random variables $\tilde{\mu}$ and $\tilde{I}$ we have to apply the method of convolution of density functions. We assume that the price expectations are uniformly distributed and that the inventories have an exponential distribution.

The exponential distribution in inventories can be explained by the assumption that the market consists of a large number of small private investors and a small number of large investors (for instance international banks) which creates the skewness of the distribution.

Thus we have:

$pb^r = \tilde{\mu} - \gamma\sigma^2 Q - 2\gamma\sigma^2 \tilde{I}$ and $f_\mu(\mu) = [1/(b-a)]$ with μ uniform over the interval [a,b]. We define $A = y + \mu = (-2\gamma\sigma^2 I) + \mu$.

The density function of I is $f_I = \lambda e^{-\lambda I}$ with a cumulative distribution function of $F = \int_0^I \lambda e^{-\lambda I} = -(\lambda/\lambda)e^{-\lambda I}\Big|_0^I = 1-e^{-\lambda I}$ on $[0,\infty)$.

We can write: $y = -2\gamma\sigma^2 I$ on $(-\infty,0]$ so y has a density

$f_I = (-y/2\gamma\sigma^2)\ (1/2\gamma\sigma^2) = \lambda e^{\lambda y/2\gamma\sigma^2}(1/2\gamma\sigma^2)$

Thus, $f_A(A) = \int_a^b (1/b-a)\ \lambda e^{\lambda/2\gamma\sigma^2 (A-\mu)}(1/2\gamma\sigma^2)d\mu$

$= [\lambda/(b-a)](1/2\gamma\sigma^2)e^{(\lambda/2\gamma\sigma^2)A} \int_a^b e^{-(\lambda/2\gamma\sigma^2)\mu}\ d\mu$

$= [\lambda/(b-a)]e^{\lambda A/2\gamma\sigma^2}(-2\gamma\sigma^2/\lambda)[-e^{-\lambda b/2\gamma\sigma^2} + e^{-\lambda a/2\gamma\sigma^2}] > 0.$

Hence A is within the interval$(-\infty,b]$.

As we have $pb^r = A - \gamma\sigma^2 Q$ we define:

$f_{pb}(pb^r) = f_A(pb^r + \gamma\sigma^2 Q)$

$= (\lambda/b-a)e^{\lambda(pb+\gamma\sigma^2 Q)/2\gamma\sigma^2}(2\gamma\sigma^2/\lambda)[e^{-\lambda a/2\gamma\sigma^2} - e^{-\lambda b/2\gamma\sigma^2}]$ which is an exponential function.

If we assume that both random variables have a uniform distribution we get a density function of the form $f_b(B) = [(pb + \gamma\sigma^2 Q)/(b-a)(4\gamma^2(\sigma^2)^2]$.

Under the assumption that the ranges of the two distributions are identical we get again a uniform distribution.

This analysis on the underlying random variables gives us a more general application of the price quotes with respect to the form of the respective distribution function.

2.4. Fragmented Markets and Trading in a Divisible Good

The main difference between this analysis and the model presented in the previous section is that we allow for the splitting of the order. Dealers know the whole order size, but they have to form some expectation about the share of the order they will get since they do not know the reservation prices and quotes of the other dealers. The trading procedure is modeled in the way described below.

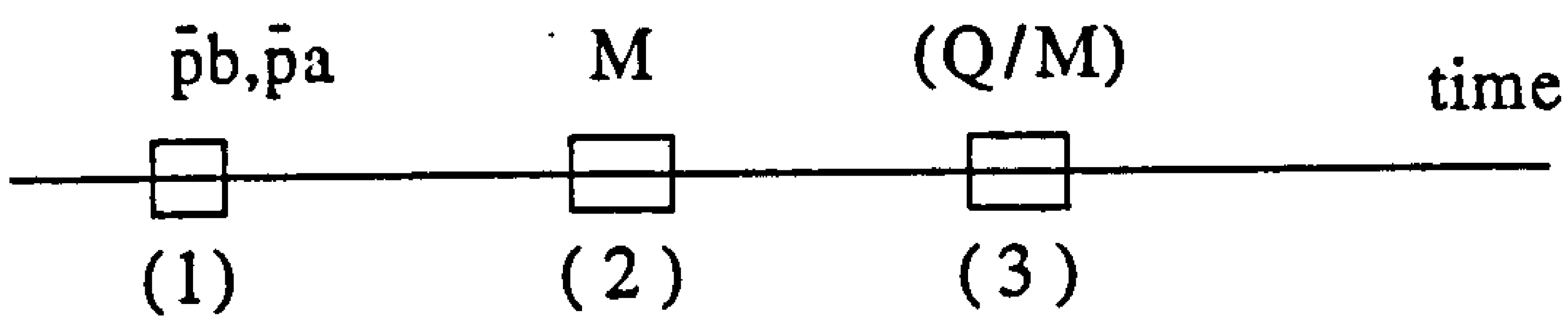

Figure 2.5. : Trading procedure

At date (1) the dealers quote their prices, i.e. submit their bids which we denote as $\bar{p}b_1$, $\bar{p}b_2$, $\bar{p}b_3$, ..., $\bar{p}b_N$.

After having received the bids, at date (2), a public authority decides how to share the order Q, i.e. how many shares she will distribute among the dealers, which determines the number of winners M. The rule of the market is that the investor has to share the order equally between the winners of the bidding procedure. The investor has to follow the rules of the market which restricts the possibilities of allocating the shares. Such a rule is known to all the participants in the market and the price quotes have to be public. Such a rule is required as the private investor does not have an incentive to share the order as for her it is optimal to give the whole order to the best quoting dealer and not to consider the other dealers at all. Hence, it will not be possible that the investor can choose the dealer with the best price from several possible auctions and give her the whole order. Finally, at date (3), the dealers get their share Q/M at their quoted prices.

Such a procedure is similar to a discriminating auction described by Harris and Raviv (1981). In our case, dealers maximise their expected gain over several possible discriminating auction procedures. The auction is chosen by the public authority and the dealers do not know ex ante which discriminating auction will be held, nor whether they will be one of the winners in the auction.

2.4.1. The Basics of the Model

We assume that the order is of size Q. There are N bidders who bid for Q/M of the order. However, the market makers do not know M and thus the share they will finally get. The N bidders (or dealers) have reservation prices denoted as pb^R_i with i=1, ..., N. Again, we assume a symmetric market and, this time, we consider the buying side only.

We can write $pb^R_i = \mu_i - 2\gamma\sigma^2 I_i - \gamma\sigma^2(Q/M)] = \alpha_i - \beta(Q/M)$ [3] and allow α_i to vary between dealers. In this way, we assume that market makers differ in their inventories and their price expectations, but that they have common risk aversions and hence β. The bidding price is a function of the reservation cost α_i which we can formulate as $\bar{p}b_i = \bar{p}b_i(\alpha_i)$.

Each dealer knows her reservation cost and assumes that each other dealer draws her reservation cost from the same distribution of α with density function g and the cumulative distribution function of $G(\alpha_i)$. The range of g is [-R, R].

If the private investor chooses M winners, each of the M best bidders get a share of Q/M at their own bid price. Then, the probability of being one of M winners if the bid is $\bar{p}b_i$ is:

$Pr(\bar{p}b_i > \bar{p}b_N, \bar{p}b_{N-1}, ..., \bar{p}b_{N-(M+1)}) = F(\alpha_i, M)$

which means that dealer i's bid has to be higher than all N-(M+1) other dealers' (who are not one of the M winners) bids. The number of winners M can range from 1 to N which is chosen by the private investor. So the dealer does not know ex ante which auction is chosen by the private investor and hence, which share of the order she will get.

---

[3]The reservation ask price is $pa^R_i = \mu_i - 2\gamma\sigma^2 I_i + \gamma\sigma^2(Q/M) = \alpha_i + \beta(Q/M)$.

2.4.2. Bid and Ask Prices in Equilibrium

We assume that the ith dealer believes that the other dealers use the common increasing strategy function [4] $\bar{p}b_j = \bar{p}b(\alpha_j)$ for $j \neq i$.

Now, we can derive the Nash equilibrium bidding strategy.

We assume that we have a uniform distribution for the selection of M which means that it is equally likely that the investor holds each of the possible auction types. Thus, we can say that the probability of winning the auction with M winners with the bid $\bar{p}b_i$ is the probability that α_i exceeds the N-M order statistic among the N-1 other bidders with M=1,..,N.

We define y to be the N-M lowest bid of the other N-1 bidders.

$\Pr[\text{winning the bid } \bar{p}b_i] = \mathrm{pr}[\alpha_i > \tilde{y}]$

$= \Pr[\alpha_i > \tilde{\alpha}_{(N-1)-(M-1)}] = F(\alpha_i, M)$ [5]

with $\alpha_{(N-1)-(M-1)}$ being the N-M order statistic among the N-1 reservation costs of the other dealers and F(.) being the cumulative distribution of (N-1)-(M-1) order statistic when N-1 random variables are drawn.

We get the cumulative distribution function by considering all possible combinations of winning the auction with M winners, i.e. all possible draws made by the private investor out of the sample of N dealers, which is

$$F(\alpha_i, M) = \sum_j \binom{N-1}{j} G(\alpha_i)^j \left[1-G(\alpha_i)\right]^{N-1-j} \tag{24}$$

with j = N-M, ..., N-1.

Thus, the expected utility of a dealer is given by the probability of winning and the gain from trading which is

---

[4]The proof of monotonicity is given in appendix B.

[5]On the ask side, the probability of winning the bid $\bar{p}a_i$ is $\Pr[\alpha_i < \tilde{y}] = \Pr[\alpha_i < \alpha_{(N-1)-(M-1)}] = F_a(\alpha_i, M) = \sum_j \binom{N-1}{j} G(\alpha_i)^{N-1-j}(1-G(\alpha_i))^j$.

$$EU^B = \left[\sum_M F(\alpha_i,M)[\alpha_i - \bar{p}b_i - \beta(Q/M)] \right](Q/M) \qquad (25)$$

with $\alpha_i = \mu_i - 2\gamma\sigma^2 I_i$ and $\beta = \gamma\sigma^2$

The expected utility on the ask side is:

$$EU^S = \left[\sum_M F_a(\alpha_i,M))[\bar{p}a_i - \alpha_i - \beta(Q/M)] \right](Q/M) \qquad (26)$$

We can then state the optimization problem for the bid side which is $\max_{\bar{p}b_i} EU^B$.

The first order condition is: [6]

$$\sum_M [(\delta F(\alpha_i,M)/\delta\alpha_i)(\delta\alpha_i/\delta\bar{p}b_i)(\alpha_i - \bar{p}b_i - \beta(Q/M))](1/M) - \sum_M F(\alpha_i,M)(1/M) = 0$$

The resulting bidding prices are [7]

$$\bar{p}b_i = \left[\sum_M F(\alpha_i,M)[\alpha_i - \beta(Q/M)](1/M) / \sum_M F(\alpha_i,M)(1/M) \right] - \left[\sum_M \int_0^{\alpha_i} F(X)dX(1/M) / \sum_M F(\alpha_i,M)(1/M) \right] \qquad (27)$$

By symmetry, the ask price is:

$$\bar{p}a_i = \left[\sum_M (F_a(\alpha_i,M))[\alpha_i + \beta(Q/M)](1/M) / \sum_M (F_a(\alpha_i,M))(1/M) \right] + \left[\sum_M \int_{\alpha_i}^{R} (F_a(X))dX(1/M) / \sum_M (F_a(\alpha_i,M))(1/M) \right] \qquad (28)$$

with $F_a(\alpha_i,M) = \sum_j \binom{N-1}{j} G(\alpha_i)^{N-1-j}(1-G(\alpha_i))^j$ being the cumulative distribution function which gives the probability that the ask price of dealer i is smaller than the ask prices of all the other dealers who are non-winners denoted by j, i.e. $Pr(pa_i < \tilde{p}a_j)$.

If we compare this result with the prices of the modified Biais' model we can observe that our reservation prices are a weighted average over the possible outcomes and also the integral term is such a weighted average.

---

[6]The second order condition is assumed to be satisfied.

[7]The mathematics can be found in appendix B.

In order to determine which bid-ask spread is smaller we have to compare the price quotes for the divisible and the indivisible good case.

The prices for trading in an indivisible good are:

$$\bar{pa}_{ind} = pa^R_i + \left[\int_{pa^R_i}^{R} (1-G(X))^{N-1}dX/(1-G(pa^R_i))^{N-1}\right] \text{ and}$$

$$\bar{pb}_{ind} = pb^R_i - \left[\int_0^{pb^R_i} G(X)^{N-1}dX/G(pb^R_i)^{N-1}\right].$$

The respective prices for the divisible good case are:

$$\bar{pa}_{div} = \left[\sum_M (F_a(\alpha_i,M))[\alpha_i + \beta(Q/M)](1/M)/ \sum_M (F_a(\alpha_i,M))(1/M)\right]$$

$$+ \left[\sum_M \int_{\alpha_i}^{R} (F_a(X))dX(1/M)/ \sum_M (F_a(\alpha_i,M))(1/M)\right]$$

$$\bar{pb}_{div} = \left[\sum_M F(\alpha_i,M)[\alpha_i - \beta(Q/M)](1/M)/ \sum_M F(\alpha_i,M)(1/M)\right]$$

$$- \left[\sum_M \int_0^{\alpha_i} F(X)dX(1/M)/ \sum_M F(\alpha_i,M)(1/M)\right] \text{ and}$$

According to our hypothesis trading in a divisible good should yield a smaller spread than trading in an indivisible good. Risk averse dealers can reduce their risk exposure by trading a smaller quantity under the assumption of decreasing returns to scale. As a consequence, they quote a lower ask price and a higher bid price which results in a smaller spread.

Thus we have to show that $\bar{pa}_{div} < \bar{pa}_{ind}$ and/or $\bar{pb}_{div} > \bar{pb}_{ind}$.
The first term of the price quotes in the divisible good case is a weighted average of the reservation prices. On the bid side with a reservation price of $pb^r = \alpha - \beta(Q/M)$, we see that the reservation price is decreasing with a higher share of the market order, i.e. (Q/M). Hence, the weighted average price is higher than the price in Biais' model with an indivisible good.

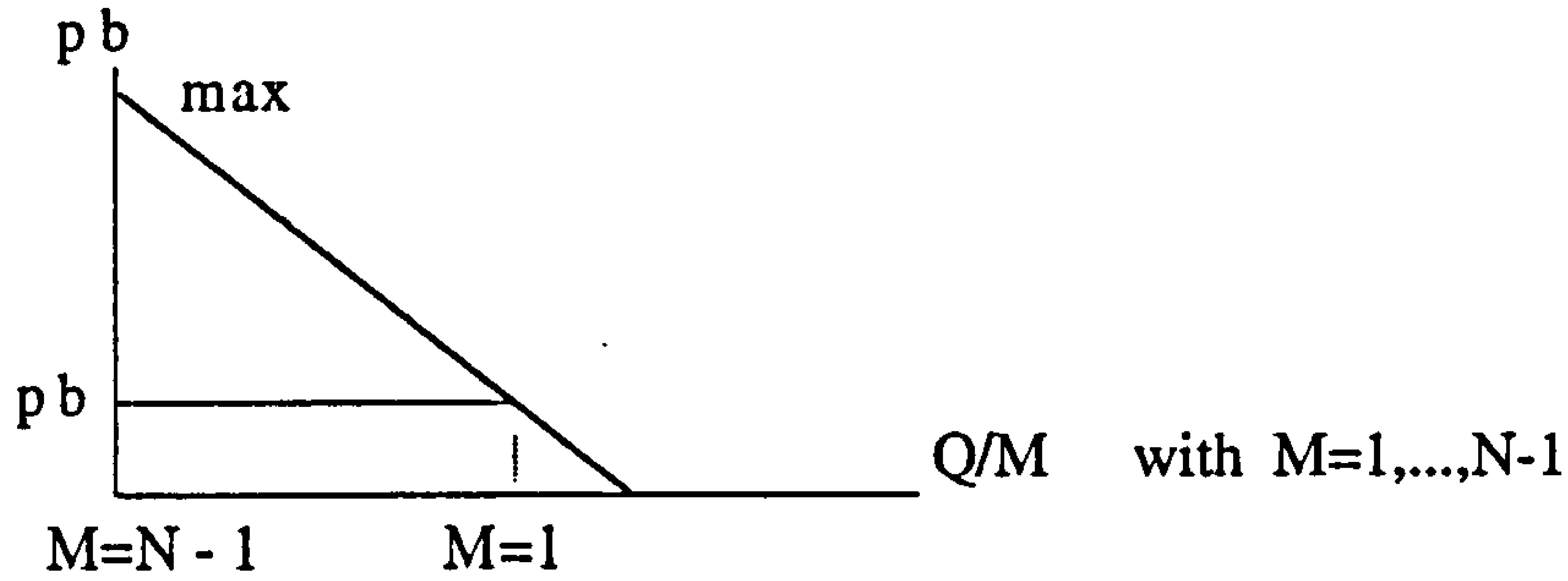

Figure 2.6. : Weighted average of prices

On the ask side, we have the opposite relation between the size of the share and the reservation price and thus the weighted average is smaller than the reservation price for trading in an indivisible good.

The direct comparison of the respective second terms of the price quotes will give us the sign of the expression and hence will tell us whether our hypothesis is confirmed.

For our analysis we can express the second term of the price quotes for the divisible case as $\sum_{M=1}^{N} a_M / \sum_{M=1}^{N} b_M$ and for the indivisible case as a_1/b_1.

We can write:

$$\sum_{M=1}^{N} a_M / \sum_{M=1}^{N} b_M - a_1/b_1 = \left[\sum_{M=2}^{N} a_M / \sum_{M=2}^{N} b_M - a_1/b_1 \right] b_1 \sum_{M=2}^{N} b_M / b_1 \sum_{M=1}^{N} b_M$$

We can show that our assumption is confirmed by assuming that N=2.

The proof for the general case is given in appendix C by following the same line of argument as we present below. With N=2 we have:

$$\text{sign}[(a_2+a_1)/(b_2+b_1) - (a_1/b_1)] = \text{sign}[(a_2/b_2) - (a_1/b_1)].$$

By differentiating the respective probabilities of a and b for the bid side we get

$$\delta/\delta j \left[[\int_0^{\alpha} G(y)^j (1-G(y))^{N-1-j}] / [G(\alpha)^j (1-G(\alpha))^{N-1-j}] \right]$$

$$= \left[[\int_0^{\alpha} G(y)^j(1-G(y))^{N-1-j}][\ln[G(y)/(1-G(y))] - \ln[G(\alpha)/(1-G(\alpha))]] \right]$$

$$/[G(\alpha)^j(1-G(\alpha))^{N-1-j}] < 0$$

and we find that the ratio (a/b) is decreasing in respect of N which we can write as

$a_1/b_1 > a_2/b_2$. Thus, the expression $[(a_2/b_2) - (a_1/b_1)]$ is negative which confirms our hypothesis. [8]

Hence, we have established that under our set of assumptions the sharing of the market order reduces the bid-ask spread in a fragmented market.

2.5. Summary and Conclusions

In section 2.2.1. we have summarized the HS inventory control model for a centralised market structure (modeled as a transparent auction with Bertrand price competition among dealers). The resulting bid and ask prices are the prices of the second best dealer. These prices serve as a benchmark for the comparison with the other models.

One of these models is our approach in section 2.2.2. which explains the bid and ask prices in a fragmented market where dealers can share the order between them.

In a centralised market the ask and the bid prices are lower respectively higher than the HS prices and the spread is smaller due to the reduced risk in trading of a smaller quantity.

An important difference to the HS model is that, in our model, the bid-ask spread depends on the inventory levels of the dealers also within a one period framework which is in accordance to the findings of O'Hara and

---

[8]The proof for the ask side can be found in appendix C.

Oldfield (1986). In our case, this comes from the fact that due to their differences mainly in inventory there are various trading patterns possible if dealers can share the market order.

In addition, we have shown that there exist several types of equilibrium as, for instance, we can have the situation that on the selling side two dealers share the market whereas on the buying side one dealer buys all of the order. In a fragmented market where dealers know the order size but not the reservation prices of their competitors the situation is different.

We have presented the model of Biais which explains the market making in a fragmented market in section 2.3..

His model is based on the same assumptions as HS which is that the order goes to the best quoting dealer who serves both sides of the market. In addition, dealers are assumed to have constant absolute risk aversion and common price expectations. The resulting bid and ask spread is higher than the HS bid-ask spread which means that dealers face an increased risk exposure due to incomplete information. In a rational expectations equilibrium the expected prices in a fragmented market are identical to the average of prices which can be observed in a centralised market. However, if we modify Biais' model and let the price expectation vary amongst dealers then the result shows that the inventory levels of the dealers influences the bid-ask spread.

Based on Biais' model the bid-ask spread is the same as in a centralised market if the number of market makers is very large. We confirm this finding based on our modified version.

Our model in section 2.4. analyses the bidding behaviour of the dealers in the case where they cannot observe the reservation prices and the inventory positions of their competitors. They know the order flow, but they do not know the share of the order they finally will execute. An important difference to Biais' model is that for a given set of quotes dealers face equal probabilities of winning the bid and being the sole seller (or buyer) or sharing the order with either one, two, or N-1 other bidders.

We modeled such a trading procedure as a set of discriminating auctions from which a private investor chooses which auction will be held and she also decides amongst which dealers she shares the total order.

The findings are that, compared to Biais' results, the bid-ask spread is smaller as the dealer takes a weighted average over the reservation prices of the M best dealers. In addition we have shown that the weighted average over the second term of the price quote (which represents an incomplete information cost) is smaller than in Biais' model which indicates that the risk is reduced by sharing the market order.

The analysis of the bid and ask prices in a centralised and a fragmented market includes the analysis about the decision of how to structure and how to organise a market.

First, the market organised as an auction may be socially more favourable [9] than a pure dealership market where dealers have incomplete information about market trading which gives the private investor a more powerful position in the trading process.

Second, based on the characteristics of the traders, i.e. whether there are institutional private investors who would like to trade large quantities, the auction rules, such as how to share an incoming order, may be an important factor in determining the equilibrium prices and hence the equilibrium bid-ask spread.

The conclusion based on above findings is that the design of markets is important and influences the bid-ask spread. We have shown that a market

---

[9] in the Pareto sense

designed as an auction with the rule that the order is split in equal parts between the best quoting dealers is socially superior to an auction where the whole order is allocated to the best quoting dealer. We have given evidence that this is true in a centralised and in a fragmented market.
However, our analysis is based on trading in one risky asset only. If we extend the analysis to several assets or to several different markets the outcome may be different.

This problem will be investigated in the subsequent chapters.

APPENDIX A

1. Bidding Price for an Indivisible Good:

We assume that dealers have different inventories and different price expectations:

We define $[U(Q,\bar{p}a_i) - U(0)]Q = (\bar{p}a_i - pa^R_i)Q$ and the probability of winning the bid is $\pi(pa^R_i) = (1-G(pa^R_i))^{N-1}$.

The optimisation problem of dealer i is to maximize

$$\max_{\bar{p}a_i} \pi(pa^R_i)[U(Q,\bar{p}a_i) - U(0)]Q.$$

By differentiation we get the first order condition which is

$$(\delta\pi(pa^R_i)/\delta pa^R_i)(\delta pa^R_i/\delta\bar{p}a_i)(\bar{p}a_i - pa^R_i) + \pi(pa^R_i) = 0$$

$$(\delta\pi(pa^R_i)/\delta pa^R_i)\bar{p}a_i + [\pi(pa^R_i)/(\delta pa^R_i/\delta\bar{p}a_i)] = pa^R_i(\delta\pi(pa^R_i)/\delta pa^R_i)$$

which we can rewrite as

$(\delta\pi(pa^R_i)/\delta pa^R_i)\bar{p}a_i + \pi(pa^R_i)(\delta\bar{p}a_i/\delta pa^R_i) = pa^R_i(\delta\pi(pa^R_i)/\delta pa^R_i)$ which is

$$\delta/\delta pa^R_i\left[\pi(pa^R_i)\bar{p}a_i\right] = pa^R_i(\delta\pi(pa^R_i)/\delta pa^R_i)$$

Next, we are integrating [10].

$\pi(pa^R_i)\bar{p}a_i\Big|_{pa^R}^{R} = \int_{pa^R}^{R} pa^R_i(\delta\pi(pa^R_i)/\delta pa^R_i) + C$ (where C is a constant)

First we have to evaluate C at $pa^R_i = R$

$$\int_{pa^R}^{R} pa^R_i[\delta(1-G(pa^R_i))^{N-1}/\delta pa^R_i]\delta pa^R_i = pa^R_i\pi\Big|_{pa^R}^{R} - \int_{pa^R}^{R}\pi\delta pa^R_i + C$$

with $\pi(R) = (1-G(R))^{N-1}$ we have $\pi(R) = 0$ and $\pi(pa^R_i) = (1-G(pa^R_i))^{N-1}$ we get

---

[10]The integration is in the interval [0,R] where R is a real number.

$= pa^R_i\pi(0) - pa^R_i[(1-G(pa^R_i))^{N-1}] - [pa^R_i\pi(0) - [pa^R_i(1-G(pa^R_i))^{N-1}] = 0.$

Hence the RHS is zero and so is C.

Thus, we have:

$$\pi\bar{p}a_i(pa^R_i)\Big|_{pa^R}^{R} = \int_{pa^R}^{R} pa^R_i(\delta\pi(pa^R_i)/\delta pa^R_i)$$

By integrating by parts we obtain

$$\pi(pa^R_i)\bar{p}a_i(pa^R_i)\Big|_{pa^R}^{R} = pa^R_i\pi(pa^R_i)\Big|_{pa^R}^{R} - \int_{pa^R}^{R} \pi(pa^R_i)\delta pa^R_i$$

$$\bar{p}a_i\pi(0) - \bar{p}a_i(1-G(pa^R_i))^{N-1}] = pa^R_i[\pi(0)-(1-G(pa^R_i))^{N-1}]$$

$$- \int_{pa^R}^{R} (1-G(X))^{N-1}dX \quad \text{which is}$$

$$- \bar{p}a_i(1-G(pa^R_i))^{N-1}) = pa^R_i(-(1-G(pa^R_i))^{N-1}) - \int_0^{pa^R} (1-F(X))^{N-1}dX$$

by multiplying by -1 and dividing by $(1-G(pa^R_i))^{N-1}$ we finally get

$$\bar{p}a_i = pa^R_i + \left[\int_0^{pa^R} (1-G(X))^{N-1}dX/(1-G(pa^R_i))^{N-1} \right]$$

2. Proof of Monotonicity:

From our result we know that $\bar{p}b_i > pb^R_i$ and that the price quote is a function of the reservation price.

The first order condition of the optimisation problem is:

$[\delta\pi(pb^R_i)/\delta pb^R_i][\delta pb^R_i/\delta\bar{p}b_i](pb^R_i - \bar{p}b_i) - \pi(pb^R_i) = 0$

$\pi(pb^R_i) = Pr(pb^R_i > \max pb^R_j) = G(pb^R_i)^{N-1}$ and we define

$G(pb^R_i)^{N-1} = H(pb^R_i)$.

$[\delta\pi(pb^R_i)/\delta pb^R_i][\delta pb^R_i/\delta\bar{p}b_i] = [\delta H(pb^R_i)/\delta pb^R_i](1/\bar{p}b_i')$

with $\bar{p}b_i'$ being the derivative of $\bar{p}b_i$ with respect to pb^R_i we get

$[\delta H(pb^R_i)/\delta pb^R_i](pb^R_i - \bar{p}b_i)(1/\bar{p}b_i') + H(pb^R_i) = 0$ which is

$[\delta H(pb^R_i)/\delta pb^R_i](pb^R_i - \bar{p}b_i)(1/\bar{p}b_i') = - H(pb^R_i)$ after rearranging we get

$\bar{p}b_i' = - [\delta H(pb^R_i)/\delta pb^R_i]/H(pb^R_i)(pb^R_i - \bar{p}b_i)$

The sign of $\bar{p}b_i'$ is positive as the expression in the first bracket is positive and the expression in the last bracket is negative. Hence the price quote is an increasing function of the reservation price. Q.E.D.

3. Expectations in the Fragmented Market:

We prove that the expected prices in the fragmented market are the same as the average of the prices in a centralised market [11], i.e. that

$$E(pb^{R^*}_{N-1}) = E\left(pb^{R^*}_N - [\int_{-R}^{pb^{R^*}_N} G(x)^{N-1}dx]/G(pb^{R^*}_N)^{N-1}\right)$$

We define pb^{R^*} as the highest reservation price with the cumulative distribution function $H_N(.)$. We have $H_N(pb^R) = G(pb^R)^N$. The expected value of pb^R is $E(pb^{R^*}_N) = \int_{-R}^{R} pb^R \, d(G(pb^R)^N)$ (A1)

If we integrate by parts we get:

$$E(pb^{R^*}_N) = [pb^R G(pb^R)^N \Big|_{-R}^{R} - \int_{-R}^{R} G(pb^R)^N \, dpb^R$$

$$= R - \int_{-R}^{R} G(pb^R)^N \, dpb^R \qquad \text{(A2)}$$

The next best, i.e. the second best dealer is j = N-1. The c.d.f. of her reservation price is $H_{N-1}(.)$ such that $H_{N-1}(x) = NG(x)^{N-1} - (N-1)G(x)^N$.

The expectation of the second highest reservation price is

$E(pb^R_{N-1}) = \int_{-R}^{R} xdH_{N-1}(x)$.

Integrating by parts yields:

---

[11]We follow the proof of Biais pp. 76 ff.

$$E(pb^{R}_{N-1}) = [xH_{N-1}(x) \mid_{-R}^{R} - \int_{-R}^{R} H_{N-1}(x)dx \quad \text{which is}$$

$$E(pb^{R}_{N-1}) = R + \int_{-R}^{R} ((N-1)G(x)^{N} - NG(x)^{N-1})dx \qquad \text{(A3)}$$

This proof shall establish that:

$$E(pb^{R^*}_{N-1}) = E\left(pb^{R^*}_{N} - [\int_{-R}^{pb^{R^*}_N} G(x)^{N-1}dx]/G(pb^{R^*}_{N})^{N-1}\right) \qquad \text{(A4)}$$

We have:

$$E\left(- [\int_{-R}^{pb^{R^*}_N} G(x)^{N-1}dx]/G(pb^{R^*}_{N})^{N-1}\right)$$

$$= - [\int_{-R}^{pb^{R^*}_N} G(x)^{N-1}dx/G(pb^{R^*}_{N})^{N-1}](dG(pb^{R^*}_{N})^{N})$$

$$= - N \int_{-R}^{R} \left[\int_{-R}^{pb^{R^*}_N} G(x)^{N-1}dx\right] g(pb^{R^*}_{N})dpb^{R^*}_{N} \text{ with } dG^{N} = gG^{N-1}dpb^{R^*}_{N}$$

If we integrate by parts, this results in:

$$- N\left[(\int_{-R}^{pb^{R^*}_N} G(x)^{N-1}dx \; G(pb^{R^*}_{N})\mid_{-R}^{R} - \int_{-R}^{R} G(pb^{R^*}_{N})^{N-1} \; G(pb^{R^*}_{N})dpb^{R^*}_{N}\right]$$

which is

$$- N\left[\int_{-R}^{R} G(x)^{N-1}dx - \int_{-R}^{R} G(x)^{N}dx\right] \qquad \text{(A2a)}$$

Using (A2) and (A2a) we get:

$$E\left(pb^{R^*}_{N} - [\int_{-R}^{pb^{R^*}_N} G(x)^{N-1}dx]/G(pb^{R^*}_{N})^{N-1}\right) = R + \int_{-R}^{R}$$

$$\left[(N-1)F(x)^{N} - NF(x)^{N-1}\right]dx = E(pb^{R^*}_{N-1})$$

Q.E.D.

APPENDIX B

1. Calculation of the Bidding Price for a Divisible Good in a Fragmented Market:

The expected utility for the bid side is :

$EU = \sum_M [F(\alpha_i, M)[\alpha_i - \beta(Q/M) - \bar{p}b_i](Q/M)$ with $\alpha_i = \mu_i - \gamma\sigma^2 I_i$ and $\beta = \gamma\sigma^2$.

The optimisation problem is $\max_{\bar{p}b_i} EU^B$

The first order condition is:

$$\sum_M [(\delta F(\alpha_i,M)/\delta\alpha_i)(\delta\alpha_i/\delta\bar{p}b_i)[\alpha_i - \beta(Q/M) - \bar{p}b_i](1/M) + \sum_M F(\alpha_i,M)(1/M) = 0$$

$$\sum_M [(\delta F(\alpha_i,M)/\delta\alpha_i)(\delta\alpha_i/\delta\bar{p}b_i)(\alpha_i - \beta(Q/M)](1/M)$$
$$- \sum_M [(\delta F(\alpha_i,M)/\delta\alpha_i)(\delta\alpha_i/\delta\bar{p}b_i)\bar{p}b_i](1/M) - \sum_M F(\alpha_i,M)(1/M) = 0$$

by dividing through $(\delta\alpha_i/\delta\bar{p}b_i)$ we can rewrite it as

$$\sum_M [(\delta F(\alpha_i,M)/\delta\alpha_i)\bar{p}b_i](1/M) + \sum_M F(\alpha_i,M)(\delta\bar{p}b_i/\delta\alpha_i)(1/M) =$$
$$\sum_M [(\delta F(\alpha_i,M)/\delta\alpha_i)(\alpha_i - \beta(Q/M)](1/M) \text{ which is}$$

$$\sum_M (\delta/\delta\alpha_i)[F(\alpha_i,M)\bar{p}b_i](1/M) = \sum_M [(\delta F(\alpha_i,M)/\delta\alpha_i)(\alpha_i - \beta(Q/M)](1/M)$$

Next, we are integrating [12].

$$\sum_M [F(\alpha_i,M)\bar{p}b_i \Big|_0^{\alpha_i}](1/M) = \sum_M [\int_0^{\alpha_i} \alpha_i(\delta F(\alpha_i,M)/\delta\alpha b_i)d\alpha_i](1/M) + C$$
$$- \sum_M [(\delta F(\alpha_i,M)/\delta\alpha_i)\ \beta(Q/M)](1/M) \quad \text{(where C is a constant)}$$

If we integrate the first term on the RHS by parts we get

$$\sum_M [F(\alpha_i,M)\bar{p}b_i \Big|_0^{\alpha_i}](1/M) = \sum_M \alpha_i[F(\alpha_i,M) \Big|_0^{\alpha_i}](1/M)$$

[12] The integral is over $[0, \alpha_i]$.

$$- \sum_M [\int_0^{\alpha_i} F(X)\, dX](1/M) + C - \sum_M [F(\alpha_i,M) \Big|_0^{\alpha_i} \beta(Q/M)](1/M)$$

First we have to evaluate C at $\alpha_i = 0$:

$$\sum_M [\int_0^{\alpha_i} \alpha_i(\delta F(\alpha_i,M)/\delta\alpha b_i)d\alpha_i](1/M)$$

$$= \sum_M \alpha_i[F(\alpha_i,M) \Big|_0^{\alpha_i}](1/M) - \sum_M [\int_0^{\alpha_i} F(X)\, dX](1/M) + C$$ which is

$$\sum_M[F(\alpha_i,M)\alpha_i](1/M) = \sum_M[F(\alpha_i,M)\alpha_i](1/M)$$

$$- \sum_M[F(\alpha_i,M)\alpha_i](1/M) + C$$ which results in (with $F(0) = 0$)

$\sum_M (0)\alpha_i - (0)\alpha_i = 0$ and so is the RHS. Hence, also C is zero at $\alpha_i = 0$.

Thus, we have:

$$\sum_M [F(\alpha_i,M)\bar{p}b_i \Big|_0^{\alpha_i}](1/M) = \sum_M \alpha_i[F(\alpha_i,M) \Big|_0^{\alpha_i}](1/M)$$

$$- \sum_M [\int_0^{\alpha_i} F(X)\, dX](1/M) - \sum_M [F(\alpha_i,M) \Big|_0^{\alpha_i} \beta(Q/M)](1/M)$$

$$\sum_M [F(\alpha_i,M)\bar{p}b_i](1/M) = \sum_M [F(\alpha_i,M)(\alpha_i - \beta(Q/M)](1/M)$$

$$- \sum_M [\int_0^{\alpha_i} F(X)\, dX](1/M)$$

and finally: $\bar{p}b_i = \sum_M [F(\alpha_i,M)(\alpha_i - \beta(Q/M)](1/M)/\sum_M [F(\alpha_i,M)(1/M)$

$$- \sum_M [\int_0^{\alpha_i} F(X)\, dX](1/M)/\sum_M [F(\alpha_i,M)(1/M)$$

By following the same procedure, the ask price is:

$$\bar{p}a_i = \sum_M [(F_a(\alpha_i,M))(\alpha_i + \beta(Q/M)](1/M)/\sum_M [(F_a(\alpha_i,M))(1/M)$$

$$+ \sum_M [\int_{\alpha_i}^{R} (F_a(X))\, dX](1/M)/\sum_M [(F_a(\alpha_i,M))(1/M)$$

with $F_a(\alpha_i,M) = \sum_j \binom{N-1}{j} G(\alpha_i)^{N-1-j}(1-G(\alpha_i))^j$

2. Proof of the Monotonicity of the Bidding Function:

We conduct the proof for the bid side.

We rewrite the price quote of (27) as

$$\bar{p}b = \left[\sum_M F(\alpha,M)(1/M)pb^R\right] / \left[\sum_M F(\alpha,M)(1/M)\right]$$

$$+\left[\sum_M (1/M)\int_{-R}^{\alpha} F(\alpha,M)dy\right] / \left[\sum_M F(\alpha,M)(1/M)\right]$$

with M = 1, ... , N

The derivative of the price quote with respect to the reservation price, resp. α (with $pb^R = \alpha - \beta(Q/M)$) is

$$\delta\bar{p}b/\delta\alpha = \left[(\sum_M (1/M)(\delta F/\delta\alpha)pb^R)(\sum_M (1/M)F) - (\sum_M (1/M)(\delta F/\delta\alpha))(\sum_M (1/M)Fpb^R)\right]$$

$$/\left[(\sum_M (1/M)F)^2\right]$$

$$+\left[(\sum_M (1/M)\int_{-R}^{\alpha} F(y)dy)(\sum_M (1/M)(\delta F/\delta\alpha))\right] / \left[(\sum_M (1/M)F)^2\right] \qquad (A5)$$

with $F(\alpha,M) = \sum_j [(N-1)!/(N-1-j)!j!]G(\alpha)^j \left[1-G(\alpha)\right]^{N-1-j}$

and $F > 0$, $(\delta F/\delta\alpha) > 0$.

We want to prove that the price quote is an increasing function in the reservation price. The second term in (A5) is composed of probabilities only and it is evident that this expression is positive. Thus, it is sufficient to show that the first term of (A5) is zero or positive in order to establish the proof.

Next, we substitute the values of the probabilities into our expression which is

$$(\delta F/\delta\alpha) = \sum_j [(N-1)!/(N-1-j)!j!]gG(\alpha)^{j-1} \left[1-G(\alpha)\right]^{N-2-j}(j-(N-1)G)$$

where g is the derivative of G with respect to α.

$$(\delta F/\delta\alpha) = g\sum_j [(N-1)!/(N-1-j)!j!]j\ G(\alpha)^{j-1} \left[1-G(\alpha)\right]^{N-2-j} - g[(N-1)F]/(1-G)$$

We define $K_M = \sum_j [(N-1)!/(N-1-j)!j!]j\ G(\alpha)^{j-1} \left[1-G(\alpha)\right]^{N-2-j}$ and thus

$$(\delta F/\delta\alpha) = gK_M - g[(N-1)F]/(1-G) \tag{A6}$$

Now we use the numerator of the first term of (A5) (which we denote as F5) and (A6) and we get:

$$F5 = [\sum_M (1/M)(gK_M - g[(N-1)F]/(1-G))pb^R][\sum_M (1/M)F]$$

$$- [\sum_M (1/M)(gK_M - g[(N-1)F]/(1-G))][\sum_M (1/M)Fpb^R]$$

$$= [\sum_M (1/M)gK_M pb^R][\sum_M (1/M)F] - [\sum_M (1/M)g[(N-1)F/(1-G)]pb^R][\sum_M (1/M)F]$$

$$- [\sum_M (1/M)gK_M][\sum_M (1/M)Fpb^R] + [\sum_M (1/M)g[(N-1)F/(1-G)]][\sum_M (1/M)Fpb^R]$$

$$= [\sum_M (1/M)gK_M pb^R][\sum_M (1/M)F] - g[(N-1)/(1-G)]\sum_M (1/M)Fpb^R][\sum_M (1/M)F]$$

$$- [\sum_M (1/M)gK_M][\sum_M (1/M)Fpb^R] + g[(N-1)/(1-G)]\sum_M (1/M)Fpb^R][\sum_M (1/M)F]$$

$$F5 = [\sum_M (1/M)gK_M pb^R][\sum_M (1/M)F] - [\sum_M (1/M)gK_M][\sum_M (1/M)Fpb^R]$$

We can write K_M as:

$F = \sum_j \delta_j$ and $K_M = \sum_j \delta_j\ j[1/G(1-G)]$ and $pb^R = \alpha - \beta(Q/M)$

We make the respective substitutions and get:

$$F5 = \left[\sum_M (1/M)g\{\sum_j \delta_j\ j[1/G(1-G)\}[\alpha - \beta(Q/M)]\right][\sum_M (1/M)F]$$

$$- \left[\sum_M (1/M)g\{\sum_j \delta_j\ j[1/G(1-G)\}\right][\sum_M (1/M)F[\alpha - \beta(Q/M)]]$$

$$= [g\beta Q/G(1-G)]\Big[[\sum_M (1/M)\sum_j \delta_j j](\alpha - (1/M))[\sum_M (1/M)F]$$

$$- [\sum_M (1/M)\sum_j \delta_j j][\sum_M (1/M)F(\alpha - (1/M))]\Big]$$

$$= [g\beta Q/G(1-G)]\Big[[\sum_M (1/M)\sum_j \delta_j j][\sum_M (1/M^2)F] - [\sum_M (1/M^2)\sum_j \delta_j j][\sum_M (1/M)F]\Big]$$

We know that $F = \sum_j \delta_j$ which we substitute into above equation and this is

$$F5 = [g\beta Q/G(1-G)] \Big[[\sum_M(1/M)\sum_j \delta_j j][\sum_M \sum_j \delta_j(1/M^2)]$$

$$- [\sum_M(1/M^2)\sum_j \delta_j j][\sum_M (1/M)\sum_j \delta_j]\Big]$$

We can write the terms of the sequences as follows:

$$F5 = [g\beta Q/G(1-G)] \Big[[(1/1)\delta_{N-1}(N-1)+(1/2)(\delta_{N-2}(N-2)+\delta_{N-1}(N-1))+...]$$

$$[(1/1)\delta_{N-1}+(1/4)(\delta_{N-1}+ \delta_{N-2})+(1/9)(\delta_{N-1}+\delta_{N-2}+\delta_{N-3})+...]$$

$$-[(1/1)\delta_{N-1}(N-1)+(1/4)(\delta_{N-1}(N-1)+\delta_{N-2}(N-2))+(1/9)(\delta_{N-1}(N-1)+\delta_{N-2}(N-2))+..]$$

$$[(1/1)\delta_{N-1}+(1/2)(\delta_{N-1}+\delta_{N-2})+(1/3)(\delta_{N-1}+\delta_{N-2}+\delta_{N-3})+...]$$

$$= 0$$

Hence F5 equals zero and we have proven that the price function is monotonic and increasing in the reservation price for the bid side.

Q.E.D.

APPENDIX C

1. Comparison of the Integral Terms:

Comparison of the integral terms ("information cost term"):

The difference of the integral terms for the divisible good case and the indivisible good case can be written as

$\sum_{M=1}^{N} a_M / \sum_{M=1}^{N} b_M - a_1/b_1$. We want to show that this expression is < 0 which should be valid for both sides of the market.

We defined

$$\sum_{M=1}^{N} a_M / \sum_{M=1}^{N} b_M - a_1/b_1 = \left[\sum_{M=2}^{N} a_M / \sum_{M=2}^{N} b_M - a_1/b_1 \right] b_1 \sum_{M=2}^{N} b_M / b_1 \sum_{M=1}^{N} b_M$$

We have shown that a_M/b_M is decreasing on the bid side. On the ask side we get:

$$\delta/\delta j = \left[[\int_{\alpha}^{R} G(y)^{N-1-j}(1-G(y))^j]/[G(\alpha)^{N-1-j}(1-G(\alpha))^j] \right]$$

$$= \left[[\int_{\alpha}^{R} G(y)^{N-1-j}(1-G(y))^j][\ln[(1-G(y))/G(y)] - \ln[(1-G(\alpha))/G(\alpha)]] \right]$$

$$/[G(\alpha)^{N-1-j}(1-G(\alpha))^j] < 0$$

which means that the ratio is also decreasing on the ask side.

In the general case with M = 1, ..., N we have

$(a_1/b_1) > (a_2/b_2) > (a_3/b_3) > ...> (a_{N-1}/b_{N-1}) > (a_N/b_N)$.

We can now formulate the problem for M= 1, ..., N as follows:

$$\sum_{M=1}^{N} a_M / \sum_{M=1}^{N} b_M - a_1/b_1 = \left[\sum_{M=2}^{N} a_M / \sum_{M=2}^{N} b_M - a_1/b_1 \right] b_1 \sum_{M=2}^{N} b_M / b_1 \sum_{M=1}^{N} b_M$$

$$= \left[(\sum_{M=2}^{N} a_M / \sum_{M=2}^{N} b_M - a_2/b_2) + (a_2/b_2 - a_1/b_1) \right] \sum_{M=2}^{N} b_M / \sum_{M=1}^{N} b_M$$

$$= \left[\left(\sum_{M=3}^{N} a_M / \sum_{M=3}^{N} b_M - a_2/b_2 \right) \sum_{M=3}^{N} b_M / \sum_{M=2}^{N} b_M + (a_2/b_2 - a_1/b_1) \right] \sum_{M=2}^{N} b_M / \sum_{M=1}^{N} b_M$$

$$= \left(\sum_{M=3}^{N} a_M / \sum_{M=3}^{N} b_M - a_2/b_2 \right) \sum_{M=3}^{N} b_M / \sum_{M=1}^{N} b_M + (a_2/b_2 - a_1/b_1) \Big] \sum_{M=2}^{N} b_M / \sum_{M=1}^{N} b_M \text{ which}$$

is equivalent to:

$$= \left(\sum_{M=3}^{N} a_M / \sum_{M=3}^{N} b_M - a_3/b_3 \right) \sum_{M=3}^{N} b_M / \sum_{M=1}^{N} b_M + (a_3/b_3 - a_2/b_2) \sum_{M=3}^{N} b_M / \sum_{M=1}^{N} b_M$$

$$+ (a_2/b_2 - a_1/b_1) \Big] \sum_{M=2}^{N} b_M / \sum_{M=1}^{N} b_M \text{ and next we have}$$

$$= \left[\left(\sum_{M=4}^{N} a_M / \sum_{M=4}^{N} b_M - a_3/b_3 \right) \sum_{M=4}^{N} b_M / \sum_{M=3}^{N} b_M \right] \sum_{M=3}^{N} b_M / \sum_{M=1}^{N} b_M$$

$$+ (a_3/b_3 - a_2/b_2) \sum_{M=3}^{N} b_M / \sum_{M=1}^{N} b_M + (a_2/b_2 - a_1/b_1) \Big] \sum_{M=2}^{N} b_M / \sum_{M=1}^{N} b_M$$

which we can write as:

$$= \sum_{k=0}^{N-1} \left[(a_{N-k}/b_{N-k}) - (a_{N-k-1}/b_{N-k-1}) \right] \sum_{i=N-k}^{N} b_i \Big/ \sum_{k=1}^{N-1} b_k \, .$$

Knowing that the ratios a_M/b_M are decreasing we can determine the sign of the last expression which is smaller than zero. Hence we showed that our hypothesis is true also for the general case.

Q.E.D.

CHAPTER THREE

INTERACTION OF SPOT AND FUTURES MARKETS: THE THEORY OF STORAGE, FORECAST POWER, AND RISK PREMIUMS

3.1. Introduction

In recent years we have experienced a fast change in the markets whether these may be production or financial markets. Such changes of, for instance, market mechanisms or market structure were caused by changes in production technology, and the growing concern about the increasing risk of trading due to high price volatility. Information as such gets more and more important in order to remain competitive in this ever changing environment.

In general, in such markets, the risk is increased by the high price volatility and the need for insurance is growing.

The above scenario may be a reason of the existence and the importance of futures markets. Trading in futures started a long time ago primarily in agricultural commodities. Since then, the volume of trading in futures is a multiple of what it was at the beginning and there is a variety of commodities and markets for futures trading today. With the increase of trading in futures and the innovation of other financial instruments the financial market place gets more complicated and the interaction of the different markets whether spot, futures, or options, is quite complex.

An important part of the concern of people is the pervasive role of speculation. Nevertheless, it is accepted that futures market provide a hedging opportunity for the price risk of commodities [Cootner (1960), Houthakker (1968), Telser (1958) and others]. If we take speculation into account the cost of hedging may be affected by speculators in the market.

The question of whether the cost of hedging is increased or decreased with speculators in the market is not resolved yet.

Anderson and Danthine (1983) provide an interesting analysis of the role and influence of the various participants in the futures market and the influence of trading on the prices. Another argument is that of asymmetry of information in the market. The most prominent model in this respect is the model of Kyle (1985). Kumar and Seppi (1992) used Kyle's model and examined the problem of manipulating prices by traders who have superior information and who trade in the spot and the futures markets. Their results are that informed traders make positive profits. In addition, their findings are that manipulation generates liquidity transfers from the futures market to the cash market which benefit the informed traders and the spot noise traders.

However up-to-date, there exist a few papers only analysing the influence of futures trading on the spot market bid-ask spread. For example the paper by Holden (1990) which is referred to in Tuckman and Villa (1992). Very little is known about the effects or spill overs from other markets on the spot market bid-ask spread. The interaction of markets has not yet been analysed in respect of effects on trading strategies of market makers and the influence on the cost of trading. This will be the subject of the subsequent chapter.

This chapter contains an overview of all the relevant issues related to futures trading and the respective influence of it on the spot market prices. In section 3.2., we start with a historical summary of the evolution of futures markets and the early studies of the relation between futures and spot market prices (backwardation and contango).

It also includes a description of the various theories of the term structure of interest rates. Based on such early work, several theories have evolved which are described in the subsequent sections which are: in section 3.3. the theory of storage, in section 3.4. the concept of risk premium, and in section 3.5. the forecast power of prices.

The latter section gives the basis for the next theoretical chapter. In section 3.5.1., we analyse the model of Anderson and Danthine (1983) which explains the determinants of the spot market price considering the influence of futures trading.

We summarise some empirical studies about futures trading in section 3.6. and the conclusions are given in the final section 3.7. .

3.2. The Evolution of Futures Markets

The beginning of futures or better forward markets (as a forerunner of futures markets) has taken place with agricultural commodities.

Let us consider a merchant who ships grain abroad. She knows that the payment is made upon delivery of the commodity when the ship arrives let say in three months time. We also assume that the price for grain varies a lot and therefore, the merchant prefers to fix the price for the commodity already today rather than in three months time. She would even be prepared to lower the current price and sell the grain at a price below the current price in order to avoid the price risk involved of waiting until the delivery date.

On the other side, there may be a miller who wants to buy grain and she is willing to fix the price today for delivery of the commodity in three months time, again, the reason is to avoid any price risk. The instrument which has been created to meet such a demand and supply is a forward contract.

Such a forward contract is an agreement between a seller and a buyer that calls for the seller to deliver to the buyer a specified quantity and grade (quality) of an identified commodity, at a fixed time in the future, and at a price agreed to when the contract is first entered into.

It is, of course, not very likely that a seller (also called short hedger) and a buyer (called a long hedger) meet at the same time demanding and supplying the identical commodity for the same delivery date.
Therefore, a middleman takes the role of matching demand and supply by assuming the price risk and also the default risk. There is a third type of participant in the market, a speculator. The speculator expects to make a profit from the variation in the price. She is not interested in the physical delivery of the merchandise. She trades a forward contract today and offsets the position by trading the opposite contract at maturity.

3.2.1. Institutional Aspects

Forward markets have developed for several agricultural commodities, but for some commodities, the implementation of a forward market failed. The reason of such failure could be that the underlying commodity was not readily and continuously available to write enough forward contracts and to keep the market liquid. Another reason may be that the price variance was too small in order to create the need for risk layoff. In addition, with a small price variance, the speculator is not willing to participate in the market and hence, the market is less liquid.
In order to make the forward market more liquid and to facilitate the trading the forward contracts were standardized. These standardized contracts, called futures contracts, are traded at organized exchanges which are regulated. The contracts are of fixed size and they have standard maturity dates. The qualities of the commodity are agreed and standardized as well. Payments are made in form of margins which have to be paid when a contract is traded.
The middlemen (or brokers) do not have to deal with the default risk of all

the various market participants, because all the trades are done through a clearing house. The clearing house can be part of the futures exchange body or it can be a totally separate entity. All futures trades occurring at the exchange are reported to the clearing house.

Each "member firm" who is allowed to trade with the floor broker (who execute the order [1] at the exchange) has an account with the clearing house. After the trades are reported to the clearing house, each member firm is requested to pay the margin requirements according to the balance of the account. An initial margin has to be paid upon trading a futures contract and a variation margin has to be paid which is calculated and adjusted on a daily basis.

The settlement of a futures contract can be done in different ways. Among other possibilities there is the physical delivery of the commodity at maturity. Another way is "offsetting" which means that the liquidation of the open futures position is to effect an offsetting futures position which is the reverse of the initial transaction. The latter is the most common one today. The question of regulating the futures market is still not resolved, but it is agreed that due to the high leverage and risk in the market there could be notorious defaults by traders who are unable to fulfill their commitments. In order to create liquid markets futures exchanges are regulated. There are, for example, centralized trading in a limited number of contracts, and clearing associations guaranteeing contract performance (based on the system of margin requirements, capital requirements, and mark-to-market accounting procedures).

---

[1]There are different kinds of orders which can be given at an exchange. We do not discuss the differences of these orders as for our analysis we always consider the market order only.

3.2.2. The Imbalance of Hedging

In this section, we are concerned with the question of whether and why a difference arises between the current spot price at time T and the futures price of a maturing contract at time T (called the "basis"). Under the assumption of rational expectations of the market participants we would expect that in equilibrium the spot price at time T equals the futures price at T for a futures contract maturing at T.

In our example above, we described the situation of a merchant who is willing to sell the commodity today rather than to wait until the ship arrives in three months time. The only problem which exists is that the merchant has to find someone who is willing to buy exactly this quantity and at this quality level. Hence, the terms of the contracts have to be identical.

Therefore, let us assume that the participants in our example prefer to trade standardised contracts which are easier to trade. Hence, we consider the trades in a futures market rather than in a forward market.

It may be that there is an imbalance of the short hedgers and the long hedgers in the futures market [2] then the futures price will be different. One case is that if a speculator [3] expects that the current futures price will be higher in three months time compared to the futures price of today then she will buy a futures contract. If the expectation of the speculator is correct and we assume that the spot price remains about the same for

---

[2]Short hedgers commit to sell an asset at a future point in time whereas long hedgers agree to buy. Both carry inventories of the commodity.

[3]A speculator is defined to be a trader who does not hold any physical stock of the commodity and her interest is purely in price differences.

this period then we have the situation that in three months time the future spot price is below the futures price at that point in time (which is the price agreed between the speculator and the merchant). This situation where the futures price of a maturing contract is above the spot price is called **contango.**

In the other case the speculator will sell a futures contract if she expects the future futures price to be lower than the current futures price.

Hence, if the expectation of the speculator is correct the futures price of a maturing contract is below the future spot price at this future point in time. This situation is called **backwardation.**

Empirical studies about spot and futures prices reveal that, generally, the situation of backwardation is observed in the market.

One reason for this finding is that hedgers hold large inventories and therefore the short hedgers are predominant which results in backwardation.

Thus, under "normal" conditions the spot price is above the futures price because hedgers who are risk averse and hold large inventories would like to hedge their inventories, i.e. go short in futures.

To make it attractive for speculators to be long in futures the futures price has to be below the cash price. This fact has been recognized already by Keynes (1930) and has been pointed out by Hicks (1939) as follows:

> In "normal" conditions, when demand and supply conditions are expected to remain unchanged, and therefore the spot price is expected to be about the same in a month's time as it is today, the future price for one month's delivery is bound to be below the spot price now ruling. The difference between these two prices (the current spot price and the currently fixed futures price) is called by Mr Keynes 'normal backwardation'.
> (Hicks, 1939, pp. 138)

The existence of normal backwardation was examined further by Houthakker (1968). He examined the imbalance of hedging in detail.

He argues that in a market without any hedgers and only with speculators such a bias between spot and futures prices cannot exist. A bias arises where hedgers are in the market whose position in the futures market is balanced by an opposite position in the cash market.

In order to have a situation of normal backwardation, i.e. that the spot price is above the futures price, hedgers must be net short, so that speculators will be net long. But the fact that, in general, hedgers are net short in the futures market has first to be proven. The Keynes/Hicks theory does not give any satisfactory answer to that problem.

One argument is that the producer of a commodity needs to look much further ahead than the consumer and Hicks argued that the entrepreneur is less constrained by the acquisition of inputs than by the completion of the output. Hence, the hedge of planned purchases is less important than the hedge of planned sales.

However, this argument is not valid for merchants who are independent of any technological considerations and who are the middlemen in most futures markets.

Telser (1958) argued that competition and free market entry result in reducing the difference between spot and futures price, even bringing it down to zero upon expiration of the futures contract.

This statement has been criticized by Cootner (1960). He rejected the assumption of Telser that the "net open position" (X) of maturing futures of a speculator becomes infinite whenever the futures price falls below the expected spot price, i.e. $X \rightarrow \infty$ when $(p-\bar{p}')<0$ (p=futures price at time T, p'=weighted average spot price expected by speculators at time T).

By introducing time preferences of speculators the demand of speculators for futures would not be infinite anymore when the futures price falls below the spot price. The time preference acts as a transaction cost and speculators would trade only if the difference between the future spot and the future futures price is large enough to give them an adequate return which means that the return should be at least as high as the adequate rate of interest of an alternative investment.

This argument leads us to the phenomenon of the **term structure of interest rates.**

3.2.3. The Term Structure of Interest Rates

The term structure of interest rates describes the relationship between interest rates and loan maturity or in other words the relationship between the yield-to-maturity and the maturity of a given fixed-income security. The usual way of presentation of the term structure is by a plot of the yields on default free government securities with different terms to maturity, at a given moment in time. Another expression for this yield-maturity relationship is the **yield curve.**

From such a yield curve we can see, for instance, that the annual interest rate of a security is not the same for each year.

The level and the shape of a yield curve may change even from day to day which depends mainly on economic factors. Normally, a security with short term maturity carries a lower return than a security with long term maturity which results in an upward sloping yield curve. However, the yield curve may also be downward sloping or almost horizontal.

There exist three theories which explain the term structure or the relationship between the short term and the long term interest rates: the expectations theory, the liquidity preference theory, and the market segmentation theory.

The expectation theory :

The expectation theory seems to be the dominant theory of the term structure and it says that the expected futures interest rate on a long term bond is the same as the observed short term forward rates on a bond over the same period. Usually, it is assumed that the investors are risk neutral and that under this assumption, the outcome of investing in a long term bond is equivalent to investing in a short term bond which will be rolled over (renewed for another period) until the same maturity date is reached as under the long term bond investment. These two investment strategies should give the same return at the end of the period. However, due to uncertainty and fluctuation of the interest rates there arise arbitrage opportunities which are exploited by professional investors.

We give a simple example to illustrate the situation [4].

We assume that an investor has the choice to buy either a two year bond of £ 100 with a yield of 9 percent per year or a one year bond with a yield of 8 percent per year. She can then reinvest at the end of one year the £ 108 in another one year bond. The end of the period yields (Y) of the two cases are:

case one (two year bond): Y = £ 100 (1.09)(1.09) = £ 118.81.

case two : This yield depends on the investor's expected future rate on the one year bond for the second year denoted $E(r_2)$.

$Y = £\ 100\ (1.08)[1 + E(r_2)]$

According to the expectation theory $E(r_2)$ is:

$£118.81 = £\ 108[1 + E(r_2)]$ which is $1 + E(r_2) = 1.1001$ and we get

$E(r_2) = 0.1001 = 10.01\ \%$

[4]The example is taken from Weston and Copeland (1988)

Hence, the current future rate is used to infer the "future" forward rate (in our case the forward rate of the second year). If, for instance, the actual observed one year forward rate in the second year is 10.5 % the investor would be better off to invest in the one year bond than in the two year bond which gives a lower pay-off. Hence, the expectation theory predicts that market competition forces forward rates to be equal to expected future rates over the same period.

When the long term future rates are above the short term forward rates then we have an upward sloping yield curve and short term interest rates are expected to rise. The downward sloping yield curve implies that the futures rates are below the forward rates and thus, the short term interest rates are expected to fall. In reality, yield curves are very often upward sloping which is not explained by the expectation theory. We have to analyse the term structure within a different environment which is the done in the next description of the liquidity preference or liquidity premium theory.

The liquidity preference theory :

The expectation theory is modified by taking into account the uncertainty inherent in the future. With uncertainty the investor prefers to buy a short term bond rather than a long term bond. This can be explained by the fact that the short term bond is more liquid and the near future seems to be easier to predict than the long term future. Hence, a liquidity premium must be paid to the investor to induce her to buy a long term bond.

On the other hand, the borrower is interested in long term borrowing as the longer the period the lower the risk of having to repay at an unfavourable moment in time under averse conditions. Thus, the borrower is also prepared to pay a risk premium to the investor for long term bonds.

Therefore, the yield curve is not flat anymore under constant expected returns, but upward sloping. We can say that we have an upward biased yield curve under the expectation theory. However, it seems that there is still another influence on the term structure of interest rates which is the market segmentation hypothesis.

The market segmentation hypothesis :

This theory is also called the hedging pressure theory. The assumption implied in the expectation theory is that the investor is indifferent between the short term bond and the long term bond, except for any differences of expected yield based on maturity.

The liquidity preference theory assumes that investors prefer short term bonds and borrowers favour long term bonds due to the uncertainty involved.

The market segmentation theory argues that there exist some investors, for example insurance companies, who prefer long term investment due to their long term liabilities. Also in case of the borrowers, they adjust their borrowing requirements according to the maturity structure of their assets.

Thus, the market is segmented with participants who exhibit strong maturity preferences whith each maturity as a separate segment. The market segmentation theory implies that the bonds with different maturities are not substitutes and the expectations play no role.

Researchers agree that all these different theories are important to explain the term structure of interest rate.

We now turn to the analysis of the effects of the interaction of markets on the prices. Today, we can divide the theory about futures markets into three major categories which we just briefly summarize. We will analyze them in detail thereafter. One of them is the theory of storage. It analyses the role of the futures market which provides an opportunity to manage risk and hedge a commodity position which has been acquired in the cash market. The marginal cost of hedging includes the interest forgone, for the period between purchase and sale, the marginal warehousing cost, and the marginal convenience yield. This theory contradicts the notion of "normal backwardation". Hedgers are always prepared to reduce the risk of their inventory position by accepting that the current futures sale price is lower than the current spot sale price.

A different aspect of the futures market is that of risk shifting. Through futures markets hedgers are able to shift some of the risk involved of holding a position to speculators who are willing to take that risk. The futures price is a composite of the price expectation and a risk premium which the hedger pays to be able to shift the risk to the speculator.

Another view is the discovery role of futures prices which examines the forecast power of futures prices and how far futures prices improve the information available in the cash market. There exists a separate strand of papers examining the information aggregation process in a rational expectations equilibrium. However, we restrict our analysis to the determinants of cash and futures prices by examining the role of the different agents in the market, particularly the speculators who convey information into the market place.

3.3. The Theory of Storage

All concepts mentioned above try to explain the difference between the current spot price and the futures price, i.e. the spot price at time t and the maturing futures contract price at time t. Throughout our analysis, we assume, unless it is mentioned otherwise, that market participants are risk averse.

The early works about the theory of storage are Keynes (1930), Kaldor (1939), Hicks (1946), Working (1948), Brennan (1958), Telser (1958), and Cootner (1960).

The theory of storage predicts that the total return from a purchase of a commodity at t and selling it for delivery at T equals the interest forgone for the period (T-t), plus marginal storage cost, less the marginal convenience yield from an additional unit of inventory. We assume that the futures contract matures after one period. Hence we can write:

$$F(p_T^T) - S(p^T) = S(p^T)R + MS - C \qquad (1)$$

where:

$(F(p_T^T) - S(p^T))$:	total return
$S(p^T)R$:	interest forgone
MS:	marginal storage cost
C:	marginal convenience yield

and $F(p_T^T)$ = futures price at time T for delivery at T

$S(p^T)$ = spot price at time T

$F(p_T^T) - S(p^T)$ is the basis.

The convenience yield can be explained as being a return for holding inventory, for example, to cover unexpected demand.

The theory makes it evident that there is a negative relationship between the size of inventories and the marginal convenience yield.[5]

However, the theory of storage does not take into account the activities of the speculators in the market and therefore, does not explain the determinants of the futures price in an equilibrium.

---

[5]Studies of Brennan (1958) and Telser (1958) give some evidence about that for several agricultural commodities.

3.4. The Concept of Risk Premium

The existence of futures markets enables hedgers to manage their price risk. Together with the notion of risk premium goes the definition of the "expectation" of the future cash price. The expected spot price is formed by the expectations of each dealer about the cash price which will prevail at some point in the future. Our exposition about the expected spot price and the concept of risk premium is mainly based on the description of Edwards and Ma (1992).[6] The first major contribution in this area were Dusak (1973), and Breeden (1980).

As a starting point, we assume that there are no speculators in the market. The relationship between spot and futures price can be explained by examining the respective demand and supply functions of the short and long hedgers. As a reminder, short hedgers are traders who wish to sell futures (supply) and long hedgers wish to buy futures (demand).

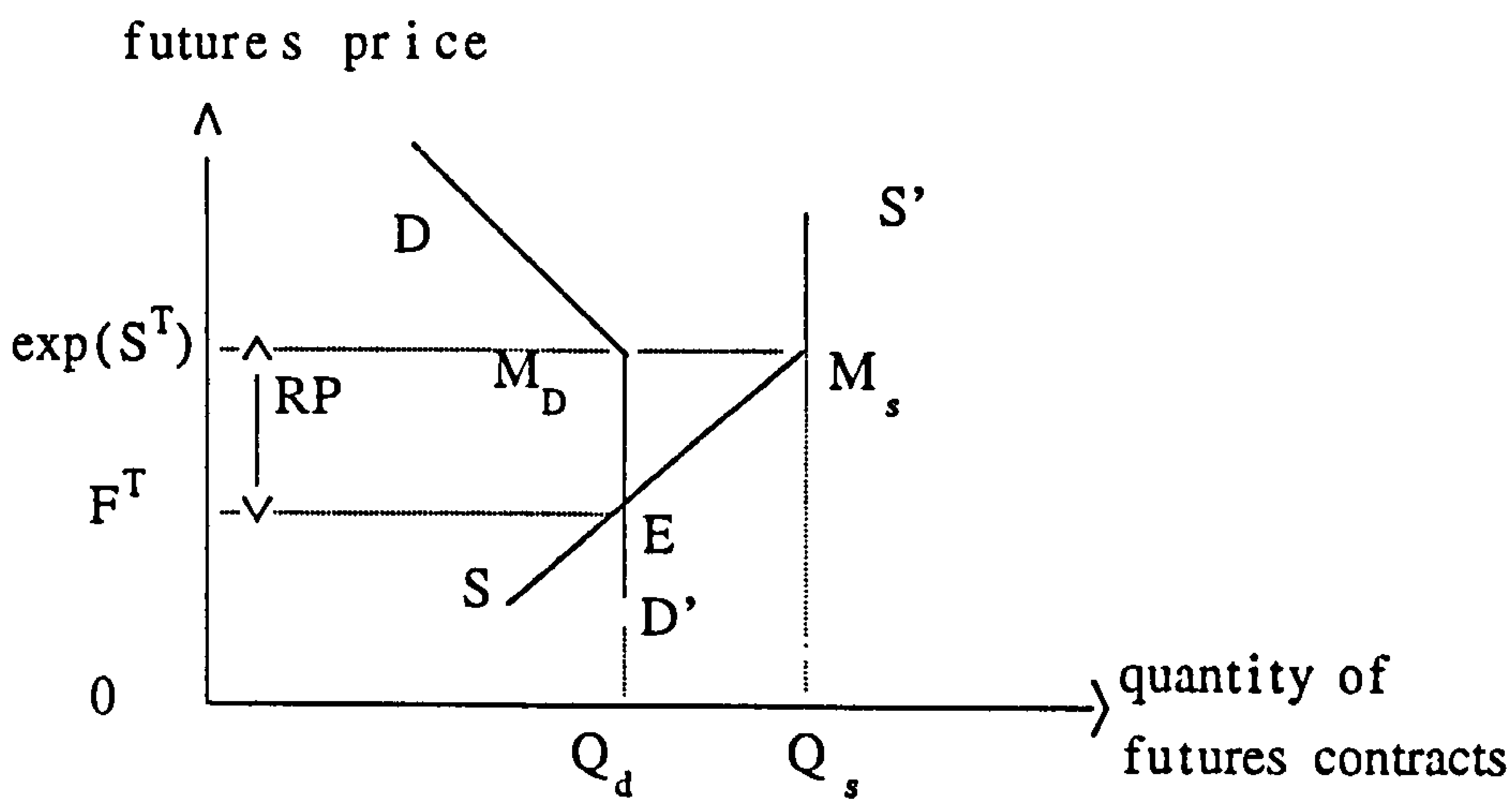

Figure 3.1. Net short hedging imbalance

In figure 3.1. an exogenously given net short hedging imbalance is illustrated which shows the demand (DD') and the supply (SS') with the

---

[6]pages 106-113

quantities of futures contracts on the X-axis and the futures prices on the Y-axis.

This figure depicts the situation of a net short hedging imbalance (Q_s-Q_d), where the supply of futures contracts (Q_s) exceeds the demand of futures contracts (Q_d). (Q_s) and (Q_d) are exogenously given.

The short hedgers total inventory is Q_s which they are willing to sell. However, if the futures price falls below the spot price they are more and more reluctant to sell which is depicted in M_sS.

On the other hand, long hedgers, with a total of future commitments of Q_D, purchase futures as long as the futures price is below the spot price. However, they are less willing to buy futures if the futures price is above the spot price which can be seen along line M_DD.

The equilibrium and thus the futures price is determined by the intersection of SS' and DD' at E where the futures price is below the spot price.

Figure 3.2. below illustrates the opposite case, a long hedging imbalance (Q_d-Q_s) with the futures price exceeding the expected price.[7]

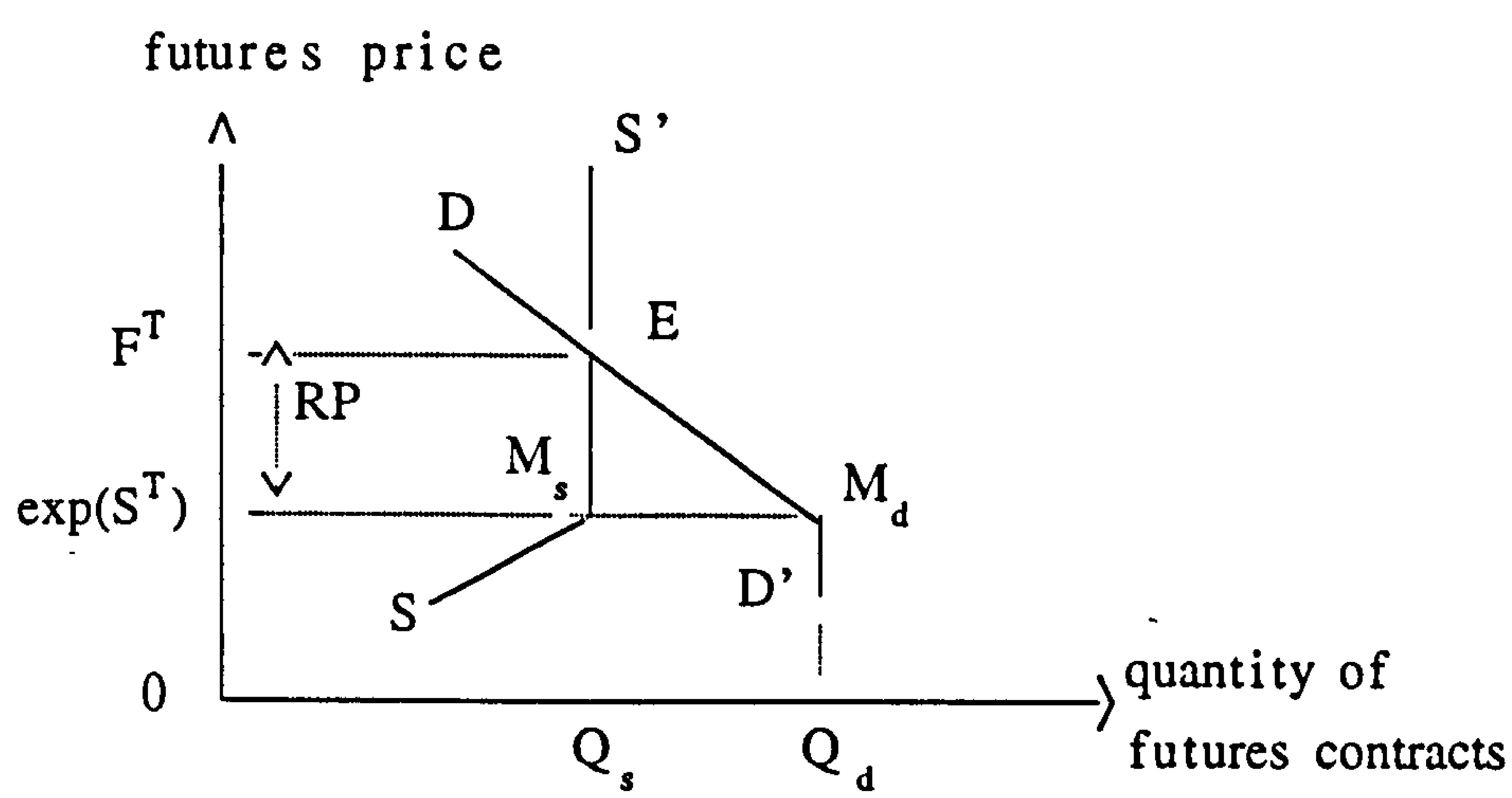

Figure 3.2. Net long hedging imbalance

---

[7]We still assume that there are no speculators.

As we can see, in both cases there exists a risk premium which either short hedgers or long hedgers have to pay. It is evident that the situation of no imbalance and zero risk premium is very unlikely, because the futures price must then be equal to the expected future spot price.

If we take into account that speculators are also in the market we can show that the risk premium decreases as, for example in the net short balance case, the demand function will be shifted from DD' to DD" (see figure 3.3.) by adding the demand of the speculators. The speculators enter the market when the futures price is below the expected spot price which results in an aggregation of the demand functions of the hedgers and the speculators below M_D from D' to D". Hence, the risk premium is reduced to RP'.

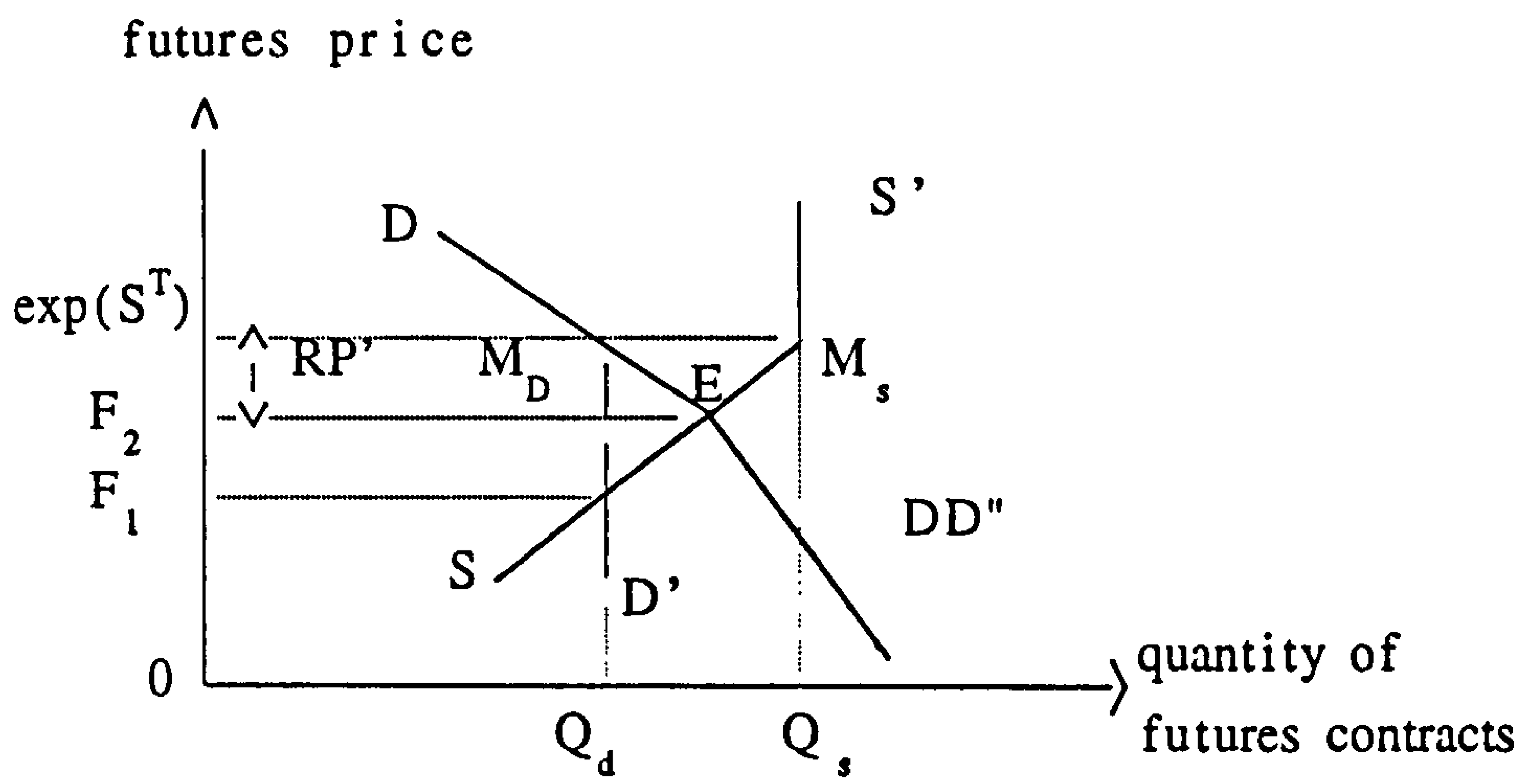

Figure 3.3. Short hedging imbalance with speculation

Similarly, the analogous process can be applied for the case of the net long hedging imbalance and the results is that the risk premium is reduced as well. The speculators believe that the expected price really occurs in the future, or better at some time in the future, and therefore, overall, they expect to make profits from their trading.

In contrast to the theory of storage, the concept of risk premium does account for the influence of speculation in the futures market. One drawback is that the analysis is based on an imbalance of hedging which is exogenously given. In our opinion, there are interactions between the determination of the futures price and the optimal futures positions of the short and long hedgers and thus, the imbalance of hedging should be treated as being endogenous.

3.4.1. Risk Premium and the Capital Asset Pricing Model

The existence of a risk premium has been analysed by Dusak (1973) within the framework of the capital asset pricing model (CAPM).

Her risk premium depends on the extent to which the variation in prices are systematically related to variations in the return of total wealth. Hence, an important determinant is the degree of correlation between the markets.

This is different from the earlier literature where the risk premium depends on the price variability only. The idea behind her model is that futures can be included in a portfolio like any other asset. Hence, in the CAPM framework, returns on any risky asset, including futures market assets, are governed by that asset's contribution to the risk of a large and well diversified portfolio of assets.

The basic formula is:

$$E(\tilde{R}_i) = R_f + \left[\frac{E(\tilde{R}_w - R_f)}{\sigma(\tilde{R}_w)}\right] \frac{\delta\sigma(\tilde{R}_w)}{\delta x_i} \qquad (2)$$

where:

$\tilde{R}_i$:	random rate of return on asset i
$E(\tilde{R}_i)$:	mathematical expectation of $\tilde{R}_i$
R_f :	pure time return to capital (riskless rate of interest)
$\tilde{R}_w$:	random rate of return on a portfolio

$E(\tilde{R}_w)$:	mathematical expectation of $\tilde{R}_w$
$\sigma(\tilde{R}_w)$:	standard deviation of $\tilde{R}_w$
x_i :	proportions of all the existing assets
$[\delta\sigma(\tilde{R}_w)]/\delta x_i$:	marginal contribution of asset i to the risk of the return of the portfolio which can also be expressed as $cov(R_i,R_w)/var(R_w)$

The influence of the risk factor can be shown in the following way:

$$\sigma(\tilde{R}_w) = \left[\sum_{i=1}^{N} \sum_{j=1} x_i x_j \, cov(\tilde{R}_i,\tilde{R}_j)\right]$$

$$\delta\sigma(\tilde{R}_w)/\delta x_i = \left[\sum_{j=1}^{N} x_j \, cov(\tilde{R}_i,\tilde{R}_j)\right] \text{ which equals}$$

$$1/\sigma(\tilde{R}_w)\left[x_i\sigma^2(\tilde{R}_i) + \sum_{j\neq i}^{N} x_j \, cov(\tilde{R}_i,\tilde{R}_j)\right]$$

with $\sum_{j\neq i}^{N} x_j \, cov(\tilde{R}_i,\tilde{R}_j)$ to be written as $cov(\tilde{R}_i,\tilde{R}_j)$.

Hence, (2) can be rewritten as

$$E(\tilde{R}_i) = R_f + \left[\frac{E(\tilde{R}_w - R_f)}{\sigma(\tilde{R}_w)}\right] \frac{cov(\tilde{R}_i,\tilde{R}_w)}{\sigma(\tilde{R}_w)}$$

or equivalent to the CAPM formulation:

$$E(\tilde{R}_i) - R_f = [E(\tilde{R}_w) - R_f]\beta_i \text{ where } \beta_i = cov(\tilde{R}_i,\tilde{R}_w)/\sigma^2(\tilde{R}_w).$$

Thus, in equilibrium, the expected rate of return on any asset i will be equal to the riskless rate of interest plus a risk premium proportional to the contribution of the asset to the risk of the return on the portfolio.

The crucial point is which asset to choose as a benchmark to the futures asset. The problem is that there is no capital investment in trading a futures contract (besides a margin which is relatively small) that could be interpreted as a "rate of return".

In the case of a futures contract, the return is the percentage change in the futures price. The corresponding return to $\tilde{R}_i$ is the return (net of storage cost) that a holder of a spot commodity would earn. Such a return includes the interest on the capital invested in the commodity plus a return (which could be positive or negative) over and above pure interest due to the unanticipated change in the price of the commodity. The expected return on any asset i can be expressed as:

$$E(\tilde{R}_i) = (1 - \beta_i)R_f + \beta_i E(\tilde{R}_w) \qquad (3)$$

where: $\beta_i = \mathrm{cov}(\tilde{R}_i,\tilde{R}_w)/\sigma^2(\tilde{R}_w)$

If we rewrite $E(\tilde{R}_i)$ in terms of period 0 and period 1 prices for the asset as

$E(\tilde{R}_i) = [E(\tilde{P}_{i,1}) - P_{i,0}]/P_{i,0}$, the equilibrium risk-return relation is:

$$P_{i,0} = \left[E(\tilde{P}_{i,1}) - [E(\tilde{R}_w) - R_f]P_{i,0}\beta_i\right]/(1 + R_f) \qquad (4)$$

It means that the current price of any asset (assuming no storage cost) i is the discounted value (at the riskless rate of interest) of its expected period one price. This value is adjusted downward for risk by the factor $[E(\tilde{R}_w) - R_f]P_{i,0}\beta_i$.

For a futures contract with no payment at time 0, but with a commitment for period 1, the current price for the futures is given by $P_{i,0}\ (1 + R_f)$ which means that the purchaser must pay a one-period interest rate of $P_{i,0}\ R_f$ in addition to $P_{i,0}$ (on a credit so to speak) because the transaction is made at time 0, but consummated at time 1.

If we multiply both side of equation (4) by $(1 + R_f)$ we get:

$$P_{i,0}\ (1 + R_f) = \left[E(\tilde{P}_{i,1}) - [E(\tilde{R}_w) - R_f]P_{i,0}\beta_i\right] \qquad (5)$$

which represents a futures contract where asset i refers to the spot commodity. We can see that the expression on the right hand side of (5) is the current futures price for delivery and payment of the spot commodity one period later, and $E(\tilde{P}_{i,1})$ can be seen as the spot price expected to prevail at time 1.

Setting $P_{f,0} = P_{i,0}$ (1 + R_f) for the futures value and rearranging terms, we get

$$[E(\tilde{P}_{i,1}) - P_{f,0}]/P_{i,0} = \beta_i[E(\tilde{R}_w) - R_f] \qquad (6)$$

Equation 6 states the risk premium on the spot commodity which is expressed as the deviation of the expected futures price from the current futures price divided by the period 0 spot price.

The essential point in Dusak's model is that buying a futures contract is like buying a capital asset on credit where the capital asset in this case is the spot commodity. By hedging a commodity position the holder converts the position into a riskless asset on which she earns the riskless rate R_f only.

On the other side, the speculator, who takes over the risk, does not invest any capital in the futures contract earns only the return over and above pure interest which is ($\tilde{R}_i$-R_f). Several studies support the view that speculation plays an important role in today's futures markets. Edwards and Ma show speculation as a percentage of open interest [8] in various markets.

For instance, there is speculation of 32.8 % at the Chicago Mercantile Exchange in Eurodollars, 53.5 % at the Chicago Board of Trade in oats, and 58 % at the New York Cotton Exchange in NYSE composite index.[9]

These findings lead us to the third theory about the futures prices, the interaction of markets and the influence of speculation on the market prices.

---

[8]Speculative open interest as a percentage of total month-end open interest: mean of monthly percentages over the year.

[9]Figures are taken from the table at pp. 466

3.5. Interaction of Markets and Forecast Power of Prices

Very often the term "the price discovery role of the futures markets" is used. Speculators are believed to provide information about the future cash price which will lead to an efficient market place. As we have shown above, speculators help to reduce the cost of hedging by reducing the risk premium.

At the same time, they convey information into the market as they speculate on the expected cash price, based on their private information. Thus, the price expectations are more informative with speculators than without them. Some researchers criticise that speculators do not base their expectations on the fundamentals of the commodity, but on other more short-term "chartist" facts. However, Froot et al. (1992) show that, although this argument may be true, prices are still more informative than without speculators in the long run.

The informational aspect includes a variety of different issues. One main area of research is the analysis of the information aggregation process of futures trading in a rational expectations equilibrium.

The most prominent papers are, Grossman (1976), Kyle (1985), and Bray (1981). The question to be answered in these models is whether speculation and noise trading in the futures market affects the efficiency of the financial markets.

We do not analyse these kind of models in detail as we are particularly interested in the determinants of the spot and futures prices. However, for reasons of completeness, we include a brief overview of the literature.

An important feature in these models is that, in equilibrium the prices are not fully revealing. Hence, an informed trader can make profits by trading in futures markets. These profits, of course, depend on how well the futures price interacts and predicts the expected spot price, i.e. it is important to what extent the markets are correlated. The correlation between the spot and futures prices leads us to another problem. That is whether futures trading stabilizes or destabilizes spot prices. This question has been studied, amongst others, by Danthine (1978), Turnovsky (1979), and Newbery (1987).

There is no clear answer to that. If we consider that speculators assume risks by trading futures the consequence is that the risk is reduced for hedgers and hence, the price variability in the spot market may decrease.

However, there exists a counterargument that hedgers, by being able to shift some risk, are less reluctant to undertake riskier transactions and thus increase the overall risk in the market and also the price variance.

On the other hand, if we look at the informational role of futures prices we can see that speculation may cause a high volatility in futures prices and, depending on the correlation of spot and futures market, consequently may induce higher volatility in the spot market.

There is an interesting paper of Kumar and Seppi (1992) regarding manipulation of prices. Their model combines various aspects mentioned earlier. They examine manipulation of prices by using a modification of Kyle's model (1985). In their two period model with asymmetric information amongst dealers, trade occurs first in the futures market and subsequently in the spot market.

Four types of investors are defined:

- a strategic risk-neutral informed trader
- a group of uninformed noise traders
- a group of risk-neutral floor traders and specialists who set competitive futures and spot prices

- a risk-neutral uninformed manipulator who *only if it is expected to be profitable* strategically submits an order in the futures market and in the spot market.

Manipulation is interpreted as a form of endogenous noise trading. Exogenous spot noise trading is not essential for the functioning of the spot market, but futures noise trading is needed. The reason is the same as in Kyle (1985) as futures noise trading is used as a camouflage for the speculators to be able to participate in the market without full information disclosure through prices.

Within such a setting, Kumar and Seppi show that uninformed traders are still in the position to make positive profits by establishing a futures position and then trading in the spot market to manipulate the spot price which will determine the cash settlement at delivery date.

Another aspect of their model is that manipulation transfers liquidity from futures to cash market which benefits the informed traders and the spot noise traders. With more manipulators, in the limit, profits from manipulation disappear, but price liquidity effects persist.

An interesting finding is that, with imperfect information linkage between spot and futures market, spot market traders are unlikely to observe all order related futures information which creates a temporary "price pressure" in the futures price. This effect is unique in the sense that previously identified factors such as market maker risk aversion or inventory control effects are absent here.

There are some studies which analyse precisely such factors and these are described in detail in the next section.

3.5.1. Determinants of Spot and Futures Prices

We are going to describe the paper of Anderson and Danthine (1983) in which they study the relationship between spot and futures prices in a "microstructure" setting. The open argument whether there exists a bias between spot and futures prices (a so called "basis risk", or risk premium) is examined in a rational expectations equilibrium. In other words, they analyse whether the Keyne's/Hicks argument of "normal backwardation" is valid.

Anderson and Danthine's model is elaborate and allows us to analyse different aspects (within the same model) which have been examined in previous papers, such as Stein (1979), Holthausen (1979), and Rolfo (1980), but not in such a general context. They allow for price and quantity risk (like Rolfo), but they consider the bias reflected in equilibrium futures prices by analysing the expected utility maximization problems of the individual producers and users of the good traded on the futures market.

There are three types of goods in the market: primary, secondary and final where the secondary good is traded in both markets, the spot and the futures market. In addition there are two trading dates. At time t, the futures trading occurs and the spot trading of the primary good.

At time (t+1), the secondary and the final goods are traded. The unknowns, at time t, are the time (t+1) supply of the secondary good and the prices of the secondary and the final goods.

The participants in the market place are speculators, producers of the secondary good (farmers), processors of the secondary good, and when the secondary good is storable, storage companies.

We can illustrate the structure of the model as follows:

t		t+1	
spot	futures	spot	spot
p.g.	i.g.	i.g.	f.g.
		(spot i.g. if the secondary good is storable)	

Figure 3.4.: Structure of the "Anderson and Danthine model"

where:

p.g. = primary good

i.g. = intermediate or secondary good

f.g. = final good

Next, we examine the role of each participant in the market and we make the assumption that there is no basis risk.

a) Speculator:

The speculator is not active in the spot market and hence, does not trade in the primary good. She only acquires a futures position at time t and closes out the position with an offsetting trade at time (t+1) in the secondary good. The net revenue function of the speculator is given by

$$\tilde{\pi}_s = (p^f - \tilde{p})f_s \qquad (7)$$

where:

p^f : futures price of the secondary good at time t

$\tilde{p}$: spot price of the secondary good at time (t+1)

f_s : number of futures contracts sold by the speculator at time t

If $f_s > 0$ the speculators is short in futures, and vice versa.

The speculator solves her maximization problem by choosing f_s, at time t, so as to maximize expected utility of net revenue $EU(\tilde{\pi}_s)$.

b) Producer (farmer):

The primary activity of the farmer is to purchase primary goods for input at time t and she then transforms them according to a production function which results in the secondary good which is available at time (t+1). The farmer may also trade in futures this means that, at time t, the farmer is active in the spot market and trades the primary good and she is also active in futures by trading the secondary good. At time t+1, the farmer sells the secondary good in the spot market. Hence, the net revenue function is

$$\tilde{\pi}_f = \tilde{p}g(x_f,\tilde{\varepsilon}) + rx_f + (p^f - \tilde{p})f_f \qquad (8)$$

where:

$\tilde{p}$: spot price of the intermediate good at time t+1

p^f : futures price of the intermediate good at time t

x_f : position size in the primary good ($x_f<0$ for a purchase)

r : price of the primary good

g() : production function which depends on the primary input and a production shock $\tilde{\varepsilon}$

f_f : number of futures contracts at time t

The farmer's maximization problem is to choose x_f and f_f so as to maximize $EU(\tilde{\pi}_f)$.

It is evident that the farmer faces price uncertainty ($\tilde{p}$) and quantity uncertainty ($\tilde{\varepsilon}$).

c) Processor:

The processor purchases the intermediate good and transforms it into the final good. She may also trade in futures. Thus, at time t, she trades in futures for the intermediate good and, at time t+1, she trades in the spot market for the secondary and the final good. Hence, the net revenue function is

$$\tilde{\pi}_p = \tilde{q}h(\tilde{y}_p) + \tilde{p}\tilde{y}_p + (p^f - \tilde{p})f_p \qquad (9)$$

where:

$\tilde{p}$: spot price of the intermediate good at time t+1

p^f : futures price of the intermediate good at time t

q : period (t+1) price of the final good

y_p : the processor's position in the intermediate good (with $y_p<0$ for purchases)

h() : processor's production function

The processor has to deal with uncertainty in both input and output. It is important to note that the processor's purchases of the intermediate good is a random variable at time t.

This implies that the processor may determine inputs at time (t+1) given knowledge of the intermediate good price. That can result in an asymmetry in the problems of the producer and the processor.

d) Storage Company:

When the intermediate good is storable the storage company can purchase the good at time t and carry it forward until (t+1). The storage company may also trade in futures. The net revenue function is given by

$$\tilde{\pi}_c = \tilde{p}c(i_c) - Rp_ti_c + (p^f - \tilde{p})f_c \qquad (10)$$

where:

p_t : spot price of the intermediate good at time t

R : one plus the interest rate

i_c : inventory of the intermediate good held by the storage company

c() : carry out function which results in the amount of the intermediate good brought forward to time (t+1) net of wastage and cost of storage

The company's maximization problem at time t is to choose f_c and i_c so as to maximize $EU(\tilde{\pi}_c)$.

From equations (7) to (10) we see that all participants except the speculator are hedgers who have a quantity commitment. However, the speculator's problem appears in each other participants optimization problem. Thus, each hedger has a speculative term and a hedging term in their utility function.

Anderson and Danthine's analysis defines a rational expectations equilibrium for three different market structures which are:

A: There is a perishable intermediate good (without any storage) and hedger's sales and purchases decisions are made under spot price risk (no quantity uncertainty). Processors are assumed to be inflexible in their inputs.

B: In this scenario, processors are flexible in their input decision and producers have to deal with additional quantity uncertainty.

C: A storage company comes into the market with the assumption that the intermediate good is storable, otherwise the condition under A apply.

Market Equilibrium A:

We assume that there is no production and no processing uncertainty in this equilibrium scenario. Processors are technologically constrained to choose input levels at the same time as their futures choice.

In equilibrium, demand equals supply and by assuming that the intermediate good is perishable we get the following market clearing condition:

$$n_f f_f + n_p f_p + n_s f_s = 0 \tag{11}$$

with n_f identical producers, n_p identical processors, and n_s identical speculators.

A result from the structure of the model (i.e.$\tilde{\pi}_s = (p^f - \tilde{p})f_s$) is that the speculator will go long in futures when she expects that the futures price $\tilde{p}$ will rise and she will go short in futures when she expects that the futures price will fall. This is obvious as the speculator does not have any interest in physical stock and therefore trades futures based on expected price differences which can be proved in the following way. The necessary condition for the optimal f_s is

$$EU'(\tilde{\pi}_s)(p^f - \tilde{p}) = 0 \qquad (12)$$

By knowing that the utility function is concave which gives

$EU''(\tilde{\pi}_s)(p^f - \tilde{p}) < 0$ above condition (12) is also sufficient.

Hence, the second order condition corresponding to (11) implies that $f_s \gtreqless f_s^*$ as $EU'[(p^f - \tilde{p})f_s] \gtreqless 0$ with f_s^* as the optimal futures position, and by having $f_s = 0$ as a reference we get the result:

$f_s^* \gtreqless 0$ if and only if $p^f \gtreqless E\tilde{p}$.

This result implies that speculators will sell futures when a futures price exceeds the expected cash price and, speculators will buy futures when the futures price is below the expected cash price.

Now we analyse the situation for the farmer.

The assumption of no basis risk simplifies the analysis in so far that the optimal output can be chosen depending on the input and futures prices only (there is no quantity risk, i.e. $\varepsilon=0$). Thus, the producer's expectations and risk aversion are not important.

We assume that U(.) and g(.) are strictly concave which implies that the following first order conditions are necessary and sufficient for the maximization of $EU(\pi_f)$, where π_f is given by (8):

$$EU'(\tilde{\pi}_f)[\tilde{p}g_1(x_f) + r] \geq 0 \qquad (=0 \text{ if } x_f < 0) \qquad (13)$$

$$EU'(\tilde{\pi}_f)(p^f - \tilde{p}) = 0 \qquad (14)$$

Since we assume that ε is non-stochastic and that there is no basis risk ($\tilde{p}=p^f$) we obtain (for an interior solution $x_f < 0$):

$$p^f g_1(x_f) + r = 0 \tag{15}$$

with $g_1 = \delta g/\delta x_f$

This means that the farmer determines her input level by equating the input price with the marginal revenue product evaluated at the futures price p^f, without depending on her degree of risk aversion or price expectation.

However, this separation result breaks down if we assume quantity uncertainty which will be shown in the next equilibrium case.

If we define x^*_f as the optimal input of the farmer and rewrite the revenue function we get

$$\pi_f = \pi^0_f + (p^f - \tilde{p})s_f \tag{16}$$

where $\pi^0_f = p^f g(x^*_f,\tilde{\varepsilon}) + rx^*_f$, $s_f = f_f - g(x^*_f,\tilde{\varepsilon})$.

s_f can be defined as the amount by which the producer's futures position differs from a fully hedged one; it is her speculative decision variable.

Given x^*_f the producer's futures position depends on the speculative term only and the optimal s_f is the solution to the maximization of $EU(\pi_f)$ where π_f is given in (16).

Now, we can see that we have the identical problem to the speculator's case except for the presence of a non-zero hedgeable net revenue.

As a natural corollary to the speculator's result above we get

$s_f \gtreqless 0$ if and only if $p^f \gtreqless E\tilde{p}$.

If we examine the expected utility of the processor we know that she must decide at time t about the input purchase at time (t+1). It turns out that, under this assumption of inflexibility in inputs, the processor's problem is very similar to the producer's so that hedging activity tends to be very symmetrical.

The maximum can be obtained for the processor by using (9):

$$EU'(\tilde{\pi}_p)[qh'(y_p) + \tilde{p}] \geq 0 \quad (=0 \text{ if } y_p < 0) \tag{17}$$

$$EU'(\tilde{\pi}_p)(p^f - \tilde{p}) = 0 \tag{18}$$

Also in this case, with q being non-stochastic, we have the separation result like in the farmer's case. (17) and (18) imply , for an interior solution:

$$-qh'(y_p) = p^f \tag{19}$$

It is clear that the input is determined independently from the futures position. We can write

$$\tilde{\pi}_p = qh(y^*_p) + p^f y^*_p + (p^f - \tilde{p})s_p \tag{20}$$

where y^*_p is the optimal purchase of the intermediate good and $s_p=f_p-y^*_p$ is the amount of deviation from the routine hedge.

This separation result means that the choice of the optimal deviations from the routine hedge is exactly identical to the speculators problem. Consequently, we obtain as a corollary to the speculator's result:

$s_p \gtreqless 0$ if and only if $p^f \gtreqless E\tilde{p}$.

Now, we can substitute these separation results into the equilibrium condition (11) and we get:

$$n_f F(p^f/r) + n_p P(p^f/q) = -(n_f s_f + n_p s_p + n_s f_s) \tag{21}$$

F() is the solution of (15) and P() is the solution of (19).

We can define the left hand side of (21) as the total net hedging (T.N.H.) and the expression in the parenthesis at the right hand side of (21) as the total net speculation (T.N.S.). This result and the assumption of homogeneous price expectations leads to "proposition 2":

T.N.S. $\gtreqless 0$ as $p^f \gtreqless Ep$ and T.N.H. $\lesseqgtr 0$ as $p^f \gtreqless Ep$

Anderson and Danthine discuss this result as follows:

The latter part of the proposition clearly links the position of the futures price relative to the expected cash price to the next excess of producer's output plans over processor's purchase plans. Traditionally, backwardation has been interpreted as the price to be paid for risk transfer. Proposition 2 that the aptness of this interpretation depends on the direction of the transfer. When total planned output exceeds total planned input an incentive is indeed needed to induce other agents (speculators) to commit themselves to receive this excess (net) planned output. The incentive is provided by a futures price below the expected cash price, i.e. backwardation. However, when total planned output falls short of total planned input, the incentive needed to induce speculators to be net short is a futures price in excess of the expected cash price, i.e. contango. Thus, backwardation is 'normal' only in markets where hedging activity is systematically dominated by suppliers of the commodity specified in the contracts (short hedgers).[10]

As a next step we assume rational expectations among participants which means that they have complete knowledge of next date's cash market structure. In this case, the cash price equilibrium condition for time (t+1) is:

$$n_f F(p^f/r) + n_p P(p^f/q) + D(p,\tau) = 0 \qquad (22)$$

where $D(p,\tau)$ is the external net supply function which is generated by , for instance, additional participants entering the market and τ representing a sort of a noise factor. It follows that the equilibrium cash price at t+1 depends upon p^f and τ:

$$p^* = C(p^f,\tau) \qquad (23)$$

It is assumed that, in period t, external demand is unknown and therefore τ and thus p^* are random variables. Furthermore, we assume that the cumulative distribution function of τ J(.) is known. Hence, for rational participants the distribution of p^* is fully specified, conditional on p^f, by equation (23) and J(.). The final result is summarized in:

$$n_f F(p^{*f}/r) + n_p P(p^{*f}/q) = - n_f s_f[p^{*f},C(p^{*f},.),J(.)] + n_p s_p[p^{*f},C(p^{*f},.),J(.)] + n_s f_s[p^{*f},C(p^{*f},.),J(.)] \qquad (24)$$

Based on the above result we can say that if there is no excess demand or excess supply of the secondary good no additional participant is needed for reallocation of the risks of the producers and the processors. Thus the futures market provides a costless opportunity to reduce risk for both the product and the input. If there is a excess demand of the intermediary good which is unhedged producers will have a planned output larger than a

---

[10]pp. 383/384

planned input of the processors. To induce speculators to come into the market to absorb the risk of the surplus a risk premium has to be paid which will lead to backwardation. Similarly, we find a risk premium in the market if there is excess supply which leads to a "contango situation" (i.e. that the futures price is above the expected cash price).

Market Equilibrium B:

Under the assumption of quantity uncertainty ($\varepsilon > 0$) and with X as the producer's coefficient of absolute risk aversion, the first order condition is:

$$E(\tilde{p}\tilde{g}_1) + r - X\left[cov(\tilde{p}\tilde{g},\tilde{p}\tilde{g}_1) - f_f cov(\tilde{p},\tilde{p}\tilde{g}_1)\right] = 0 \qquad (25)$$

where $f_f = \bar{y}_f + [cov(\tilde{p},\tilde{p}\tilde{y})/var\tilde{p}] + [(p_f - E\tilde{p})/Xvar\tilde{p}]$ (26)

with $\tilde{g} = g(x,\tilde{\varepsilon})$, $\tilde{g}_1 = g_1(x,\tilde{\varepsilon})$, $\bar{y}_f = Eg(x,\tilde{\varepsilon})$, and

$\tilde{y} = \tilde{g} - \bar{y}_f$.

By substitution of (26) into (25) and rewriting the result is:[11]

$$E(\tilde{p}\tilde{g}_1) + [(p_f - E\tilde{p})cov(\tilde{p},\tilde{p}\tilde{g}_1)/var\tilde{p}] + r$$

$$- \left[X[cov(\tilde{p}\tilde{g},\tilde{p}\tilde{g}_1)var\tilde{p} - cov(\tilde{p},\tilde{p}\tilde{g}_1)cov(\tilde{p},\tilde{p}\tilde{g})]/var\tilde{p}\right] = 0 \qquad (27)$$

This time, in contrast to the certainty case, we have an additional term, the last one on the left hand side which is a risk premium. Now, the farmer's risk aversion together with the price expectation plays a role in the decision making process. Hence, the separation result is no longer valid.

The situation of the processor changes with the assumption of input flexibility. The decision of y_p will be made at time (t+1) so as to maximize profits given prices p and q.

---

[11]Detailed calculation is given in appendix.

We define $y_p = P(p/q)$ where $P'>0$. It implies that at time t y_p,p, and q are random variables. Using (9) we get after rearranging:

$$f_p = [(p_f - E\tilde{p})/Xvar\tilde{p}] + y^f_p + \left[cov[\tilde{p},\tilde{p}(\tilde{y}_p - y^f_p)]/var\tilde{p}\right]$$

$$+ [cov(\tilde{p},\tilde{q}\tilde{z})/var\tilde{p}] \qquad (28)$$

where $y^f_p = P(p^f/\bar{q})$, $\bar{q} = E\tilde{q}$, and $z = h(y_p)$.

If we compare the hedging position of the processor (28) with the one of the farmer (26) we can see that the processor's hedging adjustment term is composed of two parts; namely one coming from the uncertainty of the input and the other from the uncertainty of the output.

Hence, we can write the market clearing condition as:

$$n_f\bar{y}_f + n_f[cov(\tilde{p},\tilde{p}\tilde{y})/var\tilde{p}] + n_p y^f_p$$
$$+ n_p\left[cov[\tilde{p},\tilde{p}(\tilde{y}_p - y^f_p)]/var\tilde{p}\right] + n_p[cov(\tilde{p},\tilde{q}\tilde{z})/var\tilde{p}]$$
$$= -(n_f s_f + n_p s_p + n_s f_s) \qquad (29)$$

where s_f and s_p are $(p_f - E\tilde{p})/Xvar\tilde{p}$.

This expression may be interpreted as in the previous equilibrium case T.N.H. = - T.N.S. with T.N.S. having the same qualitative properties as in the previous case, but not so T.N.H.

T.N.H. can no longer be regarded as simply the excess of planned supply over planned demand.

Expectations are important now, on the left hand side as well as on the right hand side of the market clearing condition.

Like in the previous case backwardation or contango arises this time caused either from an imbalance of plans or from an asymmetry in the adjustments or a mixture of the two.

For a rational expectations equilibrium, the distribution of p must reflect the period (t+1) cash market clearing condition,

$n_f y_f + n_p y_p = 0$ which implies a link between y_f and y_p:

$n_f\tilde{y}_f = - n_p\tilde{y}_p$ which in turn implies that

$n_f cov(\tilde{p},\tilde{p}\tilde{y}_f) = -n_p cov(\tilde{p},\tilde{p}\tilde{y}_p)$

Hence, the future market clearing condition becomes:

$$n_p[cov(\tilde{p},\tilde{q}\tilde{z})/var\tilde{p}] = -(n_f s_f + n_p s_p + n_s f_s) \quad (30)$$

This time, there is backwardation, martingale, or contango according to whether $cov(\tilde{p},\tilde{q}\tilde{z}) \gtreqless 0$.

Hence, with an asymmetrical hedging problem of processors, either backwardation or contango are possible in a rational expectations equilibrium. This finding contradicts the Keynes/Hicks argument of 'normal' backwardation in this kind of market structure.

Market Equilibrium C:

If there are n_c identical storage companies the clearing condition is:

$$n_f f_f + n_p f_p + n_c f_c + n_s f_s = 0 \quad (31)$$

Considering the assumptions made we can rewrite it as

$$n_f y_f + n_p y_p + n_c c(i_c) = -(n_f s_f + n_p s_p + n_c s_c + n_s f_s) \quad (32)$$

This expression is equivalent with T.N.H. = -T.N.S. under the market equilibrium A except that this time T.N.H. is the excess of output and storage plans over input plans.

In a rational expectations equilibrium the following condition holds:

$$n_f y_f + n_p y_p + n_c c(i_c) - n_c i_{c2} + D(p,\varepsilon) = 0 \quad (33)$$

Compared to (22) there are two additional terms coming from the storage company. Again $D(p,\varepsilon)$ represents external "unhedged" net supply. Also the storage company's future demand for the commodity $- n_c i_{c2}$ is unhedged today because it will be determined by the next period's futures-cash spread. This is part of the storage company's arbitrage activity which induces a tendency towards backwardation with $i_{c2} \geq 0$.

If $E[D(p,\varepsilon)] = 0$ there is a backwardation in a rational expectation equilibrium. This conclusion is in agreement with Stein (1979) who finds that, with large inventories, short hedging will dominate long hedging and backwardation will result. However, we have no basis risk in the model which actually makes the model trivial. So there is no real evidence of "normal backwardation".

3.6. Empirical Evidence

Several studies mentioned earlier have given empirical evidence. However, some of them are controversial because of the estimation techniques and the lack of adequate data available.

Dusak (1973) examined futures commodity contracts of wheat, corn, and soybeans and she finds that returns and portfolio risk are both close to zero during the sample period. This is the case, although variability in prices and hence risk is high.

Fama and French (1987) find that the theory of storage is supported by their analysis. In the same paper, they also try to find significant coefficients indicating a risk premium. They fail to produce a result which would confirm a positive risk premium.

In contrast, Yoo and Maddala (1991) test the hypothesis that large hedgers consistently lose money in the futures market. This means that they pay a risk premium to speculators who take the risk by providing the opportunity for hedging. They find that large speculators as a whole consistently make profit on the average which supports their hypothesis. That gives a reason why large speculators continue to stay in the market. It has to be noted that the analysis is based on aggregate figures and does not say anything about individual large speculator's profits.

3.7. Summary and Conclusions

Our task has been to give an overview of the theories of futures markets and of existing work about the interaction of spot and futures markets. Most of the studies are concerned with the relation between spot and futures prices, the so called backwardation and contango. In the first case, the current futures price is below the current spot price; whereas contango means that the current futures price is above the current spot price.

We have shown that there are three main lines of argument explaining this bias. These are the theory of storage, the concept of risk premium, and the informational role of futures prices, also called the price discovery role.

All these theories capture some or all of the aspects arising from trading in spot and futures markets. The outcome of the study of all these theories is that we can not say unambigously whether trading in futures and spot market results in the situation of backwardation or contango. The interaction of the markets is quite complex and there is another line of literature which focuses on the informational asymmetry among participants. These theories examine how information is integrated in the spot market price from trading in futures.

However, there is no consent whether, in a rational expectations equilibrium, information aggregation stabilizes or destabilizes asset prices. We did not discuss this part of the literature because it is not our primary concern and, in addition, it is worth an examination on its own.

The determinants of the spot prices are the current asset holdings of the traders plus their degree of risk aversion and their price expectations.

In respect of the bid and ask prices and consequently the spread, in the spot market the determinants have been studied by Ho and Stoll within a "microstructure of market making framework" for the spot market. However, their analysis is within one market only. Although they examine the determinants of the bid-ask spread in a model with two assets [Ho and Stoll (1983)] they do not find a different result which could mean that diversification into different assets does not affect the bid-ask spread of a market maker. This is not a convincing result as most of the assets are correlated and we expect an influence coming from diversification.

The analogous case for the spot and the futures market analysis is the paper of Anderson and Danthine (1983) which we discussed in detail in this chapter.

The main difference between the studies of Ho/Stoll, and Anderson/Danthine is that the latter does not examine the effects on the spot market bid-ask spread. The analysis focuses on the determination of the bias between spot and futures prices.

There does not exist any analysis which examines the influence of the interaction between spot and futures market and the effect on the bid-ask spread.

One question which will be interesting to examine is whether market makers are in the position to reinsure their inventory of the risky asset and at what cost. We would expect that the costs in the spot market would decrease as a consequence of this reinsurance.

This shortcoming will be our concern in the next chapter.

APPENDIX

Derivation of the result in equation (27):

$\pi_f = \tilde{p}\tilde{g}(x_f,\tilde{\varepsilon}) + rx_f + (p_f - \tilde{p})f_f$

$E(\pi_f) = E(\tilde{p}\tilde{g}) + E(p_f - \tilde{p})f_f$

$var(\pi_f) = var(\tilde{p}\tilde{g}) + var(\tilde{p})f_f^2 - 2cov(\tilde{p}\tilde{g},\tilde{p})f_f$

$EU(\pi_f) = E - X\ (var)$

The first order conditions are:

$\delta EU/\delta x_f = E(\tilde{p}\tilde{g}_1) + r - X\left[\delta var(\tilde{p}\tilde{g})/\delta x_f - 2f_f(\delta cov(\tilde{p}\tilde{g},\tilde{p})/\delta x_f)\right] = 0$

with $cov(\tilde{p},\tilde{p}\tilde{Y}) = cov(\tilde{p},\tilde{p}(\tilde{g} - E\tilde{g})) = cov(\tilde{p},\tilde{p}\tilde{g}) - E(\tilde{g})cov(\tilde{p},\tilde{p})$

$= cov(\tilde{p},\tilde{p}\tilde{g}) - E(\tilde{g})var(\tilde{p})$

We can also write:

$var(\tilde{p}\tilde{g}) = E[(\tilde{p}\tilde{g}) - E(\tilde{p}\tilde{g})]^2 = \int[pg - E(pg)]^2 dF$ and

$\delta var(\tilde{p}\tilde{g})/\delta x_f = \int 2[pg - E(pg)][pg_1 - (\delta E(pg)/\delta x_f)]df$

$= \int[2p^2gg_1 - 2(E(pg))pg_1 - 2pg(\delta E(pg)/\delta x_f) + 2E(pg)(\delta E(pg)/\delta x_f]dF$

$= 2\int[p^2gg_1 - E(pg)pg_1]dF$

whereas: $cov(\tilde{p}\tilde{g},\tilde{p}\tilde{g}_1) = E\left[[\tilde{p}\tilde{g} - E(\tilde{p}\tilde{g})][\tilde{p}\tilde{g}_1 - E(\tilde{p}\tilde{g}_1)]\right]$

$= E[\tilde{p}\tilde{g} - E(\tilde{p}\tilde{g})]\tilde{p}\tilde{g}_1$

which is $\delta var(\tilde{p}\tilde{g})/\delta x_f = cov(\tilde{p}\tilde{g},\tilde{p}\tilde{g}_1) = E[\tilde{p}\tilde{g} - E(\tilde{p}\tilde{g})]\tilde{p}\tilde{g}_1$.

On the other side we have:

$\delta EU(\pi_f)/\delta f_f = E(\tilde{p}\tilde{g}_1) + r - X2f_f var(\tilde{p}) + X2cov(\tilde{p}\tilde{g},\tilde{p}) = 0$

$f_f = (cov(\tilde{p}\tilde{g},\tilde{p})/var(\tilde{p})) + ((p_f - E\tilde{p})/2Xvar(\tilde{p}))$

We want to show that this is equal to:

$E\tilde{g} + [cov(\tilde{p},\tilde{p}(\tilde{g} - E\tilde{g}))/var(\tilde{p})] + ((p_f - E\tilde{p})/Xvar(\tilde{p}))$

which is : $E\tilde{g} + [cov(\tilde{p},\tilde{p}\tilde{g}) - E\tilde{g}var(\tilde{p})]/var(\tilde{p})$

$= E\tilde{g} + cov(\tilde{p},\tilde{p}\tilde{g})/var(\tilde{p}) - E\tilde{g}var(\tilde{p})/var(\tilde{p})$

$= (cov(\tilde{p}\tilde{g},\tilde{p})/var(\tilde{p})) + ((p_f - E\tilde{p})/Xvar(\tilde{p}))$

Q.E.D.

CHAPTER FOUR

THE BID - ASK SPREAD AND FUTURES TRADING

4.1. Introduction

The aim of this chapter is to account for the shortfall in the theory about the interaction of financial markets with particular concern of the effects of such interaction on the bid-ask spread. The traditional bid-ask spread theory, for instance Ho and Stoll [HS] (1983), describe the determinants of the bid-ask spread within one market. Although HS examine the case with two risky stocks, their result shows that the spread is independent of the inventory (within a one period horizon) and hence, the spread is not affected by any diversification into several stocks as the spread is only changed by the deviation from the optimal inventory level. The deviation from the optimal inventory level is caused by the fact that dealers have to trade at their quoted prices if an order arrives. Such an inventory level is composed of different assets. As long as the transactions of the various assets are independent the spread does not change. Their model is based on an order arrival rate which is the same for each stock and the transactions (i.e. order arrivals) are independent. HS point out that the result of no influence of diversification on the bid-ask spread is no longer true if the transactions of the stocks are dependent.

Our bid-ask spread model includes trading in the spot and the futures market at the same time. Up-to-date, futures trading has been investigated along the lines which we described in the previous chapter. The most recent studies deal with the informational aspect, i.e. the discovery role of prices. Although this line of argument is, without any doubt, the most important one in the primary markets today, we like to analyse the effects of futures trading on the spot market bid-ask spread which is based on an analysis similar to Anderson and Danthine (1983), which we described in some length in the literature survey chapter.

They investigate futures trading decisions for various types of participants and they show the effects of futures trading on the spot market prices in a rational expectations equilibrium.

In our opinion, the investigation of the spot market bid-ask spread has been neglected. Especially in dealership markets, the cost of market making, i.e. the bid-ask spread, may reveal the influence of futures trading and, in general, may give some insight into the interaction of spot markets with futures markets. Furthermore, by applying a "microstructure" model in the spot market, identical to the model in a centralized market with full information for an indivisible good, like the HS model, we are able to show the determinants of the bid-ask spread of traders who also trade in futures. This kind of model is analysed in section 4.2. .

Overall, assuming no asymmetry of information among market makers, we expect that the bid-ask spread in the spot market decreases by trading in futures. This is based on the argument that by trading in futures the market maker who holds an inventory of the asset in the spot market is able to reduce the risk inherent in the price uncertainty of the asset.

Under the assumption of risk averse traders and the fact that they can reduce the risk by trading in futures, and thus hedge the risky position, the market maker is able to narrow the spread.

In section 4.3., we then change the structure of the model by assuming that the market maker in the spot market does not know the order flow and hence, the kind of transaction she enters into. Such a situation represents a dealership market. Dealers faces not only the price uncertainty, but also the transaction uncertainty. The latter kind of uncertainty leads us to considering two possible subcases.

First, we assume that the dealer knows that there is symmetry in the quantity demanded and quantity supplied. Second, we change the assumption of symmetry and we allow for differences in demand and supply. For both subcases, we derive the influence of the futures trading on the spot bid-ask spread.

Furthermore, in section 4.4., based on the symmetric model in section 4.3., we investigate whether and how the dealer's decision about the optimal futures position changes if there is the possibility of adjustment of the futures position in a subsequent second period. Thus, we analyse the intertemporal effects of futures trading on the spot bid-ask spread in a two period model. However, for simplicity, the dealer trades in the futures market only in the second period which means that after having learnt about the kind of spot market transaction (whether it is a purchase or a sale) she adjusts the futures position accordingly.

Finally, section 4.5. contains the summary of the various results and the conclusions.

4.2. The Features of a Model without Quantity Uncertainty in the Spot Market Trading

4.2.1. The Spot Market Structure

The spot market is modeled as a dealership market with Bertrand oligopolistic competition among market makers which means that all trade is at the "best" price. The market maker with the lowest ask quote gets all the sell orders and the market maker with the highest bid quote gets all the buy orders.

The dealer's initial wealth (V), consists of a cash position (c) and the holding of a risky asset (I):

$$V_t = c_t + p_t I_t \quad (1)$$

Market makers are risk averse and we assume mean variance utility. [1]

Dealers maximize expected utility of terminal wealth which is:

$$V = V_{t+1} = c_{t+1} + p_{t+1}I_{t+1} \quad (2)$$

Expected utility:

$$E(U) = E(V) - k/2(VAR(V)) \quad (3)$$

where

k: coefficient of risk aversion

U(.): Von Neumann-Morgenstern utility function

In the spot market, the dealer faces uncertainty which is the price uncertainty.

For the spot trading, we assume a probability λ that a buy order arrives in the market in one period. An equal probability of λ is assumed for the arrival of a sell order. The respective probability that no order arrives is $(1-2\lambda)$.

However, we assume that the dealers know about the public orders before they have to quote their prices and, in addition, we assume that the dealers know each others' reservation prices. Thus, we base on the same structure as the one of the centralised market model with trading in a divisible good, which is described in chapter two above.

All orders that arrive are of a fixed size X.

The market maker sets her bid and ask prices knowing the current futures price. The dealer optimizes the futures position (in our case N), after learning the kind of order, i.e whether it is a sale or a purchase, in period t, given the current cash (C_t) and current inventory consisting of a risky asset (I_t). Then, the dealer, taking into account the futures

---

[1]This could be justified as an approximation of a Taylor's series expansion of the utility of a risk averse dealer with constant absolute risk aversion. We also assume that the utility is normally distributed.

trading, sets her bid-ask spread accordingly.

We can illustrate this sequence of events as follows:

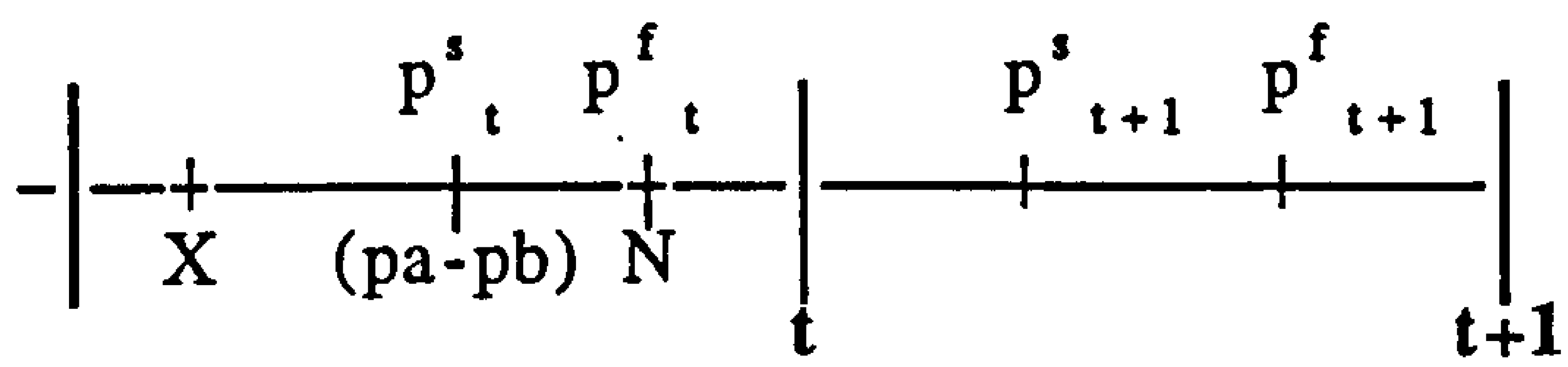

Figure 4.1.: Sequence of events with known X in period t

Based on above assumptions we can define the expected utilities for the dealers.

4.2.2 Determination of the Spot Bid-Ask Spread with Futures Trading

For simplicity, and to facilitate the analysis, we consider the case where the same dealer is in the position to quote the best price whichever order arrives. We call her a monopoly dealer. The dealer maximizes expected utility of terminal wealth.

The dealer is willing to trade if there is a profit making opportunity; that is if her terminal expected utility is at least equal to her initial utility:

$$E(U)_1 \geq E(U)_0 \qquad (4)$$

The terminal wealth for a dealer can be written as:

$W_1 = V +$ futures which is:

$$W_1 = (1+r)c_{t+1} + p^s_{t+1} I_{t+1} + (p^f_{t+1} - p^f_t)N \qquad (5)$$

with r = risk free interest rate; c_{t+1} = cash holding in period t+1; p^s_{t+1} = spot asset price at t+1; I_{t+1} = the inventory of the risky asset at t+1; p^f_{t+1} = futures price at t+1; p^f_t = current futures price at t; N = net purchase futures position committed at t and maturing at t+1.

We will write the subscripts t+1 as 1 and t as 0.

Furthermore, we define pa as the dealer's ask price quote, pb as the dealer's bid price quote, and X as the fixed order size.

Hence for a spot purchase we can write:

$$E(W_1) = (c_0\text{-}pbX) + \mu_s(I_0+X) + (\mu_f\text{-}p^f_0)N \qquad (6)$$

where $\mu_s = E(p^s_1)$ and $\mu_f = E(p^f_1)$ and

$$V(W_1) = V(p^s_1)(I_0+X)^2 + V(p^f_1\text{-}p^f_0)N^2$$
$$+ 2\text{cov}[p^s_1,(p^f_1\text{-}p^f_0)](I_0+X)N \qquad (7)$$

We define: $V(p^s_1)=\sigma_{p_s}^2$. As p^f_0 is known the variance of $(p^f_1\text{-}p^f_0)$ is $V(p^f_1)=\sigma_{p_f}^2$.

Consequently, expected utility coming from a spot purchase of X at pb is

$$EU_B(W_1) = (c_0\text{-}pbX) + \mu_s(I_0+X) + (\mu_f\text{-}p^f_0)N$$
$$-k/2\left[\sigma_{p_s}^2(I_0+X)^2+\sigma_{p_f}^2N^2+2\text{cov}[p^s_1,(p^f_1\text{-}p^f_0)](I_0+X)N\right] \qquad (8)$$

The optimal futures position is found by maximizing expected utility over N:

$$\delta EU_B(W_1)/\delta N = (\mu_f\text{-}p^f_0) - k\,\sigma_{p_f}^2N - k\text{cov}()(I_0+X)=0$$

where $\text{cov}() = \text{cov}[p^s_1,(p^f_1\text{-}p^f_0)]$ from equation (8)

Thus, $N_B = (\mu_f\text{-}p^f_0)/k\sigma_{p_f}^2 - \text{cov}()(I_0+X)/\sigma_{p_f}^2$ (9)

The "no trade" expected utility $EU(W_0)$ includes the initial wealth of the dealer plus an optimal trading position in futures:

$$EU(W_0) = c_0 + \mu_s(I_0) - (k/2)\,\sigma_{p_s}^2(I_0)^2$$
$$+ 1/2[(\mu_f\text{-}p^f_0)^2/k\sigma_{p_f}^2] - (\mu_f\text{-}p^f_0)\text{cov}()(I_0)/\sigma_{p_f}^2$$
$$+ (1/2)\,k\text{cov}()^2(I_0)^2/\sigma_{p_f}^2 \qquad (10)$$

Similar to Anderson and Danthine (1983) we have two different terms in N; the first one on the RHS of (9) is a pure speculative term, and the second one is a hedging term. By substituting the optimal N back into (8) we get an expected utility of a buyer which is: [2]

$$EU_B(W_1) = (c_0\text{-}pbX) + \mu_s(I_0+X) - (k/2)\,\sigma_{p_s}^{\ 2}(I_0+X)^2 + 1/2[(\mu_f\text{-}p^f_0)^2/k\sigma_{p_f}^{\ 2}]$$
$$- (\mu_f\text{-}p^f_0)\mathrm{cov}()(I_0+X)/\sigma_{p_f}^{\ 2} + (1/2)\ k\mathrm{cov}()^2(I_0+X)^2/\sigma_{p_f}^{\ 2} \qquad (11)$$

If we set $EU_B(W_1) = EU(W_0)$ we obtain the reservation bid price which is

$$pb = \mu_s - k\ \sigma_{p_s}^{\ 2}[I_0 + (1/2)X] - (\mu_f\text{-}p^f_0)\mathrm{cov}()/\sigma_{p_f}^{\ 2}$$
$$+ k\mathrm{cov}()^2[I_0 + (1/2)X]/\sigma_{p_f}^{\ 2} \qquad (11a)$$

The first two expressions on the RHS of (11a) come from the spot market activity. The other terms on the RHS come from futures trading. Here again, we can distinguish a pure speculative term (the third), and a hedging term (the fourth).

Similarly, the expected utility of a spot seller is:

$$EU_S(W_1) = (c_0+paX) + \mu_s(I_0\text{-}X) + (\mu_f\text{-}p^f_0)N$$
$$- k/2\left[\sigma_{p_s}^{\ 2}(I_0\text{-}X)^2+\sigma_{p_f}^{\ 2}N^2+2\mathrm{cov}[p^s_1,(p^f_1\text{-}p^f_0)](I_0\text{-}X)N\right] \qquad (12)$$

Again, if we maximize expected utility over N the optimal futures position is:

$$N_S = (\mu_f\text{-}p^f_0)/k\sigma_{p_f}^{\ 2} - \mathrm{cov}()(I_0\text{-}X)/\sigma_{p_f}^{\ 2} \qquad (13)$$

where $\mathrm{cov}() = \mathrm{cov}[p^s_1,(p^f_1\text{-}p^f_0)]$ from equation (12)

By substituting the optimal futures position into expected utility we get:[3]

$$EU_S(W_1) = (c_0+paX) + \mu_s(I_0\text{-}X) - (k/2)\ \sigma_{p_s}^{\ 2}(I_0\text{-}X)^2$$
$$+ 1/2[(\mu_f\text{-}p^f_0)^2/k\sigma_{p_f}^{\ 2}] - (\mu_f\text{-}p^f_0)\mathrm{cov}()(I_0\text{-}X)/\sigma_{p_f}^{\ 2}$$
$$+ (1/2)\ k\mathrm{cov}()^2(I_0\text{-}X)^2/\sigma_{p_f}^{\ 2} \qquad (14)$$

---

[2]Detailed calculation can be found in appendix A

[3]Details see appendix A

By setting $EU_S(W_1) = EU(W_0)$ we get

$$pa = \mu_s - k\sigma_{p_s}^2[I_0 - (1/2)X] - (\mu_f - p_0^f)cov()/\sigma_{p_f}^2$$
$$+ kcov()^2[I_0 - (1/2)X]/\sigma_{p_f}^2 \quad (14a)$$

In order to calculate the reservation spread we have

$$EU(W_1) = EU_S(W_1) + EU_B(W_1) \quad (15)$$

By substituting (10) and (13) into (14) we get [4]

$$EU(W_1) = (pa-pb)X - k\sigma_{p_s}^2X^2 + kcov()^2X^2/\sigma_{p_f}^2$$
$$+ c_0+\mu_s(I_0)-(k/2)\sigma_{p_s}^2(I_0)^2 + 1/2[(\mu_f-p_0^f)^2/k\sigma_{p_f}^2]$$
$$- (\mu_f-p_0^f)cov()(I_0)/\sigma_{p_f}^2 + 1/2\ kcov()^2(I_0)^2/\sigma_{p_f}^2 \quad (16)$$

According to our assumption that $E(U)_1 \geq E(U)_0$ the spread equation is:

$$(pa-pb) = k\sigma^2_{p_s}X - kcov()^2X/\sigma^2_{p_f} \quad (17)$$

which we can express as:

$$(pa-pb) = kX\sigma^2_{p_s}\left[1 - \rho\right] \quad (18)$$

with $\rho = cov()^2/\sigma_{p_s}^2\sigma_{p_f}^2$

This spread is a kind of counterfactual spread as the dealer quotes only one side of the market at the time (under the assumption that the dealer knows the order at the time she has to quote her price).

In this situation, we see that the bid-ask spread depends on the degree of risk aversion, the fixed order size of spot trading, the price variance of the spot price and the correlation of the spot and the futures prices.

Proposition I : The spot market bid-ask spread is always smaller if a market maker in the spot market can trade in futures as well.

---

[4]The detailed mathematics are given in appendix B.

It is evident that with futures trading the spot spread is smaller as long as the covariance between the spot and futures prices is not equal to zero. This finding confirms our intuitive assumption that the dealer, by trading in futures, is in the position to hedge her inventory and thus to reduce the price risk which enables her to narrow the spot spread.

If we change the assumption of a monopoly dealer and assume that there are two dealers in the market, we have to enlarge our analysis and take into account that the dealers may have differences in inventory levels, in the degree of risk aversion, in price expectations, and consequently in futures positions.

First, we assume the market situation where each dealer is the sole active market maker on each side of the market, i.e. dealer one (1) only sells and dealer two (2) only buys without any sharing of the orders.

Furthermore, the dealers are not identical which means that they have different inventory levels, different degrees of risk aversion, heterogeneous price expectations, and different futures positions.

Then, we can express the expected utility of the spot seller as:

$$EU_S(W_1) = (c_0^1 + paX) + \mu_{s_1}(I_1 - X) + (\mu_{f_1} - p_0^f)N_1$$
$$-(k_1/2)\left\{\sigma_{p_{s_1}}^2(I_1 - X)^2 + \sigma_{p_{f_1}}^2 N_1^2 + 2cov(a)(I_1 - X)N_1\right\} \qquad (19)$$

$cov(a) = cov[p^s_1, (p^f_1 - p^f_0)]$ and $I_1 = I_0$ of dealer 1

After multiplying out and setting $EU_1(W_1) = EU_0(W_0)$ we get:

$$pa = \mu_{s_1} - k_1\sigma_{p_{s_1}}^2[I_1 - (1/2)X] - cov(a)/\sigma_{p_{f_1}}^2(\mu_{f_1} - p^f_0)$$
$$- k_1 cov(a)^2/\sigma_{p_{f_1}}^2[I_1 - (1/2)X] \qquad (20)$$

The bid price can be calculated in the analogous way which results in:

$$pb = \mu_{s_2} - k_2\sigma_{p_{s_2}}^2[I_2 + (1/2)X] - cov(b)/\sigma_{p_{f_2}}^2(\mu_{f_2}-p^f_0)$$

$$+ k_2 cov(b)^2/\sigma_{p_{f_2}}^2[I_2 + (1/2)X] \qquad (21)$$

with $cov(b) = cov[p^s_1,(p^f_1-p^f_0)]$ and $I_2 = I_0$ of dealer 2

This time, the bid-ask spread is:

$$(pa-pb) = (\mu_{s_1} - \mu_{s_2}) + X/2\left[k_1\sigma_{p_{s_1}}^2+k_2\sigma_{p_{s_2}}^2\right] - (k_1I_1\sigma_{p_{s_1}}^2-k_2I_2\sigma_{p_{s_2}}^2)$$

$$- cov(a)/\sigma_{p_{f_1}}^2\left[(\mu_{f_1}-p^f_0)-cov(a)k_1I_1+cov(a)k_1X/2\right]$$

$$+ cov(b)/\sigma_{p_{f_2}}^2\left[(\mu_{f_2}-p^f_0)-cov(b)k_2I_2-cov(b)k_2X/2\right] \qquad (22)$$

In contrast to the bid-ask spread in the monopoly case, this time, the bid ask spread depends also on the dealers price expectations, for spot prices and for futures prices, and on their inventory levels.

4.2.2.1. Comparative Statics Analysis

In order to examine the influence of changes in the various parameters on the bid-ask spread we carry out a comparative statics analysis for the ask price (20) and for the bid price (21).

I is defined as I_0 of the respective dealer.

The first order derivatives for the ask side:

1. Effects of the risk aversion parameter on the ask price:

$$\delta pa/\delta k = [(X/2)-I]\left[\sigma_{p_s}^2 - (cov(a)^2/\sigma_{p_f}^2)\right] \qquad (23)$$

with $\left[\sigma_{p_s}^2 - (cov(a)^2/\sigma_{p_f}^2)\right] = \sigma_{p_s}^2\left[1-cov(a)^2/\sigma_{p_s}^2\sigma_{p_f}^2\right]$

which is $\sigma_{p_s}^2\left[1-r^2\right] > 0$

where r^2 is the correlation coefficient of p_f and p_s.

Hence, the last term in the square brackets on the RHS of (23) is positive. But still, equation (23) does not tell us unambiguously the sign of the right hand side. We know that X is positive, but the dealer can be short or long in inventory.

The influence of the risk aversion of a dealer depends on the sign of [(X/2)-I]. Hence theoretically, we can get a positive or a negative relationship between the k and pa, i.e.

if [(X/2)-I] > 0, with increasing k pa increases as well.

If [(X/2)-I] < 0, with increasing k pa decreases.

Hence, a dealer with increased risk aversion and a large inventory position is in a worse position of hedging her inventory and as a result she wants to sell and, therefore, she reduces her ask price.

2. Ask price moves with changing inventory I:

$$\delta pa/\delta I = -k\left[\sigma_{P_s}^{\ 2} - (cov(a)^2/\sigma_{P_f}^{\ 2})\right] \tag{24}$$

We see that $\sigma_{P_s}^{\ 2} - (cov(a)^2/\sigma_{P_f}^{\ 2})$ can be positive or negative which means that with a large spot price variance the expression is positive and hence the relationship between the ask price and a change in inventory is negative.

3. Order size influence:

$$\delta pa/\delta X = k/2\left[\sigma_{P_s}^{\ 2} - (cov(a)^2/\sigma_{P_f}^{\ 2})\right] \tag{25}$$

Again, the expression in the square bracket can be positive or negative. If the spot price variance is large compared to the covariance and the variance in the futures price we get a positive relationship of the order size and the ask price. The dealer does not like to take the risk of a large transaction size and therefore increases the ask price in order to avoid large transactions.

4. Effects of the spot price variance:

$$\delta pa/\delta\sigma_{P_s}^{\ 2} = k\left[(X/2) - I\right] \tag{26}$$

If the order size is relatively small compared to the inventory position, i.e. $\left[(X/2) - I\right] < 0$ we get a negative influence of the price variance on the ask price.

This makes sense as a risk averse dealer with a high inventory position wants to sell the asset if the spot price variance increases (given the same hedging conditions in the futures market) and therefore she lowers the ask price.

5. Effects of the futures price variance:

$$\delta pa/\delta\sigma_{p_f}^{\ 2} = -cov(a)\left[(\mu^f - p^f_0) + kcov(a)[(X/2)-I]\right] \qquad (27)$$

The influence of a change in the futures price variance is more difficult to evaluate and the analysis is not so straightforward compared to the derivation of the other parameter effects.

What we can say is that the covariance of spot and futures prices and the expectation about the future price are important. In addition, the sign of the second term in the bracket on the right hand side of (27) changes depending whether the dealer is short or long in the inventory.

If we assume that cov(a),$(\mu^f-p^f_0)$, and $[(X/2)-I]$ are positive then the ask price decreases with an increase in the futures price variance.

This is in line with the findings above. With an increase in the variance the dealer prefers to sell today rather than to wait and bear the increased futures price risk.

6. Influence of the future price expectation:

$$\delta pa/\delta\mu^f = -cov(a)/\sigma_{p_f}^{\ 2} \qquad (28)$$

It is clear that the price expectation depends on the covariance between the spot and the futures prices and the futures price variance.

If we have a negative covariance between spot and futures prices then we get a positive relationship between the futures price expectation and the ask price.

In other words, if the dealer expects the future price to rise, with the negative covariance the spot price is expected to rise and hence, the dealer increases the ask price as she is not willing to sell at a lower spot price. Similarly, if she expects the future price to rise and assuming a positive covariance, the spot price is expected to fall. Now, the dealer likes to sell and therefore lowers her ask price.

7. Role of the covariance of spot and futures prices:

$$\delta pa/\delta cov(a) = -(1/\sigma_{p_f}^{2})(\mu^f - p^f_0) - 2k cov(a)[(X/2)-I] \quad (29)$$

This first order derivative does not give any conclusive answer. The influence on the ask price depends on the sign of the covariance and the signs of $(\mu^f - p^f_0)$ and $[(X/2)-I]$. As we generally expect that cov > 0 and $(\mu^f - p^f_0) > 0$ and by assuming that $[(X/2)-I] > 0$ we get a negative dependence between the ask price and the covariance.

The first order derivatives for the bid side:

1. Changes in the risk aversion parameter:

$$\delta pb/\delta k = [(X/2)+I]\left[(cov(b)^2/\sigma_{p_f}^{2}) - \sigma_{p_s}^{2}\right] \quad (30)$$

with $\left[(cov(b)^2/\sigma_{p_f}^{2}) - \sigma_{p_s}^{2}\right] = \sigma_{p_s}^{2}\left[(cov(b)^2/\sigma_{p_f}^{2}\sigma_{p_s}^{2}) - 1\right]$

which is $\sigma_{p_s}^{2}\left[r^2 - 1\right] < 0$

We face the same situation compared to the ask side analysis. If $[(X/2)+I] > 0$, with increasing k pb decreases.

If $[(X/2)+I] < 0$, with increasing k pb increases as well.

By analysing the first case, where the dealer holds a positive inventory position, we show that the dealer lowers the bid price with increasing risk aversion as she does not want to buy more of the risky asset. If the dealer is short in inventory she increases the bid price with increasing risk aversion in order to buy now rather than to wait any longer and hence bear the price risk.

2. Effects of changing inventory position:

$$\delta pb/\delta I = k\left[(cov(b)^2/\sigma_{p_f}^{\ 2}) - \sigma_{p_s}^{\ 2}\right] \tag{31}$$

The expression in the square bracket can be positive or negative. If we assume that the sign of this term is negative we get a negative relationship between inventory and bid price. It implies that with increasing inventory, the dealer lowers the bid price because she does not want to buy anymore.

3. Order size changes:

$$\delta pb/\delta X = k/2\left[(cov(b)^2/\sigma_{p_f}^{\ 2}) - \sigma_{p_s}^{\ 2}\right] \tag{32}$$

Again this time the sign of the term in the square bracket can either be positive or negative. If the spot price variance is relatively large the expression is negative and we obtain a negative relationship which means that with increasing order size the dealer lowers the bid price in order to induce less trade.

4. Influence of changes in the spot price variance:

$$\delta pb/\delta\sigma_{p_s}^{\ 2} = -k\left[(X/2) + I\right] \tag{33}$$

If the dealer is long in the inventory position an increase in the spot price variance results in a decrease in the bid price as the dealer does not want to buy and does not want to take the spot price risk.

However, if the dealer has a short inventory position which is large enough that $\left[(X/2)+I\right]<0$, the dealer can profit from buying with an increase in the spot price variance and hence, the bid price increases.

5. Effects of the futures price variance:

$$\delta pb/\delta\sigma_{p_f}^{\ 2} = -cov(b)\left[(\mu^f - p_0^f) - kcov(b)[(X/2)+I]\right] \tag{34}$$

We encounter a similar problem as in the ask side case. The influence of a change in the futures price variance on the bid price is inconclusive and depends on the sign of the covariance, the sign of $(\mu^f - p^f_0)$, and the sign of $[(X/2)+I]$.

6. Changes of the future price expectation:

$$\delta pb/\delta\mu^f = -cov(b)/\sigma_{p_f}^2 \quad (35)$$

Under the assumption of a positive covariance, we get a negative relationship between the bid price and the future price expectation. If the dealer expects a high future price the spot price fall. The dealer is not willing to buy in the spot market and sell the asset in a future point in time as it is not profitable. Hence the dealer lowers her spot bid price.

7. Influence of changes in the covariance of prices:

$$\delta pb/\delta cov(b) = -(1/\sigma_{p_f}^2)(\mu^f - p^f_0) + 2k cov(b)[(X/2)+I] \quad (36)$$

Again, we cannot say anything conclusive about the influence. It depends on the signs of the covariance, $(\mu^f - p^f_0)$, and $[(X/2)+I]$.

We are particularly interested in the situation where the bid-ask spread may be widened through trading in futures. Therefore, we analyse the changes in the futures price variance $\sigma_{p_f}^2$ and the futures price expectations μ^f.

Proposition II : The ask price will be increased if either:

- there is an increase in $\sigma_{p_f}^2$ and cov < 0, $(\mu^f - p^f_0) > 0$, $[(X/2)-I] < 0$

or there is a decrease in $\sigma_{p_f}^2$ and cov > 0, $(\mu^f - p^f_0) < 0$, $[(X/2)-I] < 0$.

The bid price will be decreased if either:

- there is an increase in $\sigma_{p_f}^2$ and cov < 0, $(\mu^f - p^f_0) < 0$, $[(X/2)-I] < 0$

or there is a decrease in $\sigma_{p_f}^2$ and cov < 0, $(\mu^f - p^f_0) > 0$, $[(X/2)-I] > 0$.

At the same time, the effects of changes of the futures price expectation give us the following results:

Proposition III : There is an increase in the ask price if either:

- the dealer has an optimistic futures price expectation and cov < 0

or the dealer expects the futures price to fall and cov >0.

There is a decrease in the bid price if either:

- the dealer's futures price expectation is optimistic with cov >0

or the dealer expects the futures price to fall with cov < 0.

If we take proposition II, both the first cases for the ask price and the bid price we have the following situation.

The selling market maker who is long in the inventory expects an increase in the futures price variance. At the same time, she expects the futures price to rise. With a negative covariance the spot price is expected to fall which means that the dealer expects to make a gain in a distant point in time by not selling now. Hence she is not prepared to sell today which results in the fact that she increases her ask price.

On the other side, the buyer who is short in inventory expects the futures price variance to rise. In addition, the dealer expects the futures price to fall. The result is that the dealer would make a loss by buying now and therefore she lowers her bid price. In this situation just described, the crucial difference is that the selling dealer is long in inventory and expects the futures price to rise whereas the buying dealer is short in inventory and she expects a decrease in the futures price.

4.2.3. Equilibrium Conditions

In the following section we consider the two dealer case with dealers who are not identical and differ in their degrees of risk aversion, their inventory levels and their price expectations.

4.2.3.1. Spot Market Equilibrium

The determination of the equilibrium in the spot market is given by the individual prices which we derived earlier from equation (20) and (21) which are

$$pa = \mu_{s_1} + k_1\sigma_{p_{s_1}}^2[(1/2)X - I_1] - [cov(a)/\sigma_{p_{f_1}}^2](\mu_{f_1}-p^f_0)$$
$$- [k_1 cov(a)^2/\sigma_{p_{f_1}}^2][(1/2)X - I_1] \text{ and}$$

$$pb = \mu_{s_2} - k_2\sigma_{p_{s_2}}^2[I_2 + (1/2)X] - [cov(b)/\sigma_{p_{f_2}}^2](\mu_{f_1}-p^f_0)$$
$$+ [k_2 cov(b)^2/\sigma_{p_{f_2}}^2][I_2 + (1/2)X]$$

We can rewrite these prices as

$$pa = \mu_{s_1} + [(1/2)X - I_1](k_1\sigma_{p_{s_1}}^2 - k_1 cov(a)^2/\sigma_{p_{f_1}}^2)$$
$$- [cov(a)/\sigma_{p_{f_1}}^2](\mu_{f_1}-p^f_0) \text{ and}$$

$$pb = \mu_{s_2} - [I_2 + (1/2)X](k_2\sigma_{p_{s_2}}^2 - k_2 cov(b)^2/\sigma_{p_{f_2}}^2)$$
$$- [cov(b)/\sigma_{p_{f_2}}^2](\mu_{f_2}-p^f_0)$$

Under the assumption that market makers are homogeneous except in their inventory levels, dealers differ in the second expression on the RHS of the price equations only. Thus, if we assume the there are several competing dealers in the spot market we get the same Bertrand type equilibrium as HS which is that the market prices are determined by the second best dealer's prices (bearing in mind that the market makers know each others' reservation prices and that the order is indivisible).

Although the prices themselves are different compared to "pure" spot market trading we can still model the process of reaching this equilibrium as a second price or Vickrey auction.

4.2.3.2. Futures Market Equilibrium

In order to determine the equilibrium we have to add the analysis of a fourth type of market participants which are the speculators who only trade in the futures market.

The speculator also maximizes expected utility of terminal wealth which is:

$$EU(W_{SP}) = (\mu_{SP} - p^f_0)N - k/2(\sigma^2_{f_{SP}})N^2 \qquad (37)$$

Maximizing expected utility over N results in an optimal futures position of the speculator of:

$$N_{SP} = (\mu_{SP} - p^f_0)/k_{SP}\sigma^2_{f_{SP}} \qquad (38)$$

We assume that the market makers active and non-active in the spot market are the only participants in the futures market together with the speculators.

In equilibrium there is no excess demand or supply and therefore, in the futures market, the following condition must hold:

$$VN_B + WN_{NT} + YN_S + ZN_{SP} = 0 \qquad (39)$$

with V_B, W_{NT}, Y_S, Z_{SP}, being the number of buyers, non-traders, and sellers in the spot market and the number of the speculators active in the futures market respectively.

N_B, N_{NT}, and N_S, are the optimal futures positions, derived earlier, for the buyers, the non traders, and the sellers in the spot market.

Substituting the optimal futures positions of equations (9), (13), and the optimal futures position of "no trade" which is:

$$N_{NT} = (\mu^f_{NT} - p^f_0)/k_{NT}\sigma_{f_{NT}}^{\ 2} - cov(NT)I/\sigma_{f_{NT}}^{\ 2}$$

and the optimal futures position of a speculator into equation (39), we finally get the current futures price p^f_0: [5]

[5]Calculations see appendix C.

$$p^f_0 = S\Big[(V/k_b\sigma_{f_B}{}^2)\big[(\mu^f{}_B)-k_b cov(b)(I_B+X)\big]$$
$$+ (W/k_{NT}\sigma_{f_{NT}}{}^2)\big[(\mu^f{}_{NT})-k_{NT}cov(NT)(I_{NT})\big]$$
$$+ (Y/k_S\sigma_{f_S}{}^2)\big[(\mu^f{}_S)-k_S cov(a)(I_S-X)\big]$$
$$+ (Z/k_{SP}\sigma_{f_{SP}}{}^2)(\mu^f{}_{SP})\Big] \qquad (40)$$

with $S = \left[(V/k_b\sigma_{f_B}{}^2)+(W/k_{NT}\sigma_{f_{NT}}{}^2)+(Y/k_S\sigma_{f_S}{}^2) + (Z/k_{SP}\sigma_{f_{SP}}{}^2)\right]^{-1}$

Let us define:

$M = V_B + W_{NT} + Y_S + Z_{SP}$ (total number of agents)

Furthermore, we assume that the dealers have the same degree of risk aversion and have homogeneous expectations. The current futures price p^f_0 is determined by: [6]

$$p^f_0 = \mu^f - (k cov()/M)\big[V(I_B+X)+WI_{NT}+Y(I_S-X)\big] \qquad (41)$$

Under our assumption for the spot market that only one order arrives at the time, equation (41) simplifies as either V or Y is zero depending whether a sell order or a buy order is executed in the spot market.

If we look at today's financial markets with professional market makers with sophisticated screen trading and instantaneous information about price changes the above assumptions seem reasonable.

Hence, equation (41) shows us that the futures price is mainly a function of the number of dealers in the market, their degree of risk aversion, the variance of the futures price, the covariance of the spot and futures prices, and the inventory positions of the hedgers.

---

[6]Details see appendix C.

4.2.4. Robustness of the Futures Market Equilibrium

We conduct an analysis of the equilibrium conditions which are given in equation (41). The aim is to examine how robust these equilibrium conditions are in respect of the number of participants, the differences in the inventories of the hedgers, and the influence of the price variances and the covariance of spot and futures prices.

Comparative statics results:

We define $[Y(I_s-X)+V(I_B+X)+WI_{NT}]$ to be [I].

1. Influence of risk aversion:

$$\delta p^f_0/\delta k = - \text{cov}()[I]/M \qquad (42)$$

The result is not very clear. The influence of the degree of risk aversion on the futures price depends on the sign of the covariance and whether the overall inventory position of all dealers in the market is long or short.

2. Changes in the number of market participants:

$$\delta p^f_0/\delta M = - k\text{cov}()[I] \qquad (43)$$

Also this time, the dependence is subject to the sign of the covariance and the inventory term which includes the proportion of the number of sellers,buyers, and non-traders.

If $\text{cov}() > 0$ and $[I] > 0$ we have a negative relationship between M and p^f_0. This finding can be explained by the increased possibility of sharing the risk among a higher number of market participants and as a result the lower cost of trading.

3. Influence of changes in the covariance:

$$\delta p^f_0/\delta \text{cov}() = - k[I]/M \qquad (44)$$

The change in the covariance may have a positive or negative impact on the futures price depending on the sign of the inventory term.

If [I] > 0 we get a negative correlation between the covariance and the futures price which is reasonable because with an increase in the covariance dealers are in a better position to hedge their long inventory and can thereby reduce the risk of holding the asset.

As a consequence of the higher demand in futures the futures price is reduced.

4. The effect of the inventory positions:

$$\delta p^f_0/\delta[I] = -\ kcov()/M \qquad (45)$$

The sign of the covariance determines the relationship between the futures price and the inventory.

With a positive covariance we get a negative dependence which means that with an increasing inventory (assuming the same degree of risk aversion) there is an increased demand of hedging by trading futures and hence the futures price decreases.

The results of our comparative statics analysis show that the futures price is determined by the covariance between the spot and futures price, the degree of risk aversion of dealers, the respective individual (in terms of buyers, sellers, "not active traders in the spot market", and speculators) inventory positions, as well as the number of market participants.

However, we have not shown yet how important the respective proportion of the various participants is, especially in respect of the speculators. This is done in the next analysis.

Such an analysis examines what happens in equilibrium if one or several groups of market participants disappear from the market.

The tables with the results of the spread sheet analysis are shown in appendix D. [7]

---

[7]The analysis is based on a spread sheet calculation of the equilibrium equation (41).

The basis of the analysis is the situation with an equal number of each type of agent. If we eliminate one group of participants we can see the influence on the current futures price.

1. If we eliminate the speculators, the current futures price is lowered. With an increase in the number of speculators the futures price is increased.

2. The elimination of the buyers (buyers in the spot market) increases the futures price. On the other hand, if we increase the number of buyers the futures price decreases.

3. If we assume that there are no traders in the futures market who also do not trade in the spot market then we can observe a slight decrease of the futures price. With an increasing number of this type of traders the futures price increases.

4. The absence of any sellers in the futures market results in an increase of the futures price and the opposite is true that with an increase of the number of sellers the futures price falls.

5. If we increase the number of speculators and buyers simultaneously the futures price is increased. On the other hand, if we increase the number of speculators and sellers, at the same time, then the the sign of the futures price is the same and the size of it is hardly changed.

These findings show that the proportion of the number of the various agents does influence the determination of the equilibrium futures price.

4.3. The Structure of a Model with Uncertainty in the Order Flow in the Spot Market

The second part of this chapter deals with a different kind of "microstructure model" for the spot market. So far, we assumed that the market makers in the spot market have full information about the inventory positions of their competitors and that they also know about the order flow before they have to quote their prices. This type of model applies to a centralized market.

Now, we consider an over the counter market where dealers compete without complete information about the public orders. We can think of a telephone market where dealers quote their prices by telematic circuit. Such price quotes are binding, but the trades are done over the phone with subsequent adjustment of the screen prices.

We can illustrate the sequence of events as follows:

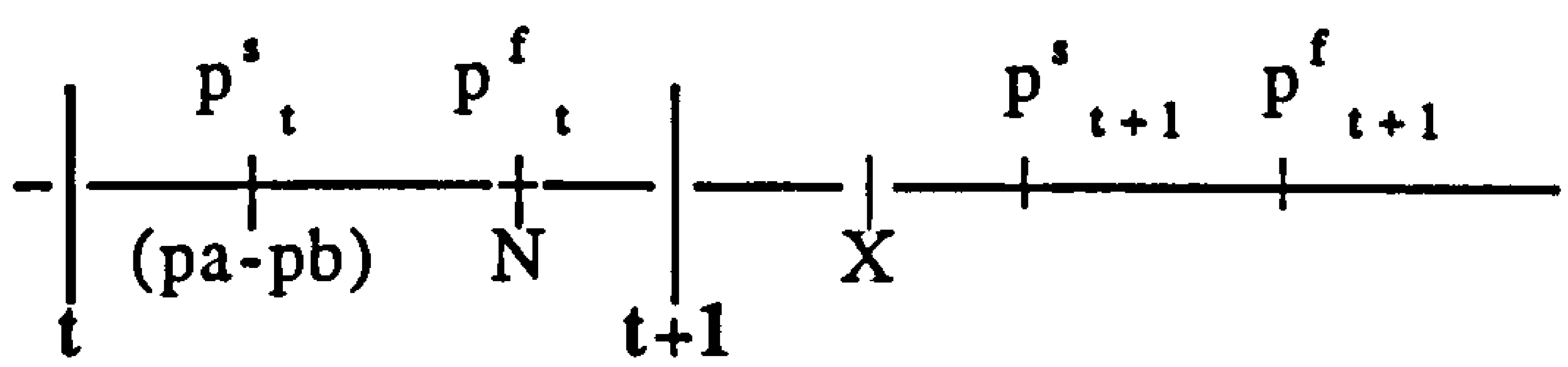

Figure 4.2.: Sequence of events with unknown X

4.3.1. Determinants of the Bid-Ask Spread with Symmetric Demand and Supply

Initially we assume that traders know that the demand and supply functions are symmetric.

The basics of the spot market model are identical to the features we discussed in the previous section, except that, for simplicity reasons, we assume that the order cannot be split between dealers.

Also in this case, dealers maximize expected utility of terminal wealth.

Terminal wealth is composed of:

$W_1 = c_1 + p_1^s I_1 + (p_1^f - p_0^f)N$

with c_1 : cash holding at time t+1

I_1 : inventory position of the risky asset at t+1

N : net futures purchases at t+1

p_1^s, p_1^f: spot and futures prices respectively at t+1

Again, we use a mean variance utility with risk averse dealers.

Expected utility is :

$$E(U) = E(W_1) - k/2(VAR(W_1))$$

We assume independence between p_1^s and I_1. Hence we have:

$$E(W_1) = E(c_1) + E(p_1^s)E(I_1) + E(p_1^f - p_0^f)N \tag{46}$$

where p_0^f is the futures price at time t. The variance is

$$\begin{aligned} VAR(W_1) = {} & Var(c_1) + Var(p_1^s I_1) + Var(p_1^f - p_0^f)\, N^2 \\ & + 2cov(c_1, p_1^s I_1) + 2cov(c_1, (p_1^f - p_0^f))N \\ & + 2cov(p_1^s I_1, (p_1^f - p_0^f))N \end{aligned} \tag{47}$$

Still based on the assumptions that the order size is X and the order arrival rate is λ, we calculate the expectations, the variances, and the covariances and we get: [8]

$$E(c_1) = c_t + \lambda X(pa-pb)$$

$$E(p_1^s) = \mu_s,\ var(p_1^s) = \sigma^2_{p_s}$$

$$E(I_1) = I_0$$

$$E(p_1^f) = \mu_f,\ var\ (p_1^f) = \sigma^2_{p_f}$$

$$Var(c_1) = \lambda X^2\left[(1-\lambda)(pa^2+pb^2)+2\lambda papb\right]$$

$$Var(I_1) = 2\lambda X^2$$

$$Var(p_1^s I_1) = \sigma^2_{p_s}(2\lambda X^2 + I_0^2) + (\mu_s)^2 2\lambda X^2$$

$$cov(c_1, p_1^s I_1) = -\lambda X^2 \mu_s (pa+pb)$$

$$cov(c_1, (p_1^f - p_0^f)) = 0$$

$$cov(p_1^s I_1, (p_1^f - p_0^f)) = cov(p_1^s, p_1^f) I_0$$

Now, we can write the expected utility as:

---

[8]The mathematics can be found in appendix E.

$$EU(W_1) = c_0 + \lambda X(pa-pb) + \mu_s I_0 + (\mu_f - p_0^f)N$$
$$- k/2\left[\lambda X^2\left[(1-\lambda)(pa^2+pb^2)+2\lambda papb\right]\right]$$
$$- k/2\left[\sigma^2_{p_s}(2\lambda X^2 + I_0^2) + (\mu_s)^2 2\lambda X^2\right]$$
$$- k/2(\sigma^2_{p_f} N^2) + k\lambda\mu_s X^2(pa+pb) - kcov(p_1^s, p_1^f)I_0 N \qquad (48)$$

The dealer sets her bid and ask prices knowing that she will optimize the futures position. Thus we first optimize over N.

The first order condition of $EU(W_1)$ is:

$\delta EU(W_1)/\delta N = (\mu_f - p_0^f) - k\sigma^2_{p_f} N - kcov(p_1^s, p_1^f)I_0 = 0$ As a result,

$$N = (\mu_f - p_0^f)/k\sigma^2_{p_f} - cov(p_1^s, p_1^f)I_0/\sigma^2_{p_f} \qquad (49)$$

Compared to the optimal futures position in the previous model, we get an identical result.

If we substitute the optimal position in (49) back into the expected utility function given in (48) and rearrange and simplify we end up with: [9]

$$EU(W_1) = c_t + \lambda X(pa-pb) + \mu_s I_0 - (k/2)\left[\lambda X^2\left[(1-\lambda)(pa^2+pb^2)+2\lambda papb\right]\right]$$
$$- (k/2)\left[\sigma^2_{p_s}(2\lambda X^2 + I_0^2) + (\mu_s)^2 2\lambda X^2\right] + \lambda\mu_s X^2(pa+pb)$$
$$+ (\mu_f - p_0^f)^2/2k\sigma^2_{p_f} + kcov(p_1^s, p_1^f)^2 I_0^2/2\sigma^2_{p_f}$$
$$- \left[(\mu_f - p_0^f)cov(p_1^s, p_1^f)I_0\right]/\sigma^2_{p_f} \qquad (50)$$

Under the zero profit condition (in a competitive market) the trader sets the expected utility $EU(W_1)$ equal to the expected utility of terminal wealth without any trading in the spot market $EU(W_0)$ which is:[10]

$$EU(W_0) = c_t + \mu_s I_0 - (k/2)\sigma^2_{p_s} I_0^2$$
$$+ (\mu_f - p_0^f)^2/2k\sigma^2_{p_f} + kcov(p_1^s, p_1^f)^2 I_0^2/2\sigma^2_{p_f}$$
$$- \left[(\mu_f - p_0^f)cov(p_1^s, p_1^f)I_0\right]/\sigma^2_{p_f} \qquad (51)$$

---

[9]Detailed calculation see appendix E.

[10]Details see appendix E.

Hence

$$EU(W_1) = EU(W_0) = \lambda X(pa-pb) - (k/2)\left[\lambda X^2\left[(1-\lambda)(pa^2+pb^2)+2\lambda papb\right]\right]$$

$$- (k/2)\left[\sigma^2_{p_s}(2\lambda X^2) + (\mu_s)^2 2\lambda X^2\right]+\lambda\mu_s X^2(pa+pb) \qquad (52)$$

Equation (52) is purely determined by the expressions coming from the spot market trading. Thus, the result shows that there is no influence of futures trading on the spot market bid-ask spread.

Proposition IV : In a market where market makers do not know the order flow at the time they have to quote their prices and under the assumption of symmetric demand and supply, the bid-ask spread is not affected by trading in futures.

One reason of this outcome could be that the demand and supply of the orders are both of the same size X and the resulting expected inventory is I_0 which is identical to the inventory position of a dealer who is not trading in the spot market at all. Based on our assumption of mean variance preferences and that the sell order is of the same size as the buy order, the dealer does not face an increase in risk and therefore the spread is unaffected in this setting.

If it turns out that this assumption of symmetry in sell and buy orders is crucial in respect of the influence on the bid-ask spread, the traditional model may be a quite restricted versions of the real situation. Especially in dealership markets, dealers may not be in the position to close out their position (i.e. balance their books) as quickly as they wish to do so. There may be several reasons, for instance there is a thin market and the order arrival is very slow, or the competition among them is very hard. Another reason may be that there is a trend in the market. Therefore we change this restriction and introduce asymmetry in demand and supply.

4.3.2. Influence of Futures Trading with Asymmetry in Demand and Supply

We assume that a purchase order is of size X and the sale order is of size Y. Then the respective expectations, variances, and covariances are: [11]

$E(c_1) = c_t + \lambda(paY\text{-}pbX)$

$E(p_1^s) = \mu_s$

$E(I_1) = I_0 + \lambda(X\text{-}Y)$

$E(p_1^f) = \mu_f$

$Var(c_1) = (\lambda\text{-}\lambda^2)(pa^2Y^2+pb^2X^2)\text{-}2\lambda^2 paYpbX$

$Var(I_1) = \lambda(X^2+Y^2) - (\lambda X\text{-}\lambda Y)^2$

$Var(p_1^s I_1) = (\sigma^2_{p_s}+(\mu_s)^2)\left[(\lambda X+\lambda Y)^2\text{-}\lambda^2(X^2+Y^2)\right]+\sigma^2_{p_s}(I_0+\lambda(X\text{-}Y))^2$

$cov(c_1,p_1^s I_1) = \text{-}\lambda\mu_s\left[paY^2+pbX^2+(1\text{-}\lambda)[\lambda(paY^2+pbX^2)\text{-}\lambda XY(pa+pb)]\right]$

$cov(c_1,(p_1^f\text{-}p_0^f)) = 0$

$cov(p_1^s I_1,(p_1^f\text{-}p_0^f)) = cov(p_1^s,p_1^f)(I_0+\lambda(X\text{-}Y)))$

The expected utility can be written as:

$$\begin{aligned} EU(W_1) = {} & c_t + \lambda(paY\text{-}pbX) + \mu_s(I_0 + \lambda(X\text{-}Y)) + (\mu_f\text{-}p_0^f)N \\ & - (k/2)\left[(\lambda\text{-}\lambda^2)(pa^2Y^2+pb^2X^2)\text{-}2\lambda^2 paYpbX\right] \\ & - (k/2)\left[(\sigma^2_{p_s}+(\mu_s)^2)\left[(\lambda X+\lambda Y)^2\text{-}\lambda^2(X^2+Y^2)\right]+\sigma^2_{p_s}(I_0+\lambda(X\text{-}Y))^2\right] \\ & - (k/2)\sigma^2_{p_f}N^2 - k\,cov(p_1^s,p_1^f)(I_0+\lambda(X\text{-}Y)))N \\ & + k\lambda\mu_s\left[paY^2+pbX^2+(1\text{-}\lambda)[\lambda(paY^2+pbX^2)\text{-}\lambda XY(pa+pb)]\right] \end{aligned} \qquad (53)$$

If we take the first order derivative of (53) and optimize over N we get an optimal futures position of

$$N = (\mu_f\text{-}p_0^f)/k\sigma^2_{p_f} - cov(p_1^s,p_1^f)(I_0+\lambda(X\text{-}Y)))/\sigma^2_{p_f} \qquad (54)$$

---

[11]Calculations are given in appendix F.

which differs in the inventory term compared to the futures position we derived above.

Now, we substitute (54) into (53), calculate and simplify and we get an expected utility of [12].

$$
\begin{aligned}
EU(W_1) = {} & c_t + \lambda(paY-pbX) + \mu_s(I_0 + \lambda(X-Y)) \\
& - (k/2)\left[(\lambda-\lambda^2)(pa^2Y^2+pb^2X^2)-2\lambda^2 paYpbX\right] \\
& - (k/2)\left[(\sigma^2_{p_s}+(\mu_s)^2)\left[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)\right]+\sigma^2_{p_s}(I_0+\lambda(X-Y))^2\right] \\
& + k\lambda\mu_s\left[paY^2+pbX^2+(1-\lambda)[\lambda(paY^2+pbX^2)-\lambda XY(pa+pb)]\right] \\
& + (\mu_f-p_0^f)^2/2k\sigma^2_{p_f} + k\,cov(p_1^s,p_1^f)^2[I_0+\lambda(X-Y)]^2/2\sigma^2_{p_f} \\
& - \left[(\mu_f-p_0^f)cov(p_1^s,p_1^f)[I_0+\lambda(X-Y)]\right]/\sigma^2_{p_f} \qquad (55)
\end{aligned}
$$

This time, the λ-term differs compared to (48) due to the asymmetric demand and supply. The next step is to set $EU(W_1) = EU(W_0)$ which gives us the impact of futures trading on the spot market bid-ask spread.

$EU(W_0)$ is taken from equation (52) and we get [13]

$$
\begin{aligned}
& \lambda(paY-pbX) + \mu_s(\lambda(X-Y)) - (k/2)\left[(\lambda-\lambda^2)(pa^2Y^2+pb^2X^2)-2\lambda^2 paYpbX\right] \\
& - (k/2)\left[(\sigma^2_{p_s}+(\mu_s)^2)\left[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)\right]\right. \\
& \left. + \sigma^2_{p_s}[2\lambda I_0(X-Y)+\lambda^2(X-Y)^2]\right] \\
& + k\lambda\mu_s\left[paY^2+pbX^2+(1-\lambda)[\lambda(paY^2+pbX^2)-\lambda XY(pa+pb)]\right] \\
& = - k\,cov(p_1^s,p_1^f)^2[2\lambda I_0(X-Y)+\lambda^2(X-Y)^2]/2\sigma^2_{p_f} \\
& + \left[(\mu_f-p_0^f)cov(p_1^s,p_1^f)[\lambda(X-Y)]\right]/\sigma^2_{p_f} \qquad (56)
\end{aligned}
$$

The left hand side of (56) shows the terms which are generated by the activity in the spot market whereas the right hand side gives the

[12]Calculation are in appendix F.

[13]Details see appendix F.

expressions derived from futures market trading.

The spot market bid-ask spread function turns out to be quadratic in pa and pb and we find also products of pa, pb in it. Typically, if we solve the problem of the bid-ask spread we would get more than one solution.

This implies that the optimization problem of the bid price and the ask price are not independent anymore. The ask price is a function of the bid price and the order arrival rate λ, and vice versa, the bid price is a function of the ask price and the order arrival rate λ.

4.3.2.1. Analysis of Influence on the Spot Bid-Ask Spread

To facilitate our analysis we can write (56) as $F(pa,pb,\theta) = G(\theta)$. The calculation under a comparative statics analysis is not so straightforward as it has been in the previous section. First, we have to define the respective comparative statics equations.

In order to compare the influence of the various spot and futures market parameters on the spot bid and ask prices we have to evaluate the derivative of pa and pb in respect of the particular parameter.

Thus, we have $(\delta F/\delta pa)dpa + (\delta F/\delta\theta)d\theta = (\delta G/\delta\theta)d\theta$ and $(\delta F/\delta pb)dpb + (\delta F/\delta\theta)d\theta = (\delta G/\delta\theta)d\theta$ with θ taking the value of the respective parameters which are: $\sigma^2_{p_s}$, μ_s, and λ for the spot market and $\sigma^2_{p_f}$, μ_f, and $cov(p_s,p_f)$ for the futures market. We then can solve for dpa and dpb in order to get the marginal change in the spread $S = (dpa-dpb)$.

$dpa = [(\delta G/\delta\theta) - (\delta F/\delta\theta)]d\theta/(\delta F/\delta pa)$ and

$dpb = [(\delta G/\delta\theta) - (\delta F/\delta\theta)]d\theta/(\delta F/\delta pb)$ so that the change in the spread is

$dS/\delta\theta = [(\delta G/\delta\theta) - (\delta F/\delta\theta)]d\theta\ [1/[(\delta F/\delta pa) - (\delta F/\delta pb)]$

However, this analysis does not give us any clear evidence of the effects of the various parameters on the bid and ask prices. By differentiating the LHS of (56) with respect to pa and pb we get

$(\delta F/\delta pa)dpa$

$$= [\lambda Y - k\lambda Y^2 pa + k\lambda^2 Y^2 pa + k\lambda^2 YpbX + k\lambda\mu_s Y^2 + k\lambda^2\mu_s Y^2 - k\lambda^3\mu_s Y^2 - k\lambda^2\mu_s XY + k\lambda^3\mu_s XY]\ dpa \quad (57)$$

$(\delta F/\delta pb)dpb$

$$= [-\lambda X - k\lambda X^2 pb + k\lambda^2 X^2 pb + k\lambda^2 YpaX + k\lambda\mu_s X^2 + k\lambda^2\mu_s X^2 - k\lambda^3\mu_s X^2 - k\lambda^2\mu_s XY + k\lambda^3\mu_s XY]\ dpb \quad (58)$$

These two expressions are too complicated to be evaluated if we try to extract the influence, for instance, of the spot price variance on the bid-ask spread which is:

$$(\delta F/\delta\sigma^2_{P_s})d\sigma^2_{P_s} = (k/2)[1 + [2\lambda I_0(X-Y) + 2\lambda^2(X-Y)^2]]d\sigma^2_{P_s} \quad (59)$$

$dS = (dpa - dpb)$ in respect of the parameter $\sigma^2_{P_s}$ is

$$= -[(\delta F/\delta\sigma^2_{P_s})d\sigma^2_{P_s}/(\delta F/\delta pa)] + [(\delta F/\delta\sigma^2_{P_s})d\sigma^2_{P_s}/(\delta F/\delta pb)] \quad (60)$$

With the substitution of dpa and dpb of (57), (58) and (59) into (60) we see that the sign of the overall change can not be determined.
Unfortunately, this is the case for all the various parameters.

Under the assumption of an asymmetric order flow we show that the futures market trading affects the prices in the spot market. However, the influence on the bid-ask spread is to complex too evaluate within such a framework.
What we can observe is that the important determinants of the spot bid-ask spread are the inventory position, the difference between purchases and sales, the order arrival rate, the covariance between the spot and futures price, and the difference between the expected futures price and the current futures price.

4.3.3. Futures Market Equilibrium Conditions

We examine the equilibrium condition for the asymmetric case. We define:

N : number of sellers and buyers who are active in the spot and futures market.

M : number of speculators who only trade in futures

T : total number of dealers in the futures market with

T = N + M

To facilitate the analysis we assume that the dealers are homogeneous in their expectations and in their degrees of risk aversion.

In order to have an equilibrium situation the following condition must hold:

$$N\left[(\mu_f-p_0^f)/k\sigma^2_{p_f}-[cov(I_0+\lambda(X-Y)]/\sigma^2_{p_f}\right] + M\left[(\mu_f-p_0^f)/k\sigma^2_{p_f}\right] = 0 \qquad (61)$$

Under the assumption of homogeneous dealers with identical initial inventory positions we get a futures price of

$$p_0^f = \mu_f - [(kcovI_0)/T][N\lambda(X-Y)] \qquad (62)$$

Compared to the model in section 4.2., i.e. without any order flow uncertainty, we observe an additional parameter which is λ, the order flow probability.

However, we do not distinguish between buyer, sellers, or non active traders as the dealers do not know what kind of order finally arrives or whether they will trade in this period at all.

4.3.3.1. Equilibrium Analysis

We analyse the comparative statics of (62) which are:

1. Changes in the covariance:

$$\delta p_0^f/\delta cov = -(kI_0/T)[N\lambda(X-Y)] \qquad (63)$$

The sign of the impact of changes in the covariance is determined by the sign of the inventory I_0 and of (X-Y).

2. Influence of the degree of risk aversion:

$$\delta p_0^f/\delta k = - [(covI_0)/T][N\lambda(X-Y)] \qquad (64)$$

The dependence is determined by the sign of the covariance, the inventory and the difference between purchases and sales in the spot market.

3. Impact of changes in the inventory:

$$\delta p_0^f/\delta I_0 = - [(kcov)/T][N\lambda(X-Y)] \qquad (65)$$

The relationship between the changes in inventory and the futures price depends on the sign of the covariance and (X-Y).

4. Influence of (X-Y):

$$\delta p_0^f/\delta(X-Y) = - (\lambda NkcovI_0)/T \qquad (66)$$

What kind of impact (X-Y) has on the futures price is dependent on the sign of the covariance and the inventory.

Hence, the crucial parameters which determine the nature of influence on the futures price are the inventory position, the covariance, and the difference between purchases and sales in the spot market.

Until now, we examined and analysed the influence of futures trading on the spot bid-ask spread within a one period framework which makes sense as the futures trading is settled mark-to-market which means that the futures position is valued every day and margin payments have to be made to cover the daily open position.

However, if we look at our model with uncertainty in the order flow in the spot trading dealers may change their optimal futures position if they know that they can adjust it in the next period after having learnt about the spot market trade.

Therefore, we extend the model in section 4.3. to a two period model with futures trading activity in both periods in order to see whether we find some intertemporal effects on the bid-ask spread.

4.4. Intertemporal Effects of Futures Trading on the Spot Bid-Ask Spread

For exposition purposes we analyse the symmetric case of the model in section 4.3.. In addition, we introduce spot market trading in the first period only. The sequence of events will give us an overview about the structure of the model.

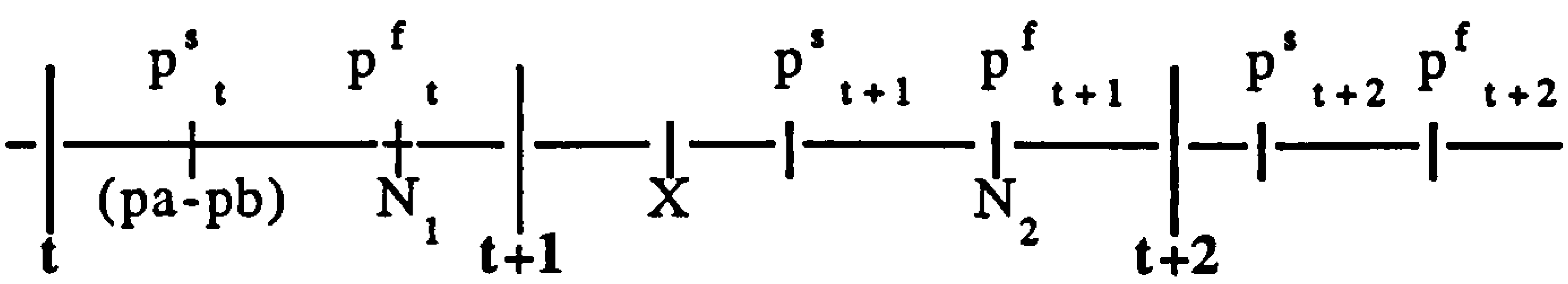

Figure 4.3.: Sequence of events with unknown X for two periods

We will solve this two period problem by the method of backward induction.

4.4.1. Influence of Futures Trading on the Spot Bid-Ask Spread

Grossman and Miller's model (1988) explains the optimal pricing strategies of the market participants over two periods. The model is equally applicable to the spot and the futures market. However, they do not try to link the two markets and to examine the interdependence of the two markets.
In our model, we investigate the pricing policies in the same line, but we take into account the interdependence of the two markets.

We can write the utility of terminal wealth at the end of the second period as:

$$W = c_1 + p_1^s I_1 + (p_1^f - p_0^f)N_1 + (p_2^f - p_1^f)(N_1 + N_2) \quad (67)$$

with N_1 = net purchase futures committed at t and maturing at t+2; N_2 = net purchase futures committed at t+1 and maturing at t+2.

In the second period the following variables are known:

p_1^f, p_1^s, X, I_1, p_0^f, c_1.

The only unknown parameter is p_2^f.

As a consequence we can write the expected utility as

$$EU_2(W) = c_1 + p_1^s I_1 + (p_1^f - p_0^f)N_1 + E_2(p_2^f - p_1^f)(N_1+N_2) - (k/2)\sigma^2_{f_2}((N_1+N_2)^2 \quad (68)$$

with:

$$var(p_2^f) = \sigma^2_{f_2}$$

First, we derive the optimal futures position N_2 which we obtain by optimizing $EU_2(W)$ over N_2:

$$\delta EU_2(W)/\delta N_2 = E_2(p_2^f - p_1^f) - k\sigma^2_{f_2}N_2 - k\sigma^2_{f_2}N_1 = 0$$

$$N_2 = [E_2(p_2^f - p_1^f)]/k\sigma^2_{f_2} - N_1 \quad (69)$$

The optimal futures position in (69) is replaced in (68) and after rearranging we get: [14]

$$U^*_2(W) = c_1 + p_1^s I_1 + (p_1^f - p_0^f)N_1 + \left[E_2(p_2^f - p_1^f)\right]^2/2k\sigma^2_{f_2} \quad (70)$$

Moving to the first period we have p_0^f, pa, and pb as the known variables and $E_2(p_2^f)$, p_1^f, p_1^s, X and I_1 are the unknown variables. We define the expected utility as: $EU_1 = E_1(U^*_2) - (k/2)Var_1(U^*_2)$

Hence, we can write

$$E_1(U^*_2) = E_1c_1 + E_1p_1^s E_1I_1 + E_1(p_1^f - p_0^f)N_1 + E_1\left[[E_2(p_2^f - p_1^f)]^2/2k\sigma^2_{f_2}\right] \quad (71)$$

---

[14]Mathematics are given in appendix G.

$$
\begin{aligned}
Var_1(U^*_2) = {} & V_1c_1 + V_1(p_1^sI_1) + V_1(p_1^f-p_0^f)N_1^2 \\
& + V_1\left[[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] + 2cov(c_1,p_1^sI_1) \\
& + 2cov[c_1,(p_1^f-p_0^f)]N_1 + 2cov\left[c_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] \\
& + 2cov[p_1^sI_1,(p_1^f-p_0^f)]N_1 + 2cov\left[p_1^sI_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] \\
& + 2cov\left[(p_1^f-p_0^f),[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]N_1 \qquad (72)
\end{aligned}
$$

Furthermore, we assume that X and $\sigma^2_{f_2}$ are independent of p_1^s and p_1^f. Also, there is no dependence between c_1 and the futures prices p_1^f and p_2^f which results in the fact that $cov[c_1,(p_1^f-p_0^f)]=0$ and $cov\left[c_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]=0$.

The resulting expected utility is:

$$
\begin{aligned}
EU_1 = {} & E_1c_1 + E_1p_1^sE_1I_1 + E_1(p_1^f-p_0^f)N_1 + E_1\left[[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] \\
& - (k/2)\left[V_1c_1 + V_1(p_1^sI_1) + V_1(p_1^f-p_0^f)N_1^2\right] \\
& - (k/2)V_1\left[[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] - kcov(c_1,p_1^sI_1) \\
& - kcov[p_1^sI_1,(p_1^f-p_0^f)]N_1 - kcov\left[p_1^sI_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] \\
& - kcov\left[(p_1^f-p_0^f),[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]N_1 \qquad (73)
\end{aligned}
$$

Now, we can optimise the expected utility over N_1 in order to get the optimal futures position.

$$
\begin{aligned}
\delta EU_1/\delta N_1 = {} & E_1(p_1^f-p_0^f) - k\sigma^2_{f_1}N_1 - kcov[p_1^sI_1,(p_1^f-p_0^f)] \\
& - kcov\left[(p_1^f-p_0^f),[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] = 0
\end{aligned}
$$

with $var(p_1^f) = \sigma^2_{f_1}$

which gives us:

$$
\begin{aligned}
N_1 = {} & E_1(p_1^f-p_0^f)/k\sigma^2_{f_1} - cov[p_1^sI_1,(p_1^f-p_0^f)]/\sigma^2_{f_1} \\
& - cov\left[(p_1^f-p_0^f),[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]/\sigma^2_{f_1} \qquad (74)
\end{aligned}
$$

If we compare the optimal futures position in (74) with the respective position in the one period model in equation (49) we observe that now we have an additional term which is the covariance between the futures price of the first period and the futures price of the second period, a pure speculative term which is not dependent of any inventory term.

If we substitute (74) back into (73) and simplify we get:[15]

$$\begin{aligned}
EU_1 = {} & E_1c_1 + E_1p_1^sE_1I_1 + E_1\Big[[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\Big] \\
& - (k/2)\Big[V_1c_1 + V_1(p_1^sI_1) + V_1\Big[[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\Big]\Big] \\
& - k\mathrm{cov}(c_1,p_1^sI_1) - k\mathrm{cov}\Big[p_1^sI_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\Big] \\
& + [E_1(p_1^f-p_0^f)]^2/2k\sigma^2_{f_1} - E_1(p_1^f-p_0^f)\mathrm{cov}[p_1^sI_1,(p_1^f-p_0^f)]/\sigma^2_{f_1} \\
& + k\mathrm{cov}[p_1^sI_1,(p_1^f-p_0^f)]\mathrm{cov}\Big[p_1^sI_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\Big]/\sigma^2_{f_1} \\
& + (k/2)\mathrm{cov}[p_1^sI_1,(p_1^f-p_0^f)]^2/\sigma^2_{f_1} \\
& - E_1(p_1^f-p_0^f)\mathrm{cov}\Big[p_1^sI_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\Big]/\sigma^2_{f_1} \\
& + (k/2)\mathrm{cov}\Big[p_1^sI_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\Big]^2/\sigma^2_{f_1} \qquad (75)
\end{aligned}$$

If we replace the expressions of expectations, variances and covariances by the terms calculated already in the symmetric one period model we can rewrite the expected utility as: [16]

$$\begin{aligned}
EU_1 = {} & c_0 + \lambda X(pa-pb) + \mu_sI_0 - (k/2)\Big[\lambda X^2\big[(1-\lambda)(pa^2+pb^2)+2\lambda papb\big]\Big] \\
& - (k/2)\Big[\sigma^2_{p_s}(2\lambda X^2 + I_0^2) + (\mu_s)^2 2\lambda X^2\Big] + \lambda\mu_sX^2(pa+pb) \\
& + E_1(p_1^f-p_0^f)^2/2k\sigma^2_{f_1} + k\mathrm{cov}(p_1^s,p_1^f)^2I_0^2/2\sigma^2_{f_1} \\
& - \Big[E_1(p_1^f-p_0^f)\mathrm{cov}(p_1^s,p_1^f)I_0\Big]/\sigma^2_{f_1} \\
& + [E_1(p_1^f-p_0^f)]^2/2k\sigma^2_{f_1} + E_1\Big[[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\Big]
\end{aligned}$$

---

[15]Details can be found in appendix G.

[16]Calculations see appendix G.

$$- (k/2)V_1 \left[[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] - kI_t cov\left[p_1^s,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]$$
$$+ (k/\sigma^2_{f_1})I_0 cov(p_1^s p_1^f) cov\left[p_1^f,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]$$
$$- [E_1(p_1^f-p_0^f)/\sigma^2_{f_1}] cov\left[p_1^f,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]^2$$
$$+ (k/2\sigma^2_{f_1}) cov\left[p_1^f,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]^2 \quad (76)$$

with: [17]

$$cov\left[p_1^s I_1,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] = I_0 cov\left[p_1^s,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] \text{ and}$$
$$cov\left[(p_1^f-p_0^f),[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right] = cov\left[p_1^f,[E_2(p_2^f-p_1^f)]^2/2k\sigma^2_{f_2}\right]$$

The last six terms on the right hand side of (79) represent the additional expressions for the second period.

If we examine the optimization problem of a dealer who decides not to be active in the spot market for the two periods considered, although holding an inventory of the risky asset, then her expected utility is:

$$EU_2(W_0) = c_1 + p_1^s I_1 + (p_1^f-p_0^f)N_1 + E_2(p_2^f-p_1^f)N_2 - (k/2)\sigma^2_{f_2}N_2^2 \quad (77)$$

We assume that the futures position N_1 taken at t matures at t+1; N_2 is assumed to be taken at t+1 and matures at t+2. We do not have the idea of any adjustment, except changes in the futures position due to changes in the price expectation, as the dealer does not face any order uncertainty without any trade in the spot market.

The change in the price expectation may come from the fact that the inactive dealer learns about the order flow which carries some information for the dealer to enable her to adjust her price expectation.

Following the same procedure as before we derive the optimal futures

---

[17]By symmetry, the result follows from the calculation of $cov(p_1^s I_1,(P_1^f-p_0^f))$ in appendix E.

position N_2, and N_1 which are: [18]

$$N_2 = E_2(p_2^f - p_1^f)/k\sigma^2_{f_2} \tag{78}$$

and
$$N_1 = E_1(p_1^f - p_0^f)/k\sigma^2_{f_1} - cov[p_1^s I_1, (p_1^f - p_0^f)]/\sigma^2_{f_1}$$
$$- cov\left[(p_1^f - p_0^f), [E_2(p_2^f - p_1^f)]^2 / 2k\sigma^2_{f_2}\right]/\sigma^2_{f_1} \tag{79}$$

The two optimal futures position are identical to the positions of the dealers who are active in the spot market except the adjustment term in N_2 which is missing in the position of the inactive dealer.

It is easy to see that also this time the futures market terms cancel out if we set the expected utility of an active trader equal to the expected utility of a trader who does not trade in the spot market.[19]

As a consequence we get the same result as in the one period model that the futures trading does not influence the spot bid-ask spread.

By modifying the model to two periods there are not any effects on the spot market bid-ask spread through trading in futures.

The modification of asymmetric purchases and sales in the spot market trading would not change the basic finding of no additional influence of futures trading with two periods as the structure of our model does not account for spot trading in the second period.

4.5. Conclusions

Our analysis of the influence of futures trading on the spot market bid-ask spread gives evidence that the organisation of the spot market is important.

---

[18]Mathematics can be found in appendix G.

[19]The respective utility $EU(W_0)$ is derived in appendix G.

If we consider a centralised market in which market makers have full information, i.e. the market is transparent, and the traders know the public orders then their futures trading may change the spot bid-ask spread.

For the case of a monopoly dealer, for instance a specialist dealer, the result of futures trading is always a reduction in the spot bid-ask spread as the only parameter coming into the function is the order size X, the spot price variance, the degree of risk aversion, and the correlation of the spot and the futures prices. The market maker is able to lay off some of the price risk, by hedging the inventory, and therefore is in the position to narrow the spread.

We can think of a market where traders are linked through a telematic circuit and they are bound to quote their prices on the screen which should be committing. Nevertheless, the actual trading occurs over the telephone. The reservation quotes on the screen enable the traders to deduce the inventory positions of their competitors.

This allows them to set their prices in a way to get the orders under a Bertrand type price competition which means that they are, depending on their inventory and risk aversion, in the position to quote the "best" prices (which means the lowest ask price and the highest bid price).

Investors are assumed to call the traders to inquire about the prices. In that way, traders learn about the public orders. Such a scenario represents our model in section 4.2. with competing dealers. If we analyse such a market the influence on the bid-ask spread through futures trading can be positive or negative depending on the covariance between the prices, the futures price expectations and the inventory positions of the dealers.

The important factors are the differences in inventories and the heterogeneous expectations among market makers.

If, for instance, the seller has a long inventory position and she expects the futures price to rise then she will increase the ask price. The dealer is less willing to sell today as she expects a higher futures price which makes it less attractive to sell today. On the other hand, if the buyer who is short in her inventory expects the futures price to fall she will decrease her bid price as it is less profitable to buy now.

If we analyse an over the counter market (which is not transparent at all and where the traders do not know about the public order flow) then the result changes.

According to the findings in our model in section 4.3., the futures trading does not have any influence on spot market trading if we assume that purchases and sales are symmetric. This comes from the fact that the expected inventory position with spot market trading is identical to the inventory position without any trading in the spot market. Hence there is no need for any hedging of the unknown spot market trading amount.

However, the symmetric order flow and the assumption of mean-variance preferences turn out to be very restrictive assumptions. The expected inventory of a dealer who is active is the same as the expected inventory of a dealer who does not trade at all. Hence, the futures market trading does not affect the spot market bid-ask spread.

In a dealership market, dealers may not be in the position to close out their position (i.e. balance their books) as quickly as they wish to do so, and therefore the model with asymmetric purchases and sales seems more applicable.

By changing this assumption for the over the counter market we again can show that futures trading influences the spot bid-ask spread.

This effect on the spread may be positive or negative which depends on the sign of the inventory, the covariance between the spot and the futures price, the futures price expectations, and the difference between purchases and sales in the spot market.

However, the model formulation does not allow us to determine unambiguously in which direction the spread is changed through futures trading.

In case of uncertainty of the order flow, the result is not changed by extending the model to two periods. There is no intertemporal effect of futures trading observed.

The conclusion we can draw from this analysis is that the influence of futures trading is subject to the organisation and market structure of the spot market trading. Depending on the information available in the market the spot market trader can set her bid and ask prices accordingly which may result in a larger spread if the dealer is also trading in futures.

We have shown that in both circumstances, full knowledge of the order flow or not knowing about the order flow, there is an influence on the bid-ask spread of the spot market. In the first case, we even showed that the bid-ask spread may be larger with trading in futures due to differences of inventories and heterogeneous expectations among competing market makers.

APPENDIX A

Non-Stochastic Order Flow: Section 4.2.

Calculation of the expected utility of a buyer with an optimal futures position of:

$$N = (\mu_f - p^f_0)/k\sigma_{p_f}^2 - \mathrm{cov}()(I+X)/\sigma_{p_f}^2$$

$$EU_B(W_1) = (c_0 - pbX) + \mu_s(I+X) - (k/2)\,\sigma_{p_s}^2(I+X)^2$$
$$+ (\mu_f - p^f_0)\left[(\mu_f - p^f_0)/k\sigma_{p_f}^2 - \mathrm{cov}()(I+X)/\sigma_{p_f}^2\right]$$
$$- k/2\;\sigma_{p_f}^2\left[(\mu_f - p^f_0)/k\sigma_{p_f}^2 - \mathrm{cov}()(I+X)/\sigma_{p_f}^2\right]^2$$
$$- k\mathrm{cov}()(I+X)\left[(\mu_f - p^f_0)/k\sigma_{p_f}^2 - \mathrm{cov}()(I+X)/\sigma_{p_f}^2\right]$$

which is:

$$EU_B(W_1) = (c_0 - pbX) + \mu_s(I+X) - (k/2)\,\sigma_{p_s}^2(I+X)^2$$
$$+ (\mu_f - p^f_0)^2/k\sigma_{p_f}^2 - (\mu_f - p^f_0)\mathrm{cov}()(I+X)/\sigma_{p_f}^2$$
$$- k\sigma_{p_f}^2(\mu_f - p^f_0)^2/2k^2(\sigma_{p_f}^2)^2$$
$$+ 2k\sigma_{p_f}^2(\mu_f - p_0)\mathrm{cov}()(I+X)/2k(\sigma_{p_f}^2)^2$$
$$- k\sigma_{p_f}^2\mathrm{cov}()^2(I+X)^2/2(\sigma_{p_f}^2)^2$$
$$- k(\mu_f - p^f_0)\mathrm{cov}()(I+X)/k\sigma_{p_f}^2 + k\mathrm{cov}()^2(I+X)^2/\sigma_{p_f}^2$$

After rearranging expected utility of a buyer is:

$$EU_B(W_1) = (c_0 - pbX) + \mu_s(I+X) - (k/2)\,\sigma_{p_s}^2(I+X)^2$$
$$+ 1/2[(\mu_f - p^f_0)^2/k\sigma_{p_f}^2] - (\mu_f - p^f_0)\mathrm{cov}()(I+X)/\sigma_{p_f}^2$$
$$+ 1/2\;k\mathrm{cov}()^2(I+X)^2/\sigma_{p_f}^2$$

Similarly the calculation of the expected utility of a seller is:

$$EU_S(W_1) = (c_0 + paX) + \mu_s(I-X) - (k/2)\,\sigma_{p_s}^2(I-X)^2$$
$$+ (\mu_f - p^f_0)\left[(\mu_f - p^f_0)/k\sigma_{p_f}^2 - \mathrm{cov}()(I-X)/\sigma_{p_f}^2\right]$$

$$- k/2\ \sigma_{p_f}^{\ 2}\left[(\mu_f - p^f_0)/k\sigma_{p_f}^{\ 2} - cov()(I-X)/\sigma_{p_f}^{\ 2}\right]^2$$

$$- kcov()(I-X)\left[(\mu_f - p^f_0)/k\sigma_{p_f}^{\ 2} - cov()(I-X)/\sigma_{p_f}^{\ 2}\right]$$

which is:

$$EU_s(W_1) = (c_0+paX) + \mu_s(I-X) - (k/2)\ \sigma_{p_s}^{\ 2}(I-X)^2$$
$$+ (\mu_f - p^f_0)^2/k\sigma_{p_f}^{\ 2} - (\mu_f - p^f_0)cov()(I-X)/\sigma_{p_f}^{\ 2}$$
$$- k\sigma_{p_f}^{\ 2}(\mu_f - p^f_0)^2/2k^2(\sigma_{p_f}^{\ 2})^2$$
$$+ 2k\sigma_{p_f}^{\ 2}(\mu_f - p^f_0)cov()(I-X)/2k(\sigma_{p_f}^{\ 2})^2$$
$$- k\sigma_{p_f}^{\ 2}cov()^2(I-X)^2/2(\sigma_{p_f}^{\ 2})^2$$
$$- k(\mu_f - p^f_0)cov()(I-X)k\sigma_{p_f}^{\ 2} + kcov()^2(I-X)^2/\sigma_{p_f}^{\ 2}$$

Finally we can simplify and we get:

$$EU_S(W_1) = (c_0+paX) + \mu_s(I-X) - (k/2)\ \sigma_{p_s}^{\ 2}(I-X)^2$$
$$+ (1/2)(\mu_f - p^f_0)^2/k\sigma_{p_f}^{\ 2} - (\mu_f - p^f_0)cov()(I-X)/\sigma_{p_f}^{\ 2}$$
$$+ (1/2)\ kcov()^2(I-X)^2/\sigma_{p_f}^{\ 2}$$

APPENDIX B

Known Order Flow: Spread Equation

The expected utility of a monopoly dealer is$EU_S(W_1)$ for selling and $EU_B(W_1)$ for buying. In order to calculate the spread we have $EU(W_1) = EU_S(W_1) + EU_B(W_1)$
which we can rewrite as:

$$EU(W_1) = (c_0+paX)+\mu_s(I-X)-(k/2)\sigma_{p_s}^{\ 2}(I-X)^2$$
$$+ (1/2)(\mu_f - p^f_0)^2/k\sigma_{p_f}^{\ 2} - (\mu_f - p^f_0)cov()(I-X)/\sigma_{p_f}^{\ 2}$$
$$+ (1/2)\ kcov()^2(I-X)^2/\sigma_{p_f}^{\ 2}$$
$$+ (c_0-pbX)+\mu_s(I+X)-(k/2)\sigma_{p_s}^{\ 2}(I+X)^2$$
$$+ (1/2)(\mu_f - p^f_0)^2/k\sigma_{p_f}^{\ 2} - (\mu_f - p^f_0)cov()(I+X)/\sigma_{p_f}^{\ 2}$$
$$+ (1/2)\ kcov()^2(I+X)^2/\sigma_{p_f}^{\ 2}$$

With $EU(W_1) \geq EU(W_0)$ and

$$EU(W_0) = 2\left[(c_0+\mu_s(I_0)-(k/2)\sigma_{p_s}^{2}(I_0)^2 + (1/2)(\mu_f-p^f_0)^2/k\sigma_{p_f}^{2} - (\mu_f-p^f_0)cov()(I_0)/\sigma_{p_f}^{2} + (1/2)\, kcov()^2(I_0)^2/\sigma_{p_f}^{2}\right]$$

we can simplify and get

$$EU(W_1) = (pa-pb)X - k\sigma_{p_s}^{2}X^2 + kcov()^2X^2/\sigma_{p_f}^{2}$$

APPENDIX C

Model of Section 4.2.: Futures Market Equilibrium

Calculation of the futures price:

In equilibrium:

$$VN_B + WN_{NT} + YN_S + Z_{SP} = 0$$

By substituting the optimal futures positions into the equilibrium equation above and solving for p^f_0, we get:

$$V\left[(\mu^f_B-p^f_0)/k_b\sigma_{f_B}^{2} - cov(b)(I_B+X)/\sigma_{f_B}^{2}\right]$$
$$+ W\left[(\mu^f_{NT}-p^f_0)/k_{NT}\sigma_{f_{NT}}^{2} - cov(b)(I_{NT})/\sigma_{f_{NT}}^{2}\right]$$
$$+ Y\left[(\mu^f_S-p^f_0)/k_S\sigma_{f_S}^{2} - cov(a)(I_S-X)/\sigma_{f_S}^{2}\right]$$
$$+ Z\left[(\mu^f_{SP}-p^f_0)/k_{SP}\sigma_{f_{SP}}^{2}\right]=0$$

which is:

$$p^f_0\left[(V/k_b\sigma_{f_B}^{2})+(W/k_{NT}\sigma_{fNT}^{2})+(Y/k_S\sigma_{f_S}^{2})+(Z/k_{SP}\sigma_{f_{SP}}^{2})\right]$$
$$= (V/k_b\sigma_{f_B}^{2})\left[(\mu^f_B)-k_bcov(b)(I_B+X)\right]$$
$$+ (W/k_{NT}\sigma_{f_{NT}}^{2})\left[(\mu^f_{NT})-k_{NT}cov(NT)(I_{NT})\right]$$
$$+ (Y/k_S\sigma_{f_S}^{2})\left[(\mu^f_S)-k_Scov(a)(I_S-X)\right] + (Z/k_{SP}\sigma_{f_{SP}}^{2})(\mu^f_{SP})$$

which we can rewrite as:

$$p^f_0 = S\Big[(V/k_b\sigma_{f_B}^{2})\left[(\mu^f_B)-k_bcov(b)(I_B+X)\right]$$
$$+ (W/k_{NT}\sigma_{f_{NT}}^{2})\left[(\mu^f_{NT})-k_{NT}cov(NT)(I_{NT})\right] + (Y/k_S\sigma_{f_S}^{2})\left[(\mu^f_S)-k_Scov(a)(I_S-X)\right]$$
$$+ (Z/k_{SP}\sigma_{f_{SP}}^{2})(\mu^f_{SP})\Big]$$

with

$$S = \left[(V/k_b\sigma_{f_B}{}^2)+(W/k_{NT}\sigma_{fNT}{}^2)+(Y/k_S\sigma_{f_S}{}^2) +(Z/k_{SP}\sigma_{f_{SP}}{}^2)\right]^{-1}$$

Derivation of p^f_0 under homogeneous beliefs among dealers and identical degrees of risk aversion:

$$p^f_0 = S /k\sigma^2_{p_f}\Big[V\mu^f - Vkcov()(I_B+X) + W\mu^f - Wkcov()I_{NT}$$
$$+ Y\mu^f - Ykcov()(I_S-X) + Z\mu^f\Big]$$

with $S = 1/(V+W+Y+Z)k\sigma^2_{p_f} = 1/Mk\sigma^2_{p_f}$, hence

$$p^f_0 =1/M \left[\mu^f M-kcov()\left[V(I_B+X)+WI_{NT}+Y(I_S-X)\right]\right]$$

and finally

$$p^f_0 = \mu^f - (kcov()/M)\left[V(I_B+X) + WI_{NT} + Y(I_S-X)\right]$$

APPENDIX D

Robustness Analysis

Futures price and varying numbers of agents:

V	1	1	0	1	1
W	1	1	1	0	1
Y	1	1	1	1	0
Z	1	0	1	1	1
k	0.5	0.5	0.5	0.5	0.5
Is	0.65	0.65	0.65	0.65	0.65
Ib	0.2	0.2	0.2	0.2	0.2
I NT	0.15	0.15	0.15	0.15	0.15
X	0.1	0.1	0.1	0.1	0.1
E(pf)	90.1	90.1	90.1	90.1	90.1
var(pf)	0.03	0.03	0.03	0.03	0.03
cov()	0.16	0.16	0.16	0.16	0.16
M	4	3	3	3	3
pf0	90.08	90.07333	90.08133	90.07733	90.088

Futures price and varying numbers of agents:

V	2	1	1	1	1
W	1	2	1	1	1
Y	1	1	2	1	2
Z	1	1	1	2	2
k	0.5	0.5	0.5	0.5	0.5
Is	0.65	0.65	0.65	0.65	0.65
Ib	0.2	0.2	0.2	0.2	0.2
I NT	0.15	0.15	0.15	0.15	0.15
X	0.1	0.1	0.1	0.1	0.1
E(pf)	90.1	90.1	90.1	90.1	90.1
var(pf)	0.03	0.03	0.03	0.03	0.03
cov()	0.16	0.16	0.16	0.16	0.16
M	5	5	5	5	6
pf0	90.0792	90.0816	90.0752	90.084	90.07933

APPENDIX E

Stochastic and Symmetric Order Flow: Section 4.3.

$$E(c_1) = \lambda(c_0+paX) + \lambda(c_0-pbX) + (1-2\lambda)c_0$$
$$= \lambda c_0+\lambda Xpa+\lambda c_0-\lambda Xpb+c_0 -2\lambda c_0 = c_0 + \lambda X(pa-pb)$$

$$E(I_1) = \lambda(I_0+X)+\lambda(I_0-X)+(1-2\lambda)I_0 = \lambda I_0+\lambda X+\lambda I_0-\lambda X+I_0-2\lambda I_0 = I_0$$
$$E(p_1^s) = \mu_s, \; E(p_1^f) = \mu_f$$

$$V(c_1) = E[c_1 - E(c_1)]^2 = \lambda\left[paX - \lambda X(pa-pb)\right]^2$$
$$+ \lambda\left[-pbX - \lambda X(pa-pb)\right]^2 + (1-2\lambda)\left[-\lambda X(pa-pb)\right]^2$$
$$= \lambda\Big[pa^2X^2-2\lambda pa^2X^2+\lambda^2pa^2X^2+2\lambda papbX^2-2\lambda^2papbX^2+\lambda^2pb^2X^2$$
$$+\lambda^2pb^2X^2-2\lambda pb^2X^2+pb^2X^2-2\lambda^2papbX^2+2\lambda papbX^2+\lambda^2pa^2X^2$$
$$+\lambda pb^2X^2-2\lambda papbX^2+\lambda pa^2X^2+4\lambda^2papbX^2-2\lambda^2pb^2X^2-2\lambda^2pa^2X^2\Big]$$
$$= \lambda\left[pa^2X^2+pb^2X^2-\lambda pa^2X^2+2\lambda papbX^2-\lambda pb^2X^2\right]$$
$$= \lambda X^2\left[(pa^2 + pb^2) - \lambda(pa - pb)^2\right]$$

$$V(I_1) = E[I_1 - E(I_1)]^2 = \lambda\left[(I_0-X)^2-I_0^2 + (I_0+X)^2-I_0^2\right]$$
$$+ (1-2\lambda)(I_0-I_0)$$
$$= 2\lambda X^2$$

We assume that p_1^s is independent of I_1. Thus we have

$$V(p_1^sI_1) = E\left[(p_1^s)^2(I_1)^2 - (Ep_1^sEI_1)^2\right]$$
$$= E(p_1^{s^2} I_1^2) - E(p_1^s)^2E(I_1)^2$$
$$= E(p_1^{s^2})E(I_1^2) - E(p_1^s)^2E(I_1)^2$$
$$= \left[var(p_1^s)+(E(p_1^s)^2)\right]\left[var(I_1)+(E(I_1)^2)\right] - (E(p_1^s)^2)(E(I_1)^2)$$
$$= var(p_1^s)var(I_1)+(E(I_1)^2)var(p_1^s)+var(I_1)(E(p_1^s)^2)$$
$$= \sigma^2_{p_s}(2\lambda X^2) + (I_0^2\sigma^2_{p_s}) + (2\lambda X^2)\mu_s^2$$
$$= \sigma^2_{p_s}(2\lambda X^2 + I_0^2) + \mu_s^2(2\lambda X^2)$$

$$\operatorname{cov}(c_1, p_1^s I_1) = E\left[(c_1 - E(c_1))(p_1^s I_1 - E(p_1^s)E(I_1))\right]$$

$$=E\Big\{\lambda paXp_1^s(I_0-X)-\lambda paX\mu_s I_0-\lambda^2(pa-pb)Xp_1^s(I_0-X)+\lambda^2(pa-pb)X\mu_s I_0$$

$$-\lambda pbXp_1^s(I_0+X)+\lambda pbX\mu_s I_0-\lambda^2(pa-pb)Xp_1^s(I_0+X)+\lambda^2(pa-pb)X\mu_s I_0$$

$$-\lambda(pa-pb)Xp_1^s(I_0)+\lambda(pa-pb)X\mu_s(I_0)+2\lambda^2(pa-pb)Xp_1^s(I_0)$$

$$-2\lambda^2(pa-pb)X\mu_s(I_0)\Big\}$$

Take expectations over p_1^s:

$$= \lambda paX\mu_s(I_0-X)-\lambda paX\mu_s(I_0)-\lambda pbX\mu_s(I_0+X)+\lambda pbX\mu_s(I_0)$$

$$-\lambda^2(pa-pb)X\mu_s(I_0-X)-\lambda^2(pa-pb)X\mu_s(I_0+X)+2\lambda^2(pa-pb)X\mu_s(I_0)$$

$$= -\lambda paX^2\mu_s - \lambda pbX^2\mu_s = -\lambda X^2\mu_s(pa+pb)$$

$$\operatorname{cov}(c_1,(p_1^f-p_0^f)) = E\left[(c_1 - E(c_1))[(p_1^f-p_0^f) - E(p_1^f-p_0^f)]\right]$$

$$= E\Big\{\lambda paXp_1^f-\lambda paXE(p_1^f) -\lambda^2X(pa-pb)p_1^f+\lambda^2X(pa-pb)E(p_1^f)$$

$$-\lambda pbXp_1^f+\lambda pbXE(p_1^f) -\lambda^2X(pa-pb)p_1^f+\lambda^2X(pa-pb)E(p_1^f)$$

$$-\lambda X(pa-pb)p_1^f+\lambda X(pa-pb)E(p_1^f)+2\lambda^2X(pa-pb)p_1^f$$

$$-2\lambda^2X(pa-pb)E(p_1^f)\Big\}$$

Take expectations over p_1^f:

$$= \lambda paXE(p_1^f)-\lambda paXE(p_1^f) -\lambda^2X(pa-pb)E(p_1^f)+\lambda^2X(pa-pb)E(p_1^f)$$

$$-\lambda pbXE(p_1^f)+\lambda pbXE(p_1^f) -\lambda^2X(pa-pb)E(p_1^f)+\lambda^2X(pa-pb)E(p_1^f)$$

$$-\lambda X(pa-pb)E(p_1^f)+\lambda X(pa-pb)E(p_1^f)+2\lambda^2X(pa-pb)E(p_1^f)$$

$$-2\lambda^2X(pa-pb)E(p_1^f) = 0$$

which is a result we could expect as X and p_1^f are the only random variables which are independent of each other.

We still have $E(I_1) = I_0$ and thus

$$\operatorname{cov}(p_1^s I_1,(p_1^f-p_0^f)) = E\left[(p_1^s I_1 - E(p_1^s)E(I_1))[(p_1^f-p_0^f) - E(p_1^f-p_0^f)]\right]$$

$$= E\Big\{\lambda(I_0-X)p_1^s p_1^f-\lambda(I_0-X)p_1^s E(p_1^f)-E(p_1^s)I_0p_1^f$$

$$+ \quad E(p_1^s)I_0E(p_1^f)+\lambda(I_0+X)p_1^s p_1^f-\lambda(I_0+X)p_1^s E(p_1^f)$$

$$- E(p_1^s)I_0p_1^f+E(p_1^s)I_0E(p_1^f)+p_1^sI_0p_1^f-p_1^sI_0E(p_1^f)$$
$$- E(p_1^s)I_0p_1^f+E(p_1^s)I_0E(p_1^f)-2\lambda p_1^sI_0p_1^f+2\lambda p_1^sI_0E(p_1^f)$$
$$+ 2\lambda E(p_1^s)I_0p_1^f-2\lambda E(p_1^s)I_0E(p_1^f)\Big\}$$
$$= E\left\{\left[p_1^sp_1^f-p_1^sE(p_1^f)\right]\left[\lambda(I_0-X)+\lambda(I_0+X)+I_0-2\lambda I_0\right]\right.$$
$$\left.- \left[E(p_1^s)p_1^f-E(p_1^s)E(p_1^f)\right]\left[3I_0-2\lambda I_0\right]\right\}$$

Take expectations over p_1^s and p_1^f:

$$= \left[E(p_1^sp_1^f)-E(p_1^s)E(p_1^f)\right]I_0 - \left[E(p_1^s)E(p_1^f)-E(p_1^s)E(p_1^f)\right]\left[I_0(3-2\lambda)\right]$$

we can rewrite $E(p_1^sp_1^f) = cov(p_1^sp_1^f)+E(p_1^s)E(p_1^f)$ and finally we get:

$cov(p_1^sI_1,(p_1^f-p_0^f)) = cov(p_1^sp_1^f)I_0$

The expected utility is:

$$EU(W_1) = c_0 + \lambda X(pa-pb) + \mu_sI_0 + (\mu_f-p_0^f)N$$
$$- (k/2)\lambda X^2\left[(1-\lambda)(pa^2 + pb^2) + 2\lambda papb\right]$$
$$- (k/2)\left[\sigma^2_{p_s}(2\lambda X^2 + I_0^2) + \mu_s^2(2\lambda X^2)\right] - (k/2)\sigma^2_{p_f}N^2$$
$$+ k\lambda X^2\mu_S(pa+pb) - kcov(p_1^sp_1^f)I_0N$$

Substituting the optimal futures position N which is

$N = (\mu_f-p_0^f)/k\sigma^2_{p_f} - cov(p_1^sp_1^f)I_0/\sigma^2_{p_f}$ into $EU(W_1)$:

$$EU(W_1) = c_0 + \lambda X(pa-pb) + \mu_sI_0 - (k/2)\lambda X^2\left[(1-\lambda)(pa^2 + pb^2) + 2\lambda papb\right]$$
$$- (k/2)\left[\sigma^2_{p_s}(2\lambda X^2 + I_0^2) + \mu_s^2(2\lambda X^2)\right] + k\lambda X^2\mu_S(pa+pb)$$
$$+ (\mu_f-p_0^f)^2/k\sigma^2_{p_f} - (\mu_f-p_0^f)cov(p_1^sp_1^f)I_0/\sigma^2_{p_f}$$
$$- (k/2(\sigma^2_{p_f})^2)\sigma^2_{p_f}(\mu_f-p_0^f)^2 + k\sigma^2_{p_f}(\mu_f-p_0^f)cov(p_1^sp_1^f)I_0/k(\sigma^2_{p_f})^2$$
$$- (k/2)\sigma^2_{p_f}cov(p_1^sp_1^f)^2I_0^2/(\sigma^2_{p_f})^2$$
$$- k(\mu_f-p_0^f)cov(p_1^sp_1^f)I_0/k\sigma^2_{p_f} + kcov(p_1^sp_1^f)^2I_0^2/\sigma^2_{p_f}$$
$$= c_0 + \lambda X(pa-pb) + \mu_sI_0 - (k/2)\lambda X^2\left[(1-\lambda)(pa^2 + pb^2) + 2\lambda papb\right]$$
$$- (k/2)\left[\sigma^2_{p_s}(2\lambda X^2 + I_0^2) + \mu_s^2(2\lambda X^2)\right] + k\lambda X^2\mu_S(pa+pb)$$
$$+ (\mu_f-p_0^f)^2/k\sigma^2_{p_f} - (\mu_f-p_0^f)cov(p_1^sp_1^f)I_0/\sigma^2_{p_f}$$

$$- (\mu_f - p_0^f)^2/2k\sigma^2_{p_f} + (\mu_f - p_0^f)cov(p_1^s p_1^f)I_0/\sigma^2_{p_f}$$
$$+ (k/2)cov(p_1^s p_1^f)^2 I_0^2/\sigma^2_{p_f} - (\mu_f - p_0^f)cov(p_1^s p_1^f)I_0/\sigma^2_{p_f}$$

Finally, we can simplify and get:

$$EU(W_1) = c_0 + \lambda X(pa-pb) + \mu_s I_0 - (k/2)\lambda X^2\left[(1-\lambda)(pa^2 + pb^2) + 2\lambda papb\right]$$
$$- (k/2)\left[\sigma^2_{p_s}(2\lambda X^2 + I_0^2) + \mu_s^2(2\lambda X^2)\right] + k\lambda X^2\mu_S(pa+pb)$$
$$+ (\mu_f - p_0^f)^2/2k\sigma^2_{p_f} + (k/2)cov(p_1^s p_1^f)^2 I_0^2/\sigma^2_{p_f}$$
$$- (\mu_f - p_0^f)cov(p_1^s p_1^f)I_0/\sigma^2_{p_f}$$

If we set the dealers expected utility equal with the expected utility of no trade we obtain the influence of futures trading on the spot spread.

The expected utility of a trader who is not active in the spot market simplifies a lot as there is no variance in the inventory and the cash position.

Hence $V(c_1)$, $V(I_1)$, $cov(c_1,p_1^s I_1)$, and $cov(c_1,(p_1^f - p_0^f))$ are zero. $V(p_1^s I_1)$ is reduced to $I_0^2\sigma^2_{p_s}$.

The covariance $cov(p_1^s I_1,(p_1^f - p_0^f))$is:

$$cov(p_1^s I_1,(p_1^f - p_0^f)) = E\left[(p_1^s I_1 - E(p_1^s)E(I_1))[(p_1^f - p_0^f) - E(p_1^f - p_0^f)]\right]$$
$$= p_1^s I_0 p_1^f - p_1^s I_0 E(p_1^f) - E(p_1^s)I_0 p_1^f + E(p_1^s)I_0 E(p_1^f)$$

if we take expectations over p_1^s and p_1^f we get

$$= E(p_1^s p_1^f)I_0 - E(p_1^s)E(p_1^f)I_0 - I_0[E(p_1^s)E(p_1^f) - E(p_1^s)E(p_1^f)]$$
$$= I_0[E(p_1^s p_1^f) - E(p_1^s)E(p_1^f)]$$
$$= I_0[cov(p_1^s p_1^f) + E(p_1^s)E(p_1^f) - E(p_1^s)E(p_1^f)]$$
$$= cov(p_1^s p_1^f)I_0$$

Thus the expected utility of a non-active trader in the spot market is:

$$EU(W_0) = c_0 + \mu_s I_0 + (\mu_f - p_0^f)N - (k/2)\sigma^2_{p_s} I_0^2 - (k/2)\sigma^2_{p_f} N^2$$
$$- k\,cov(p_1^s p_1^f)I_0 N$$

Derivation of the optimal futures position:

$$\delta EU(W_0)/\delta N = (\mu_f - p_0^f) - k\sigma^2_{p_f} N - k\,cov(p_1^s p_1^f) I_0 = 0$$

$$N = (\mu_f - p_0^f)/k\sigma^2_{p_f} - cov(p_1^s p_1^f) I_0/\sigma^2_{p_f}$$

substituting back into $EU(W_0)$:

$$\begin{aligned} EU(W_0) = {} & c_0 + \mu_s I_0 - (k/2)\sigma^2_{p_s} I_0^2 + (\mu_f - p_0^f)^2/k\sigma^2_{p_f} \\ & - (\mu_f - p_0^f) cov(p_1^s p_1^f) I_0/\sigma^2_{p_f} - k\sigma^2_{p_f}(\mu_f - p_0^f)^2/2k^2(\sigma^2_{p_f})^2 \\ & + 2k\sigma^2_{p_f}(\mu_f - p_0^f) cov(p_1^s p_1^f) I_0/2k^2(\sigma^2_{p_f})^2 \\ & - k\sigma^2_{p_f} cov(p_1^s p_1^f)^2 I_0^2/2(\sigma^2_{p_f})^2 \\ & - k(\mu_f - p_0^f) cov(p_1^s p_1^f) I_0/k\sigma^2_{p_f} + k\,cov(p_1^s p_1^f)^2 I_0^2/\sigma^2_{p_f} \end{aligned}$$

$$\begin{aligned} EU(W_0) = {} & c_0 + \mu_s I_0 - (k/2)\sigma^2_{p_s} I_0^2 + (\mu_f - p_0^f)^2/2k\sigma^2_{p_f} \\ & - (\mu_f - p_0^f) cov(p_1^s p_1^f) I_0/\sigma^2_{p_f} + (k/2) cov(p_1^s p_1^f)^2 I_0^2/\sigma^2_{p_f} \end{aligned}$$

$EU(W_1) = EU(W_0)$:

$$\begin{aligned} & c_0 + \lambda X(pa-pb) + \mu_s I_0 - (k/2)\lambda X^2\left[(1-\lambda)(pa^2 + pb^2) + 2\lambda papb\right] \\ & \quad - (k/2)\left[\sigma^2_{p_s}(2\lambda X^2 + I_0^2) + \mu_s^2(2\lambda X^2)\right] + k\lambda X^2 \mu_S (pa+pb) \\ & \quad + (\mu_f - p_0^f)^2/2k\sigma^2_{p_f} + (k/2) cov(p_1^s p_1^f)^2 I_0^2/\sigma^2_{p_f} \\ & \quad - (\mu_f - p_0^f) cov(p_1^s p_1^f) I_0/\sigma^2_{p_f} \\ & \quad = c_0 + \mu_s I_0 - (k/2)\sigma^2_{p_s} I_0^2 + (\mu_f - p_0^f)^2/2k\sigma^2_{p_f} \\ & \quad - (\mu_f - p_0^f) cov(p_1^s p_1^f) I_0/\sigma^2_{p_f} + (k/2) cov(p_1^s p_1^f)^2 I_0^2/\sigma^2_{p_f} \\ & \quad = \lambda X(pa-pb) - (k/2)\lambda X^2\left[(1-\lambda)(pa^2+pb^2)+2\lambda papb\right] \\ & \quad - (k/2)\left[\sigma^2_{p_s}(2\lambda X^2) + \mu_s^2(2\lambda X^2)\right] + k\lambda X^2 \mu_S (pa+pb) \end{aligned}$$

APPENDIX F

Stochastic and Asymmetric Order Flow: Section 4.3.

$$\begin{aligned} E(c_1) &= \lambda(c_0 + paY) + \lambda(c_0 - pbX) + (1-2\lambda)c_0 \\ &= \lambda c_0 + \lambda Ypa + \lambda c_0 - \lambda Xpb + c_0 - 2\lambda c_0 = c_0 + \lambda(paY - pbX) \end{aligned}$$

$$E(I_1) = \lambda(I_0+X)+\lambda(I_0-Y)+(1-2\lambda)I_0 = \lambda I_0+\lambda X+\lambda I_0-\lambda Y+I_0-2\lambda I_0$$
$$= I_0 + \lambda(X-Y)$$

$$E(p_1^s) = \mu_s \ , \ E(p_1^f) = \mu_f$$

$$V(c_1) = E[c_1 - E(c_1)]^2 = \lambda\Big[paY - \lambda(paY-pbX)\Big]^2$$
$$+ \lambda\Big[-pbX - \lambda(paY-pbX)\Big]^2 + (1-2\lambda)\Big[-\lambda(paY-pbX)\Big]^2$$
$$= \lambda\Big[pa^2Y^2-2\lambda pa^2Y^2+\lambda^2pa^2Y^2+2\lambda paYpbX-2\lambda^2paYpbX+\lambda^2pb^2X^2$$
$$+\lambda^2pb^2X^2-2\lambda pb^2X^2+pb^2X^2-2\lambda^2paYpbX+2\lambda paYpbX+\lambda^2pa^2Y^2$$
$$+\lambda pb^2X^2-2\lambda paYpbX+\lambda pa^2Y^2+4\lambda^2paYpbX-2\lambda^2pb^2X^2-2\lambda^2pa^2Y^2\Big]$$
$$= \lambda\Big[pa^2Y^2+pb^2X^2-\lambda pa^2Y^2+2\lambda paYpbX-\lambda pb^2X^2\Big]$$
$$= \lambda(pa^2Y^2 + pb^2X^2) - \lambda^2(pa^2Y^2 + pb^2X^2) + 2\lambda^2paYpbX$$
$$= (\lambda - \lambda^2)(pa^2Y^2 + pb^2X^2) + 2\lambda^2paYpbX$$

$$V(I_1) = E[I_1 - E(I_1)]^2 = \lambda\Big[-Y-\lambda(X-Y)\Big]^2 + \lambda\Big[X-\lambda(X-Y)\Big]^2 + (1-2\lambda)(-\lambda(X-Y))^2$$
$$= \lambda\Big[\lambda^2Y^2-2\lambda Y^2+Y^2-2\lambda^2XY+2\lambda XY+\lambda^2X^2+X^2-2\lambda X^2+\lambda^2X^2+2\lambda XY$$
$$-2\lambda^2XY+\lambda^2Y^2+\lambda Y^2-2\lambda XY+\lambda X^2-2\lambda^2Y^2+4\lambda^2XY-2\lambda^2X^2\Big]$$
$$= \lambda X^2+\lambda Y^2-\lambda^2X^2-\lambda^2Y^2+2\lambda^2XY$$
$$= \lambda(X^2 + Y^2) - (\lambda X - \lambda Y)^2$$

$$V(p_1^sI_1) = E\Big[(p_1^s)^2(I_1)^2 - (Ep_1^sEI_1)^2\Big] = E(p_1^{s^2} I_1^2) - E(p_1^s)^2E(I_1)^2$$
$$= E(p_1^{s^2})E(I_1^2) - E(p_1^s)^2E(I_1)^2$$
$$= \Big[var(p_1^s)+(E(p_1^s)^2)\Big]\Big[var(I_1)+(E(I_1)^2)\Big] - (E(p_1^s)^2)(E(I_1)^2)$$
$$= var(p_1^s)var(I_1)+(E(I_1)^2)var(p_1^s)+var(I_1)(E(p_1^s)^2)$$
$$= \sigma^2_{p_s}\Big[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)+I_0^2)+2\lambda I_0(X-Y)+(\lambda X-\lambda Y)^2$$
$$+ \mu_s^2[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)]\Big]$$

$$V(p_1^sI_1) = (\sigma^2_{p_s} + \mu_s^2)\Big[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)\Big] + \sigma^2_{p_s}(I_0+\lambda(X-Y))^2$$

$$cov(c_1,p_1^sI_1) = E\Big[(c_1 - E(c_1))(p_1^sI_1 - E(p_1^s)E(I_1))\Big]$$
$$= E\ \Big\{\lambda p_1^spaY(I_0-Y)-\lambda^2p_1^spaY(I_0-Y)+\lambda^2p_1^spbX(I_0-Y)$$
$$- \mu_spaY(I_0+\lambda(X-Y))+\lambda^2\mu_spaY(I_0+\lambda(X-Y))$$

$$- \lambda^2\mu_s pbX(I_0+\lambda(X-Y))-\lambda p_1{}^s pbX(I_0+X)$$

$$- \lambda^2 p_1{}^s paY(I_0+X)+\lambda p_1{}^s pbX(I_0+X)+\lambda\mu_s pbX(I_0+\lambda(X-Y))$$

$$+ \lambda^2\mu_s paY(I_0+\lambda(X-Y))-\lambda^2\mu_s pbX(I_0+\lambda(X-Y))-\lambda p_1{}^s paYI_0$$

$$+ \lambda p_1{}^s pbXI_0+\lambda\mu_s paY(I_0+\lambda(X-Y))-\lambda\mu_s pbX(I_0+\lambda(X-Y))$$

$$+ 2\lambda^2 p_1{}^s paYI_0-\lambda^2 p_1{}^s pbXI_0-2\lambda^2\mu_s paY(I_0+\lambda(X-Y))$$

$$+ 2\lambda^2\mu_s pbX(I_0+\lambda(X-Y))\Big\}$$

Take expectations over $p_1{}^s$:

$$= \lambda\mu_s paY\ (I_0-Y-I_0+I_0+\lambda X-\lambda Y-I_0-\lambda X+\lambda Y)$$
$$- \lambda\mu_s pbX\ (I_0+\lambda X-\lambda Y-I_0-X+I_0-I_0-\lambda X+\lambda Y)$$
$$+ \lambda^2\mu_s paY\ (I_0+\lambda X-\lambda Y-I_0+Y-I_0-X+I_0+\lambda X-\lambda Y+I_0-I_0-\lambda X+\lambda Y)$$
$$+ \lambda^2\mu_s pbX\ (I_0-Y-I_0-\lambda X+\lambda Y+I_0+X-I_0-\lambda X+\lambda Y-I_0+I_0+\lambda X-\lambda Y)$$
$$= -\lambda\mu_s paY^2 - \lambda\mu_s pbX^2 + (\lambda^2\mu_s paY-\lambda^2\mu_s pbX)[(Y-X)+\lambda(X-Y)]$$

$$= -\lambda\mu_s(paY^2+pbX^2)+\lambda^2\mu_s(paY-pbX)[(Y-X)+\lambda(X-Y)]$$

$$= -\lambda\mu_s\Big[paY^2+pbX^2+(1-\lambda)\Big[\lambda paY^2+\lambda pbX^2-\lambda paXpbY-\lambda pbXY\Big]\Big]$$

$$= -\lambda\mu_s\Big[paY^2+pbX^2+(1-\lambda)\Big[\lambda(paY^2+pbX^2)-\lambda XY(pa+pb)\Big]\Big]$$

$$cov(p_1{}^s I_1,(p_1{}^f-p_0{}^f)) = E\Big[(p_1{}^s I_1 - E(p_1{}^s)E(I_1))[(p_1{}^f-p_0{}^f) - E(p_1{}^f-p_0{}^f)]\Big]$$

$$= E\ \Big\{\lambda(I_0-Y)p_1{}^s p_1{}^f-\lambda(I_0-Y)p_1{}^s E(p_1{}^f)-E(p_1{}^s)I_0 p_1{}^f$$

$$+ \qquad E(p_1{}^s)I_0E(p_1{}^f)+\lambda(I_0+X)p_1{}^s p_1{}^f-\lambda(I_0+X)p_1{}^s E(p_1{}^f)$$

$$- E(p_1{}^s)I_0p_1{}^f+E(p_1{}^s)I_0E(p_1{}^f)+p_1{}^s I_0 p_1{}^f-p_1{}^s I_0E(p_1{}^f)$$

$$- E(p_1{}^s)I_0p_1{}^f+E(p_1{}^s)I_0E(p_1{}^f)-2\lambda p_1{}^s I_0 p_1{}^f+2\lambda p_1{}^s I_0E(p_1{}^f)$$

$$+ 2\lambda E(p_1{}^s)I_0p_1{}^f-2\lambda E(p_1{}^s)I_0E(p_1{}^f)\Big\}$$

$$= E\ \Big\{\Big[p_1{}^s p_1{}^f-p_1{}^s E(p_1{}^f)\Big]\Big[\lambda(I_0-Y)+\lambda(I_0+X)+I_0-2\lambda I_0\Big]$$

$$-\Big[E(p_1{}^s)p_1{}^f-E(p_1{}^s)E(p_1{}^f)\Big]\Big[2\lambda(I_0+\lambda(X-Y))+(1-2\lambda)(I_0+\lambda(X-Y))\Big]\Big\}$$

Take expectations over $p_1{}^s$ and $p_1{}^f$:

$$= \Big[E(p_1{}^s p_1{}^f)-E(p_1{}^s)E(p_1{}^f)\Big](I_0+\lambda(X-Y))$$

$$- \Big[E(p_1{}^s)E(p_1{}^f)-E(p_1{}^s)E(p_1{}^f)\Big]\Big[(I_0+\lambda(X-Y))\Big]$$

we can rewrite $E(p_1^s p_1^f) = cov(p_1^s p_1^f)+E(p_1^s)E(p_1^f)$ and finally we get:

$cov(p_1^s I_1,(p_1^f - p_0^f)) = cov(p_1^s p_1^f)(I_0+\lambda(X-Y))$

The respective expected utility is:

$$
\begin{aligned}
EU(W_1) = {} & c_0 + \lambda(paY-pbX) + \mu_s(I_0+\lambda(X-Y)) + (\mu_f-p_0^f)N \\
& - (k/2)\left[(\lambda-\lambda^2)(pa^2Y^2+pb^2X^2)+2\lambda^2 paYpbX\right] \\
& - (k/2)\left[(\sigma^2_{p_s}+\mu_s^2)[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)]+\sigma^2_{p_s}(I_0+\lambda(X-Y))^2\right] \\
& - (k/2)\sigma^2_{p_f}N^2 - k\,cov(p_1^s p_1^f)(I_0+\lambda(X-Y))N \\
& - k\mu_S\left[paY^2+pbX^2+(1-\lambda)\left[\lambda(paY^2+pbX^2)-\lambda XY(pa+pb)\right]\right]
\end{aligned}
$$

Substituting the optimal futures position N which is

$N = (\mu_f-p_0^f)/k\sigma^2_{p_f} - cov(p_1^s p_1^f)(I_0+\lambda(X-Y))/\sigma^2_{p_f}$

into $EU(W_1)$:

$$
\begin{aligned}
EU(W_1) = {} & c_0 + \lambda(paY-pbX) + \mu_s(I_0+\lambda(X-Y)) \\
& - (k/2)\left[(\lambda-\lambda^2)(pa^2Y^2+pb^2X^2)+2\lambda^2 paYpbX\right] \\
& - (k/2)\left[(\sigma^2_{p_s}+\mu_s^2)[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)]+\sigma^2_{p_s}(I_0+\lambda(X-Y))^2\right] \\
& - k\mu_S\left[paY^2+pbX^2+(1-\lambda)\left[\lambda(paY^2+pbX^2)-\lambda XY(pa+pb)\right]\right] \\
& + (\mu_f-p_0^f)^2/k\sigma^2_{p_f} - (\mu_f-p_0^f)cov(p_1^s p_1^f)(I_0+\lambda(X-Y))/\sigma^2_{p_f} \\
& - (k/2)\sigma^2_{p_f}\left[(\mu_f-p_0^f)^2/k^2(\sigma^2_{p_f})^2\right. \\
& - 2(\mu_f-p_0^f)cov(p_1^s p_1^f)(I_0+\lambda(X-Y))/k(\sigma^2_{p_f})^2 \\
& \left. + cov(p_1^s p_1^f)^2(I_0+\lambda(X-Y))^2/(\sigma^2_{p_f})^2\right] \\
& - k(\mu_f-p_0^f)cov(p_1^s p_1^f)(I_0+\lambda(X-Y))/k\sigma^2_{p_f} \\
& + k\,cov(p_1^s p_1^f)^2(I_0+\lambda(X-Y))^2/\sigma^2_{p_f}
\end{aligned}
$$

We can rewrite and get:

$$
\begin{aligned}
EU(W_1) = {} & c_0 + \lambda(paY-pbX) + \mu_s(I_0+\lambda(X-Y)) \\
& - (k/2)\left[(\lambda-\lambda^2)(pa^2Y^2+pb^2X^2)+2\lambda^2 paYpbX\right] \\
& - (k/2)\left[(\sigma^2_{p_s}+\mu_s^2)[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)]+\sigma^2_{p_s}(I_0+\lambda(X-Y))^2\right]
\end{aligned}
$$

$$- k\mu_S\Big[paY^2+pbX^2+(1-\lambda)\Big[\lambda(paY^2+pbX^2)-\lambda XY(pa+pb)\Big]\Big]$$
$$+ (\mu_f-p_0^f)^2/2k\sigma^2_{p_f}+(k/2)cov(p_1^sp_1^f)^2(I_0+\lambda(X-Y))^2/\sigma^2_{p_f}$$
$$- (\mu_f-p_0^f)cov(p_1^sp_1^f)(I_0+\lambda(X-Y))/\sigma^2_{p_f}$$

The expected utility of no trade is:

$$EU(W_0) = c_0 + \mu_sI_0 - (k/2)\sigma^2_{p_s}I_0^2 + (\mu_f-p_0^f)^2/2k\sigma^2_{p_f}$$
$$- (\mu_f-p_0^f)cov(p_1^sp_1^f)I_0/\sigma^2_{p_f} + (k/2)cov(p_1^sp_1^f)^2I_0^2/\sigma^2_{p_f}$$

Hence $EU(W_1) = EU(W_0)$:

$$c_0 + \lambda(paY-pbX) + \mu_s(I_0+\lambda(X-Y)) - (k/2)\Big[(\lambda-\lambda^2)(pa^2Y^2+pb^2X^2)+2\lambda^2paYpbX\Big]$$
$$- (k/2)\Big[(\sigma^2_{p_s}+\mu_s^2)[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)]+\sigma^2_{p_s}(I_0+\lambda(X-Y))^2\Big]$$
$$- k\mu_S\Big[paY^2+pbX^2+(1-\lambda)\Big[\lambda(paY^2+pbX^2)-\lambda XY(pa+pb)\Big]\Big]$$
$$+ (\mu_f-p_0^f)^2/2k\sigma^2_{p_f}+(k/2)cov(p_1^sp_1^f)^2(I_0+\lambda(X-Y))^2/\sigma^2_{p_f}$$
$$- (\mu_f-p_0^f)cov(p_1^sp_1^f)(I_0+\lambda(X-Y))/\sigma^2_{p_f}$$
$$= c_0 + \mu_sI_0 - (k/2)\sigma^2_{p_s}I_0^2 + (\mu_f-p_0^f)^2/2k\sigma^2_{p_f}$$
$$- (\mu_f-p_0^f)cov(p_1^sp_1^f)I_0/\sigma^2_{p_f} + (k/2)cov(p_1^sp_1^f)^2I_0^2/\sigma^2_{p_f}$$

It is obvious that some terms cancel out and we end up with:

$$\lambda(paY-pbX) + \mu_s(\lambda(X-Y)) - (k/2)\Big[(\lambda-\lambda^2)(pa^2Y^2+pb^2X^2)+2\lambda^2paYpbX\Big]$$
$$- (k/2)\Big[(\sigma^2_{p_s}+\mu_s^2)[(\lambda X+\lambda Y)^2-\lambda^2(X^2+Y^2)] + \sigma^2_{p_s}(2\lambda I_0(X-Y)+\lambda^2(X-Y)^2)\Big]$$
$$- k\mu_S\Big[paY^2+pbX^2+(1-\lambda)\Big[\lambda(paY^2+pbX^2)-\lambda XY(pa+pb)\Big]\Big]$$
$$= - kcov(p_1^sp_1^f)^2[2\lambda I_0(X-Y)+\lambda^2(X-Y)^2]/2\sigma^2_{p_f}$$
$$+ (\mu_f-p_0^f)cov(p_1^sp_1^f)(\lambda(X-Y))/\sigma^2_{p_f}$$

APPENDIX G

Two Period Model: Section 4.4.

$$U_2(W) = c_1 + p_1^s I_1 + (p_1^f - p_0^f)N_1 + E_2(p_2^f - p_1^f)(N_1 + N_2)$$
$$- (k/2)\sigma^2_{f_2}(N_1 + N_2)^2$$

optimal futures position $N_2 = E_2(p_2^f - p_1^f)/k\sigma^2_{f_2} - N_1$

Substitution of N_2 into $U_2(W)$:

$$U^*_2(W) = c_1 + p_1^s I_1 + (p_1^f - p_0^f)N_1 + E_2(p_2^f - p_1^f)N_1$$
$$- E_2(p_2^f - p_1^f)N_1 + E_2(p_2^f - p_1^f)^2/k\sigma^2_{f_2} - (k/2)\sigma^2_{f_2}N_1^2$$
$$- k\sigma^2_{f_2}E_2(p_2^f - p_1^f)N_1/k\sigma^2_{f_2} + (k/2)\sigma^2_{f_2}N_1^2$$
$$- k\sigma^2_{f_2}E_2(p_2^f - p_1^f)^2/2k^2(\sigma^2_{f_2}) + k\sigma^2_{f_2}E_2(p_2^f - p_1^f)N_1/k\sigma^2_{f_2} - k\sigma^2_{f_2}N_1^2/2$$

we can simplify and get:

$$U^*_2(W) = c_1 + p_1^s I_1 + (p_1^f - p_0^f)N_1 + E_2(p_2^f - p_1^f)^2/2k\sigma^2_{f_2}$$

$$EU_1 = E_1c_1 + E_1p_1^s E_1 I_1 + E_1(p_1^f - p_0^f)N_1$$
$$+ E_1\left[E_2(p_2^f - p_1^f)^2/2k\sigma^2_{f_2}\right] - (k/2)V_1c_1 - (k/2)V_1(p_1^s I_1)$$
$$- (k/2)V_1(p_1^f - p_0^f)N_1^2 - (k/2)V_1\left[E_2(p_2^f - p_1^f)^2/2k\sigma^2_{f_2}\right]$$
$$- k\,cov(c_1, p_1^s I_1) - k\,cov(p_1^s I_1, (p_1^f - p_0^f))N_1$$
$$- k\,cov(p_1^s I_1, [E_2(p_2^f - p_1^f)^2/2k\sigma^2_{f_2}])$$
$$- k\,cov((p_1^f - p_0^f), [E_2(p_2^f - p_1^f)^2/2k\sigma^2_{f_2}])N_1$$

The optimal futures position N_1:

$$N_1 = E_1(p_1^f - p_0^f)/k\sigma^2_{f_1} - cov(p_1^s I_1, (p_1^f - p_0^f))/\sigma^2_{f_1}$$
$$- cov((p_1^f - p_0^f), [E_2(p_2^f - p_1^f)^2/2k\sigma^2_{f_2}])/\sigma^2_{f_1}$$

Replacing N_1 in EU_1 by the optimal futures position:

$$EU_1 = E_1c_1 + E_1p_1^s E_1 I_1 + E_1(p_1^f - p_0^f)^2/k\sigma^2_{f_1}$$

$$- E_1(p_1^f-p_0^f)\mathrm{cov}(p_1^sI_1,(p_1^f-p_0^f))/\sigma^2_{f_1}$$

$$- E_1(p_1^f-p_0^f)\mathrm{cov}((p_1^f-p_0^f),[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])/\sigma^2_{f_1}$$

$$+ E_1\left[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}\right]-(k/2)V_1c_1-(k/2)V_1(p_1^sI_1)$$

$$- k\sigma^2_{f_1}E_1(p_1^f-p_0^f)^2/2k^2(\sigma^2_{f_1})^2$$

$$+ 2k\sigma^2_{f_1}E_1(p_1^f-p_0^f)\mathrm{cov}(p_1^sI_1,(p_1^f-p_0^f))/2k(\sigma^2_{f_1})^2$$

$$- k\sigma^2_{f_1}\mathrm{cov}(p_1^sI_1,(p_1^f-p_0^f))^2/2(\sigma^2_{f_1})^2$$

$$+ 2k\sigma^2_{f_1}E_1(p_1^f-p_0^f)\mathrm{cov}((p_1^f-p_0^f),[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])/2k(\sigma^2_{f_1})^2$$

$$- 2k\sigma^2_{f_1}\mathrm{cov}(p_1^sI_1,(p_1^f-p_0^f))$$

$$(\mathrm{cov}((p_1^f-p_0^f),[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}]))/2(\sigma^2_{f_1})^2$$

$$- k\sigma^2_{f_1}\mathrm{cov}((p_1^f-p_0^f),[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])^2/2(\sigma^2_{f_1})^2$$

$$- (k/2)V_1\left[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}\right]-k\mathrm{cov}(c_1,p_1^sI_1)$$

$$- k\mathrm{cov}(p_1^sI_1,[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])$$

$$- kE_1(p_1^f-p_0^f)\mathrm{cov}(p_1^sI_1,(p_1^f-p_0^f))/k\sigma^2_{f_1}$$

$$+ k\mathrm{cov}(p_1^sI_1,(p_1^f-p_0^f))^2/\sigma^2_{f_1}$$

$$+ k\mathrm{cov}(p_1^sI_1,(p_1^f-p_0^f))\mathrm{cov}((p_1^f-p_0^f),[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])/\sigma^2_{f_1}$$

$$- kE_1(p_1^f-p_0^f)\mathrm{cov}((p_1^f-p_0^f),[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])/k\sigma^2_{f_1}$$

$$+ k\mathrm{cov}((p_1^f-p_0^f),[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])(\mathrm{cov}(p_1^sI_1,(p_1^f-p_0^f)))/\sigma^2_{f_1}$$

$$+ k\mathrm{cov}((p_1^f-p_0^f),[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])^2/\sigma^2_{f_1}$$

which we can rewrite as

$$EU_1 = E_1c_1 + E_1p_1^sE_1I_1 + E_1\left[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}\right]$$

$$- (k/2)V_1c_1 - (k/2)V_1(p_1^sI_1)$$

$$- (k/2)V_1\left[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}\right] - k\mathrm{cov}(c_1,p_1^sI_1)$$

$$- k\mathrm{cov}(p_1^sI_1,[E_2(p_2^f-p_1^f)^2/2k\sigma^2_{f_2}])+E_1(p_1^f-p_0^f)^2/2k\sigma^2_{f_1}$$

$$- E_1(p_1{}^f-p_0{}^f)cov(p_1{}^sI_1,(p_1{}^f-p_0{}^f))/\sigma^2_{f_1}$$

$$+ kcov(p_1{}^sI_1,(p_1{}^f-p_0{}^f))cov((p_1{}^f-p_0{}^f),[E_2(p_2{}^f-p_1{}^f)^2/2k\sigma^2_{f_2}])/\sigma^2_{f_1}$$

$$+ kcov(p_1{}^sI_1,(p_1{}^f-p_0{}^f))^2/2\sigma^2_{f_1}$$

$$- E_1(p_1{}^f-p_0{}^f)cov((p_1{}^f-p_0{}^f),[E_2(p_2{}^f-p_1{}^f)^2/2k\sigma^2_{f_2}])/\sigma^2_{f_1}$$

$$+ k[cov((p_1{}^f-p_0{}^f),[E_2(p_2{}^f-p_1{}^f)^2]/2k\sigma^2_{f_2}])^2/2\sigma^2_{f_1}$$

The next step is to replace the expectations, variances and covariances by the respective values which have been calculated under section 4.2. except:

$cov(p_1{}^sI_1,[E_2(p_2{}^f-p_1{}^f)^2/2k\sigma^2_{f_2}])$, and

$cov((p_1{}^f-p_0{}^f),[E_2(p_2{}^f-p_1{}^f)^2/2k\sigma^2_{f_2}])$

If we consider the calculation of $cov(p_1{}^sI_1,(p_1{}^f-p_0{}^f))$ then by symmetry we have:

$$cov(p_1{}^sI_1, [E_2(p_2{}^f-p_1{}^f)^2/2k\sigma^2_{f_2}]) = cov(p_1{}^s, [E_2(p_2{}^f-p_1{}^f)^2/2k\sigma^2_{f_2}])I_0$$

The utility of terminal wealth of a trader who is not active in the spot market, but does trade in futures is:

$$W_0 = c_1 + p_1{}^sI_1 + (p_1{}^f-p_0{}^f)N_1 + (p_2{}^f-p_1{}^f)N_2$$

There is no adjustment term in the second futures position.

$$EU(W_0) = c_1+p_1{}^sI_1+(p_1{}^f-p_0{}^f)N_1+E_2(p_2{}^f-p_1{}^f)N_2-(k/2)\sigma^2_{f_2}N_2{}^2$$

$$\delta EU(W_0)/\delta N_2 = E_2(p_2{}^f-p_1{}^f) - k\sigma^2_{f_2}N_2 = 0$$

$N_2 = E_2(p_2{}^f-p_1{}^f)/k\sigma^2_{f_2}$, substituting back we get

$$U^*_2 = c_1 + p_1{}^sI_1 + (p_1{}^f-p_0{}^f)N_1 + E_2(p_2{}^f-p_1{}^f)^2/k\sigma^2_{f_2}$$

$$- k\sigma^2_{f_2}E_2(p_2{}^f-p_1{}^f)^2/2k^2(\sigma^2_{f_2})^2$$

$$U^*_2 = c_1 + p_1{}^sI_1 + (p_1{}^f-p_0{}^f)N_1 + E_2(p_2{}^f-p_1{}^f)^2/2k\sigma^2_{f_2}$$

which turns out to be identical to the solution for a trader who is active in the spot market.

The optimization problem of N_1 is identical to the active dealer's problem. As a result, all the terms coming from the futures market activity are identical and cancel out if we set $EU_1 = EU(W_0)$.

PART TWO: EMPIRICAL PART

ANALYSES BASED ON DATA OF THE ITALIAN SECONDARY MARKET OF GOVERNMENT BONDS

CHAPTER FIVE

MARKET SEGMENTATION AND MARKET SPECIALISATION

5.1. Introduction

The aim of this chapter is to analyse the structure of the Italian secondary market for government bonds and to explore whether the market is segmented with different dealers specialising in different assets. If this is the case then there may be opportunities for profitable arbitrage between different bonds.

Using the cross-bond, cross-dealer data we investigate whether assets can be grouped by level of activity in it and by their returns based on the price quotes into distinct homogeneous groups. Similarly, dealers may be grouped according to the frequency of the quotes and the returns they make in various bonds.

A summary of the frequency of quotes of the market makers is given in table 5.5. in the appendix.

From table 5.5. we can see that some of the market makers are very active, i.e. quote prices on many assets and they quote prices continuously. Others are less active which means that they quote prices over a wide range of assets, but not continuously. There is however a third type of market maker to be distinguished which is specialised in a few assets only, but on which prices are quoted on a continuous basis.

We assume that all dealers know about the order flow of the various assets. As the data reveal, some of the dealers are very active in one or several bonds and they quote continuously over time in that asset.

Our hypothesis is that market makers who are very active in a particular bond gain better knowledge about the future price of the bond which we can regard as superior information compared to the other dealers in the market. This fact of asymmetry of information is recognized by all market makers and, as a consequence, the bid-ask spread is larger in such a bond as uninformed market makers face increased transaction costs due to this asymmetry of information. The "specialised" market maker with lower reservation prices than the uninformed dealers is in the position to undercut the uninformed dealers' prices and still make a profit. Such a profit comes from the difference between the spread quoted under such asymmetry of information and the spread without any asymmetric information in the market.

The better informed dealers should exhibit a different quoting pattern than the less informed and hence the dealers may be grouped accordingly.

This analysis of the structure of the Italian bond secondary market is carried out with daily time series data on bid and ask prices. Quotes for 15 Italian government bonds are available. This market consists of 18 market makers, so called primary dealers, who are obliged by regulation to quote their bid and ask prices for every day.

The quotes have been collected for the period from May 1988 until April 1989.

The types of government bonds are:

a) Floating rate credit certificates (CCT) with an annual coupon indexed on the base of treasury bills with a maturity of 12 months and without any tax levied on them. (asset numbers:12838, 12805, 12811, 12812, 12817, 12859, 12825).

b) Floating rate notes subject to a tax rate of 12.5% (asset numbers:13009, 13011, 13013), others are subject to a tax rate of 6.25% (12879,12882).

c) Treasury bonds (BTP) with semi annual coupons at a fixed rate (asset numbers:12616 and 12610 with a tax rate of 6.25%, and 12628, with a tax rate of 12.5%).

In the first part of this chapter (section 5.2.), we identify a market segmentation and a market specialization pattern based on a measure of the frequency of trade quotes. An alternative approach is taken in section 5.3. where the market segmentation is derived through the technique of cluster analysis.

Finally, in section 5.4., we test the hypothesis that bonds which are traded by "specialised" market makers should have a higher return (larger spread) than other bonds due to the asymmetry of information. So differences in asset returns for informed and uninformed dealers should give the same grouping as grouping by frequency of trading.

If we find confirmation of such grouping, we observe a market inefficiency and informed investors may profit from arbitrage opportunities.

An analysis of variance using regression methods on the returns of price quotes will provide further evidence of the accuracy of the segmentation derived in the previous sections based on the frequency of quoting. Section 5.5. contains the conclusions.

5.2. Market Analysis using the Frequency of Quotes

5.2.1. Market Segmentation

One criterion for segmenting the market is the frequency of price quotes for various assets[1]. The reason behind such a segmentation is that the market makers, by quoting their bid and ask prices, commit themselves to accepting trades and take a certain risk as there is uncertainty about the future price of the asset. Hence, market makers who actively quote prices for a particular asset are specialised in such an asset and they are

[1]The quoting pattern of the dealers is given in table 5.1. in the appendix.

assumed to have superior knowledge about the asset fundamentals and hence the future price. We can form the following groups of market makers, based on this criterion of "frequency of quotes".

Group I :(very actively quoting market makers)

If we try to identify a measure for activity, we can say that these market makers quote prices on more than 200 days for at least four assets and, in addition, that they quote prices on more than 100 days for at least two assets. Such activity can be seen not as trading activity but as quoting with the purpose of conveying information to the market. Such information can be to signal that the dealer has superior information about this particular bond.

Market makers: ICCS, ROLB, CIMI, SIGB.

Group II:

This time the measure is that the market maker quotes prices on more than 200 days for three or more assets:

Market makers: CRRU, BSMT, SPTR, BRCB, CTOS, BNAT, CRMU, MPSG.

If we look at the less active market makers who maybe concentrate their activity on a few assets only we can define the third group as follows:

Group III

These market makers quote their prices on more than 200 days for less than three assets:

Market makers: BCMT, BSSE, BNLT, BPNQ, NAPQ, BPMQ.

In order to define the activities in a particular asset we examine the number of active dealers in the market for one asset. The measure for the first asset group is that more than four market makers quote their prices on this asset on more than 200 days.

GROUP I: Asset number: 12882, 12805, 12812, 12628, 13009, 13013.

We can define a typical asset traded by most of the market makers throughout the whole period as an asset on which prices have been quoted by at least three market makers for more than 200 days and by at least three market makers for more than 100 days.

GROUP II:

asset number: 12817, 12811, 12879, 12610 (with differences in ask and bid)

For the rest of the assets we can say that they are not very actively traded, although some of them are traded by almost every market maker, but not continuously over time.

GROUP III:

asset number: 12825, 12859, 13011, 12838, 12616.

A summary of the classification is shown in table 5.1. below:

Table 5.1.: Dealer and asset specification

GROUPS OF MARKET MAKERS		GROUPS OF ASSETS	
GROUP I		GROUP I	
ICCS		12882	12628
ROLB		12812	13013
CIMI		13009	
SIGB		12805	
GROUP II		GROUP II	
CRRU	CTOS	12817	
BSMT	BNAT	12811	
SPTR	CRMU	12879	
BRCB	MPSG	12610	
GROUP III		GROUP III	
BCMT	NAPQ	12825	12859
BSSE	BPMQ	13011	
BNLT		12838	
BPNQ		12616	

5.2.2. Market Specialisation

In this section, we focus on the question of which market maker is specialized in which asset. Specialisation means that a dealer quotes prices for more than 200 days in a particular asset. As a result, we get the combination:

for asset 12838: market makers BCMT, BRCB, SIGB

for asset 12805: market makers BNAT, BRCB, CTOS, ICCS, MPSG

for asset 12811: market makers BNAT, NAPQ, BSMT,

for asset 12812: market makers BPNQ, NAPQ, CRMU, ICCS, MPSG

for asset 12817: market makers BCMT, CRMU, CTOS, ICCS, SPTR, and MPSG (but only on the ask price!)

for asset 12859: market makers BRCB, CRRU, SPTR, SIGB

for asset 13009: market makers BSSE, BSMT, CRMU, ROLB, ICCS,

for asset 13011: market makers BNLT, CIMI, ROLB,

for asset 13013: market makers BNLT, BSSE, CIMI, ROLB, MPSG

for asset 12825: market makers BPNQ, CRRU, SIGB

for asset 12879: market makers BPMQ, BSMT, CRRU,

for asset 12882: market makers BNAT, CRRU, CTOS, ROLB, SPTR,

for asset 12628: market makers BSMT, CIMI, ROLB, ICCS, SIGB

for asset 12616: market maker BPNQ

for asset 12610: market makers BPNQ, CIMI, BCMT, MPSG

As mentioned earlier, we have two types of assets, one is the floating rate credit certificate (CCT), the other one is the treasury bond (BTP).

The market makers active in CCT's taxed with a rate of 6.25% are:

BNAT, CRRU, CTOS, ROLB, SPTR, BPMQ, BSMT

Market makers dealing actively in CCT's with a tax rate of 12.5% are:

BSSE, BSMT, CRMU, ROLB, ICCS, BNLT, CIMI, MPSG

The dealers trading in CCT's which are not subject to any tax are:

BPNQ, NAPQ, CRMU, ICCS, MPSG, BNAT, BRCB, CTOS, BCMT, SPTR, CRRU, SIGB, NAPQ, BSMT

Dealers BSMT, CIMI, ROLB, ICCS, SIGB ,BPNQ, BCMT, and MPSG are active in the treasury bonds.

It is evident that more dealers trade in the tax free asset than in the taxable assets [2]. This is also true for the treasury bonds which are also subject to taxes where we find only few dealers trading actively. This is not surprising as the taxable bonds represent higher transaction costs for the market makers. To be active in the trading of an asset a dealer has to keep the asset on stock to meet the market demand. However, if a market maker holds the bond in her inventory she is subject to pay tax on the bond which increases her trading costs. Thus, market makers prefer to trade in a tax free bond instead.

5.3. Cluster Analysis

Section 5.2. found groupings of dealers and assets based on quote frequency; the aim of this section is to define a market segmentation and specialization based on cluster analysis.

Techniques for cluster analysis seek to separate a set of data into groups or clusters.

There are various cluster techniques developed, such as:

- Hierarchical techniques, in which the classes themselves are classified into groups, the process being repeated at different levels.
- Optimization-partitioning techniques, in which the clusters are formed by

---

[2] Particular bonds are subject to a tax on the yield.

optimization of a 'clustering criterion'; furthermore, the classes are mutually exclusive, thus forming a partition of the set of entities.

- Density or mode-seeking techniques, in which clusters are formed by searching for regions containing a relatively dense concentration of entities.
- Clumping techniques, in which the classes are clumps and can overlap.

There are more techniques, but these are the most developed ones. We have applied a hierarchical clustering technique which, in this case, seemed to be the most appropriate method as we do not have an apparent grouping of market makers and assets to start with.

Hierarchical clustering techniques may be subdivided into agglomerative methods which proceed by a series of successive fusions of N entities into groups, and divisive methods which partition the set of N entities successively into finer partitions.

The results of both agglomerative, and divisive techniques may be presented in the form of a dendrogram, which is a two-dimensional diagram illustrating the fusions or partitions which have been made at each successive level.

As mentioned before, the aim of cluster analysis is to arrange the N sampling units into more or less homogeneous groups. How this is done can vary. The general strategy is best appreciated in geometrical terms, with the N sampling units represented by points in a multidimensional space.

In agglomerative methods, these points initially represent N separate clusters, each containing one member. At each of N-1 stages, two clusters are fused into one bigger cluster, until at the final stage all units are fused into a single cluster.

a) The single link or nearest neighbour method

This method can be used both with similarity measures and with distance measures. Groups initially consisting of single individuals are fused according to the distance between their nearest members, the groups with the smallest distance being fused. Each fusion decreases by one the number of groups. For this method, then, the distance between groups is defined as the distance between their closest members.

b) The complete link or furthest neighbour method

This method is exactly the opposite of the single linkage method, in that distance between groups is now defined as the distance between their most remote pair of individuals. This method can also be used with similarity and distance measures.

There are other measures such as the centroid cluster analysis, where the distance between groups is defined as the distance between the group centroids. The procedure then is to fuse groups according to the distance being fused first.

However, a disadvantage of the centroid method is that if the sizes of the two groups to be fused are very different the centroid of the new group will be very close to that of the larger group and may remain within that group; the characteristics of the smaller group are then virtually lost.

Another technique, the median cluster analysis, tries to overcome this problem. The strategy can be made independent of group size, the apparent position of the new group will then always be between the two groups to be fused. Although this method could be made suitable for both similarity and distance measures, it should be regarded as incompatible for measures such as correlation coefficients, since interpretation in a geometrical sense is no longer possible.

c) The group average method

One last method to be mentioned is the group average method. This method defines the distance between groups as the average of the distances between all pairs of individuals in the two groups.

Various methods have been used to evaluate the groups according to the cluster analysis.

The analysis of the Italian data starts by using hierarchical cluster analysis and we have compared the results from the single link method, the complete link method and the group average method.

All these methods have been carried out on the basis of a symmetric matrix obtained by measuring the distances between the number of market maker's quotations. The measure of the distance was the squared Euclidean measure summed over all the assets:

$$D_{ij} = 1 - \sum_{k} \left(\frac{x_{ik} - x_{jk}}{\text{range}} \right)^2 \quad (1)$$

where i and j are dealers and $i \neq j$

and k are the assets with $k = 1,...,15$ and

x are the number of price quotes for an asset.

Results from the hierarchical cluster analysis are shown in table 5.2..

The figures following the market maker's name indicate the average similarity of each group member with the other group members relative to a measure of dispersion within the group: this is the mean of the D_{ij} across j for members i and j of the same group. It helps to identify typical members of each group.

Table 5.2. : Cluster grouping

single link	complete link	group average
GROUP I	GROUP I	GROUP I
CTOS 76.6	CTOS 80.3	CTOS 80.3
SPTR 72.3	CRMU 76.1	CRMU 76.1
CRMU 72.2	MPSG 75.3	MPSG 75.3
BNAT 72.2	ICCS 74.7	ICCS 74.7
ICCS 70.7	BCMT 73.5	BCMT 73.5
BCNT 70.4	BRBC 73.4	BRBC 73.4
NAPQ 70.4	NAPQ 72.7	NAPQ 72.7
BRBC 70.0	BNAT 71.3	BNAT 71.3
BPMQ 69.3	SPTR 71.2	SPTR 71.2
MPSG 68.1	BPNQ 68.3	BPNQ 68.3
CRRU 65.4		
SIGB 65.1		
BSMT 64.0		
GROUP II	GROUP II	GROUP II
BPNQ 100	ROLB 83.5	BNLT 76.5
	BNLT 82.1	CIMI 76.4
	CIMI 81.0	ROLB 76.1
	BSSE 76.7	BSSE 75.4
		BPMQ 71.1
		BSMT 70.9
GROUP III	GROUP III	GROUP III
ROLB 83.5	BPMQ 74.7	CRRU 80.1
BNLT 82.1	CRRU 74.0	SIGB 80.1
CIMI 81.0	SIGB 68.7	
BSSE 76.7	BSMT 67.4	

From the table above we can see that the single link method leads to a totally different result compared with the other two methods. In addition, the single link method is creating one group with one member only, i.e. group II with the market maker BPNQ.

We can examine which of the dealer groups are determined by each variate (asset). We can get a frequency table for each variate showing the frequency with which each dealer quotes in each variate. Each table is classified by the grouping factor and the different values of the variate between group members.

For each group we then get an interaction statistic (chi-squared) which draws attention to groups for which the distribution within groups is markedly different from the overall distribution.

Only a few assets show a high chi-squared value which means that they belong to a typical group and that the group behaves differently to others. We list the results of the complete link and the group average link only and skip the results from the single link method.

There is a specialization of different groups for various assets.

The analysis based on the complete link method shows the following:

Group I is specialized in assets 12812, and 12817

Group II is specialized in assets 13009, 13013, and 13011

Group III is specialized in asset 12879

It is encouraging that these groupings make sense:

group I specializes in non-taxable assets whereas groups II and III are specialized in taxable assets.

The result of the group average method is:

Group I is specialized in assets 12812, 12817, and 12805

Group II is specialized in assets 13009, 13013, and 13011

For group III there is no conclusive answer.

Also in this case the tax status of the asset appears crucial in determining the groupings.

We tried to get a better result by applying the non-hierarchical cluster analysis. However, by comparing the outcome of such a cluster analysis with the results discussed above we could not find any improvement in the formation of groups and the significance level of specialization.

The non-hierarchical classification methods differ according to the criterion that they optimize and in the algorithm used to search for an optimum value of the chosen criterion.

5.4. Market Segmentation Measured by the Returns Based on Price Quotes

5.4.1. Methodology

In this section, we try to find a grouping pattern of the various market makers by analyzing their price quotes for different assets.

Such spread differences could arise from information or risk aversion differences between dealers.

In order to distinguish the quoting behaviour of the dealers we regress the returns based on their quotes for each asset on dealer specific variables and a trend variable which is different for each dealer.

The return (r) is defined to be the return to an investor who buys the asset i at time t-1 from dealer j and sells it at time t to dealer j at the quoted prices, i.e.

$$r_{(ij)} = (pa_{t,ij} - pb_{t-1,ij})/pb_{t-1,ij} \qquad (2)$$

The general model includes all the dealer specific variables and the asset specific variables which can be expressed as

a) General model:

$$r_{(ij)t} = (\beta_i + \beta_j)\ \delta + (c_i + c_j) + \varepsilon_{(ij)t} \qquad (3)$$

where δ is a trend variable and c is a constant term

i = 1,...,15 (number of assets)

j = 1,...,18 (number of dealers)

We encounter a problem with the price quotes of the 15 dealers over 18 assets which is that for some assets or some dealers there may be no quote at all for consecutive days. The reason behind this is that the dealers are obliged to quote prices for only any five assets on any day. The five assets chosen by a dealer may vary from day to day.

In order to avoid any bias coming from all the zeros in the time series (no quotes) we multiply all the regressors and the regressand by a dummy variable δ which is zero if the dealer does not quote and which is 1 if the particular dealer quotes a price in the particular asset at this day and then correct the degrees of freedom for the number of effective observations. [3]

We then test restrictions on the model by assuming that

b) all dealers are identical: this reduces the number of regressors by the dealer specific variables both for the trend and the constant term: $\beta_j = \beta$ and $c_j = c$.

with:

i = 1,...,15 (number of assets)

j = 1 (common dealers)

c) all assets are common. This time, the number of the asset specific variables is reduced: $\beta_i = \beta$ and $c_i = c$

with:

i = 1 (common assets)

j = 1,...,18 (number of dealers)

Based on the cluster analysis (complete link method) we obtained groups of

---

[3]In order to have accurate OLS estimations we have actually written programs in Gauss for the various models.

dealers and assets.

In order to test the hypothesis that dealers who are very active in a particular asset have an information advantage and their price quotes exhibit a similar pattern which allows us to group them together, we analyse the same groups of dealers and assets which resulted from the cluster analysis, this time, using the returns based on the quotes instead of the frequency of quotes.

d) assets are grouped into the three following groups by restricting the constant terms and the trend variables to be equal within each group:

A : assets 12812 and 12817 (with $\beta_i = \beta_a$ and $c_i = c_a$)

B: assets 13009, 13011, and 13013 (with $\beta_i = \beta_b$ and $c_i = c_b$)

C: assets 12838, 12805, 12811, 12859, 12825, 12879, 12882, 12610, 12628 and 12616 (with $\beta_i = \beta_c$ and $c_i = c_c$)

e) dealers are grouped into three groups by assuming that each group is different in the trend and the constant variable:

A: CTOS, CRMU, MPSG, ICCS, BCMT, BRBC, NAPQ, BNAT, SPTR, and BPNQ
with $\beta_i = \beta_a$ and $c_i = c_a$

B; BPMQ, CRRU, SIGB, and BSMT
with $\beta_i = \beta_b$ and $c_i = c_b$

C; ROLB, BNLT, CIMI, and BSSE
with $\beta_i = \beta_c$ and $c_i = c_c$

f) Restricted model:

We assume that all the dealers and all the assets are homogeneous and have a common constant term and a common trend
with $\beta_i = \beta$ and $c_i = c$.

5.4.2. Results

For testing we assume that ε in equation (3) is normally distributed. The detailed results of the respective coefficients and the t-statistics for each individual model are listed in tables 5.6.a)-f) in the appendix. With the F-tests we test the restrictions of the various models based on tables 5.6.a)-f). The outcome of the various regressions are listed in table 5.3. and the results of the F-tests are shown in table 5.4. below. They show that for both the dealers and the assets the unrestricted version which allows for heterogeneous dealer behaviour and heterogeneous assets is clearly rejected.

Table 5.3.: Results of the OLS regression on returns

	SSE	n	k
General model a)	0.041477	476	64
restricted dealers model b)	0.045814	476	30
restricted assets model c)	0.044578	476	36
grouped dealers model e)	0.045513	476	34
grouped assets model d)	0.044201	476	40
all restricted model f)	0.050597	476	2

Total sum of squares: 0.13678324
with: SSE=residual sum of squares
n=number of observations
k= number of regressors

Table 5.4. : F-tests on the returns

Tests:	restrictions	n	F-value
model a) versus e)	30	476	1.3363
model e) versus b)	4	476	0.7308
model a) versus d)	24	476	1.1988
model d) versus c)	4	476	0.5016
model b) versus f)	28	476	1.6630
model c) versus f)	34	476	1.3511
model a) versus f)	62	476	1.4612

It looks like all sets of restrictions are accepted. The final result is that the restricted model is the superior one which implies that we have homogeneous dealer returns and homogeneous assets. Both the constant term and the trend variable are highly significant in this restricted model. Hence, we have to reject the hypothesis of market segmentation of the Italian secondary bond market.

5.5. Conclusions

When we compare the introspective analysis of the data structure with the cluster analysis we observe quite different results.

The grouping made by each of the analyses has turned out not to be congruent. What we can say is that it seems that group II of the first analysis is very similar to group I formed under the cluster analysis.

These groups consisting of 7 members based on the introspective analysis, (respectively 10 members based on the cluster analysis), have 5 members in common. In the same way we can say that group I resembles group II as out of 4 members 2 are common.

However, in respect of specialisation we cannot observe such a trend as for the grouping of market makers.

One reason may be that the specialisation of market makers for a particular asset has been considered from different points of view under the two analyses.

With the first analysis each asset has been examined separately whereby, under the cluster analysis, the significance of the specialisation has been evaluated over the range of all the assets.

The same may apply to the differences in grouping of the market makers.

It can be argued that the results obtained for the grouping based on the frequency of quotes have shown that the Italian secondary market for government bonds can be divided into at least three segments.

The results from the OLS analysis based on the returns computed from price quotes give a totally different result. The grouping of the dealers and the asset is not confirmed by the findings under the OLS regression.

In each version of the model whether restricted or unrestricted we find significant dealer specific or asset specific variables. Also some of the trend variables are significant. However, in the end, the F-tests point out that the model in which the assets and the dealers are homogeneous is the most appropriate model.

As a result our hypothesis of a segmented market has to be rejected under the OLS analysis. This means that, based on the analysis of returns, the primary dealers in the market do not earn a monopoly profit which could have arisen from information asymmetry in the market and furthermore, it shows that there are no arbitrage opportunities for investors in the Italian secondary market of government bonds.

APPENDIX : Table 5.5a. : Dealer and asset classification

NUMBER OF OBSERVATIONS OF EACH ASSET AND EACH MARKET MAKER

(first line: bid quotes; second line: ask quotes)

ASs. / MARKET MAKERS:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1:	216	2	1	119	216	1	97	1	31	1	5	3	110	4	123
	216	2	1	119	216	1	97	1	31	1	5	3	110	4	123
2:	0	217	219	40	25	13	1	0	0	0	118	203	47	0	7
	0	217	219	40	25	13	1	0	0	0	118	203	7	0	7
3:	0	5	5	6	7	1	166	218	217	0	6	4	4	4	116
	0	5	5	6	7	1	166	218	217	0	6	4	4	4	116
4:	38	8	0	4	0	30	19	0	0	0	213	105	16	7	9
	38	8	0	4	0	30	0	0	0	0	213	105	12	7	9
5:	0	114	0	213	0	0	0	0	1	197	2	4	100	1	204
	0	114	0	213	0	0	0	0	1	197	2	4	208	1	204
6:	2	8	214	213	1	3	0	0	0	4	0	17	104	0	1
	2	8	214	213	1	3	0	0	0	4	0	17	8	0	1
7:	216	211	183	161	168	214	109	111	99	52	10	23	25	26	63
	216	211	183	161	168	214	109	111	99	52	10	23	27	26	63
8:	1	5	0	0	2	189	216	1	215	6	101	178	22	0	6
	1	5	0	0	2	189	216	1	215	6	101	178	22	0	6
9:	0	5	217	2	1	3	217	1	4	0	215	86	8	210	6
	0	5	217	2	1	3	217	1	4	0	215	86	3	210	6
10:	10	47	3	211	214	3	217	83	2	0	8	179	8	40	88
	10	47	3	211	214	3	217	83	2	0	8	179	8	40	88
11:	1	99	7	42	14	209	3	0	1	212	210	214	101	8	7
	1	99	7	42	14	209	3	0	1	212	210	214	2	8	7
12:	16	209	140	91	213	19	65	39	7	67	7	215	48	0	83
	16	209	140	91	213	19	65	39	7	67	7	215	12	0	83
13:	73	70	85	103	105	37	104	209	214	102	73	104	211	212	69
	73	70	85	103	105	37	104	209	214	102	73	104	211	212	69
14:	0	2	2	0	4	1	215	217	215	0	2	198	116	216	6
	0	2	2	0	4	1	215	217	215	0	2	198	6	216	6
15:	2	207	112	217	199	1	216	7	18	0	3	68	9	208	10
	2	207	112	217	199	1	216	7	18	0	3	69	6	208	10
16:	175	169	156	176	217	213	92	45	51	101	162	213	116	129	37
	175	169	156	176	217	213	92	45	51	101	162	213	26	129	37
17:	6	213	16	213	14	1	4	5	207	8	3	18	123	16	92
	6	213	16	213	217	1	4	5	207	8	3	18	208	16	92
18:	211	19	32	26	126	216	20	5	14	196	52	200	11	205	7
	211	19	32	26	126	216	20	5	14	196	52	200	4	205	7

Table 5.5b.: List of assets and market makers

LIST OF ASSETS AND MARKET MAKERS:

Assets:		Market Makers	
1	12838	1	BCMT
2	12805	2	BNAT
3	12811	3	BNLT
4	12812	4	BPMQ
5	12817	5	BPNQ
6	12859	6	NAPQ
7	13009	7	BRCB
8	13011	8	BSSE
9	13013	9	BSMT
10	12825	10	CRMU
11	12879	11	CRRU
12	12882	12	CTOS
13	12610	13	CIMI
14	12628	14	ROLB
15	12616	15	ICCS
		16	SPTR
		17	MPSG
		18	SIGB

Table 5.6.a) : Results of the unrestricted model a)

Asset specific coefficients:

trend variables	T-statistics	constant term	T-statistics
2.7159725E-006	1.0322	0.0028028123	12.0256
1.0300362E-007	0.0507	0.0022097187	12.1039
3.3349355E-006	1.2763	0.0019414606	8.5239
2.0296713E-006	0.9578	0.0018291213	9.6790
6.0828326E-006	2.4485	0.0024452832	11.7676
1.5450941E-005	6.7209	0.0012844033	6.4504
1.3543880E-005	5.9535	0.0018817065	9.6671
1.2994738E-005	5.4410	0.0020205631	9.6910
1.5185683E-005	6.9189	0.0013535967	6.9562
1.4411179E-005	6.7120	0.0014285247	7.9716
1.7805725E-005	8.1519	0.0021842952	11.5081
4.6448275E-006	2.0403	0.0022778084	11.4871
8.3186173E-006	3.1388	0.0024088852	10.4797
2.8415512E-006	1.1613	0.0030613679	14.1693
2.8786882E-006	1.1815	0.0036540984	16.7380

Dealer specific coefficients:

trend variables	T-statistics	constant term	T-statistics
4.1037781E-006	1.7211	-0.0012604551	-6.1221
-1.6786992E-006	-0.6355	8.8667934E-005	0.3905
1.0239567E-005	3.8341	0.00022558285	0.9669
1.1633029E-005	3.7925	-0.0013346248	-4.6919
-5.9088708E-006	-2.3558	0.00094921759	4.3547
-8.6491902E-006	-2.6312	0.0013819417	4.8494
5.4068531E-006	2.5312	-0.00092673021	-5.0087
1.0166504E-006	0.3834	-0.00049686797	-2.1221
7.1005621E-006	2.8399	-0.00053330963	-2.4087
-1.1409209E-006	-0.4849	-0.00093462983	-4.6519
3.5274543E-006	1.4634	2.6135795E-005	0.1274
-2.1628471E-006	-0.8945	8.9617854E-005	0.4307
-1.7638918E-007	-0.0732	-0.00051835180	-2.3194
2.5521628E-006	1.0393	-0.00027193165	-1.2662
1.0329598E-005	4.4581	-0.00078068114	-3.9809
-2.1634421E-006	-0.9987	-0.00046052945	-2.3629
4.5317069E-006	1.7373	0.00080189248	3.5055

residual sum of squares:	0.041477466
number of observations:	476.000
number of regressors:	64.000000
total sum of squares:	0.13678324

Table 5.6.b) : Results of the restricted model f)

coefficients:	value	T-values:
Trend variable:	9.8510808E-006	22.781987
Constant term:	0.0018397406	48.563281

residual sum of squares:	0.050597732
number of observations:	476.000
number of regressors:	2.0000000
total sum of squares:	0.13678324

Table 5.6.c) : Results of the restricted model b)

(dealers are homogeneous)

Asset specific coefficients:

trend variables	T-statistics	constant term	T-statistics
7.1048071E-006	2.4964636	-0.0017351251	-6.8198500
1.3583494E-006	0.55390980	-0.0021387252	-9.5108969
4.9325009E-006	1.7597726	-0.0022278615	-8.8303915
4.3582284E-006	1.6787910	-0.0025956479	-10.918253
1.2528320E-005	4.6389434	-0.0019407858	-8.1897088
1.8435474E-005	6.4601831	-0.0034553126	-13.463551
1.5391125E-005	5.9944624	-0.0022312899	-9.7532969
1.3027410E-005	4.9682653	-0.0020404604	-8.6675468
1.5081128E-005	6.0885966	-0.0027182889	-12.001237
1.5861581E-005	6.1877867	-0.0031268645	-13.585077
1.9235037E-005	7.0348860	-0.0021771098	-8.8337978
9.1740350E-006	3.6243021	-0.0023016540	-10.091524
1.3393944E-005	4.7413169	-0.0020882416	-8.3308229
5.5753926E-006	2.1117749	-0.0010837387	-4.5455521

Dealer specific coefficients:

trend variables	T-statistics	constant term	T-statistics
-1.3892241E-007	-0.066157411	0.0040185271	20.865947

residual sum of squares:	0.045814524
number of observations:	476.000
number of regressors:	30.000000
total sum of squares:	0.13678324

Table 5.6.d) : Results of the restricted model c)

(assets are homogeneous)

Asset specific coefficients:

trend variables	T-statistics	constant term	T-statistics
8.6493335E-006	5.4151806	0.0021123502	15.430249

Dealer specific coefficients:

trend variables	T-statistics	constant term	T-statistics
6.3626291E-006	2.7948048	-0.0016561353	-8.4035619
-1.8639254E-006	-0.73679491	9.2337451E-005	0.42865728
7.0126967E-006	2.8844012	0.00066260338	3.1380719
7.2895260E-006	2.4844572	-0.0010349481	-3.8962371
-6.6829028E-006	-2.8032572	0.0011068566	5.3363735
-3.3068438E-006	-1.0624686	0.00094791217	3.5457792
9.1041207E-006	4.2878050	-0.0011258171	-6.1685598
-2.1038480E-006	-0.82340859	-9.4290448E-005	-0.42252808
2.8940832E-006	1.2104462	-0.00038027305	-1.8239929
-4.5118541E-006	-1.9872140	-0.00092900559	-4.8765494
2.0169936E-006	0.84290238	0.00041940408	2.0792355
-1.4343813E-006	-0.61516908	-4.2825841E-005	-0.21489437
-3.6154741E-006	-1.5860024	-0.00035646937	-1.7137160
-2.9245122E-006	-1.2602849	5.1497408E-006	0.025739424
9.8054423E-006	4.3851664	-0.00095418088	-5.0826363
4.5961909E-007	0.21083002	-0.00071223827	-3.6380204
4.4189361E-006	1.8158030	0.00077804729	3.6681785

residual sum of squares:	0.044578128
number of observations:	476.000
number of regressors:	36.000000
total sum of squares:	0.13678324

Table 5.6.e) : Results of the restricted model e)

(dealers are grouped accoding to the complete link cluster analysis)

Asset specific coefficients:

trend variables	T-statistics	constant term	T-statistics
5.7850633E-006	2.0307283	-0.0016827208	-6.6203099
1.3171573E-006	0.53384317	-0.0020871686	-9.1993050
6.6371467E-006	2.3025621	-0.0021825163	-8.3612636
4.9338473E-006	1.8807979	-0.0026136579	-10.841367
1.3929061E-005	5.0105812	-0.0018601391	-7.5660026
1.8412503E-005	6.4254772	-0.0033483105	-12.937115
1.6720041E-005	6.3374386	-0.0021374555	-9.0465373
1.3651908E-005	5.1357664	-0.0019318504	-8.0482392
1.6454876E-005	6.4334893	-0.0026106094	-11.109881
1.7338904E-005	6.6219219	-0.0030536479	-12.968874
1.8803694E-005	6.8801551	-0.0021201150	-8.5931857
1.0177612E-005	3.8999933	-0.0022952112	-9.7170520
1.4971159E-005	5.0182069	-0.0021331253	-8.0078993
7.3893382E-006	2.6342952	-0.0011462518	-4.4828284

Dealer specific coefficients:

trend variables	T-statistics	constant term	T-statistics
-1.6430902E-006	-0.74237702	0.0039023698	19.093351
-1.9161228E-006	-0.78601443	0.0041438440	18.403874
2.4821673E-006	1.1481022	0.0039989130	20.435626

residual sum of squares:	0.045513544
number of observations:	476.000
number of regressors:	34.000000
total sum of squares:	0.13678324

Table 5.6.f) : Results of the restricted model d)

(assets are grouped accoding to the complete link cluster analysis)

Asset specific coefficients:

trend variables	T-statistics	constant term	T-statistics
1.5243968E-005	8.0463558	0.0014878133	9.2823753
6.4654011E-006	3.2168291	0.0026408503	15.192649
7.5234444E-006	4.7080281	0.0022416774	16.297738

Dealer specific coefficients:

trend variables	T-statistics	constant term	T-statistics
4.5102876E-006	1.9685172	-0.0015411593	-7.7991342
-1.2394240E-006	-0.49158055	1.1132447E-005	0.051810132
8.9620521E-006	3.4622381	0.00023548104	1.0455396
8.4154151E-006	2.8760246	-0.0011642753	-4.3906976
-7.2678800E-006	-3.0560226	0.0011453224	5.5416711
-6.0372820E-006	-1.9272581	0.0011936060	4.4484433
9.0902567E-006	4.2775611	-0.0012009640	-6.5814243
4.6166864E-008	0.017483832	-0.00046336326	-2.0009486
4.3864418E-006	1.8270425	-0.00061133627	-2.9144366
-5.3898451E-006	-2.3264475	-0.00094256189	-4.8457964
2.5259690E-006	1.0594495	0.00035540874	1.7672901
-2.3250791E-006	-0.99488230	5.3880586E-006	0.027110293
-2.1003468E-006	-0.90170231	-0.00065541808	-3.0763642
-9.7251340E-007	-0.40045494	-0.00038034498	-1.8068664
8.8990553E-006	3.9272727	-0.00094286025	-4.9755380
-8.8268302E-007	-0.40427375	-0.00063115667	-3.2309011
4.0306528E-006	1.6472282	0.00073740171	3.4638737

residual sum of squares:	0.044201292
number of observations:	476.000
number of regressors:	40.000000
total sum of squares:	0.13678324

CHAPTER SIX

DETERMINANTS OF THE BID-ASK SPREAD IN A DEALERSHIP MARKET

6.1. Introduction

This chapter is closely related to chapter two in which we analysed the determinants of the bid-ask spread for different market structures. Our aim is to see how these frameworks can be used to explore the trading and quoting pattern across dealers.

We base our analysis on data of the Italian secondary market for government bonds. Initially, (after a major reorganisation of this market) 22 government bonds are traded by 18 recognized market makers who are linked by an electronic circuit. Of the market makers 17 are banks of varying size and specialisation and 1 is a nonbank financial intermediary.

The market is organised so that the market makers quote their prices which are binding for a certain period of time. These quotes are valid for up to 5 mio lire. The price quotes are displayed on a screen of a computerised information system. The actual trade with the public is executed on the telephone. After a trading period, the aggregate volume, but not its distribution which was traded is made public.

The data used are daily observations starting from 16 May 1988 until 10 April 1989 on one of the assets: a floating rate credit certificate whose annual coupon rate is indexed by the rate on 12 month Treasury bills. The maturity date of the certificate is 1/4/1997 and the returns on the asset are liable to 6.25% tax.

The quotes have been collected every day between 12 o'clock and 1 o'clock p.m. which is the most active trading time of the day and therefore seems to represent the pricing strategies of the market makers in the most accurate way.

From Tables 6.1., 6.2. and 6.3. it is evident that there is considerable diversity of trading pattern.

On some days there is a single dealer quoting both the best ask and bid price. On other days there are several different dealers on each side of the market and none simultaneously on both sides.

Under the assumption that the market order is executed by the best quoting dealer or, in case of several best quoting dealers, is shared among them, we find that several dealers are actually trading on one or the other side of the market.

This suggests that theories based on dealer costs which increase with the size of the order, which implies that market makers are interested in sharing the order, are actually observed in the real world.

In table 6.1., we list the number of days on which only one dealer quotes the best price and on which more than one dealer quote the best price (on the horizontal the bid side and on the vertical the ask side).

Table 6.1. : Trading pattern observations

bid side / ask side	1	> 1	total
1	114	53	167
> 1	20	21	41
total	134	74	208

In table 6.2., we list the number of dealers who quote the best price for different periods of consecutive days for the ask side. The same is illustrated in table 6.3. for the bid side.

Table 6.2. : Best quoting days ask side

no. of days	no. of dealers
< 10 days	9
10-20 days	3
20-40 days	4
> 40 days	2

Table 6.3. : Best quoting days bid side

no. of days	no. of dealers
< 10 days	8
10-20 days	4
20-40 days	4
> 40 days	2

With a sample of 208 daily price quotes in the single asset, we have analysed the 10 most commonly quoting dealers.

In this chapter, we analyse empirically whether the inventory control argument is valid which means that we have to find some supportive evidence. We also investigate the question whether the best quoting dealer takes the price of the next best dealer into account by setting her price quote.

Basically, we carry out two analyses. The first one is an ordinary least squares analysis where we test the hypothesis of the inventory control aspect and the next best dealer argument. In order to do that we compare the dealers' quotes with the variables which determine the reservation price.

The second analysis investigates the pricing strategies of the market makers. This empirical study is based on a probit analysis which evaluates the probability that the dealer quotes the best price based on the variables which determine the reservation price.

According to our theoretical model in chapter 2 (section 2.2.2.), considering the ask side, if $pa^R_i > p$, where p is the market price (i.e. the best price), the dealer either quotes her reservation price pa^R_i (for the quantity Q/k where (k-1) is the number of best quoting dealers) or she does not quote at all. In either case she does not get any trade. If $pa^R_i[Q(k+1)/k] > p > pa^R_i(Q/k)$ she quotes p where she shares the market order. If $pa^R_i[Q(k+1)/k] < p$ she quotes p-ε to get the whole order.

Thus, it is possible that the dealer, knowing that her reservation price is above the market price, decides not to quote at all instead of quoting her reservation price as she knows that she is not competitive and that she will not get any trade.

In principle, we analyse the following two hypotheses:

(i) the determinants of the expected utility and hence reservation prices of each dealer are price expectations; the degree of risk aversion and the inventory level;

(ii) the strategic price quoting behaviour in the market: each dealer i knows the reservation prices of each of their rivals j: (p_j). From these each dealer forms expectations of the pattern of trades that can emerge under the next dealer argument , i.e. the best dealer does not quote her own reservation price, but the entry-limiting price of the next best dealer.

If in equilibrium i is not best on a side of the market then either i does not quote or quotes their own reservation price, not expecting to do any trade at that price. On the other hand, if i is best on one side of the market then i quotes the next best dealer's reservation price.

We assume that the quoted prices are in fact transaction prices. As we do not have any information about the individual trade we have to assume that the price quotes are the actual prices of the deals. We justify this assumption by the fact that the quotes were taken from the screen during a period when trading was very active which means that the quotes were updated very quickly. This in turn implies that the quotes and the actual transaction prices were very close.

In section 6.2. we analyse the data based on a situation as in our model presented in chapter two, (section 2.2.2.) for a centralised market structure. In such an environment, market makers know each others reservation prices and the order flow and they can share the market order.

The second analysis, in section 6.3., investigates the trading strategies of dealers which means we predict which dealer quotes the best price based on the variables which determine the reservation price.
Section 6.4. contains a discussion of the results and the conclusions.

6.2. Analysis of the Price Quotes in a Centralised Market

Our data on the Italian secondary market for government bonds show ask and bid quotes by individual dealers (not all dealers are quoting on any given day) and the volume that is traded. We can deduce the inner spread and the dealers who are quoting the lowest ask price and the highest bid price for each day. We can also observe the total volume that is traded on the day although we do not know its division between buy and sell orders.
To identify the active dealers and the volumes that are traded on both sides of the market we make the nontestable assumptions that:
(1) the daily volume is equally divided between buy and sell orders;
(2) the total volume on one side of the market is divided equally between all dealers who have the most competitive quote.
For each day we can then identify the selling dealers, the buying dealers and the volume that each trading dealer trades.
We can also calculate the quantity traded by each dealer, by dividing the total amount traded by the number of best quoting dealers and so get the ex post share or quantity traded by each best quoting dealer which is Q/M.
The inventory of each dealer is not observable; but within the sample the past trading history of each trader is observable and the current inventory level is defined by

$$I_{it} = I_{i0} + \sum_{s<t} D^{b}_{is} \frac{Q_s}{M_s} - \sum_{s<t} D^{s}_{is} \frac{Q_s}{M_s} \qquad (1)$$

where D^{b}_{is} is a dummy variable that is 1 if i bought on date s; D^{s}_{is} is a dummy that is 1 if i sold on date s.

The sample contains every trading day since the organised dealership market was created; it is thus not unreasonable to take the opening inventory as 0. The trading pattern also appears to justify this since in most cases dealers buy on days before they sell.

Both main assumptions of our model, i.e. the determinants of the reservation price (i) and the strategic price quoting behaviour (ii) are testable.

For (i) we can test (on the bid side)

$$p_i = \alpha_i + \mu_i^e + \gamma_i(Q/M + 2I_i) + \varepsilon_i \tag{2}$$

where p_i is the price quote against

$$p_i = \alpha_i + \mu_i^e + \gamma_i Q/M + \delta_i SA_i + \eta_i SB_i + \varepsilon_i \tag{3}$$

where SA and SB are the gross totals of past sales and purchases respectively.

For (ii) we adopt the alternative that dealers quote their own reservation price; the minimum profit price at which trade is undertaken. In this setting, dealers ignore the competitive bidding process and they do not try to maximize their profits. However, each dealer has to know the number of best quoting other dealers or has to form some expectation about it in order to determine the reservation prices and hence the spread. The two alternatives for the actual quote of i (given that dealers know their rivals reservation prices) are

$$p_i = \alpha_i + \mu_i^e + \gamma_i Q/M + \delta_i SA_i + \eta_i SB_i + \varepsilon_i \tag{4}$$

as contrasted with

$$p_i = \alpha_i + \mu_i^e + \gamma_i Q/M + \delta_i SA_i + \eta_i SB_i + \varepsilon_i \quad \text{if i is not best}$$

$$p_i = E_i p_j^R = \alpha_j + \mu_j^e + \gamma_j Q/M + \delta_j SA_j + \eta_i SB_j + \varepsilon_i \text{ if i is best} \tag{5}$$

where j is the index of the critical entry limited inactive dealer, and p_i is the actual price quote of the individual dealer.

μ_i^e is the individually expected market price for the next period taking account of the assumption that each dealer knows the distribution of the market orders (i.e. Q), Q/M are the market shares of the orders if the dealer is a buyer or a seller, and γ represents the risk aversion of the dealer.

If DUM_i is a dummy variable that is 1 if i is best on a side of the market and 0 otherwise (5) becomes

$$p_i=(1-DUM_i)(\alpha_i + \mu_i^e + \gamma_i Q/M + \delta_i SA_i + \eta_i SB_i) + DUM_i(\alpha_j + \mu_j^e + \gamma_j Q/M + \delta_j SA_j + \eta_i SB_j) + \varepsilon_i \quad (6)$$

where j is the index of the critical entry limited inactive dealer.

Dealers are assumed to have rational expectations of the future unit value of inventory, μ_i, which we instrument by the first two price lags.

The random error term is assumed to be normally distributed $\varepsilon \sim N(0;\sigma^2)$.

We can interpret such a random error as some randomness in the measurement of the market prices. We can also interpret it as a deviation between the quotes and the transaction prices as we assumed above that the quotes are transaction prices.

We first test the strategic behaviour assumptions in (6) against (4) for each dealer separately. By Ericsson's (1983) nonnested hypothesis test, (6) was accepted. So there is evidence for the next best dealers argument and strategic price quoting. The results of the test are listed in table 6.4. below.

Table 6.4. : Non-nested test Ericsson:

	Ask side		Bid side	
dealers	(4)vs(6)	(6)vs(4)	(4)vs(6)	(6)vs(4)
2	--	--	4.873	4.067
4	1.681	0.425	3.267	1.634
8	1.719	0.828	62.431	3.221
10	15.216	2.116	1.932	0.029
11	3.633	0.352	nested m.	nested m.
12	5.256	1.121	47.716	5.174
14	--	--	nested m.	nested m.
15	--	--	4.316	0.923
16	3.822	1.955	4.171	2.439
18	3.709	-0.003	8.539	-0.343

However for some dealers there is some heteroscedasticity (using White's test (1980)) and autocorrelation:(L.G.Godfrey (1978)); although the Reset tests give little evidence of mispecification.

We accept the hypothesis that dealers have rational expectations of the future unit value of inventory which we instrumented by the first two price lags.The IV estimates for the future price give the coefficients which are all very close to one which indicates that the instruments chosen approximate the variable very well. The results are listed in table 6.5.:

Table 6.5.: IV estimates for the futures price

	Bid side		Ask side	
dealers	coefficient	t-value	coefficient	t-value
2	1.0063	118.00	0.9923	130.69
4	1.0323	101.92	0.9380	47.79
8	1.0039	115.61	0.9790	111.29
10	0.9960	67.43	0.9770	76.84
11	1.0261	89.00	0.9790	177.52
12	1.0850	72.72	0.8420	78.81
14	1.0280	96.83	0.9740	146.18
15	0.9890	29.50	0.9650	79.49
16	1.0210	94.18	0.9820	122.33
18	1.0350	83.84	0.9999	156.93

Although we have found significant evidence of the next best dealer argument we still have to find the appropriate form of the model for each dealer separately as mentioned under (i) above.

We can test (2) against (3) by imposing restrictions on (6). If $\eta = -\delta$ only the opening inventory of a dealer enters into the trading decision; we call this model 2. Past buying and selling must have a common effect but the effect may be different to that of current trading.

If in addition $\gamma=\delta$ we have the full model of (2). So we test the assumptions in (i) by successively testing $\eta=-\delta$ (Model 2) and then $\gamma=\delta$ (Model 3); the unrestricted model of (6) is Model 1. In order to determine which model is valid for each dealer we conducted an F-test. The results (in table 6.6. and 6.7.) show that the unrestricted version is the appropriate model for most of the dealers. The restricted version only applies for two dealers on the bid side.

Table 6.6.: Model form evaluation for the ask side

Ask side

dealers	no.obs.	SSEr	SSEm	SSEu
2	196	7.678	7.371*	7.281
4	197	21.349	20.563*	20.499
8	173	5.728	5.591*	5.570
10	175	6.524	6.324	6.018*
11	201	5.755	5.477*	5.395
12	203	12.656	12.166	11.092*
14	200	6.549	6.133	5.302*
15	69	1.104	0.894*	0.865
16	200	5.981	5.613	5.477*
18	158	3.798	3.597	3.4878

Table 6.7.: Model form evaluation for the bid side

Bid side

dealers	no.obs.	SSEr	SSEm	SSEu
2	196	11.157	11.016	10.702*
4	198	10.674*	10.638	10.541
8	173	9.102	8.622	7.928*
10	175	7.958	7.668*	7.694
11	201	9.628*	9.590	9.484
12	202	13.158	12.684	12.317*
14	199	9.231	9.085	8.521*
15	68	2.630	2.627	2.261*
16	200	9.567	9.391	9.038*
18	158	7.788	7.743	7.437*

where: no.obs. = number of observations,

SSEr = residual sum of squares of the restricted model (model 1)

SSEm = RSS of the less restricted model (model 2)

SSEu = RSS of the unrestricted model (model 3)

The asterisks indicate the superior model according to an F-test on the restrictions.

The diagnostic tests of the OLS regression of the reduced form are slightly improved by using the appropriate version for each dealer compared to the unrestricted version.

Table 6.8.: Diagnostic tests on OLS regressions

	Bid side			Ask side		
dealers	autocorr.	hetero	RESET	autocorr.	hetero	RESET
2	0.000	4.844**	0.324	0.150	1.530	0.739
4	1.560	0.717	2.221	0.400	5.346**	1.084
8	0.410	0.430	0.141	0.340	2.980	0.025
10	-0.070**	0.322	0.532	0.970	2.302*	0.985
11	5.990*	1.539	0.169	5.330*	0.289	1.721
12	20.310**	0.309	2.300	110.160**	4.234**	4.855*
14	8.26**	0.038	2.016	10.340**	2.225*	0.890
15	0.350	4.926**	0.324	10.860**	3.182*	0.016
16	0.470	3.290*	1.258	0.830	0.501	0.098
18	11.740**	2.293	1.585	2.980	11.011**	7.201**

(Asterisks mean that the null hypotheses are rejected.)

The final models, i.e. the respective version of the restrictions and the reduction to significant coefficients, give us some insights to the theory.

On the bid side we observe mainly two strategies which on the one side is that the determinants of the reservation price are the next best critical dealers inventory positions. This is true for 6 dealers out of ten. On the other side, some of the dealers (three of them) consider their own inventory positions only. One dealer does rely on the volume traded only.

On the ask side we observe that only one dealer takes into account the next best critical dealer's inventory. Six dealers rely on their own inventory position and the volume traded. Two dealers rely on the volume traded only. The results are listed in tables 6.14. and 6.15. in the appendix.

However, we have to mention that a weakness of this approach is that we have to identify the next best critical dealer by the actual quote (which includes measurement error) and not by the expected dealer quote.

6.3. Investigation of Pricing Strategies

In this analysis we compare the observed market price with the determinants of each dealers reservation price to predict which dealers quote the best price. For example if p^a is the market ask price and (3) gives the reservation price, the probability of the dealer quoting the best ask is $Pr(p^R_i < p^a)$. The probability of the dealer not quoting is the probability that her reservation ask price is above the market ask, since the market ask is set strategically just below the critical dealers reservation price.

So for an equilibrium:

(i) for any trader i who is buying on day t

$$- p^b_t (1+r_t) + \alpha_i + \mu^e_i + \gamma_i Q/M + \delta_i SA_i + \eta_i SB_i + \varepsilon_i > 0$$

(ii) for any dealer i who is selling on day t

$$p^a_t (1+r_t) - a_i - \mu^e_i - c_i Q/M - d_i SA_i - e_i SB_i + \varepsilon_i > 0$$

(iii) for any dealer i who is not selling on day t

$$p^a_t (1+r_t) - a_i - \mu^e_i - c_i Q/M - d_i SA_i - e_i SB_i + \varepsilon_i < 0$$

(iv) for any dealer i who is not buying on day t

$$- p^b_t (1+r_t) + \alpha_i + \mu^e_i + \gamma_i Q/M + \delta_i SA_i + \eta_i SB_i + \varepsilon_i < 0$$

Hence the probability that i buys is

$$Pr(- p^b_t (1+r_t) + \alpha_i + \mu^e_i + \gamma_i Q/M + \delta_i SA_i + \eta_i SB_i + \varepsilon_i > 0) \qquad (7)$$

and the probability that i sells is

$$Pr(p^a_t (1+r_t) - a_i - \mu^e_i - c_i Q/M - d_i SA_i - e_i SB_i + \varepsilon_i > 0) \qquad (8)$$

The likelihood function for the observed pattern of trading for a single dealer i over the sample of 208 trading days for one side of the market is then

$$Pr(D_{i208},\dots D_{i1}) = Pr(D_{i208} | D_{i207},\dots D_{i1}) Pr(D_{i207} | D_{i206},\dots D_{i1}) \dots\dots\dots Pr(D_{i2} | D_{i1}) Pr(D_{i1}) \qquad (9)$$

where D_{it} is a dummy that is 1 if trader i is active, i.e. quotes the best price, on that side on day t and 0 otherwise. In each conditional probability, we can treat the past history of trading, given by the dummies, as predetermined and so perform a probit analysis for each trader across time[1].

The results of the probit analysis will predict which traders should be observed to be selling or buying on particular days.

Again, as with the OLS regressions, we can test (2) against (3) by imposing restrictions on (7)-(9). If $\eta = -\delta$ or $e = -d$ only the opening inventory of a dealer enters into the trading decision; we call this model 2. Past buying and selling must have a common effect but the effect may be different to that of current trading.

If in addition $\gamma = \delta$ or $c = d$ we have the model of (2). So we test the assumptions in (i) by successively testing $\eta=-\delta$ (Model 2) and then $\gamma=\delta$ (Model 3); the unrestricted model of (7)-(9) is Model 1.

We try modeling price expectations extrapolatively:

$$p_t^a (1+r_t) - \mu_{it} = \gamma_0 + \gamma_1(1+r)p_t + \gamma_2(1+r)^2 p_{t-1} + \gamma_3(1+r)^3 p_{t-2} + u_t$$

where p_t is the best price quote on the relevant side of the market. This may induce heteroscedasticity in the probit with the variance of the error ε related to the interest rate. We test for this.

We estimate models 1-3 for each dealer on each side of the market and use likelihood ratio tests to derive a preferred form of model. The results are listed in table 6.9. below.

Here χ^2 is a likelihood ratio test of joint significance of the regressors; with an asterisk it is significant at the 5% level. Generally the probits are significant. The next step consists of reducing the chosen equation by

---

[1]In addition there is the question of possible correlation between the errors ε^a and ε^b; bivariate probits revealed that for most dealers there was insignificant correlation.

eliminating the insignificant variables.

The final version ("model 4") is the one with the highest possible overall significance level with significant t-statistics for all variables included.

Table 6.9.: Likelihood ratio tests

Dealers	model 1	LRT	model 2	LRT	model 3
		(1 vs 2)		(2 vs 3)	
	Log LH		Log LH		Log LH
A 2	-63.487*	0.014	-63.480	0.03	-63.495
B 2	-74.996*	0.054	-74.969	5.954	-71.992
A 4	-72.232	8.246	-68.109*	0.12	-68.049
B 4	-85.878*	0.314	-85.721	0.626	-85.408
A 8	-53.553*	0.274	-53.416	1.686	-52.573
B 8	-70.757	3.978	-68.768	8.104	-64.716*
A 10	-93.769	5.550	-90.994*	2.382	-89.803
B 10	-106.35	16.842	-97.929	14.09	-90.884*
A 11	-40.302	3.062	-38.771	4.572	-36.485*
B 11	-63.228*	0.020	-63.218	1.948	-62.244
A 12	-67.567	3.652	-65.741*	0.100	-65.691
B 12	-50.581	4.222	-48.470	3.054	-46.943*
A 14	-32.235*	0.102	-32.184	0.068	-32.150
B 14	-56.136*	0.056	-56.108	0.150	-56.033
A 15	-32.770*	0.942	-32.299	0.042	-32.278
B 15	-58.700	3.780	-56.810*	1.326	-56.147
A 16	-108.34	4.600	-106.04*	0.02	-106.03
B 16	-104.75	10.628	-99.436*	1.038	-98.917
A 18	-72.441*	0.920	-71.981	0.786	-71.588
B 18	-93.815*	2.266	-92.682	4.352	-90.506

LRT=Likelihood ratio test and Log LH=log likelihood value. "A" stands for ask quotes and "B" for bid quotes. The asterisks indicate the model selected to be reduced further.

If we examine the accuracy of predicted trading days and non-trading days we can say that the non-trading days are correctly predicted for almost 100%. However, there is a much lower percentage for the trading days.

The results of the probits also support partially the theory. The findings are listed in tables 6.16. and 6.17. in the appendix.

For the bid side results are mixed: 6 dealers show significant coefficients for inventory as well as past prices. Only one dealer supports the inventory control model whereas 3 dealers rely on past prices only. The ask side turns out to be different. 2 dealers are mixed, 4 dealers rely on past prices only, and the remaining 4 dealers support the inventory control argument.

The respective results are listed in table 6.10. below.

Table 6.10. : Overview of the Probit results

	Bid side:	Ask side:
	number of dealers	
	PROBIT	PROBIT
inventory	1	4
past prices	3	4
mixed	6	2

The results of the diagnostic tests are much better than for the OLS regressions. The test for normality has been rejected for only one dealer on the ask side. Heteroscedasticity is found on both sides of the market, but for one dealer on each side only. The missspecification tests show a slightly different result. Three dealers on the buying side and one dealer on both sides of the market fail the test.

However, the majority of the tests pass as can be seen in table 6.11. below.

The tests used are Lagrange multiplier variants of tests for omitted variables, heteroscedasticity and normality, see Ch. Orme (1988).

From the probit equations we can identify the critical dealer on each side of the market on each day; this is the dealer with the predicted reservation price closest to the market price.

Table 6.11. : Diagnostic tests for the probits

dealer	normality	omitted var.	hetero
Bid side			
2	3.046	1.082	2.044
4	4.293	4.293 **	8.018 **
8	0.728	3.904 **	0.004
10	2.977	3.386 **	1.708
11	3.405	0.776	0.057
12	0.962	1.861	0.814
14	3.486	9.808 **	0.206
15	0.103	0.013	1.241
16	1.102	0.365	0.069
18	4.452	0.143	1.057
Ask side			
2	2.119	0.976	0.033
4	1.124	3.310 **	1.086
8	4.486	0.339	1.132
10	2.179	0.668	0.066
11	1.133	0.283	0.766
12	1.343	1.312	1.769
14	0.405	1.397	0.216
15	0.987	0.787	0.477
16	6.704 **	0.244	10.270 **
18	2.622	0.143	0.871

We expect that the market ask price should be above the active dealer's reservation prices but close to the lowest inactive dealer's reservation ask. Also the market bid price should be below the reservation bid price of the best quoting dealer, but close to the maximum of the reservation bids of the inactive dealers. The means and standard deviations over time of these prices are given in Table 6.12. (based on the probit analysis).

Table 6.12. : Analysis of the market prices

	Mean	σ
Market Ask	93.264	0.993
Critical Ask	93.881	1.063
Market Bid	93.252	1.049
Critical Bid	92.645	1.145

The relative means are consistent with the theory if around 2% is interpreted as small. The critical prices display more variability than the best market prices; in part this may reflect the estimation error that is incorporated within them.

The number of days on which particular dealers are critical one are listed in table 6.13. below.

Table 6.13. : Probit predictions

dealers:	2	4	8	10	11	12	14	15	16	18
			number of days							
bid	25	26	8	40	4	6	1	7	48	40
ask	9	21	8	73	7	20	3	0	58	6

From this table it is evident that there is a strong concentration amongst dealers with dealer 16 dominating the bid side and dealer 10 dominating the ask side.

6.4. Conclusions

In this chapter we have analysed two issues. First, we investigated whether the inventory control argument is valid for the Italian secondary market and, secondly, whether dealers exhibit a strategic price quoting behaviour.

The findings show that we have differences between the bid and ask side.

The OLS estimations reveal that, on the bid side, market makers rely heavily on the inventories. Out of ten dealers, six of them seem to take into account the next best dealer's inventory position whereas three market makers quote their price related to their own inventory position. On the ask side, we find that the inventory control argument is confirmed for seven market makers of which six base on their own inventory position and a single dealer relates the quote to the next best dealer's inventory position. Hence we can say that we do have evidence that the inventory is important in determining the price quote. In addition, the strategic pricing behaviour is also confirmed, but it seems to be more applicable on the bid than on the ask side.

We also get some evidence on the inventory control assumption under the probit analysis. This time, the bid side shows somewhat different results as only one dealer quotes her price based on the inventory position. Three dealers take into account the past prices only whereas six dealers balance their quoting on past prices and their inventory positions.

On the ask side, we do not get any conclusive evidence as there are four dealers who rely on past prices and four dealers who consider their inventory positions for quoting their prices. Two market makers take into account both the past prices and their inventories.

Finally, we can summarise our findings by saying that we do have evidence of the inventory control argument and the strategic behaviour.

We have presented an analysis for the Italian secondary market for government bonds which takes into account the possibility of sharing of the market order which seems to be the case in this market.

APPENDIX

The variables are:

CONST=constant term,

EXP = current market price either the ask price or the bid price depending whether the current inventory is positive or negative respectively,

INVA = [Q/N-2I] where Q=volume and I=approximation of inventory (as described above)

INVB = [Q/N+2I],

IN_i = (sb_i- sa_i) where sb_i is the trade history of dealer i on the bid side and sa_i is the trade history of dealer i on the ask side.

NAV=Q/N (number of sellers) and NBV=Q/M (number of buyers)

duma (for the ask side) and dumb (for the bid side) are dummy variables defined as : d=1 if dealer is best and d=0 if dealer is not best.

1-d is a dummy representing the reverse case.

1-dINVA / 1-dINVB = INVA / INVB multiplied by the dummy 1-d.

1-dsa/1-dsb = sa /sb multiplied by the dummy 1-d.

dcrsa / dcrsb = the next best dealers sa /sb multiplied by the dummy d.

1-dvol = [Q/N] multiplied by the dummy (1-d).

dsa / dsb = sa / sb multiplied by the dummy d.

dvol = [Q/N] multiplied by d.

The following tables show the coefficients and the respective t-values for the OLS and the PROBIT regressions:

Table 6.14. : Results of the OLS estimation on the ask side

dealers	variable	coeff.-value	t-statistics
2	constant 1-dIN2 1-dNAV dNAV	0.140 -0.001 -0.002 -0.004	5.248 -2.507 -2.780 -2.821
4	constant 1-dIN4	0.067 0.002	2.286 4.954
8	constant 1-dIN8	0.072 -0.002	4.006 -4.488
10	constant 1-dsa dcrsa duma	0.070 0.0003 0.001 -0.177	2.802 2.510 2.239 -2.159
11	constant 1-dNAV dNAV	0.192 -0.001 -0.003	10.537 -2.966 -3.422
12	constant 1-dsa 1-dNAV duma	0.120 0.001 -0.001 -0.143	2.749 5.014 -1.890 -2.213
14	constant 1-dsb 1-dNAV dNAV duma	0.324 -0.121 -0.001 0.005 -0.520	13.062 -4.596 -2.944 2.454 -5.272
15	constant 1-dIN15 1-dNAV	0.087 0.002 -0.002	3.409 3.101 -2.880
16	constant 1-dNAV dNAV duma	0.174 -0.002 -0.002 -0.090	8.158 -2.806 -1.663 -2.177
18	constant 1-dsa 1-dNAV	0.062 0.0004 -0.001	2.606 5.087 -2.232

Table 6.15. : Results of the OLS estimation on the bid side

dealers	variable	coeff.-value	t-statistics
2	constant	-0.066	-3.664
	dcrsa	0.002	2.326
	dcrsb	-0.003	-2.089
4	1-dINVB	-0.001	-3.355
8	constant	-0.065	-3.355
	dcrsa	0.002	4.225
	dcrsb	-0.001	-1.863
10	1-dNBV	-0.001	-1.988
11	constant	-0.048	-2.803
	1-dINVB	-0.001	-3.370
12	constant	-0.106	-5.859
	dcrsa	0.001	3.557
14	constant	-0.078	-5.100
	dcrsb	-0.003	-4.087
	dNBV	0.002	2.423
15	1-dsb	0.011	1.713
	dcrsb	-0.001	-3.285
16	constant	-0.074	-2.812
	1-dsa	0.0002	1.738
	dcrsa	0.001	3.029
18	constant	-0.156	-4.109
	1-dsb	0.0003	2.557
	dumb	0.204	3.883

TEXT CUT OFF IN ORIGINAL

Table 6.16. : Results of the Probit analysis (Bid side)

D	CONST.	EXP	EXP[-1]	EXP[-2]	INVB	IN	SA	SB	NBV
2			-1.31 (-2.70)	1.15 (2.64)					
4		1.77 (2.74)		-1.49 (-2.79)					
8	42.27 (2.58)	1.72 (2.81)		-2.58 (-3.08)				11.71 (2.8)	-4.62 (-2.4
10		-5.97 (-3.50)	5.40 (3.48)				-18.11 (-5.1)	7.33 (4.6)	-6.95 (-3.9
11	19.40 (1.60)		-0.61 (-1.70)		-2.58 (-1.4)				
12		-1.94 (-2.95)		1.57 (2.88)				-22.7 (-2.5)	
14		0.76 (1.95)		-0.65 (-2.06)					
15	25.64 (1.84)	1.53 (3.06)		-1.98 (-3.62)		35.34 (2.70)			-3.35 (-1.8
16	-15.76 (-1.71)		0.45 (1.67)						-4.54 (-2.9
18	-1.05 (-8.52)				4.37 (2.20)				

Table 6.17. : Results of the Probit analysis (Ask side)

D	CONST.	EXP	EXP[-1]	EXP[-2]	INVA	IN	SA	SB	NAV
2	19.55 (1.86)		-0.63 (-2.02)		6.11 (2.35)				
4	-0.51 (-2.55)		-0.63 (-2.02)			-10.40 (-3.2)			-4.3[illegible] (-2.[illegible]
8	-34.13 (-2.56)	1.06 (2.47)							
10		-0.96 (-3.15)		0.75 (3.06)		6.80 (3.19)			-3.4[illegible] (-2.[illegible]
11	-1.17 (-3.08)						6.80 (1.83)	-13.5 (-2.5)	3.5[illegible] (2.1[illegible]
12		1.62 (3.52)		-1.35 (-3.59)					-2.6[illegible] (-1.[illegible]
14	29.42 (1.76)	1.26 (1.63)		-1.86 (-3.14)					
15	-1.04 (-3.67)				14.86 (2.46)				
16	-0.62 (-4.15)					-4.27 (-1.9)			-2.5[illegible] (-2.[illegible]
18									

CHAPTER SEVEN

A MEASURE OF THE BID-ASK SPREAD IN THE SECONDARY MARKET

7.1. Introduction

Since the publication of the paper by Roll (1984), various researchers have investigated the effect of price changes on the effective bid-ask spread. Roll created a measure which allows the effective bid-ask spread to be inferred directly from a time series of market prices. His method is simple in that he requires the transaction prices themselves only. However, he had to impose the assumptions of an informationally efficient market and that the probability distribution of observed price changes is stationary.

Roll derived a measure of the spread by examining the changes of prices following a transaction. Under the assumption of an efficient market, it is not possible that the change in prices occurs due to new information in the market. Hence, assuming a constant bid-ask spread over a certain period, the change in transaction prices is due to the spread only which represents a cost compensation to the market maker, (as proposed by the theory of Demsetz (1968)).

A transaction, corresponding either to a buy or sell order, occurs at the bid price or at the ask price with equal probability. By examining recorded transaction prices we cannot observe whether the preceding transaction was at the ask or the bid, hence, the probability distribution of price changes consists of two parts, i.e. the probability of a change if the transaction is at the bid or the ask. Based on the joint probability of a buy or sell order a measure for the bid ask spread can be derived which Roll shows to be:[1]

$s = 2\sqrt{(-cov)}$

whereby cov is defined as the first order serial covariance of transaction price changes. Note this has to be negative. The variance of price changes

[1]Roll: pp. 1129 resp.1135

includes new information in the market whereas on the other hand the covariance between price changes cannot be due to new information if markets are efficient. If markets are efficient we have $cov(\Delta p^*_t, \Delta p^*_{t-j})=0$ with $j \neq 0$ and where p^* is the unobserved "true" value of the asset. It is also true that $cov(\Delta p^*_t, \Delta p_{t-j})=0$. [2]

Roll supports his theory with empirical results derived from data of AMEX and NYSE listed stocks, between 1963 and 1982 on one-day and five-day returns based on transaction prices.

One of the various extensions to the model of Roll is the approach of Stoll (1989). He examines, amongst others, the relation between the quoted bid-ask spread and the serial covariance of **transaction** returns on the one hand and the serial covariance of **quoted** returns on the other hand. The determinants of the spread are expressed as a function of the probability of a price reversal, π, and the magnitude of an adverse price change, δ ($0 \leq \delta \leq 1$) which actually is a fraction of the quoted spread S.

Data from NASDAQ/NMS (National Market System) stocks are used to show the time series behaviour of the transaction prices and the quoted prices and the respective spread. The data used are over a three months period, i.e. October, November, and December 1984.

The empirical results of Stoll are twofold. First, the serial covariances of transaction returns are strongly negatively associated with the square of quoted spreads which is in accordance with the findings of Roll. Second, the results for the serial covariances of returns derived from price quotes are not conclusive. The serial covariances seem to be negatively associated with the square of quoted spreads, but the level of significance is not satisfactory for all months. In addition, the proportion of variation explained is also very small.

---

[2]See Roll (1984) pp. 1135.

In this paper, we conduct an empirical analysis within the framework of Stoll. In particular, we focus on the relation between the serial covariance of returns calculated from daily price quotes and the square of quoted spreads where the results of Stoll did not bring any conclusive evidence. We also examine the variance components. That is we allow for different intercepts of the regression lines which could lead to the findings of market imperfections [3]. Furthermore, we analyse whether there are differences in the relation between the covariances of price changes and the spread for different assets. We use daily data as well as weekly data in order to determine the serial correlation between the price changes and the bid-ask spread.

The paper is structured so that in section two, the theory of the spread and the serial covariance is presented, based on the paper of Stoll. In section three the data of the Italian bond market are described. Our empirical model formulations are given in section four and the results are presented and interpreted in section five. The final section contains the implications and conclusions.

7.2. The Theory of Spread and Serial Covariance

7.2.1. The Spread

Stoll defines the measures of the spread to be the price reversal δ (as a fraction of the spread S) and the probability of a reversal π. The spread is assumed to be constant over time.

---

[3]Both Roll and Stoll considered an efficient market. However, Roll observed that his regression line showed a significantly positive intercept which means that there are imperfections in the market.

Figure 7.1. below can be explained as follows:

If we start with a purchase (bid transaction) we expect a subsequent sale (ask transaction) with a probability of π and the size of the price reversal is $(1-\delta)S$. On the other side, the probability for a subsequent purchase (preceding transaction still assumed to be a purchase) is $(1-\pi)$ and the respective magnitude of the price reversal is $-\delta S$. In case of an initial ask transaction followed by an ask transaction the size of the reversal is δS.

If we first start with a sale (ask transaction) we have a subsequent purchase with probability π and the size of the price reversal is $-(1-\delta)S$. If the subsequent transaction is a sale (given that the first transaction is a purchase) the probability of a price reversal is $(1-\pi)$ and the size of the reversal is δS.

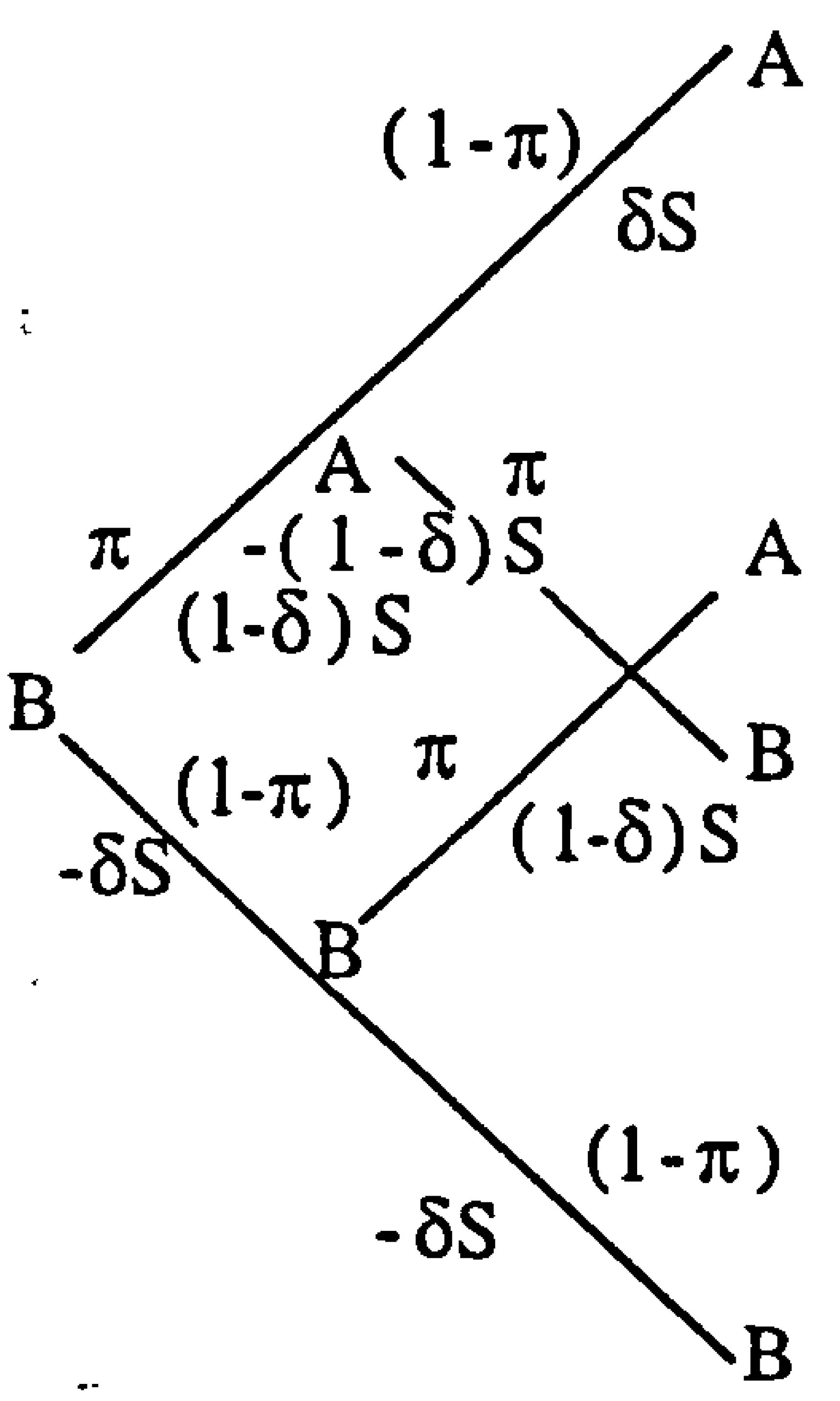

Figure 7.1.[4] : Determinants of the spread measure

where A: ask transaction; B: bid transaction; π: probability of a price reversal; $(1-\delta)S$: size of a price reversal as a function of the spread S with $0 \leq \delta \leq 1$.

[4]Stoll pp.119

Stoll gives three views on the definition of the spread with varying size of the parameters δ and π. First, he considers the case identical to Roll where the spread represents a transaction cost compensation only. In this scenario δ equals 0 and π is 0.5 as the market maker does not change his/her price relative to the "true" market price p* (the expected market price) in order to adjust inventory or as a response to asymmetric information. This is illustrated in figure 7.2.(A).

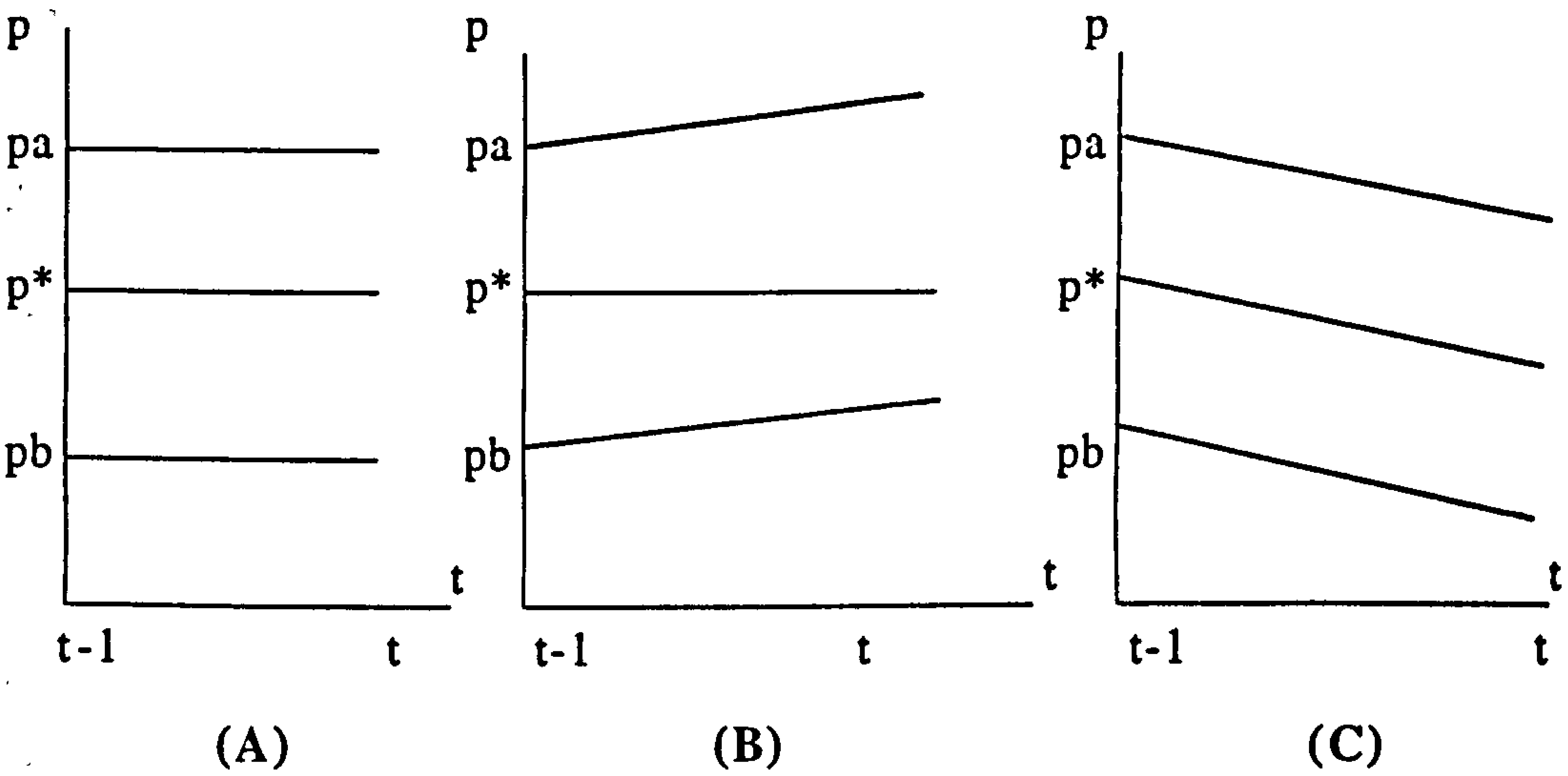

Figure 7.2. : Pricing strategies

The second view is that the market makers adjust their quotes to maintain an optimal inventory level. This is the view of the theories of Stoll (1989), Ho and Stoll (1981/1983), and Amihud/Mendelson (1980). This implies that after a sale (purchase) the price will be lowered (increased) in order to induce a purchase (sale). The respective parameters become: δ=0.5 (as the spread is twice the inventory cost - see Ho and Stoll (1983)) and $1>\pi>0.5$. The third view is dominated by adverse information costs (as defined by the theories of Copeland and Galai (1983), and Glosten/Milgrom (1985)). The reason for a price change is the revision of the equilibrium price after a transaction has taken place, based on the information obtained from the transaction price.

Under the assumption that all traders have superior information to the market maker the adjustment of the bid and ask prices is the same as the one described for the inventory control model. The respective parameters are: $\delta=0.5$ and $\pi=0.5$.

7.2.2. The Serial Covariance of Price Changes

The serial covariance of price changes can be explained by examining what the reasons are for such a change. A price change of a security (ΔS) may occur for different reasons. A part of a price change can be explained by the expected price change due to basic securities characteristics (SC). Another influence on the price change may be due to the existence of the spread (ΔP_t) combined with order reversal and a third reason may be that new information has arrived at the market place (ε_t) where $E(\varepsilon_t)=0$.

The total price change can be expressed as

$$\Delta S_t = SC + \Delta P_t + \varepsilon_t$$

Updating $\Delta S_{t+1} = SC + \Delta P_{t+1} + \varepsilon_{t+1}$ and so the serial covariance of price changes is

$$cov(\Delta S_t,\Delta S_{t+1})=cov(\Delta P_t,\Delta P_{t+1})+cov(\Delta P_t,\varepsilon_{t+1})+cov(\varepsilon_t,\Delta P_{t+1})$$
$$+cov(\varepsilon_t,\varepsilon_{t+1})$$

Due to the fact that the price changes caused by new information, in an efficient market, are serially uncorrelated, and in addition, are also uncorrelated both with lagged and leading values of the price change due to the existence of a spread, the covariance becomes:

$$cov(\Delta S_t,\Delta S_{t+1})=cov(\Delta P_t,\Delta P_{t+1}).$$

This implies that the serial covariance of price changes is caused by the covariance induced by the spread only, still assuming an efficient market.

Under the assumption of a constant spread and symmetry, the serial covariance of price changes can be obtained either by the ask quotes or by the bid quotes.

Now, considering figure 7.1., and assuming that the starting transaction is a bid, the price change can be derived to be $(1-\delta)S$ with probability π for (A_t-B_{t-1}) and $-\delta S$ with probability $(1-\pi)$ for (B_t-B_{t-1}).

By symmetry, the same is true if the transaction in period t is an ask.

Now, the possible price change ΔS can be derived as follows according to the pattern defined in figure 7.1..

We assume that the preceding transaction is a purchase:

$$\Delta S_t = \begin{cases} (A_t-B_{t-1})=(1-\delta)S & \text{with probability } \pi \\ (B_t-B_{t-1})=-\delta S & \text{with probability } (1-\pi). \end{cases}$$

This time, we start from a sale transaction:

$$\Delta S = \begin{cases} (B_t-A_{t-1})=-(1-\delta)S & \text{with probability } \pi \\ (A_t-A_{t-1})=\delta S & \text{with probability } (1-\pi). \end{cases}$$

Hence, the expected price change conditional on a purchase transaction is

$E(\Delta S_t|B_{t-1})=\pi(1-\delta)S+(1-\pi)(-\delta S)=(\pi-\delta)S.$

The respective expected price change for the sale transaction is

$E(\Delta S_t|A_{t-1})=\pi(-(1-\delta))S+(1-\pi)(\delta S)=-(\pi-\delta)S.$

The spread, as the difference between the purchase and the sale is $2(\pi-\delta)S$.

The serial covariance depends on two consecutive periods. We have already defined the respective pattern of price changes in figure I for two periods for the bid transaction.

Under the assumption of symmetry, the same is true for the ask transaction. Given the underlying joint distribution [5] of successive transaction price changes, the covariance is

[5] Stoll pp. 133 appendix A

$\text{cov}(\Delta S_t, \Delta S_{t-1}) = (1-\pi)^2\delta^2S^2 - \delta S(1-\delta)S\pi(1-\pi) + \delta S(1-\delta)S\pi(1-\pi) - (1-\delta)^2S^2\pi^2$

$\text{cov}(\Delta S_t, \Delta S_{t-1}) = S^2[\delta^2(1 - 2\pi) - \pi^2(1 - 2\delta)$

In our case, we are particularly interested in the covariance of price changes for the quoted prices. In this case the joint distribution is simpler as the value of the price change, from the initial change in a bid price (B_t-B_{t-1}), can only be either $-\delta S$ if the preceding transaction was at the bid or $+\delta S$ if the transaction in the previous period was at the ask. Also for the next period, the price changes can take on the same values only[6]. Hence the joint distribution of successive changes can be tabulated as

Table 7.1. : Joint distribution in the bid price changes [7]

		ΔB_t	
		Initial trade at bid	Initial trade at ask
		$-\delta S$	δS
ΔB_{t+1}	Next trade at bid $-\delta S$	$1-\pi$	π
	Next trade at ask δS	π	$1-\pi$

Under the assumption of a symmetric market $\text{cov}_a = \text{cov}_b = \text{cov}_q$ where cov_q is the covariance of the quoted prices.

Hence, the serial covariance of changes in quoted prices is

$\text{cov}=(1-\pi)\delta^2S^2-\pi\delta^2S^2$ which finally becomes $\text{cov}=\delta^2S^2(1-2\pi)$.

The predictions for the outcome of the value of the covariance under the different models can be summarised as

1. cov=0 under the order processing model with $\delta=0$ and $\pi=0.5$
2. cov=0 under the adverse information model with $\delta=0.5$ and $\pi=0.5$
3. $-0.25S^2 < \text{cov} < 0$ under the inventory control model with $\delta=0.5$ and $\pi>0.5$.

---

[6]Stoll pp.133 appendix B

[7]Stoll pp. 133.

7.3. Empirical Analysis of the Covariance of Price Changes

The empirical analysis is carried out with price quotes of eighteen market makers in the Italian secondary market for government bonds. Fifteen assets are examined over a period of approximately 150 consecutive trading days from May 1988 until January 1989.

The daily estimation is based on seven periods of 21 trading days. Consequently, seven covariances of price quote changes have been derived. We have examined the ask quotes (p_a)and the bid quotes (p_b) separately. As an alternative analysis weekly data from the same sample period were taken. For the weekly data we computed six covariances each ranging over a period of four weeks.

In comparison with Stoll's data set, we have seven covariances (observations) which have been computed over approximately 150 trading days instead of three covariances over a period of roughly 60 days for the daily data analysis. This may give us more accurate results. However, we only have 15 assets to compare where Stoll analyses 700-800 assets. Initially, we run a simple regression with the covariances as the dependent variable and the proportional spread of the effective bid and ask quotes as the single explanatory variable. From the individual quotes we can evaluate the market quote which is the highest bid and the lowest ask quote (under the assumption of Bertrand price competition among market makers). The proportional spread is calculated by dividing the difference between the highest ask and the lowest bid quote by the sum of the respective ask and bid quote. The squared spread has been scaled by multiplying it by 10,000. The number of days considered for the daily analysis is 147 and the number of days for the weekly analysis is 120.

Our analysis investigates several issues. The first version of the models given below regards all assets to be identical in respect of the relationship between the serial covariance and the spread. Thus the parameters α and β are common. The second version accounts for other influences than from the existence of the spread. We have argued above that in an efficient market all the information should be conveyed to the market through the occurrence of a transaction and hence be incorporated in the prices. According to the model presented in the previous section the only influence on the price changes should come from the spread (which is assumed to be constant) if the market works efficiently which means that all the information should be incorporated in the prices. Hence, if we observe another influence we can interpret it as a kind of "inefficiency" in the market. In order to distinguish this influence among the assets we let α vary between the assets. The third version presented below assumes differences between the various assets in the extent of the relationship between the serial covariance and the spread. This is expressed in β which varies among assets.

The form of the simple regression can be written as

model [1]:

$$(\mathrm{cov})_{P_a\, i,t} = \alpha_{(1)} + \beta_{(1)} S^2_{i,t} + \varepsilon_{i,t} \qquad (1)$$

$$\text{and } (\mathrm{cov})_{P_b\, i,t} = \alpha_{(2)} + \beta_{(2)} S^2_{i,t} + \varepsilon_{i,t} \qquad (2)$$

where i=1,...,15 and

t=1,...,7 for daily data

t=1,...,6 for weekly data.

The effect of market inefficiencies on the relation between the covariance of price changes and the spread can be expressed as

model [2]:

$$(\mathrm{cov})_{P_a\, i,t} = \alpha_{i(1)} + \beta_{(1)} S^2_{i,t} + \varepsilon_{i,t} \qquad (3)$$

$$\text{and } (\mathrm{cov})_{P_b\, i,t} = \alpha_{i(2)} + \beta_{(2)} S^2_{i,t} + \varepsilon_{i,t} \qquad (4)$$

where i=1,...,15 and

t=1,...,7 for daily data

t=1,...,6 for weekly data.

Another regression captures the effects for each individual asset separately. The respective form of regression is

model [3]:

$$(cov)_{p_a\,i,t} = \alpha_{(1)} + \beta_{i(1)} S^2_{i,t} + \varepsilon_{i,t} \quad (5)$$

and $$(cov)_{p_b\,i,t} = \alpha_{(2)} + \beta_{i(2)} S^2_{i,t} + \varepsilon_{i,t} \quad (6)$$

where i=1,...,15 and

t=1,...,7 for daily data

t=1,...,6 for weekly data.

All estimations have been carried out by the method of ordinary least squares, with coefficient restrictions imposed across equations according to the different models.

This is done by stacking the equations for different assets. Consequently diagnostic tests for "heteroscedasticity" are in fact tests for heteroscedasticity for the disturbance of each asset and also for a common variance of the disturbance across assets.

Similarly a functional form test would test both that each asset is well specified and that all assets have a common form.

7.4. Empirical Results

The evidence we want to show is twofold. Firstly, to get a confirmation of the Stoll theory we expect the relation between the covariance of the price changes and the spread to be negative. Secondly, the results of the daily data analysis are compared with the findings of the weekly data analysis. Generally, we can say that the restricted model with common assets (model [1]) does not show any significant influence of the spread. The serial correlation seems to be positive rather than negative for both the daily and the weekly data.

The results for model (1) are listed in table 7.2. below.

Table 7.2. : Results of model (1)

Ask side	constant	t-stat.	spread	t-stat.	n
daily	0.0176	3.1495	0.4032	0.5147	147
weekly	-0.0261	-0.6590	-0.0001	-0.3505	120
Bid side	constant	t-stat.	spread	t-stat.	
daily	0.0155	2.9277	1.0992	1.4839	147
weekly	0.0043	1.4131	-0.0001	-0.6810	120

For both sides of the market of model (1) we get very low F-values for the explanatory power. Nevertheless, the diagnostic tests which include an autocorrelation test, a test for heteroscedasticity, and a reset test, are all accepted.

The only exception is that there exists some autocorrelation on the ask and the bid side for the daily data. The test results are shown in table 7.3.

Table 7.3. : Diagnostic test results of model (1)

ask side	autoc.	hetero	RESET	bid side	autoc.	hetero	RESET
daily	8.93	0.112	0.604	daily	15.74	0.089	0.328
weekly	0.01	0.458	1.130	weekly	2.68	0.555	0.737

Our estimation of model (2) shows clearly that there are imperfections in the Italian secondary market as we have several significant coefficients for different assets on both sides of the market and based on both the daily data as well as the weekly data.

On the bid side we get a significant positive serial covariance of price changes and the spread on daily data. Although the other spread coefficients are not significant we can observe that there is a tendency for asset specific constant terms to become negative with weekly data (table 7.4.).

Table 7.4. : Results of model (2)

	Ask side				Bid side			
	daily		weekly		daily		weekly	
coeff.	value	t-stat	value	t-stat.	value	t-stat.	value	t-stat.
spread	0.281	0.321	-0.001	-0.167	1.943	2.615	-0.001	-0.282
α1	0.085	5.870	0.012	0.093	0.086	7.002	0.033	3.328
α2	0.076	5.565	0.019	0.148	0.081	6.942	0.015	1.527
α3	0.003	0.235	-0.003	-0.027	0.001	0.073	0.005	0.539
α4	0.005	0.377	0.006	0.049	0.005	0.445	0.001	0.066
α5	0.006	0.414	-0.005	-0.036	-0.005	-0.400	0.001	0.029
α6	-0.008	-0.594	-0.489	-3.677	-0.001	-0.108	0.001	0.099
α7	0.003	0.221	0.006	0.045	-0.003	-0.258	0.001	0.143
α8	0.009	0.657	0.002	0.019	-0.001	-0.099	0.001	0.045
α9	0.012	0.885	0.005	0.036	0.005	0.401	-0.008	-0.827
α10	0.002	0.132	0.007	0.055	0.001	0.076	0.001	0.068
α11	0.014	0.888	0.009	0.068	-0.001	-0.078	-0.017	-1.687
α12	0.018	1.312	0.001	0.003	0.011	0.931	0.004	0.454
α13	0.007	0.453	-0.019	-0.144	-0.003	-0.241	0.001	0.109
α14	0.009	0.653	0.010	0.074	0.001	0.056	0.006	0.527
α15	0.031	1.659	-0.008	-0.064	-0.004	-0.221	0.012	1.210

The tests carried out reveal some autocorrelation for the weekly data on the ask side and we have heteroscedastic error terms for the daily data on the bid side. The RESET test is significant on the bid side for both the daily and the weekly data; there is evidence here to reject the hypothesis.

Table 7.5. : Diagnostic test results of model (2)

ask side	autoc.	hetero	RESET	bid side	autoc.	hetero	RESET
daily	0.52	1.351	0.435	daily	1.70	5.972	17.38
weekly	4.17	1.084	0.017	weekly	0.36	1.028	4.242

The positive serial correlation between the price changes and the bid-ask spread is confirmed by model (3) in one asset on the ask side and two assets on the bid side based on the daily data. None of the negative coefficients is significant at the 95 % level.

If we examine the coefficients of the weekly data we observe that the only significant value is negative which is on the bid side. The detailed results are shown in table 7.6. below.

Table 7.6. : Results of model (3)

	Ask side				Bid side			
	daily		weekly		daily		weekly	
coeff.	value	t-stat	value	t-stat.	value	t-stat.	value	t-stat.
β1	0.261	1.254	-0.326	-0.082	0.213	1.811	0.737	1.637
β2	1.874	1.164	1.454	0.164	4.519	3.069	1.283	0.830
β3	-2.422	-0.652	3.614	0.046	-3.275	-1.116	-2.271	-0.139
β4	-0.722	-0.263	5.163	0.097	-2.686	-0.800	0.297	0.065
β5	-2.595	-1.659	0.795	0.011	-1.204	-0.733	0.062	0.135
β6	3.545	1.786	-1.313	-0.070	-0.967	-0.892	0.011	0.055
β7	-1.121	-0.595	-0.276	-0.104	-1.324	-1.085	-0.150	-0.064
β8	-0.424	-0.508	0.107	0.178	-0.201	-0.351	-0.002	-0.043
β9	-19.75	-0.302	-0.079	-0.086	-33.37	-0.901	0.012	0.180
β10	-33.95	-0.720	-0.009	-0.029	-28.31	-1.058	-0.001	-0.016
β11	-4.979	-0.129	-0.003	-0.009	-6.473	-0.298	-0.060	-2.877
β12	-19.56	-0.402	-0.058	-0.168	-10.46	-0.379	0.003	0.135
β13	-18.44	-0.462	0.002	0.006	-12.91	-0.569	0.004	0.197
β14	-17.20	-0.462	0.128	0.221	-8.322	-0.394	0.005	0.132
β15	8.986	0.345	-0.017	-0.043	2.092	0.141	0.009	0.317
const	0.025	2.536	-0.039	-0.973	0.024	3.990	0.002	0.856

The diagnostic test results are different now for model (2). We cannot find any autocorrelation or heteroscedasticity, but, on the bid side, the RESET test is significant again for both the daily and the weekly data.The test parameters are listed in table 7.7.

Table 7.7. : Diagnostic test results of model (3)

ask side	autoc.	hetero	RESET	bid side	autoc.	hetero	RESET
daily	0.86	1.493	0.000	daily	1.40	0.695	8.717
weekly	0.02	0.027	0.027	weekly	0.14	0.779	28.09

7.5. Conclusions

The empirical analysis of this paper, based on data from the Italian secondary market for government bonds, has investigated evidence of the relation between the serial covariance of price changes and the bid-ask spread in a dealership market.

The most important finding is that, in contrast to Stoll's inventory control model,[8] the serial correlation tends to be positive. However, we can observe that the serial correlation tends to get negative if we use weekly data.

It can be argued that the positive serial correlation is caused by the fact that dealers, assumed to be risk averse, do not correct the whole size of their inventory adjustment in the next transaction, but gradually adjust over several periods to their preferred inventory position. Hence, as a consequence we can observe some positive relationship based on daily data as the price reversal will not take place after one transaction or even within one day.

Another important point is the missspecification problem on the bid side. It suggests that there is another influence on the price reversal. Considering that we used data of a dealership market with several competing

[8]Note, that it is equally contrary to Roll's transaction cost approach.

market makers it is very likely that there is strategic pricing behaviour in the market. This may have an adverse influence also on the expected serial correlation and could cause the misspecification.

Finally, we showed that there is a kind of "inefficiency" in the Italian secondary market. This is confirmed by the daily data regression as well as by the weekly data investigation. However, it seems that the bid side is slightly more efficient than the ask side.

CONCLUDING REMARKS

In this thesis we have analysed the profit margins of market makers in dealership markets. Particular interest is given to the notion of risk insurance for the market makers and the investigation of the degree of competition in the market.

We investigate the bid-ask spread in a centralised market structure and in a fragmented market where dealers do not have full information. We take a new approach compared to the existing theories and assume that market orders can be split between the best quoting dealers. Hence, we assume that market makers are risk averse and that they face decreasing returns to scale. Market makers can reduce their risk exposure by sharing a market order instead of trading the whole order and thus face an increase in risk exposure.

We have shown for both market structures, the centralised and the fragmented, that the equilibrium price is lower than in the traditional setting. By allowing the splitting of the order there may be the situation in the market that there is not the same number of active dealers on the bid side and on the ask side which implies that two or more different dealers buy and sell. This situation results in the fact that the bid-ask spread depends on the inventory levels of the market makers which is in contrast to Ho and Stoll. They claim that, in a one period framework, the bid-ask spread is independent of inventory positions.

Furthermore, we find that the spread is not the same for a centralised and a fragmented market. The spread in a fragmented market is larger than in a centralised market. This can be explained by the higher risk of trading which dealers face in a fragmented market which comes from incomplete information. However, with a very large number of dealers in the market, the spreads tend to be equal.

Our investigation of risk insurance is extended by analysing the influence of diversification of a dealer's portfolio on the bid-ask spread.

We assume that a dealer in the spot market may find that she is able to hedge some of her inventory risk by trading in futures contracts.

The effects of futures trading on the spot market bid-ask spread has not so far been investigated. Our prediction is that the market maker can reduce the risk exposure by trading in futures and thus the spot market bid-ask spread is smaller than without trading. However, this result is only obtained in a centralised market with one monopoly dealer who is assumed to execute also the transaction in the subsequent period. If there are different dealers active on each side of the market the finding is different. Depending on the covariance of the spot and the futures prices, the futures price expectations, and the inventory positions of the market makers the influence can be positive or negative which means that there is the possibility that the spot bid-ask spread is larger than without futures trading.

If we change our analysis in the way that the order flow in the spot market is not known to the market makers we get a different result.

By assuming that the buy and sell order quantities are identical we do not find any influence of futures trading on the spot bid-ask spread. This comes from the fact that, due to the uncertainty of the order flow, the expected inventory is the same with trading or without trading in the spot market and therefore trading in futures does not affect the spread.

The change of this assumption and by allowing asymmetry in purchases and sales gives an effect of futures trading on the spot spread. The parameters which determine the sign of the influence on the spot bid-ask spread are the inventory position, the covariance between the spot and the futures prices, the futures price expectation, and the difference between purchases and sales in the spot market. However, the model formula does not give us an unambiguous result. This problem could be investigated by an empirical analysis which may bring evidence of sign of the influence.

Our empirical studies are aimed to support some of our theoretical findings. We have shown that under the assumption of risk averse dealers and decreasing returns to scale the inventory control arguments are still valid. We find that especially on the ask side the market makers rely on their inventory positions. We observe a somewhat different result on the bid side. Dealers seem to take into account the inventory and the past prices by quoting their prices. In addition, our empirical investigation shows that we find some evidence of the next best dealer argument in the Italian secondary market for government bonds. We analysed the parameters which determine the reservation prices and we find that several dealers rely on the next best dealer's parameters such as the inventory position by quoting their prices. This result is supported more strongly on the bid side than on the ask side.

However, by applying a different measure of the bid-ask spread which is the serial covariance of price changes we find that in the short run the inventory control aspect does not come into effect. The reason may be that there are other factors which influence the pricing strategies of market makers such as strategic behaviour. However, in the long run, which means that we base our analysis on weekly instead of daily data, the inventory control argument is supported.

In order to put all the above empirical results into perspective, we have analysed the quoting behaviour of the market makers in this Italian market. The price quotes of the dealers exhibits a distinct pattern which shows that some dealers are very active in one or several assets. This leads to the question whether some dealers can make excess profits by trading in a particular asset. If so the market can be expected to be segmented and that there are arbitrage opportunities in the market which would mean that the market is inefficient. Our finding is that there are no statistically significant differences between the returns of the various assets, although we could find some inefficiency in the market.

Our research shows that there is still scope for further analyses in this area. In our investigation of the bid-ask spread we do not include any aspects of interdealer trading.

Although it is a valid argument to include interdealer trading in the analysis, we think that in our competitive market structure the competitive pressure on price almost eliminates the need of interdealer trading in order to balance the inventory position. In addition, the possibility of sharing the market order makes interdealer trading less attractive.

Another aspect which we have not explored is the trading procedure itself. As we have pointed out the design of markets is important in respect of market liquidity. In our bid-ask spread models we assume that the market makers quote their prices at the same time and trading is executed according to an auction procedure. Such an auction procedure depends on the transparency in the market. Most the existing models consider such a framework. Scope for further research is the change of such auction procedures to a two stage bargaining situation between the market makers and the private investors.

On the empirical side, it is interesting to see how the spot market bid-ask spread is affected by futures trading. Our model in chapter four, gives testable predictions which can bring some evidence on the interaction of spot and futures market.

REFERENCES

Amihud Y. and H.Mendelson, (1980),' Dealership Market: Market Making with Inventory', *Journal of Financial Economics* 8 (March) pp. 31-53

Arisi E., R. Baccolini, G. Lusignani (1988),' Il Mercato Secondario Italiano Dei Titoli Di Stato', *Prometea* (December), Bologna

Bagehot W., (1974), 'The Only Game in Town', *Financial Analysts Journal* 27, pp.12-14

Benston G.J.,R.L. Hagerman, (1974), 'Determinants of Bid-Ask Spreads in the Over-The-Counter Market', *Journal of Financial Economics* 1 (December) pp.353-64

Biais B. (1993), 'Price Formation and Equilibrium Liquidity in Fragmented and Centralized Markets', *Journal of Finance* (March) pp. 157-185

Bray M., (1981) ,'Futures Trading, Rational Expectations, and the Efficient Market Hypothesis', *Econometrica* 49, pp. 575-96

Breeden D.T., (1980),'Consumption Risks in Futures Markets', *Journal of Finance* 35 (May), pp. 503-20

Brennan M.J., (1958),'The Supply of Storage', *American Economic Review* 48 (March), pp. 50-72

Chiang A.C., (1984), 'Fundamental Methods of Mathematical Economics', third ed., McGraw-Hill International Book Company, Singapore

Choi J.Y., D.Salandro, K. Shastri, (1988), 'On the Estimation of Bid-Ask Spreads: Theory and Evidence', *Journal of Financial and Quantitative Analysis* vol.23 (June) pp.219-30

Cohen K.J., St.F. Maier, R.A. Schwartz, D.K. Whitcomb, (1981), 'Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread', *Journal of Political Economy* vol.89 no.21, pp.287-305

Cootner P.H., (1960),'Returns to Speculators: Telser vs. Keynes', *Journal of Political Economy* 68 (August), pp. 396-404

Copeland Th.E. and D.Galai, (1983), 'Information Effects On The Bid-Ask Spread', *Journal of Finance* no.38 (December) pp. 1457-69

Cox D.R., (1961), 'Test of Separate Families of Hypotheses', Proceedings of the Fourth Berkeley Symposium on *Mathematical Statistics and Probability*, vol.1, pp. 105-123

Danthine J.P., (1978),'Information, Futures Prices, and Stabilizing Speculation', *Journal of Economic Theory* 17, pp. 79-98

--- (1983),'Hedger Diversity in futures Markets', *The Economic Journal* 93 (June), pp. 370-89

Demsetz H., (1968), ' The Cost of Transacting', *The Quarterly Journal of Ecomomics* 82 (February) pp. 33-53

Dusak K., (1973),'Futures Trading and Investor Returns: An investigation of Commodity Market Risk Premiums', *Journal of Political Economy* 81 (November/December), pp. 1387-1406

Edwards F.R. and Ma C.W., (1992),'Futures and Options', (McGraw-Hill Inc. New York)

Ericsson N.R., (1983), 'Asymptotic Properties of Instrumental Variables Statistics for Testing Non-Nested Hypotheses', *Review of Economic Studies*, no. 50, pp. 287-304

Everitt B., (1974), 'Cluster Analysis: An SSRC Review of Current Research', Heinemann Educational Books Ltd. London

Fama E.F. and K.R. French, (1987),'Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage', *Journal of Business* 60, pp.55-73

French K.R., R. Roll, (1986), 'Stock Return Variances:The Arrival of Information and the Reaction of Traders', *Journal of Financial Economics* no. 17, pp. 5-26

Froot K.A., D.S. Scharfstein, and J.C. Stein, (1992),'Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation', *The Journal of Finance* (September) pp. 1461-1484

Garbade K.D. and W.L. Silber, (1976), 'Price Dispersion in the Government Securities Market', *Journal of Political Economy* vol. 84, pp.721-40

---- (1979), 'Structural Organization of Secondary Markets: Clearing Frequency, Dealer Activity and Liquidity Risk', *The Journal of Finance* no.3 (June), pp. 577-93

---- (1983),'Price Movements and Price Discovery in Futures and Cash Markets', *The Review of Economics and Statistics*, pp. 289-297

Garman M.B., (1976), 'Market Microstructure', *Journal of Financial Economics* no.3 (June) pp. 257-75

GENSTAT 5 Reference Manual (1987), Clarendon Press Oxford

Glosten L.R. (1989), 'Insider Trading, Liquidity, and the Role of the Monopolist Specialist', *Journal of Business* vol 62. no.2, pp. 211-35

Glosten L.R., L.E. Harris, (1988), 'Estimating the Components of the Bid/Ask Spread', *Journal of Financial Economics* vol. 21, pp.123-42

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

Godfrey L.G., (1978), 'Testing Against General Autoregressive and Moving Average Erros Models when the Regressors Include Lagged Dependent Variables', *Econometrica*, 46 pp. 1293-1301

Goldman M.B., A. Beja, (1979), 'Market Prices vs. Equilibrium Prices: Returns' Variance, Serial Correlation, and the Role of the Specialist', *The Journal of Finance*, no.3, pp. 595-607

Grimmett G.R., D.R. Stirzaker, (1987), 'Probability and Random Processes', Oxford Science Publications, Oxford

Grossman S., (1977),'The Existence of Futures Markets, Noisy Rational Expectations and Informational Externalities', *Review of Economic Studies* 44, pp.431-49

Grossman S.J., M.H. Miller, (1988), 'Liquidity and Market Structure', *The Journal of Finance*, no.3, (July), pp.617-37

Hasbrouck J., (1988), 'Trades, Quotes, Inventories, and Information', *Journal of Financial Economics*, no. 22, pp. 229-52

Hicks J.R., (1946),'Value and Capital', (2nd ed.; Oxford: Clarendon Press), pp. 136-39

Ho Th., H.R.Stoll, (1979), 'Optimal Dealer Pricing under Transactions and Return Uncertainty', Working Paper, New York University

---- (1980), 'On Dealer Markets under Competition', *Journal of Finance* 35, vol.2, (May) pp.259-68

---- (1981), 'Optimal Dealer Pricing under Transactions and Return Uncertainty', *Journal of Financial Economics* 9 (March) pp. 47-73

---- (1983), 'The Dynamics of Dealer Markets under Competition', *Journal of Finance* 38, (September) pp. 1053-74

Harris M., A. Raviv, (1981), 'Allocation Mechanisms and Design of Auctions', *Econometrica* vol.49, no. 6, (November), pp. 1477-99

Holden C., (1990), 'A Theory of Arbitrage Trading in Financial Market Equilibrium', Discussion Paper No. 478, Indiana University

Holthausen D.M., (1979),'Hedging and the Competitive Firm under Price Uncertainty', *American Economic Review* 69, pp. 989-95

Houthakker H.S., (1968),in:'Value, Capital, and Growth', ed.J.N. Wolfe, (Edinburgh University Press), pp. 193-214

Jaffe J.F., R.L. Winkler, (1976), 'Optimal Speculation against an Efficient Market', *The Journal of Finance*, no.1, (March), pp. 49-61

Kaldor N., (1939),'Speculation and Economic Stability', *Review of Economic Studies* 7 (October), pp. 1-27

Keynes J.M., (1930),'A Treatise on Money', Vol II: The Applied Theory of Money, (London: Macmillan & Co.), pp. 142-47

Kumar P. and D.J. Seppi, (1992),'Futures Manipulation with "Cash Settlement"', *The Journal of Finance* (September), pp. 1485-1502

Kyle A.S., (1985),'Continuous Auctions and Insider Trading', *Econometrica* 53 (November), pp. 1315-36

--- (1989), 'Informed Speculation with Imperfect Competition', *Review of Economic Studies* (January), pp. 317-356

Madhavan A., S. Smidt (1991), 'A Bayesian Model of Intraday Specialist Pricing', *Journal of Financial Economics*, 30 pp. 99-134

Madhavan A., (1992), 'Trading Mechanisms in Securities Markets', *The Journal of Finance*, (June) pp. 607-641

Newbery D.M., (1987),'When Do Futures Destabilize Spot Prices?', *International Economic Review* 28, pp. 291-96

O'Hara M., G.S. Oldfield, (1986), 'The Microeconomics of Market Making', *Journal of Financial and Quantitative Analysis*, vol. 21, (December) pp.361-76

Orme Ch., (1988), 'The Calculation of the Information Matrix, Test for Binary Data Models', *The Manchester School*, no. 4, pp. 370-76

Pagano M., (1989), 'Endogenous Market Thinness and Stock Price Volatility', *Review of Economic Studies* 56, pp. 269-288

Pagano M.,A. Röell, (1990), 'Trading Systems in European Stock Exchanges: current Performance and Policy Options', *Economic Policy* (April), pp. 63-115

---- (1992a), 'Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading', LSE Financial Markets Group

Discussion Paper No. 150, (October)

---- (1992b), 'Auction and Dealership Markets: What is the Difference?', *European Economic Review* 36, pp. 613-23

---- (1993a), 'Shifting Gears:An Economic Evaluation of the Reform of the Paris Bourse',

---- (1993b), 'Front-Running and Stock Market Liquidity',

---- (1993c), 'Auction Markets, Dealership Markets and Execution Risk',

all in: 'Financial Markets' Liberalisation and the Role of Banks, eds. V.Conti and R.Hamaui, Cambridge University Press, pp. 152-217

Riley J.G.,W.F. Samuelson (1981), 'Optimal Auctions', *The American Economic Review*, (June) vol. 7/no.3, pp. 381-392

Rolfo J., (1980),'Optimal Hedging under Price and Quantity Uncertainty', *Journal of Political Economy* 88, pp. 100-116

Roll R., (1984), ' A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market', *Journal of Finance* 39 (September) pp. 1127-39

Stein J.L., (1979),'Spot, Forward, and Futures', *Research in Finance*,vol I, pp. 225-310

Stoll H.R., (1976), 'Dealer Inventory Behaviour: An Empirical Investigation of NASDAQ Stocks', *Journal of Financial and Quantitative Analysis* 11 (September) pp.359-80

---- (1978), 'The Supply of Dealer Services in Security Markets', *The Journal of Finance*, no. 33, (September)

---- (1989), 'Inferring The Components of the Bid-Ask Spread: Theory and Empirical Tests', *Journal of Finance* pp. 115-134

Telser L.G., (1958),'Futures Trading and the Storage of Cotton and Wheat', *Journal of Political Economy* 66 (June), pp. 233-55

Tinic S.M., (1972), 'The Economics of Liquidity Services', *Quarterly Journal of Economics*, vol. 76, pp. 79-93

Tinic S.M., R.R. West, (1972), 'Competition and the Pricing of Dealer Services in the Over-The-Counter Stock Market, *Journal of Financial and Quantitative Analysis*, no.7, (June) pp. 1707-27

Tirole J. (1988), The Theory of Industrial Organization, Cambridge, Mass., MIT Press

Tuckman and Villa, (1992), 'Arbitrage with Holding Costs; A Utility based Approach', *The Journal of Finance*, (September) pp. 1283 ff.

Turnovsky S.J., (1979),'Futures Markets, Private Storage, and Price Stabilization', *Journal of Public Economics* 12, pp. 301-27

Vickrey W., (1961), 'Counterspeculation, Auctions, and Competitive Sealed Tenders', *Journal of Finance*, pp. 8-37

Weston J.F. and Th.E. Copeland, (1988), 'Managerial Finance', second UK ed., Cassell Educational Ltd., London

White H., (1980), 'A Heteroscedasticity-consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity', *Econometrica*, 48, 4, pp. 817-838

Working H., (1948), 'Theory of the Inverse Carrying Charge in Futures Markets', *Journal of Farm Economics* 30 (February), pp. 1-28

Yoo J. and G.S. Maddala, (1991),'Risk Premia and Price Volatility in Futures Markets', *The Journal of Futures Markets* 11, pp. 165-177