Is the market microstructure of stock markets important?

Randi Næs, senior adviser, and Johannes Skjeltorp, researcher, both in the Research Department

The market microstructure literature studies how the actual transaction process – i.e. how buyers and sellers find one another and agree on a price – can affect price formation and trading volumes in a market. This article provides an introduction to the concepts, frameworks and most important themes in this literature. The market serves two functions: one is to provide liquidity for buyers and sellers; the other is to ensure that prices reflect relevant information about fundamental value. Microstructure models differ from traditional financial models by recognising that legitimate information about companies' fundamentals may be unequally distributed between, and differently interpreted by, market participants. We can therefore no longer assume that prices will reflect information immediately even if all participants are rational. The microstructure literature argues that both information risk due to asymmetric information and differences in liquidity over time and between companies impact on long-term equilibrium prices in the market.

1 Introduction

If participants in the stock market behave rationally and have the same information, share prices will at all times reflect all available information about companies' fundamental value. Since it was first advanced in the 1960s, this has been one of the most important hypotheses in financial economics. However, over the last 20 years, both the theoretical foundation for this hypothesis and the previously strong empirical support for it have been challenged.

The microstructure literature challenges the hypothesis of efficient markets by studying how prices can deviate from (or converge towards) informationally efficient equilibrium prices as a result of rational participants behaving strategically (Biais et al., 2004).[1] Strategic behaviour can be put down to unequal access to information[2] or to limited liquidity[3] in the secondary market. While the efficient market hypothesis abstracts from the actual process which leads to buyers and sellers finding one another and agreeing on a price, the microstructure literature focuses on the functions performed by the marketplace.

Themes in the microstructure literature divide naturally into three: (i) the actual transaction process, (ii) the effects of market structure and trading rules on the transaction process, and (iii) the transaction process's implications for fundamental economic decisions. This subdivision also largely reflects the chronological development of this research field.

Models of the transaction process are described in section 2 below. There are two main groups of model. The first (inventory models) studies how an intermediary (hereinafter referred to as dealers, see figure 1) can solve the problem of buyers and sellers not being present in the market simultaneously. The second

(information models) analyses how information which is asymmetrically distributed between participants in the market is reflected in the prices of securities.

Research into the significance of market structure and trading rules is the subject of section 3 below. The importance of the organisation and design of the stock market came to the fore in the wake of the crash of 1987 and the revelation of collusion among the dealers on NASDAQ in 1994. There has since also emerged a considerable body of literature on the effects of market fragmentation and competition from new electronic trading systems.

Microstructure research rejects the hypothesis that the transaction process and the organisation of markets have no effect on the prices of securities. However, this does not necessarily mean that microstructure is important for our understanding of fundamental economic decisions. In section 4 we discuss a group of studies which look at whether the stock market's microstructure can also have long-term effects on prices and returns. Section 5 then sums up the most important contributions from the literature and highlights key themes and challenges in ongoing research.

2 The transaction process

2.1 Dealer markets versus limit order markets

A fundamental function of a market is to ensure that buyers and sellers find one another and have the opportunity to trade when they want to. One way of resolving the problem of coordination between buyers and sellers is to involve a dealer who undertakes to sell when

---

[1] The behavioural finance literature challenges the assumptions that investors are rational and have unlimited arbitrage opportunities, see, for example, Schleifer (2000).

[2] It is important to stress that, when talking about unequal access to information, we are not thinking of access to illicit inside information. For natural reasons, investors will differ in their ability and inclination to collect and analyse information. For example, large institutional investors, such as banks and insurers, have far greater resources than private investors and are therefore in a better position to obtain and analyse new information.

[3] In this article we use liquidity in the sense of how readily a share can be bought and sold in the secondary market.

Figure 1 Types of markets

(a) Dealer market

Dealer

Buyers

Bid price

Ask price
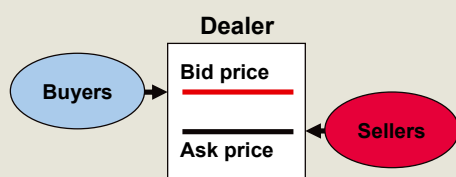
Sellers

In a **dealer market**, groups of intermediaries (dealers) are responsible for setting tradeable prices in their stocks. A dealer buys from sellers and sells to buyers on the basis of these prices. The dealer may also be required to provide a well organised/stable market in these shares.

(b) Limit order market

Order book

Sell orders

Level 3
Level 2
Level 1

Sellers

Buyers

Level 1
Level 2
Level 3
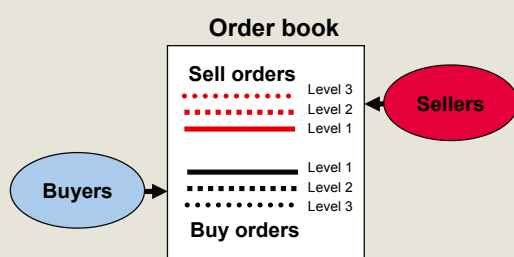
Buy orders

In a **limit order market**, investors themselves provide liquidity and set prices in the form of limit orders. A limit order is a buy or sell order for a volume and price determined by the buyer/seller. All limit orders are entered into the order book. A trade takes place when the prices of two orders cross, e.g. where a buy order is entered with a price at or above that of the sell order(s) with the lowest price (level 1). There can be more than one order at each level in the order book. Orders at the same price level are prioritised by time in most markets, i.e. the orders entered first are executed first.

somebody wants to buy, and to buy when somebody wants to sell. A trading system of this kind is illustrated in Figure 1(a).[4] To be able to perform this function, the dealer must ensure that he has an adequate inventory of shares. In return for providing this liquidity for buyers and sellers in the market, the dealer earns the difference between the bid price and the ask price (spread).

Another way of resolving the coordination problem is to gather together all buy and sell orders in a limit order book. Figure 1(b) illustrates a market of this kind. Buyers and sellers choose themselves whether they wish to provide liquidity by placing limit orders (orders to buy or sell at a given price) or demand liquidity by placing market orders (orders to buy or sell at the current price in the limit order book). In other words, a limit order market is not dependent on dealers. Trades are generated by electronically matching orders on the basis of set rules, orders typically being prioritised first by price and then by the time they were submitted to the market.

Some markets, known as hybrid markets, have come to include elements of both types of market. One example of a market of this kind is the New York Stock Exchange (NYSE), which has evolved from a dealer market into a hybrid market where the bulk of trading goes through the limit order book but where dealers (known as specialists) have to set prices if liquidity in the stocks for which they are responsible is too low. In

limit order markets, there are solutions where brokerage houses enter into agreements with listed companies to act as dealers in these companies' shares. Among other things, the broker must then ensure that the spread between bid and ask prices is not too large.[5]

2.2 Inventory models

Demsetz (1968) was the first to point out that there are costs associated with transacting shares. Besides explicit costs (such as stock exchange fees and brokers' commissions), there is also an indirect cost associated with getting to trade when you want to. As buyers and sellers do not necessarily need to trade at the same time, Demsetz argues that investors wanting to buy quickly must pay a higher price to motivate patient sellers to sell (and vice versa). Another important implication of his analysis is that the price at which you trade depends on whether or not you wish to buy or sell quickly, i.e. that there are two equilibrium prices rather than one.

The first microstructure models assumed optimal dealer behaviour. Garman (1976) looks at how a risk-neutral monopolistic dealer will set bid and ask prices in order to maximise expected profit per unit of time. The dealer wants to set prices to avoid bankruptcy, but must also ensure that prices are not set in such a way that his inventory empties. In Garman's model, the dealer sets

---

[4] There are a number of different terms for the intermediary between buyers and sellers in the stock market: market-maker, broker, dealer, specialist. We have chosen to use the term dealer. Specialist refers specifically to a dealer on the New York Stock Exchange (NYSE). Each stock on the NYSE has just one specialist, who has a duty to buy or sell up to a particular volume and is also responsible for ensuring a well organised/stable market in these shares. Dealers in other markets do not always have such stringent obligations. For example, there are often several different dealers in a particular stock.

[5] Weaver et al. (2004) look at the effect of such agreements on the Stockholm Stock Exchange and find that companies entering into such agreements see a clear improvement in liquidity in the secondary market. They also find support for a positive price effect for companies entering into these agreements which can be attributed to this improved liquidity.

prices once, after which buyers and sellers arrive in the form of two independent Poisson processes[6]. Garman shows that it is optimal for the dealer to set different bid and ask prices, and that both prices will be functions of the frequency at which buyers and sellers arrive. Thus his model explains why there is a positive spread in a dealer market.

Amihud and Mendelson (1980) expand Garman's model into a multi-period model where the dealer balances his inventory over time by changing his prices in each period. This model shows that optimal bid and ask prices fall monotonically with the size of the dealer's inventory. In other words, the dealer lowers both bid and ask price in response to a growing inventory (and vice versa when his inventory shrinks). This behaviour is known as quote shading. Thus Amihud and Mendelson's model also means that the dealer sets a positive spread; what is new in this model is that the optimal pricing strategy also takes account of the dealer wanting to keep his inventory of shares at a given level. Madhavan and Smidt (1991, 1993) and Hasbrouck and Sofianos (1993) find empirical support for dealers actually having such a desired inventory level, but also for them appearing to be willing to move away from this desired position for long periods. One empirical implication of inventory effects and quote shading is that they lead to a return towards "normal" stock returns (mean reversion).
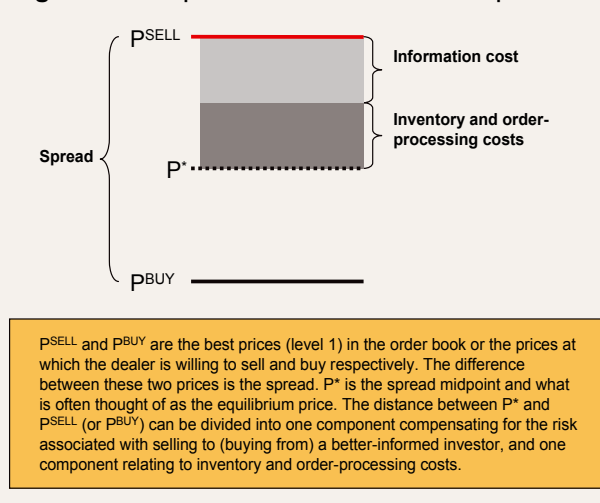
The main outcome of these inventory models is that dealers set bid and ask prices in such a way as to cover their order-processing and inventory-keeping costs.

2.3 Information models

The information models are to a great extent inspired by the insight of Bagehot (1971) that trading also entails a cost associated with some investors having better information than others. Like all other investors, informed investors can choose whether they want to trade or not, unlike the dealer who must always trade at the prices he sets. This means that, in cases where an informed investor wishes to trade, the dealer will always lose money. Copeland and Galai (1983) show that a dealer who cannot distinguish between informed and uninformed investors will always set a positive spread to compensate for the expected loss that he will incur if there is a positive probability of some investors being informed.

By expanding Copeland and Galai's model into a sequential trade framework, Glosten and Milgrom (1985) show how private information will be incorporated into prices over time. In their model, the dealer and other uninformed investors learn what the correct price is by observing the order flow. Thus the dealer takes account of information in the order flow when setting his prices. In this way, prices converge towards informationally efficient prices. However, the model

Figure 2 Composition of the dealer´s spread

P^{SELL} and P^{BUY} are the best prices (level 1) in the order book or the prices at which the dealer is willing to sell and buy respectively. The difference between these two prices is the spread. P* is the spread midpoint and what is often thought of as the equilibrium price. The distance between P* and P^{SELL} (or P^{BUY}) can be divided into one component compensating for the risk associated with selling to (buying from) a better-informed investor, and one component relating to inventory and order-processing costs.

says little about how quickly prices will converge on informational efficiency. Easley and O'Hara (1987) expand this framework to take account of a strategic element in the dealer's dilemma. In this model both informed and uninformed investors can choose between trading large or small volumes. If informed investors compete with one another, they will always want to trade large volumes in order to maximise their profit. The dealer can therefore set a different spread based on the behaviour of informed investors: investors placing small orders pay no spread, while investors wanting to make large trades have to pay a positive spread. If the informed investors know the dealer's strategy, they will want to mix their orders with those of uninformed investors (known as stealth trading). However, they will still tend towards large orders as they are also competing to exploit their private information before it is revealed and reflected in prices. In this case investors wanting to make small trades will also have to pay a positive spread, but this spread will be lower than that for large orders.

The main outcome of the early studies of the transaction process is that the spread has one component relating to information costs and one relating to inventory costs. This is illustrated in Figure 2, where P* is the equilibrium price and P^{SELL} is the price which a *buyer* has to pay to cover the dealer's two cost components. Similarly, a *seller* has to sell at a price below the equilibrium price in order to cover the dealer's cost components, which will be the difference between P* and P^{BUY}.

A more recent group of information models assumes that liquidity providers can also behave strategically as a result of having market power or access to private information. The development of these models coincides with the emergence of order-based trading systems.

Several single-period models show how prices will depart from equilibrium prices under full competition if the number of liquidity providers is limited and infor-

---

[6] A Poisson process is a random process which describes the probability of a number of events (in this case the number of arrivals of buyers and sellers) of a particular type within a given time interval.

mation costs are ignored, see Klemperer and Meyer (1989), Biais, Foucault and Salanie (1998) and Roell (1999). Calcagno and Lovo (1998) show that dealers who have private information will introduce "noise" in their quotations to avoid disclosing this information, but that quotations and trades will nevertheless reveal some information to the market.

There are also dynamic models studying optimal strategies for liquidity providers in limit order markets. Parlour (1998) shows that liquidity providers in limit order markets face a trade-off between price and time priority. Foucault (1999) shows that it will be optimal for investors to provide liquidity through limit orders when the spread is high. Conversely, it will be optimal to consume liquidity through market orders when the spread is low.

The main outcome of these recent information models is that liquidity providers with market power will earn oligopoly rents. This prediction is supported by, among others, the empirical studies of Christie and Schultz (1994) and Christie et al. (1994), which led to the revelation of price-fixing by dealers on the NASDAQ exchange.

2.4 The size of transaction costs

Keim and Madhavan (1998) divide total transaction costs into an explicit component and an implicit component. Explicit costs consist primarily of brokers' commissions, while implicit costs include the spread, possible price impacts as a result of a trade, and the opportunity cost associated with not getting to trade at the desired time. A large part of the empirical micro-structure research attempts to estimate transaction costs, especially the implicit costs of trading.

Estimating transaction costs is far from unproblem-atic. The different cost components are difficult to untangle from one another, and the data sets available typically consist of individual trades which are in many cases only parts of a larger transaction. We cannot therefore draw conclusions about total transaction costs based on estimates of unqualified cost components from different empirical studies. To be able to make a sensible estimate of transaction costs, detailed information about a trade is required right back to the time when the buyer or seller decided to trade. This is information which very few investors wish to share with the public. In recent years, there have nevertheless been several studies based on sufficiently detailed data from portfolio managers and investors. These data make it possible to estimate more precisely the cost associated with the entire transaction process, and so to obtain qualified cost estimates. The main outcome of these studies is that the implicit cost component may be considerable, relative to both explicit costs and realised portfolio returns, see Madhavan (1998).

3 The significance of market structure

On 19 October 1987, the Dow Jones index fell by 22.6 per cent without it being possible to point to any new information about companies' fundamental value.[7] This sparked off a lively debate about the significance of market structure and trading rules for price formation in the stock market. One important theme in the debate was the level of transparency, i.e. the amount of information about the transaction process to which participants in the market should have access. Christie and Schultz's revelation of price-fixing by NASDAQ dealers a few years later triggered fresh debate about market structure, this time with the emphasis on setting rules for providers of liquidity. The emergence of electronic limit order markets, where buyers and sellers provide liquidity themselves without having to go through a dealer, has contributed to further discussion of the role of dealers in the trading of securities. Another important theme has been the welfare implications of the stiff competition seen between traditional stock markets and new electronic market systems, including what stance the authorities should take on market fragmentation. One final theme is whether trading should be continuous or periodic. In a continuous trading system, participants can trade whenever they want; in a periodic system (auctions), trading is allowed only at specific points in time. Mendelson (1982) shows that periodic auctions are preferable from an efficiency perspective. This applies particularly to illiquid stocks when there is considerable uncertainty about fundamental value or the danger of a market downturn. In practice, however, it appears that demand for continuous trading is considerable.

If we look at existing stock markets around the world, there are major variations in market structure. While the US stock market consists of many different trading systems, most European countries have a centralised electronic trading system. Advances in electronic communications have meant that all major stock exchanges now operate with some form of limit order book. However, many exchanges remain dependent on dealers in various ways. When it comes to continuous versus periodic trading, the trend appears to be for markets to offer continuous trading while exploiting the efficiency of auctions at times when this is particularly important, such as at the opening or closing of the exchange and following special events when trading is suspended for a time.

3.1 Transparency

Dealer markets typically feature much lower levels of transparency than limit order markets. This goes for how much information is made public, who receives the information, and when the information is published. A

---

[7] The crash on Black Monday was almost twice as big as that on 29 October 1929, when the market fell by 11.7 per cent and triggered what would come to be known as the Great Depression.

number of theoretical studies have shown that increased transparency results in better liquidity and reduced transaction costs, see Admati and Pfleiderer (1991), Chowdhry and Nanda (1991), Forster and George (1992) and Benveniste et al. (1992). However, Madhavan (1995) shows that transparency can also reduce liquidity, because participants not wanting to reveal their interest in buying or selling will withdraw their orders from the market. Empirical and experimental studies have not come up with unambiguous results either when it comes to this issue. However, the literature is unanimous that participants trading on the basis of private information will prefer anonymous trading systems, while participants trading on the basis of pure liquidity needs – and particularly those who cannot signal this – will prefer high transparency. This means that changes in transparency will benefit some participants at the expense of others.

3.2 The dealer's role

The literature does not provide any unambiguous explanation of why so many stock exchanges are still largely based on dealers. One explanation is that it is too expensive for providers of limit orders to follow up the market. Another explanation is that dealers reduce the information costs in the market by having extensive contact with brokers (Benveniste et al. 1992). However, other studies argue that dealers increase information costs because they can trade at different times to other liquidity providers. For example, when a limit order is sent to the trading floor on the NYSE, the specialist can choose to step into the order and so stop it before it reaches the order book. A similar problem arises when opening the market, as the dealer can place his orders after all the other participants.

3.3 Market fragmentation

One noteworthy feature of many countries' stock markets is a persistently high degree of market fragmentation. In 2004, for example, NASDAQ's SuperMontage executed only around 17 per cent of the trading volume in the companies listed on NASDAQ, while the NYSE executed 78 per cent of trading by volume in the companies listed on the NYSE. In the USA, advances in electronic communications have meant that the traditional stock exchanges have run up against stiff competition from "alternative" trading systems, i.e. electronic limit order markets (ECNs) and crossing networks.

Crossing networks differ from other trading systems in that they do not contribute to price formation. Instead buyers and sellers agree to use a price from another marketplace, typically the closing price on the day the crossing transaction is carried out, or the value-weighted average price over the day.

Mendelson (1987) shows that market fragmentation can have both advantages and disadvantages. The disadvantages of fragmentation relate to reduced liquidity and increased price volatility in each submarket; the advantages relate to the increased quality of price signals. Thus the potential advantages of fragmentation do not apply to crossing networks, as they do not contribute to price formation.

Chowdry and Nanda (1991) argue that we should see the markets consolidating over time. This is because both informed and uninformed investors will benefit from flocking around a large exchange: informed investors because it is easier for them to hide their trades in a large order flow; uninformed investors because costs will be lower the more other uninformed investors there are in the order flow. Easley et al. (1996) argue that alternative marketplaces can survive in competition with a primary market by "skimming the cream" of the order flow, i.e. by offering uninformed investors a cheap alternative. A competing explanation is that a marketplace can complement the primary market by providing an opportunity for mutually beneficial transactions of large illiquid orders, see Seppi (1990). Several empirical studies find support for the hypothesis that alternative marketplaces compete with the primary market and "skim the cream" of the order flow, see Fong et al. (1999), Næs and Skjeltorp (2003) and Conrad et al. (2003). Chowdry and Nanda's arguments in favour of consolidation are based on an assumption of full competition in the market for liquidity providers. Two empirical studies find that fragmentation may be preferable for liquidity providers with market power, see Blume and Goldstein (1997) and Bessembinder and Kaufman (1997).

Næs and Skjeltorp (2003) find signs that crossing networks in the USA compete in the most liquid segment of the stock market. Simple simulations of trades carried out by the Norwegian Government Petroleum Fund in 1998 show that the transaction costs associated with crossing are very low. Based on extensive data from institutional investors in the US market, Conrad et al. (2003) came up with similar and more robust results. Transaction costs for trades through alternative trading systems are substantially lower than those for trades through traditional exchanges, especially for the most liquid shares. However, Næs and Ødegaard (2006) find that the cost savings achieved using crossing networks are partially offset by a cost associated with adverse selection. Informed investors in the network will reduce the probability of crossing good stocks, and increase the probability of crossing bad stocks. In line with this hypothesis, the authors find that the companies that cannot be bought in the network show a risk-adjusted excess return relative to the stocks that can be bought. These costs are not captured by the empirical measures of transaction costs used in the literature.

4 Long-term effects of the market's microstructure

In this section we look more closely at the literature studying the significance of market microstructure for long-term portfolio selection. This section is based largely on O'Hara (2003).

A market serves two important functions: one is to provide liquidity for buyers and sellers; the other is to ensure that new information is reflected in the prices of securities. For a market's microstructure to be able to impact on long-term portfolio selection, liquidity and/or underlying information risk must therefore impact on investors' long-term decisions.

4.1 Is there a liquidity premium?

There is an extensive body of literature concerning the relationship between transaction costs and expected stock returns. Theoretical studies generally find that transaction costs have insignificant effects on expected returns. The most cited work is Constantinides (1986). Constantinides looks at the effects of introducing a proportional transaction cost in a model where investors can invest in two assets and maximise the utility of an infinite consumption stream. Average demand for an asset falls sharply following the introduction of a transaction cost. Nevertheless, the transaction cost has only a second-order effect on the return on the asset in equilibrium. The expected utility of the future consumption stream is not sensitive to the deviations in asset allocation which the transaction cost entails.

Contrasting with these works are several studies which reveal an empirical relationship between returns and liquidity costs. The first and best-known of these studies is an article by Amihud and Mendelson dating from 1986 where they look at the relationship between stock returns, market risk (measured as beta) and spread for a selection of stocks on the NYSE during the period from 1961 to 1980. The data support the authors' hypothesis that expected return is an increasing and concave function of relative spread. The study links the liquidity premium to the level of the liquidity cost: stocks with high liquidity costs have higher returns than those with low liquidity costs. In the short term, if the cost associated with liquidity is high enough, this will obviously impact on the net return. But are these effects really big enough to influence returns in the longer term?

Amihud and Mendelson explain their results with a model where investors differ from one another by having different investment horizons. These investors buy and sell assets as part of a portfolio selection problem and have to pay transaction costs in the form of a spread. The model shows (i) that investors demand a higher return the higher the spread, and (ii) a clientele effect which moderates this excess return, especially for assets with the highest spread. Only investors with a long horizon will hold the most illiquid assets. In equilibrium, this means that return is an increasing and concave function of spread. Thus Amihud and Mendelson view spread as a kind of tax which some investors avoid by removing the stock from their portfolio, while others choose to pay this tax in return for compensation. The model also postulates that expected return net of transaction costs increases with the investment horizon, such that stocks with a high spread give their owners a higher net return. This means that investors with a long investment horizon can benefit from holding shares with a high spread.

Table 1 presents a much simplified version of Amihud and Mendelson's analysis applied to Norwegian data. The table shows the monthly return on five value-weighted portfolios of companies listed on the Oslo Stock Exchange, sorted by relative spread[8] during the period from 1980 to 2002.[9]

These data suggest that there may also be a liquidity premium in the Norwegian stock market. Average return, median return and maximum return are highest for the portfolio with the highest spread, and lowest for the portfolio with the lowest spread. Similarly, the minimum return is lowest/highest for the portfolio with the lowest/highest spread. The relationship between the standard deviation of the return and spread is less clear. Note that these figures have not been adjusted for market risk.

Amihud and Mendelson's article has been followed by numerous studies of the relationship between spread and return. Some find a link; others do not. One criti-

Table 1 The relationship between return and spread on the Oslo Stock Exchange (1980-2002)

Portfolio	Return (%)				
	Average	SD	Minimum	Median	Maximum
1 (lowest spread)	1.67	6.8	-27.3	19.93	19.9
2	2.43	7.6	-26.7	2.72	30.7
3	2.45	6.9	-18.7	2.38	25.2
4	3.07	7.8	-17.7	2.13	41.3
5 (highest spread)	3.55	7.7	-22.0	2.73	36.8

[8] Relative spread is the difference between the highest bid price and lowest ask price divided by the average of these two prices.

[9] The figures in the table were calculated by Bernt Arne Ødegaard at the Norwegian School of Management. The selection is limited to companies with a market value of more than NOK 1 million and a minimum number of days traded during the year of 20. Companies with a share price below NOK 10 have also been excluded. The portfolios have been constructed on the basis of the average relative spread the previous year.

cism that has been levelled at these studies is that the positive relationship between return and spread may be due to the return not being correctly risk-adjusted. The argument here is that the spread is derived from prices, and prices can be correlated with the asset's market risk (market beta), such that any relationship between spread and return may be due to error in measuring the company's risk.

If the level of liquidity costs is priced into the market, investors who have a long investment horizon – and are therefore less dependent on good liquidity – may earn a premium from investing in illiquid assets.

Another group of studies explores whether there is also a relationship between stock returns and fluctuations in liquidity costs, i.e. whether expected illiquidity in the market as a whole impacts on expected stock returns. The hypothesis is that liquidity costs vary with time for the market as a whole, and that investors demand compensation for carrying this market-related risk.

Amihud (2002) finds support for this view. He measures the market's liquidity as the average daily absolute return over (dollar) trading volume on the same day. Liquidity is good if this liquidity measure is low, as this means that more volume is needed to move prices, and vice versa. Amihud also assumes that investors expect this variable to follow an autoregressive process. The hypothesis is that a reduction in expected market liquidity has both an income effect and a substitution effect. All companies will see a drop in prices to compensate for reduced liquidity. However, since investors will tend to substitute away from the least liquid companies in favour of more liquid companies, there will also be an increase in some prices. Two other studies which explore whether expected return is a function of the variability in liquidity are Chordia et al. (2001) and Pástor and Stambaugh (2003). Chordia et al. test whether risk-averse investors demand a higher expected return from companies with high variability in company-specific liquidity, measured as volatility in trading volume. They do not find support for this hypothesis. In contrast, they find a significant negative relationship between return and variability in company-specific liquidity. Pástor and Stambaugh (2003) test whether systematic (rather than company-specific) liquidity risk is important for companies' expected returns. If systematic liquidity risk is priced, companies whose return is closely correlated with fluctuations in market liquidity will have a higher expected return than companies whose return has a low correlation with fluctuations in market liquidity. Pástor and Stambaugh find support for this hypothesis when volatility in market liquidity is measured as average volatility in order flows across all companies. Adjusted for market risk and exposure to other risk factors (size, book value relative to market value, and momentum), companies with the highest liquidity risk show an annual excess return of 7.5 per cent relative to companies with a low liquidity risk.

Although support has been found for a relationship between liquidity risk and expected return, these results still remain to be explained. So far, the literature has offered no simple, testable theories for how liquidity risk should be priced. Asymmetrical information can hardly be the main explanation, as it is hard to imagine any investors having private information about broad market movements. Pástor and Stambaugh (2003) suggest that investors care about liquidity risk because their wealth has a tendency to fall when the market becomes less liquid (and transaction costs rise). A drop in the value of their wealth can also result in a need to liquidate part of the portfolio at a time when it is expensive to trade. This will be particularly important where a portfolio with high liquidity exposure is debt-financed and the drop in the value of the investor's wealth forces expensive sales to cover margin requirements. One example of how important this liquidity risk can be is the collapse of Long Term Capital Management (LTCM) in 1998. LTCM had extremely high liquidity exposure in its global portfolio, which consisted of a short position in liquid instruments and a long position in less liquid instruments. The Russian debt crisis led to the company's portfolio falling dramatically in value, which resulted in the company being forced to liquidate an increasingly large part of its portfolio to meet margin requirements. Ultimately, the whole portfolio was liquidated. In the years prior to the debt crisis, this liquidity position yielded an extremely high realised return, which Pástor and Stambaugh believe to be a reflection of the high liquidity risk at LTCM.

4.2 The significance of information risk

Known pricing models, such as the capital asset pricing model (CAPM), consumption-based CAPM and arbitrage pricing theory (APT), all build on investors having symmetric information about expected return and risk for all assets in the market. One justification for this simplification has been that information is only important for the market as a whole. The price of a share is determined by the return's covariance with the return on all shares, as no investor will hold idiosyncratic risk in equilibrium. It can therefore be assumed that all participants have symmetric information, even though this is not necessarily the case for individual stocks.

One problem with this justification is that the balance between expected return and risk is dependent on being able to calculate the market's expectations. What if we do not have the same information? Whose expectations are we then calculating? O'Hara (2003) shows that, if information is asymmetrically distributed, and if those who do not have information know that others know more, we will not get an equilibrium where everyone holds the market portfolio. Uninformed investors will hold a larger share of assets which informed investors

expect to perform poorly. They will demand compensation for this, and we will then no longer have a situation where idiosyncratic risk is not priced.

There are several empirical studies which attempt to estimate the effect of information costs. Brennan and Subrahmanyam (1996) find a positive relationship between return and information costs, where information costs are measured as the coefficient in a regression which relates price changes to the size and sign of order flows (Kyle's lambda). In Kyle's model, lambda arises on account of informed investors behaving strategically, i.e. lambda is a measure of adverse selection. Brennan and Subrahmanyam argue that unfavourable selection is the primary cause of illiquidity, and use Kyle's measure as a proxy for these costs. Easley, Hvitkjaer and O'Hara (2003) look at the relationship between return and an estimate of the probability of informed trading (PIN). The PIN is estimated by looking at the relationship between the number of buy and sell orders during the course of a day. If there are no informed trades, this relationship should be close to 50/50. An excess of trades on one side of the market suggests informed trading. This measure proves to have an economically and statistically significant effect on return, even after correction for beta and Fama/French risk factors (size and book value relative to market value).[10]

It has long been claimed in investment theory that unequal access to information about a company is important for the company's capital costs, see, for example, Mayer and Majluf (1984). This does not tie in well with the pricing models from financial theory, because the distribution of information about individual companies represents idiosyncratic risk which is not supposed to play any role in expected return in equilibrium. O'Hara's model can explain this apparent discrepancy. The model can also help to explain the equity premium puzzle: if uninformed investors demand company-specific compensation for risk, they will hold more bonds in equilibrium. A third application is the home bias puzzle: investors in one country are poorly informed about assets abroad, and therefore demand compensation for holding foreign securities.

5 Summary

This article provides an overview of the financial literature which argues that market microstructure plays a role in the pricing of securities.

Studies of the transaction process and market structure have provided a significant insight into the composition and significance of transaction costs. The first theoretical models showed that the spread between bid and ask prices is determined by inventory costs and costs associated with asymmetric information. More recent models show that the spread will also reflect oligopoly rents if liquidity providers have market power. Market access and competition between providers of liquidity may impact on costs relating to market power, while transparency and equal terms between liquidity providers are important for information costs. However, there is no single market structure which is best for all participants. Empirical studies show that implicit cost components – including spread costs, price effects from trading, and opportunity costs from not trading – are both statistically and economically significant.

This section of the literature is of obvious interest to market participants wishing to minimise the cost of trading, and for authorities responsible for regulating the securities markets. Innovation in communications and computer technology has led to the rapid emergence of order-driven trading systems. This trend has spurred new theoretical research based on strategic liquidity providers with private information or monopoly powers. The emergence of transparent order-based trading systems has also meant that significantly better data has become available for research purposes, opening the door for numerous empirical studies of the efficiency and cost of order-based trading systems in the future.

Another – and very active – section of the literature argues that microstructure is also important for our understanding of fundamental economic decisions. This will be the case if information risk and/or differences in liquidity over time and between companies impact on long-term equilibrium prices in the market. Several empirical studies find a positive relationship between stock returns and various measures of liquidity costs. Some find that it is the level of liquidity which impacts on the return, while others find that it is systematic fluctuations in liquidity which are priced in the market. One theoretical study argues that idiosyncratic risk will be priced if information is asymmetrically distributed and rational participants demand compensation for informational disadvantages. This breaks radically with the well-known result of financial theory that idiosyncratic risk can be diversified away, and allows for the possibility of microstructure playing an important role in long-term returns on the stock market.

Contributing more knowledge about how private information affects long-term expected return is one of the most important challenges facing microstructure research. One important element of this work will be to find good empirical measures which can be used to test the hypothesis that the market prices information risk.

---

[10] A ten per cent increase in PIN gives a 2.5 per cent increase in return.

130

References

Admati, A.R. and P. Pfleiderer (1991): "Sunshine Trading and Financial Market Equilibrium", *Review of Financial Studies*, 4, pp. 443-481

Amihud, Y. and H. Mendelson (1980): "Dealership Market: Market Making with Inventory", *Journal of Financial Economics*, 8, pp. 31-53

Amihud, Y. and H. Mendelson (1986): "Asset Pricing and the Bid-Ask Spread", *Journal of Financial Economics,* 17, pp. 223-249

Amihud, Y. (2002): "Illiquidity and stock returns: cross section and time series effects", *Journal of Financial Markets,* 5, pp. 31-56

Anand, A.; C. Tangaard and D.G. Weaver (2004): "Paying for Market Quality", 2004, *Working Paper Rutgers Business School*

Bagehot, W. (1971): "The Only Game in Town", *Financial Analysts Journal*, 27, pp. 12-14

Benveniste, L.; A. Marcus and W. Wilhelm (1992): "What's special about the specialist?" *Journal of Financial Economics,* 32, pp. 61-86

Bernhardt, D. and E. Hughson (1997): "Splitting orders", *Review of Financial Studies,* 10, pp. 69-101

Bessembinder, H. and H. Kaufman (1997): "A cross-exchange comparison of execution costs and information flow for NYSE-listed stocks", *Journal of Financial Economics*, 46, pp. 293-319

Biais, B.; T. Foucault and F. Salanie (1998): "Floors, dealer markets and limit order markets", *Journal of Financial Markets,* 1, pp. 253-284

Biais, B.; D. Martimort and J. Rochet (2000): "Competing mechanisms in a common value environment", *Econometrica*, 68, pp. 799-838

Biais, B.; L. Glosten and C. Spatt (2004): "Market microstructure: A survey of microfoundations, empirical results, and policy implications", *CEPR Discussion Paper* No. 3288

Blume, M. and M. Goldstein (1997): "Quotes, order flow, and price discovery", *Journal of Finance*, 52, pp. 221-244

Brennan, M. and A. Subrahmanyam (1996): "Market microstructure and asset pricing: On the compensation for illiquidity in stock returns", *Journal of Financial Economics,* 41, pp. 441-464

Calcagno, R. and S. Lovo (1998): "Bid-ask price competition with asymmetric information between market makers". Working paper, CORE

Chordia, T.; R. Roll and A. Subrahmanyam (2000): "Commonality in liquidity", *Journal of Financial Economics,* 56, pp. 3-28

Chordia, T.; A. Subrahmanyam and V.R. Anshuman (2001): "Trading activity and expected stock returns", *Journal of Financial Economics,* 59, pp. 3-32

Chowdhry, B. and V. Nanda (1991): "Multimarket trading and market liquidity", *Review of Financial Studies,* 4, pp. 483-511

Christie, W.G. and P.H. Schultz (1994): "Why do Nasdaq market makers avoid odd-eighth quotes?" *Journal of Finance,* 49, pp. 1813-1840

Christie, W.G. and P.H. Schultz (1999): "The initiation and withdrawal of odd-eighth quotes among Nasdaq stocks: an empirical analysis", *Journal of Financial Economics*, 52, pp. 409-442

Conrad, J.; K.M. Johnson and S. Wahal (2003): "Institutional trading and alternative trading systems", *Journal of Financial Economics,* 70, pp. 99-134

Constantinides, G. (1986): "Capital market equilibrium with transaction costs", *Journal of Political Economy,* 94, pp. 842-862

Copeland, T. and D. Galai (1983): "Information Effects and the Bid-Ask Spread", *Journal of Finance*, 38, pp. 1457-1469

Demsetz, H. (1968): "The Cost of Transacting", *Quarterly Journal of Economics*, 82, pp. 33-53

Easley, D. and M. O'Hara (1987): "Price, Trade Size, and Information in Securities Markets", *Journal of Financial Economics,* 19, pp. 69-90

Easley, D.; S. Hvidkjaer and M. O'Hara (2002): "Is information risk a determinant of asset returns?", *Journal of Finance,* 57, pp. 2185-2222

Easley, D.; N. Kiefer and M. O'Hara (1996): "Creamskimming or profit-sharing? The curious role of purchased order flow", *Journal of Finance,* 51, pp. 811-833

Easley, D. and M. O'Hara (2003): "Microstructure and Asset Pricing", *Handbook of the Economics of Finance*

Fama, E.F. and K.R. French (1993): "Common Risk

Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics,* 33, pp. 3-56

Fong, K.Y.L.; A. Madhavan and P.L. Swan (2001): "Why do Markets Fragment? A Panel-Data Analysis of Off-Exchange Trading", Working Paper EFA 0729

Forster, M. and T. George (1992): "Anonymity in Securities Markets", *Journal of Financial Intemediation,* 2, pp. 168-206

Foucault, T. (1999): "Order flow composition and trading costs in a dynamic limit order market", *Journal of Financial Markets,* 2, pp. 99-134

Garman, M. (1976): "Market Microstructure", *Journal of Financial Economics,* 3, pp. 257-275

Glosten, L. and P. Milgrom (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders", *Journal of Financial Economics,* 13, pp. 71-100

Glosten, L. (1994): "Is the electronic open limit order book inevitable?", *Journal of Finance,* 49, pp. 1127-1161

Harris, L. (1990): "Liquidity, Trading Rules, and Electronic Trading Systems", *Monograph Series in Finance and Economics,* 4, pp. 1-62

Harris, L. (2003): *Trading and Exchanges. Market Microstructure for Practitioners,* Oxford University Press

Hasbrouck, J. and D. Seppi (2001): "Common factors in prices, order flows, and liquidity", *Journal of Financial Economics,* 59, pp. 383-411

Hasbrouck, J. and G. Sofianos (1993): "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists", *Journal of Finance,* 48, pp. 1565-1594

Heaton, J. and D. Lucas (1996): "Evaluating the effects of incomplete markets on risk sharing and asset prices", *Journal of Political Economy,* 104, pp. 443-487

Huberman, G. and D. Halka (2001): "Systematic liquidity", *Journal of Financial Research,* 24, pp. 161-178

Klemperer, P. and M. Meyer (1989): "Supply function equilibria in oligopoly under uncertainty", *Econometrica,* 57, pp. 1243-1277

Madhavan, A. (1995): "Consolidation, Fragmentation, and the Disclosure of Trading Information", *The Review of Financial Studies*, 8, pp. 579-603

Madhavan, A. (2000): "Market Microstructure: A Survey", *Journal of Financial Markets,* 3, pp. 205-258

Madhavan, A. and S. Smidt (1991): "A Bayesian Model of Intraday Specialist Pricing", *Journal of Financial Economics,* 30, pp. 99-134

Madhavan, A. and S. Smidt (1993): "An Analysis of Daily Changes in Specialists' Inventories and Quotations", *Journal of Finance,* 48, pp. 1595-1628

Mendelson, H. (1982): "Market behaviour in a clearing house", *Econometrica,* 50, pp. 1505-1524

Mendelson, H. (1987): "Consolidation, Fragmentation and Market Performance", *Journal of Financial and Quantitative Analysis,* 22, pp. 189-207

Næs, R. and J.A. Skjeltorp (2003): "Equity trading by institutional investors: Evidence on order submission strategies", *Journal of Banking and Finance,* 27, pp. 1779-1817

Næs, R. and B.A. Ødegaard (2006): "Equity trading by institutional investors. To cross or not to cross?" Forthcoming in *Journal of Financial Markets*

O'Hara, M. (2003): "Presidential address: Liquidity and price discovery", *Journal of Finance,* 58, pp. 1335-1354

Parlour, C. (1998): "Price dynamics in limit order markets", *Review of Economic Studies,* 11, pp. 789-816

Pastor, L. and R. Stambaugh (2003): "Liquidity risk and expected stock returns", *Journal of Political Economy,* 111, pp. 642-685

Roëll, A. (1999): "Liquidity in limit order book markets and single price auctions with imperfect competition", Working Paper, Princeton University

Seppi, D. (1990): "Equilibrium block trading and asymmetric information", *Journal of Finance,* 45, pp. 73-94

Vayanos, D. (1998): "Transaction costs and asset prices: a dynamic equilibrium model", *Review of Financial Studies,* 11, pp. 1-58

Vayanos, D. and J. Vila (1999): "Equilibrium interest rates and liquidity premium with transaction costs", *Economic Theory*, 13, pp. 509-539

132