ABSTRACT

In this paper, I analyze the economic forces that drive liquidity in stock markets. Liquidity is a key element for well-functioning stock markets as it has important repercussions for traders, trading venues (stock exchanges or alternative trading systems) and listed firms. Moreover, also the stability of the financial system as whole benefits from liquidity. I first provide a definition of liquidity. Subsequently, several dimensions of liquidity within one market are analyzed. Next to order processing costs, inventory and asymmetric information, the relation between market design and liquidity is studied as well. Finally, I investigate the impact of intermarket competition on liquidity. I not only consider competition between stock exchanges, but also include alternative trading systems.

* This paper is based on the first chapter of my Ph.D. dissertation. I would like to thank HansDegryse, Frank de Jong, Hans Dewachter, Dirk Heremans, Patrick Van Cayseele and Marno Verbeek for valuable comments.
I. INTRODUCTION

This paper aims to analyze the forces that drive liquidity in stock markets, as well as of the implications of liquidity for different agents in the market, i.e. traders/investors, stock exchanges and listed firms. Although essentially the same asset is traded, a stock, the institutional setup of the trading process itself differs considerably across stock markets around the world. Examples include whether or not market makers or a trading floor are present, rules on transparency, .... Given this myriad of possible organizations of exchanges and the recent evolution of them, the performance of the various types of markets is an important issue. A key aspect in evaluating the performance of a particular trading system is its liquidity\(^1\). A clear illustration of the fact that the trading environment and liquidity have an impact can be found in a recent study by Harris, Panchapagesan and Werner (2006). They investigate delistings from NASDAQ to the pink sheets. These were accompanied by a large decline in liquidity since spreads almost tripled, as did volatility. This decline in liquidity is associated with a significant wealth-loss to shareholders of about 19%. Liquidity is of crucial importance to a number of agents. It is obviously a concern to traders, since it determines their cost of buying and selling assets. However, it is also an important aspect for stock exchanges themselves in order to attract order flow from traders and in the competition for order flow with other exchanges or alternative trading systems. This may create a liquidity externality: a liquid market attracts more order flow, creating in this way even more liquidity. Liquidity is also an argument in convincing firms to list on their exchange, as it is a determinant of their cost of capital and their decision about the optimal capital structure.

In this paper, after providing a brief description of stock markets in Section II, I tackle different issues. First, in Section III, I provide a definition of liquidity. Although at first sight a simple and straightforward concept, it is often not entirely clear what precisely comprises liquidity. This lack of a uniform definition has lead to different interpretations of liquidity and consequently to different views. Therefore, I clearly state in Section III the definition of liquidity to which I adhere in the paper. This definition accounts for the fact that liquidity is not a monolithic concept, but implies different interacting dimensions.

Secondly, Section IV investigates the question whether liquidity is a desirable feature. This is indeed a rather common view today, not only
among academics, but also among practitioners and in the financial press. But this has not always been the case, and a priori, one could argue that it is not altogether straightforward why liquidity would be beneficial. In Section IV, I confront two views of liquidity. The “dark” view of liquidity starts from the fact that low trading costs might benefit the individual, but may impose a cost to the market. In some cases, this might even lead to financial instability in the financial system. The policy implication in this case would be to restrict liquidity. In contrast, the “bright” view of liquidity counters these arguments and highlights the positive effects of liquidity. Rather than imposing a cost, it attracts traders to the market by facilitating and benefiting the trading process. One could say that in this way, a positive liquidity externality is created. In this view, liquidity should be enhanced by authorities and not restrained. The latter, bright view of liquidity is also termed by O’Hara (1995) the market microstructure view to liquidity. It is this approach that will be taken in this paper. Market microstructure is the part of the literature in financial economics that focusses on the economic forces underlying the trading process and price formation. Therefore, it is particularly well suited for studying the impact of the design of stock markets on liquidity. Moreover, it starts from the behavior of individual agents, be this a trader, an exchange, or a dealer.

Subsequently, in Section V, I discuss the different dimensions of liquidity in quote driven (dealer) and order driven (limit order) markets. Also, I investigate the impact of other elements in market design, i.e. transparency, anonymity, tick size and floor versus electronic trading on liquidity.

Until that point in the paper, the focus is entirely on one single market. However, nowadays trading of a particular stock often takes place simultaneously in a number of trading venues. This competition between trading venues, sometimes with different institutional characteristics, can be expected to have an impact on liquidity. This topic is studied in Section VI. Clearly this section is relevant for stock exchanges. Finally, all main results are summarized in Section VII.

This paper is not to first to provide an overview of the market microstructure literature. An excellent study of the early theoretical literature can be found in O’Hara (1995). Madhavan (2000) and Biais, Glosten and Spatt (2005) provide surveys of both theoretical and empirical work. Lyons (2001) provides an accessible review of the microstructure of foreign exchange markets. This paper differs from these surveys in its scope. I focus on a crucial property of a financial markets: its
ability to provide liquidity. This focus allows for an in-depth discussion of both determinants and implications. The determinants include trader behavior, but also the design of a market. Moreover, I show that inter-market competition and interaction have important implications. Furthermore, I do not limit the implications of liquidity to traders, but also show the importance for exchanges and listed firms. To my knowledge, this is the first paper to provide a comprehensive and integrated overview of all these various aspects of liquidity.3

II. INSTITUTIONAL SET-UP OF STOCK MARKETS

Throughout the world, a wide variety of trading systems exists4, allowing agents to trade financial assets, such as stocks, bonds, foreign exchange and derivatives. Each system has its own properties and rules. The differences between them relate to a number of market characteristics such as the presence of market makers, pre- and post-trade transparency or the type of orders allowed. For instance, because they do not provide extensive information about order flow, or about binding bid and ask quotes, foreign exchange markets offer in general less transparency than stock exchanges (see Lyons (2001)).

In this paper I restrict the attention to the analysis of stock markets. But even between stock markets, large differences exist, see e.g. Domowitz and Steil (1999) or Jain (2002) for a survey. In general, stock markets can be classified along two main lines: order driven versus quote driven systems, and periodic versus continuous systems. In a quote driven or dealer market, a market maker posts bid and ask prices at which he wants to buy or sell and takes the opposite side of each trade. In an order driven or limit order market, traders interact directly with each other without intermediation. Periodic systems only allow trading at specific points in time, while in continuous markets, trading can occur anytime when the market is opened. In practice, often hybrid forms are in place. For instance, on NYSE, market makers (specialists) compete with a public limit order book. On Euronext (a limit order market), market makers exist for some very small stocks.

In addition, stock exchanges often use a batch auction (periodic trading) after a pre-opening or closing period during which orders can be submitted but no trading takes place, while relying on continuous trading during the day. Next to these two main characteristics, each market has its own trading rules, including those on the type of orders
allowed, minimum price variation, opening procedure, transparency, and a floor versus screen based trading system. Moreover, it goes without saying that the set-up of an exchange is not static. For instance, the London Stock Exchange, originally a pure dealer market, introduced a limit order book for stocks in the FT100 index from 20 October 1997 onwards (the SETS-system). Also rules about transparency, anonymity, ..., have been prone to change. Finally, a recent development is the emergence of alternative trading systems (ATS). These compete with the “traditional” exchanges for order flow. In general ATS can be divided into two main groups: electronic communication networks (ECNs) and crossing networks (CNs)\(^5\).

Table 1 presents an overview of a number of financial markets and their main characteristics. It demonstrates some of the key differences between major financial centres as NYSE or Euronext. Crossing networks and to a lesser extend FOREX markets are opaque trading venues since pre- and post-trade transparency are low.

<table>
<thead>
<tr>
<th>Trading Systems: Overview</th>
<th>Euronext</th>
<th>Euronext Pre-Opening</th>
<th>NYSE</th>
<th>CNs</th>
<th>ECNs</th>
<th>FOREX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Floor-Based</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limit Order Book</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Dealer</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Pre-trade Prices</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Post Trade Info</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

*Note:* Euronext may have a specialist for very small stocks to guarantee liquidity. This is however a decision at the discretion of the listed firm. FOREX markets have indicative pre-trade prices but these are not binding, in contrast to a dealer's quotes on NYSE. Crossing networks (CNs) have a book in which orders are stored, but this book only contains orders specifying a quantity and not a price in contrast to limit orders in a limit order book.
III. DEFINITION OF LIQUIDITY

Despite the omnipresence of the concept “liquidity” in the finance literature of several domains and despite being seemingly a simple concept, there is surprisingly little agreement about its definition. O’Hara (2004) makes a comparison with pornography: it is hard to define, but you know it, when you see it. An early definition of liquidity can be found in Keynes (1930), who considers an asset as more liquid if it is more certainly realizable at short notice without loss. Within the context of this paper, I start from the following definition, which goes along the same lines: “a market is liquid if traders can quickly buy or sell large numbers of shares without large price effects”. It refers to the willingness of a market participant, a liquidity supplier, to take the opposite side in a transaction initiated by another trader, the liquidity demander. Note that a supplier can be a dealer but also another trader, as in limit order markets. Although having the advantage of simplicity, the definition is also general and hard to implement in practice when analyzing liquidity. In order to make the definition more specific, Black (1971), O’Hara (1995) and Harris (1990) identify several dimensions of liquidity. Harris (1990) distinguishes four. The first one is width, referring to the bid-ask spread for a given number of shares and commissions and fees to be paid per share. Secondly, depth is the number of shares that can be traded at given bid and ask prices. The third one, immediacy, refers to how quickly trades of a given size can be done at a given cost. The final aspect is resiliency. It characterizes how fast prices revert to former levels after they changed in response to large order flow imbalances initiated by uninformed traders. It is clear that these different dimensions do not stand independently on their own, but may interact with each other. For instance, if a trader is patient and does not need to trade immediately, she may obtain better prices and/or be able to trade a larger amount at given prices. In this case width and depth depend on immediacy.

IV. LIQUIDITY: A TWO-SIDED TALE

After having defined the concept of liquidity, the natural question arises whether liquidity is important and desirable? This question has attracted quite some controversy over the years. O’Hara ((1995), (2004)) summarizes the two opposing sides under a “dark” view and a “bright”
view on liquidity. The dark side points to the dangers of liquidity and is illustrated by Keynes ((1935), p. 155). He writes: “Of all the maxims of orthodox finance none, surely, is more anti-social than the fetish of liquidity, the doctrine that it is a positive virtue on the part of investment institutions to concentrate their holdings of ‘liquid’ securities. It forgets that there is no such thing as liquidity of investment for the community as a whole”. In this view, liquidity is a source of destabilization in markets. The reasoning is that liquid markets are focused mainly on the short term and investors do no longer consider fundamentals when making their investment decision. The resulting instability can affect other markets and this contagion might lead to instability in the financial system as a whole. Moreover, liquidity can be linked to corporate governance problems. Coffee (1991) notes: “The liquidity promoted by U.S. policies has obvious benefits: investors can encash their assets quickly and diversify cheaply. The same policies, however, impair corporate governance by encouraging diffuse stockholding and discouraging active investing. Diffuse stockholders face more serious collective action problems …”. Adherents to this view thus recommend to restrict liquidity. The costs of trading should be increased to force traders to internalize the social cost of liquidity. Measures consistent with this view include circuit breakers and short sale restrictions. It is also the reasoning behind the famous suggestion of Tobin when suggesting a tax on financial speculation: “My proposal is to throw some sand in the wheels of our excessively efficient international markets”.

Contrary to the dark side just discussed, the “bright” view stresses the importance and the benefits of liquidity for a number of agents in financial markets, including traders, stock exchanges and listed firms. The reasoning is that investors are more likely to participate in the market if they are able to buy and sell stocks easily, quickly and at low costs, or in other words: when liquidity is high. This greater number of participants limits the price impact of trades and thus increases the stability in the market. The obvious policy implication is then exactly the opposite of the one before. Authorities should enhance liquidity as much as possible and any “sand in the wheels” of the financial system, will impede this objective. This bright view is rooted in the market microstructure literature. This literature investigates the forces affecting trades, quotes and prices. These forces can be economic motives of agents, but also the organization of the markets in which trading takes place. It demonstrates clearly the beneficial effects of additional liquidity for a number of parties and provides ample support
for the bright view. First, as mentioned, traders benefit from more liquidity since it allows them to buy and sell assets at a lower cost. In addition, it affects their portfolio strategy. Vayanos (2004) shows in a model that periods of high volatility are associated with a flight to liquidity in the sense that the risk premium investors require per unit of volatility increases. Moreover, illiquid assets become riskier since their market betas increase. The fact that liquidity is a risk factor that is priced in the market is confirmed in recent empirical research by Pastor and Stambaugh (2003), among others. This means that liquidity can be viewed as risk-reducing, and investors will therefore be more willing to hold assets that have greater liquidity. An extensive overview of the literature on liquidity, liquidity risk and asset pricing can be found in Amihud, Mendelson and Pedersen (2005). Secondly, liquidity is important for stock exchanges. It is an argument in attracting firms to (cross-)list. Pagano, Randl, Roell and Zechner (2001) show that firms are more likely to cross-list in liquid markets. Moreover, it is a key variable in the competition with other exchanges for order flow as shown in Parlour and Seppi (2003). Finally, given that liquidity is a determinant of asset returns, it influences also decisions of firms on the optimal capital structure. Whether a firm finances investments by means of issuing shares, bonds or via debt or internal finance, is then likely to depend, among other things, on liquidity in stock and bond markets. Ellul and Pagano (2005) demonstrate that liquidity risk in the secondary market is a determinant of the underpricing of initial public offerings (IPO). Lipson and Mortal (2006) show that firms that have more liquid equity, tend to have lower leverage and are more likely to choose equity over debt when they need new capital.

While this bright view highlights a number of benefits of liquidity, it does not completely rule out the occurrence of global financial instability, which can still arise if the liquidity in different assets moves together. This so-called commonality in liquidity is documented in, among others, Chordia, Roll and Subrahmanyam (2000) and Hasbrouck and Seppi (2001) for the NYSE. Brockman and Chung (2002) find evidence for commonality in an order driven market (the Hong Kong Stock Exchange). However, as O’Hara (2004) points out, in general this does not need to lead to financial instability. She gives two arguments. First, if investors drop one asset, but stay in the market in other assets, this “flight to quality” (see also Vayanos (2004) discussed earlier) means instability remains local and not global. Secondly, commonality might
induce investors to enter the market in the first place. Because the number of potential buyers and sellers increases, stability is improved. As an example, O’Hara refers to the use of derivatives that allow investors to hedge credit risk on bonds. By separating the risk-bearing and investing roles, these linked markets provide a mechanism to provide liquidity to the market as a whole. This possibility is not considered by Keynes and other adherents of the dark view above.

V. THE MARKET MICROSTRUCTURE APPROACH TO LIQUIDITY

As argued, the discussion of liquidity in this paper is grounded in the market microstructure literature on financial markets. The terminology “market microstructure” stems from the title of an article by Garman (1976). He investigates market making and inventory costs. Since then, it has been a collective term for the (financial) literature describing economic forces affecting trades, quotes and prices, i.e. the process through which investor demands are translated into transactions. A summary of the complete literature however, is outside the scope of this paper. Instead, I focus on one particular topic: I analyze the insights that the literature on market microstructure offers on liquidity in stock markets. As argued in Section II, a wide variety of trading mechanisms for stocks exist. A main distinction is the one between quote driven (dealer) and order driven (limit order) markets. I therefore start this section with a discussion of different aspects of liquidity in both market types. In a third subsection, I compare dealer and limit order markets with respect to liquidity. Next to the main distinction between quote and order driven, markets also differ in a number of other aspects, e.g. transparency, anonymity and tick size among others. The impact of these market characteristics on liquidity is investigated in a fourth subsection.

A. Liquidity in Dealer Markets

Recall that Harris (1990) defines four dimensions of liquidity. The first one, width, concerned the bid-ask spread for a given number of shares and commissions and fees to be paid per share. In a first subsection, I analyze the determinants of the spread in a dealer market. Secondly, I turn to the issue of depth. In a third subsection I turn to
another dimension: immediacy and investigate whether a trade-off exist between execution cost and execution time. A fourth subsection deals with resiliency.

1. Components of the Bid-Ask Spread

In a quote driven market, the bid-ask spread is obviously determined by the bid and ask prices that are set by the dealer. In the literature, three broad categories of theoretical models have been proposed to explain dealers’ quoting behavior:

1) order handling costs,
2) inventory models and
3) asymmetric information models.

Order handling costs comprise costs of maintaining a continuous market, as well as costs of matching and clearing orders. Roll (1984) develops, under certain assumptions8, a measure for the order handling cost component of the bid-ask spread, based only on data of transaction prices:

\[
\text{Spread} = 2\sqrt{-\text{cov}(P_{t+1} - P_t, P_t - P_{t-1})}
\]

where \(P_t\) denotes the transaction price at time \(t\) (i.e. the price at which the shares are bought or sold). This equation says that one can derive an estimate for the bid-ask spread, just by looking at the covariance between subsequent changes in transaction prices. Extensions of this measure can be found in Choi, Salandro and Shastri (1988), who allow for correlation in the direction of trades, and George, Kaul and Nimalendran (1991), who correct for time varying expected returns. However, neither measure takes into account inventory nor asymmetric information.

A second class of models analyzes the impact of inventory. Since order flow is uncertain for market makers, they have to deal with this uncertainty when managing their inventory while setting their prices. A seminal model is Garman (1976). In his model, market makers set lower prices to sell and higher prices at which they want to buy to avoid certain failure (similar to the famous “gambler’s ruin” problem). Amihud and Mendelson (1980) explicitly incorporate inventory in the pricing decision of a market maker. The dealer’s bid and ask prices (his decision variables) depend on inventory (the state variable). The
optimal bid and ask prices are a monotone decreasing function of the inventory position, and exhibit a positive spread. In the model of Stoll (1978), who focuses on portfolio risk, risk averse dealers have a desired portfolio composition. They allow actual positions to deviate from this to accommodate trading, but want to be compensated for this risk. This compensation is obtained by a (strictly) positive bid-ask spread. Ho and Stoll (1981) extend this model to a multivariate framework where order flow and portfolio returns are stochastic.

The origin of information models is usually attributed to Bagehot (1971). He makes a distinction between liquidity traders or uninformed traders and informed traders, possessing private information. Liquidity traders (also called noise traders) do not have private information. They trade for reasons exogenous to the model, e.g. portfolio rebalancing motives or simply because they believe to have information. Informed traders have private information, e.g. about the value of the asset, and want to use this information when trading. The spread reflects a balancing by the market maker between the losses of trading with informed traders and the gains of trading with uninformed traders. Models building upon this idea are Glosten and Milgrom (1985) and Kyle (1985). Glosten and Milgrom (1985) show that a bid-ask spread can be induced by asymmetric information. It may exist even if market makers have no costs, are risk neutral and face competition. Secondly, they find that transaction prices form a martingale. In the model of Kyle (1985), there is one informed trader, next to liquidity traders. Traders place their orders simultaneously and the market maker observes net order flow and clears all trades at one price. So in contrast with Glosten and Milgrom (1985), there is no bid-ask spread in his model. He shows that, in a multi-period framework, information is gradually incorporated in transaction prices. Prices still follow a martingale in the sense that an uninformed trader’s expectation of the price tomorrow is today’s price. Holden and Subrahmanyam (1992) extend this model introducing multiple informed traders. They find more trading and a more rapid revelation of information, compared to Kyle (1985), due to competition between informed traders.

Easley and O’Hara (1987) discuss the learning of a market maker. Informed traders will in general trade at one side of the market. The direction of trades (buy or sell) and trade size then provide a signal to market makers, who update their beliefs after each trade. In particular, two equilibria exist. In a separating equilibrium informed traders
only use large transaction sizes (uninformed use both small and large trade sizes). In this case, the market maker sets no spread for small trades, but does set one for large trades. In a pooling equilibrium, informed traders use both small and large trade sizes. Consequently, market makers set a spread for both small and large trades, although the large-trade spread is larger.

Admati and Pfleiderer (1988) investigate the strategic interaction between informed and uninformed traders. If uninformed traders have some discretion about when to trade, they show that trading will be concentrated in certain periods of the trading day. This may provide an explanation for the more intense trading at the opening and closing time.

One of the first empirical tests of microstructure models is presented by Glosten and Harris (1988). They decompose the bid-ask spread into two parts: one part being due to asymmetric information, while the remainder can be attributed to inventory costs, market makers’ risk aversion, ... . They find that the adverse selection component of the spread is economically not significant for small trades, but increases with trade size. Huang and Stoll (1997) propose a three-way decomposition of the spread and find a small, but significant asymmetric information component. The same holds for the inventory component. The order processing component is larger. The spread components also differ with trade size. The asymmetric information component is smaller for large trades than for medium and small trades. One reasoning for this result might be that large trades are prenegotiated such that the trade price reflects the information, conveyed by the trade.

2. Prices and Depth

So far, I focused on how the dealer sets bid and ask prices. In practice however, dealer submit price schedules, i.e. combinations of bid and ask prices and associated bid and ask depth at these prices. Kavajecz and Odders-White (2001) investigate how these price schedules are selected. Interestingly, they find that dealers revise their prices and depth in response to different events. Depths are revised in response to transaction of any size, while prices are revised only when transaction size exceeds quoted depth. Kavajecz and Odders-White (2001) also find that dealers respond strongly to changes in the limit order book9. In fact, economically, this turns out to be the most important factor for revising price schedules.
3. Immediacy

Trading costs and prices obtained can be considered as one dimension of the execution quality of orders. Another dimension is the speed of execution. One reason why traders prefer faster execution is that slower execution increases uncertainty over the execution price. Boehmer, Jennings and Wei (2006) show that markets tend to receive more order flow if either execution costs decline or the execution speed increases. Boehmer (2005) finds a negative trade-off between execution speed and execution cost. He finds that the difference between execution cost on NASDAQ and NYSE are inversely related to execution speed, i.e. execution is more costly on NASDAQ but also faster. Moreover, the difference in cost decreases monotonically with order size, while the difference in speed increases monotonically with order size.

4. Resiliency

Resiliency in a dealer market, i.e. how fast prices revert to former levels after they changed in response to large order flow imbalances initiated by uninformed traders, is analyzed theoretically in Cordella and Foucault (1999). In their model, dealers are uncertain about when the next order will arrive in the market. Assume that at some point in time, the dealer can revise (or set) his quotes. If he observes that the current best quote in the market, set by another dealer, is above the competitive price of the asset, the dealer is faced with a trade-off between two possible choices. First, he can opt for quoting the competitive price. He then executes the next order with certainty and has a profit equal to the tick size. Secondly, he may decide to undercut the current quote by only one tick. In case of executing the next order, the dealer has a larger profit, but on the other hand, he runs the risk of being undercut subsequently by another dealer. As a consequence of these two possibilities, execution prices in the market may deviate from the competitive price. The trading costs faced by traders thus depend on the speed of convergence of the best quotes to the competitive price. This speed is governed by two main determinants. First, there is a negative relation between the time for the best price to adjust to the competitive price and the tick size. The intuition is that a larger tick size creates a bigger difference between the expected asset value and the competitive price. If the dealer then
quotes this competitive price, he obtains a larger profit when the tick size is larger. This faster convergence implies also that a larger tick size does not necessarily result in larger trading costs for traders. Secondly, a negative relation exists between expected trading costs and the frequency of quote revisions by dealers. If dealers frequently revise quotes, the risk is larger that a dealer who does not post the competitive price is subsequently undercut. The frequency of quote revisions in turn depends on the market monitoring costs, faced by the dealer.

Empirically, Bhattacharya and Spiegel (1998) investigate NYSE trading suspensions. They define resiliency as the ability to absorb very large shocks. A cross-sectional analysis of all trading suspensions during the period 1974-1988 shows that the various dimensions of liquidity are substitutes: large-cap stocks have lower bid-ask spreads at the open but halt more often. In particular, the bid-ask spread of stocks in the largest decile have a spread at the open that is one-seventh of the spread of stocks in the smallest decile. However, the probability of a trading halt in the largest stocks is three times the one of the smallest stocks. They also find that the NYSE has become more resilient over their sample period, i.e. the time needed to absorb unusually large shocks has decreased.

B. Liquidity in Limit Order Markets

1. Introduction

In an order driven or limit order market, no dealers or market makers are present which have an obligation to supply liquidity. Liquidity is only provided by traders and their unexecuted limit orders. Hence, if the supply of limit orders dries up, trading breaks down. Therefore, it is important to study whether such market mechanism is able to provide liquidity in all circumstances and especially around shocks. In a first subsection, I discuss the components of the bid-ask spread, similar to the ones analyzed above for a dealer market. Secondly, I turn to the determinants of liquidity supply and demand. Ultimately, in a limit order market, all dimensions of liquidity are determined by the interaction between market orders, demanding liquidity, and limit orders, supplying liquidity. The factors underlying this choice are then clearly crucial for liquidity in an order driven market. Finally, I analyze the issue of resiliency.
2. Components of the Bid-Ask Spread

Recall that in a dealer market, three reasons were given for the existence of the spread: inventory costs, order handling and processing costs and asymmetric information. In an order driven market, traders providing liquidity can still be expected to require compensation for order handling costs. On the other hand, as no trader has an obligation to make the market and take the opposite side of a trade, inventory is less likely to be important. From a theoretical point of view, Glosten (1994) shows that the limit-order market will have a positive bid-ask spread arising from the possibility of trading on private information. Handa, Schwartz and Tiwari (2003) develop a model of an order driven market where investors differ in their share valuations and the arrival of news in the market is not common knowledge. They show that the size of the spread is a function of the differences in valuation among investors and of adverse selection.

Empirically de Jong Nijman and Roell (1995) find that on the Paris Bourse, order processing cost are an important determinant of spreads. de Jong Nijman and Roell (1996) show that the price impact of trades increases with trade size. Moreover, inventory control is unimportant. Ahn, Cai, Hamao and Ho (2002) find that both the adverse selection and order-processing cost components exhibit U-shaped intraday patterns in Tokyo. This contrasts with the decline over the day of the adverse selection component and the increase of the order-process component during the day on the NYSE (a hybrid market, see Madhavan, Richardson and Roomans (1997)). They also find that adverse selection costs increase with trade size while order-processing costs decrease with it. This is also in contrast with NYSE where medium trades contain more information than large trades (see e.g. Huang and Stoll (1997)). These studies imply that the process of how information is incorporated into stock prices through trading on a limit order market is different from a quote driven or hybrid system.

3. Liquidity Supply and Demand in Order Driven Markets

In a limit order market, all dimensions of liquidity — the bid-ask spread, depth, immediacy and resiliency — are ultimately determined by the interaction between market orders, demanding liquidity, and limit orders, supplying liquidity. Therefore, the choice by traders between market and limit order is crucial in this respect. Several dynamic
models\textsuperscript{11} analyze this choice, each model focusing on different determinants. These include depth (Parlour (1998)), volatility of the asset (Foucault (1999)), the order arrival rate and the composition of traders (Foucault, Kadan and Kandel (2005) and Rosu (2006)). I now discuss each of these determinants more in detail.

Parlour (1998) presents a one-tick dynamic model of a limit order market where agents choose between market and limit orders. The execution probability of a limit order is endogenous. She shows that traders look at both sides of the market (depth at bid and ask side) when deciding upon their order type, not only to their own side. In equilibrium, this generates systematic patterns in transaction prices and order placement strategies even with no asymmetric information.

Foucault (1999) analyzes the impact of the risk of being picked off\textsuperscript{12} and execution risk on traders’ order placement strategies and trading costs. A main determinant of the mix between market and limit orders is the volatility of the asset. The higher volatility, the larger the proportion of limit orders in order flow. The reasoning is that when the asset’s volatility increases, limit order traders face a higher risk of being picked off. Therefore, their reservation spreads enlarge. This increases the cost of market order trading, making it more likely that limit orders are the optimal trading strategy. The fill rate (the ratio of filled limit orders to total number of limit orders) is negatively related to asset volatility.

Foucault, Kadan and Kandel (2005) focus on yet another determinant of a trader’s choice between market and limit orders. They develop a dynamic model of a limit order market with strategic liquidity traders having different degrees of impatience. In equilibrium, patient traders tend to submit limit orders, whereas impatient traders submit market orders. Limit order book dynamics in equilibrium are mainly determined by the proportion of patient traders and the order arrival rate. They show that the resiliency of the limit order book increases in the proportion of patient traders and the waiting cost, while it decreases in the order arrival rate.

Rosu (2006) develops a continuous-time model of price formation in a limit order market. The determinants of liquidity in his model are, as in Foucault, Kadan and Kandel (2005), the arrival rate of agents in the market, and the ratio of patient to impatient traders. These variables then also determine the bid-ask spread and the price impact function. He finds that in equilibrium, impatient agents always submit market orders, while patient agents submit limit orders except for the states
where the limit order book is “full”. In those states a patient trader either places a market order, or submits a quick (fleeting) limit order. This order is then immediately accepted by a trader from the other side of the book. In states where the book is not full, new limit orders are always placed inside the bid-ask spread. The point where the limit order book is full coincides with the time when the bid-ask spread is at its minimum. This implies that there exists an optimal minimum spread, despite the fact that tick size is zero in his model. Moreover, he shows that if multi-unit market orders arrive with probabilities which do not decrease too fast with order size, then the book exhibits a hump shape, i.e. limit orders will cluster away from the bid and the ask. Furthermore, after a market sell order both the bid and ask prices decrease, with the bid decreasing more than the ask. As a result, the spread itself widens. Biais, Hillion and Spatt (1995) explain this phenomenon by asymmetric information. Rosu shows, however, that the decrease in the ask need not to reflect an information effect. It can also result from an adjustment by limit order sellers. After they have observed a decrease in the bid, they realize that the time to execution of their orders might increase. For this reason, they lower their ask. In empirical work however, it may be difficult to disentangle waiting costs and asymmetric information.

Next to theory, a wide body of empirical literature has also emerged. Biais, Hillion and Spatt (1995) present an elaborate analysis of the limit order book and order flow on the Paris Bourse (nowadays Euronext Paris). They find that the conditional probability on a limit order is larger when the spread is large or the order book is thin (i.e. depth is low). This means that liquidity is supplied when it is valuable and consumed when it is abundant. Moreover, after a market order, the probability is relatively high that the next order will provide liquidity, pointing to an interaction between buy and sell sides of the market. The market responds quickly to market orders, suggesting competition in liquidity supply. These results also suggest the presence of traders monitoring the book and waiting for profitable trading opportunities. Order flow placement is concentrated inside or at the best prices in the book. A large fraction improves the best bid or ask. This is especially the case when depth at the best quotes is large. But despite the concentration of order submissions at the best prices, depth in the book concentrates behind them, a consequence of the interaction with trading at the best prices. Finally, they also find evidence for information effects. After large sells, decreasing the best bid, often a
new sell order is placed, decreasing the best ask and reflecting the adjustment in expectations based on the information content of a trade. Moreover, large sells tend to occur in succession quickly, consistent with insider trading in Easley and O’Hara ((1987), (1992)).

There exists a large number of other empirical studies, investigating other order driven markets. Ahn, Bae and Chan (2001) find for the Stock Exchange of Hong Kong that market depth rises after an increase in transient volatility and that volatility falls following a rise in depth. Moreover, transient volatility affects the mix between market and limit orders. If it increases at the ask (bid) side, more limit sell (buy) orders are submitted, relative to market orders. As in Biais, Hillion and Spatt (1995) this means that liquidity is provided when needed. Lehmann and Modest (1994) and Hamao and Hasbrouck (1995) investigate in detail the Tokyo Stock Exchange. They conclude that, despite the absence of market makers, sufficient liquidity is provided by the limit order book. Hollifield, Miller and Sandas (2004) study liquidity supply and demand in limit order markets, focusing on the economic determinants of the traders’ optimal order submissions. The type and timing of order submissions depend on the one hand on the trade-off between price improvement, execution probability and picking off risk and on the other hand on the trader arrival rates and trader heterogeneity. The expected payoffs depend on conditioning information such as information about the limit order book and past order submission activity.

Ranaldo (2004) investigates the choice of an order type in the Swiss Stock Exchange (a pure limit order market). He analyzes how the state of the limit order book affects a trader’s order submission strategy. He finds that patient traders become more aggressive when the own (opposite) side book is thicker (thinner), the spread wider, and the temporary volatility increases. These results show that both sides of the book are important for a trader when she determines her strategy, as predicted by Parlour (1998). Finally, he finds systematic differences between the buy and sell side of the order book. Prior to the submission of any type of sell orders, the book always shows a larger spread and a thicker sell side. Also, buy orders are more autocorrelated than are sell orders.

4. Resiliency

Theoretically, resiliency is studied in Foucault, Kadan and Kandel (2005). As mentioned above, they show that the resiliency of the limit order book increases in the proportion of patient traders and the
waiting cost, while it decreases in the order arrival rate. Furthermore, there is a positive relation between the duration until a transaction, conditional on the quoted spread for the prior transaction, and market resiliency. Finally, other things being equal, the resiliency of the limit order market is always larger when there is a minimum price variation than in the absence of a minimum price variation.

Empirically, Degryse, de Jong, van Ravenswaaij and Wuyts (2005) analyze the resiliency of a pure limit order market by investigating the limit order book (bid and ask prices, spreads, depth and duration), order flow and transaction prices in a window of best limit updates and transactions around aggressive orders. Such orders are defined as large orders that move prices. Aggressive orders take place when spreads and depths are relatively low, and they induce bid and ask prices to be persistently different after the shock. Depth and spread remain also higher than just before the order, but do return to their initial level within 20 best limit updates after the shock. Relative to the sample average, depths stay around their mean before and after aggressive orders, whereas spreads return to their mean after about twenty best limit updates. The initial price impact of the aggressive order is partly reversed in the subsequent transactions. However, the aggressive order produces a long-term effect as prices show a tendency to return slowly to the price of the aggressive order.

Wuyts (2007) develops an econometric framework to study various dimensions of liquidity (prices, depth and duration) and for capturing the interactions between them. In addition, he investigates resiliency, i.e. how fast best prices, depths and duration recover to their initial, pre-shock level after the market has been hit by a liquidity shock. The results clearly demonstrate the importance of incorporating different dimensions of liquidity in the analysis. In case of a negative liquidity shock (a shock increasing the spread), he finds a permanent effect on prices, with returns (in absolute value) ranging from 0.06 to 0.16%, depending on size and tick size of the stock. Also, he finds an initial widening of the spread, but it becomes smaller again in subsequent periods. On the other hand, depth at the best prices increases, initially with up to 20%. A second main conclusion is that an analysis of liquidity should also allow for asymmetries in dynamics at bid and ask side of the market, while at the same time accounting for the existence of a relationship between them.

Wuyts (2006) extends the analysis to liquidity behind the best limits in the order book. When subsequent prices are close to the best
ones and depth at them is substantial, larger orders can be executed without an extensive price impact and without deterring liquidity. The results show a somewhat less favorable image of liquidity than often found in the literature. After a liquidity shock (in the spread or depth or in the book beyond the best limits), several dimension of liquidity deteriorate at the same time. Not only does the inside spread increase, and depth at the best prices decrease, also the difference between subsequent bid and ask prices may become larger and depth provided at them decreases. The impacts are often both econometrically and economically significant. Also, his findings point to an interaction between different measures of liquidity, between liquidity at the best prices and beyond in the book, and between ask and bid side of the market.

Coppejans, Domowitz and Madhavan (2004) study the resiliency of the Swedish stock index futures market (OMX). They find that aggregate market liquidity exhibits considerable variation. Strategic traders time their trades accordingly, reinforcing the concentration of volume and liquidity in time. Shocks to volatility reduce liquidity and impair price efficiency, suggesting that automated auctions might be vulnerable to periodic liquidity crises following sharp market movements. However, a time-series analysis shows that these effects dissipate quickly, e.g. they find that shocks to depth are restored in less than 60 minutes. This indicates that the market is resilient.

C. A Comparison of Quote and Order Driven Markets

Although I already pointed to differences between order and quote driven markets, the results so far where not based on a direct comparison of both markets. In this subsection, such direct comparison is made. Biais, Foucault and Salanie (1998) show theoretically that a dealer market has a larger spread and inefficient risk sharing, while a limit order market results in the competitive outcome. This finding is supported by Theissen (2000) who finds in experiments that continuous auctions have more efficient transaction prices than a dealer market. Moreover, the latter faces higher trading costs. The results in both papers are also consistent with empirical studies comparing both trading systems. Christie and Schultz (1994) find evidence for collusive pricing by dealers in NASDAQ but not in NYSE\textsuperscript{14}, and Huang and Stoll (1996) report larger spreads for NASDAQ stocks, compared to a matched sample of NYSE stocks. Degryse (1999) compares an order
driven and a dealer market (Brussels and London respectively) for a sample of cross-listed Belgian stocks. He finds that the order driven market outperforms the dealer market with respect to trading costs. More specifically, quoted and effective bid-ask spreads are smaller in Brussels. Although the market in Brussels is thus tighter, London, however, provides a deeper market. Finally, he shows that total trading costs are lower in Brussels for small trade sizes, while the inverse holds for large trades.

D. Market Design and Liquidity

Next to the distinction between quote and order driven, markets can differ in a number of design features. I now review the influence of a number of these characteristics on liquidity. More in particular, I discuss transparency, anonymity, tick size and the presence of a trading floor.

1. Transparency

Transparency can be defined as the ability of market participants to observe information about the trading process. Information here refers to knowledge on prices, order flow, .... It can be divided in pre- and post-trade transparency. The former refers to the dissemination of bid and ask prices, depths, .... Post-trade transparency denotes the publication of trades and their details (e.g. prices, size). In recent years, there has been a global trend in financial markets towards more pre- and post-trade transparency. Different degrees of both will obviously affect order submission strategies of traders and hence liquidity. More transparency is on the one hand associated with more informative prices, but on the other hand can also hamper liquidity because traders might be unwilling to reveal their trading intentions. Moreover, it is unlikely to have the same effect on all market participants. For instance, informed traders prefer less transparency, while liquidity traders prefer a larger disclosure. In a study on limit order markets, Pagano and Roell (1996) find that in general, more pre-trade transparency increases liquidity. The implicit bid-ask spread is narrower since price setters, who know more, can protect themselves better against losses to insiders. Moreover, they find that an auction market offers lower trading costs than a dealer market. This is consistent with the fact that the limit order book (which traders can scan) offers more
transparency than a dealer (where only the quotes for a certain size are given, making it difficult to infer the trade price). Rindi (2004) also finds that for a given proportion of informed traders, liquidity improves if markets become more transparent. However, if information acquisition becomes endogenous, i.e. when agents can choose to become informed at a cost, the effect may be reversed. The intuition is that uninformed agents might be reluctant to supply liquidity for large orders, as these might be information driven. Informed traders do not face this problem. When (pre-trade) transparency is higher however, agents have less incentives to become informed. In this way, the proportion of informed traders decreases, as well as the liquidity provided by them.

Another aspect of pre-trade transparency is investigated by Boehmer, Saar and Yu (2005). They study the consequences of making the NYSE limit order book public16, thereby increasing pre-trade transparency. They find that traders changed their strategies in response to this event. More specifically, they submit smaller limit orders and cancel limit orders in the book more quickly and more often. Moreover, trading shifts away from floor brokers towards electronic trading. Also NYSE specialists change their quoting strategies and add less depth to the quote. The introduction of a public limit order book also leads to an improvement in informational efficiency, an increase in displayed liquidity in the book, and a decline in the price impact of trades and marketable orders. However, the public limit order book is not beneficial for all market participants since their results demonstrate that welfare is redistributed. Because of the smaller price impact of trades and marketable orders, compensation for liquidity provision is reduced. This benefits market order traders (liquidity demanders), but hurts submitters of limit order and specialists (liquidity suppliers). Also, floor brokers face declining revenues since they intervene in less trades. On the other hand, OpenBook generates additional revenues for NYSE.

While indeed there exists a clear tendency towards more transparency, there is also a trend to provide market participants with a possibility to limit their exposure. The use of hidden orders (also called iceberg orders), where only part of an order is displayed to the market and the rest remains hidden, has become widespread. De Winne and D’hondt (2004) investigate how the presence of hidden depth in the limit order book affects both implicit transaction costs and traders’ behavior on Euronext. They show that ignoring hidden quantities in the order book substantially underestimates actual liq-

300
liquidity on Euronext and overestimates transaction costs incurred by traders. Moreover, the presence of hidden orders at the best opposite quote significantly increases the aggressiveness of submitted orders. This means that traders seize the opportunity to benefit from reduced implicit transaction costs due to some hidden depth at the best opposite quote.

2. Anonymity

Related to transparency is the issue of anonymity, i.e. the degree to which the identity of market participants is revealed. As with transparency, informed traders will prefer anonymous trading, while liquidity traders do not. A distinction can be made between demand and supply side anonymity. Concealing information about the identity of liquidity demanders in general increases the bid-ask spread, since suppliers can less easily make a distinction between informed and uninformed traders, see e.g. Benviste, Marcus and Wilhelm (1992).

Foucault, Moinas and Theissen (2006) analyze the effect of anonymity at the supply side on liquidity, both theoretically and empirically. They show that, in an order driven market, concealing the identity of liquidity suppliers might reduce bid-ask spreads. In their model, some “expert” traders can better assess their exposure to informed trading, and hence of the cost of providing liquidity. Without anonymity, non-expert traders derive information from their quotes. If expert traders bid cautiously, non-expert traders will defer from improving their quotes. However experts also behave strategically and sometimes bluff, i.e. set quotes as if costs of liquidity supply are high. If anonymity is introduced, traders do no longer know who placed the quotes and bluffing becomes less effective. Expert traders will bid more aggressively in this case. It is then clear that spread and depth will be affected. The ultimate effect depends on the proportion of expert traders, but there are cases when the spread is reduced. They test their model by considering the introduction of anonymity on Euronext Paris. They find that when the limit order book became anonymous, average spreads declined significantly, and depth decreased.

3. Tick Size

A number of empirical papers have investigated the impact of tick size changes on market quality. Bacidore (1997), Ahn, Cao and Choe
(1998) and Griffiths et al. (1998) consider the April 1996 reduction in tick size on the TSE, while Goldstein and Kavajecz (2000) and Chordia and Ball (2001) (among others) deal with the changes in tick size and the liquidity provision on the NYSE. In general, these papers show that after the reduction in tick size, the inside spread significantly decreased, but depth at the best bid and ask also decreased. This means that one dimension of liquidity improves (spread) while another (depth) deteriorates. Bourghelle and Declerck (2004) investigate the market quality of the Paris Bourse following the introduction of the Euro. This changed the tick size compared to prices in French Francs. Interestingly, they find that only the depth at the best prices is significantly affected whereas the spreads remain unaltered. Stocks obtaining a decrease (increase) in tick size experience a decrease (increase) in the depth at the best prices.

4. Floor versus Screen-Based Trading Systems

A final property of trading systems is whether or not trading floor is present. While NASDAQ or Euronext operate without a floor, the NYSE maintains the floor for executing trades. Theoretically, Biais, Foucault and Salanie (1998) show that large spreads and inefficient risk sharing occur in floor markets (and also in dealer markets) but not in screen based limit order markets.

Empirically, evidence is mixed. Venkataraman (2001) compares the NYSE (which has a trading floor) with Euronext Paris (fully screen based) for a sample of similar stocks and find that spreads are lower on a floor based exchange than on an electronic exchange. Theissen (2002) provides somewhat more direct evidence by comparing the floor and the screen-based trading system of the Frankfurt Stock Exchange, which operated in parallel. He finds that an electronic (screen-based) trading system offers low spreads for liquid stocks, while the floor is more competitive for less liquid stocks. Jain (2005) investigates 120 stock exchanges world wide and finds that a change from floor to electronic trading has a number of beneficial effects in the long run. He finds that the equity premium is reduced significantly after the switch. In addition, excess-over-the-world abnormal returns (AR) in the announcement month is 8.99%. Cumulative AR two years after the announcement are 28.69%. Moreover, the cost of capital of listed firms also declined and monthly trading turnover increased.
VI. INTERMARKET COMPETITION AND LIQUIDITY

A. Introduction

In Section V, liquidity was related to different elements of market design but the focus remained on one market. Nowadays however, stocks may trade simultaneously on different trading venues. Stocks are not only cross-listed on different exchanges, they also trade in so-called alternative trading systems (ATS) that emerged recently. This evolution implies that traders must decide where to submit an order, taking into account differences in trading costs or execution time, but also in market design across trading venues.

In general, the fact that one asset trades in different places potentially induces two opposite effects on liquidity. On the one hand, competition between trading venues for order flow may force exchanges or ATSs to decrease fees, thus resulting in lower trading costs. On the other hand, fragmentation of the order flow between trading venues may reduce liquidity in each venue. Moreover, in a fragmented market, orders may execute at a price worse than the best quoted due to violations of price priority (so called trade-throughs).

In a theoretical model Parlour and Seppi (2003) show that competition between trading venues, either because new ones open or because of cross-listings, can increase or decrease aggregate liquidity, relative to a single market environment. The actual outcome depends e.g. on the trading costs faced by traders. Biais (1993) finds in a model without asymmetric information, that the spread in fragmented and centralized markets is equal, but it is more volatile in centralized ones. In centralized markets, the agent with the lowest reservation value sets his ask price just below the ask of his next-best competitor. In fragmented markets, he sets his ask price below the expected next best price. His expectations are less volatile than the variable itself. In the model of Pagano (1989), informed traders have an incentive to strategically split up their orders between markets. This however at the same time induces liquidity traders to concentrate trading at one place, since this will result in a drop of the proportion of informed trading, and in narrower spreads.

In an empirical analysis, Hasbrouck (1995) moreover shows that often one market is the source of all price discovery, while others are rather matching the quotes. An extreme case is provided by crossing networks, who do not contribute at all to price discovery but derive the price at which orders are matched from the main markets (see further).
Boehmer, Jennings and Wei (2006) find that routing decisions of orders by traders are related to execution quality: markets reporting low execution costs and fast fills for orders subsequently receive more orders.

B. Dealer versus Limit Order Book

Seppi (1997) presents a theoretical model where a dealer with market power competes with a public limit order book. The dealer must respect the price, public and time priority of other traders, in other words he must undercut other liquidity providers to obtain order flow. The equilibrium liquidity of a market depends on the interaction between the specialist’s strategy, the limit order book, and latent competition from a trading crowd. In some cases, the dealer is simply a trader of last resort on larger trades, while he is the primary source of liquidity for very small orders. Seppi relates these findings to optimal market design. He shows that large institutional traders (submitting large orders) have a larger optimal tick size than small retail investors, but both prefer a tick size strictly greater than zero. Moreover, a hybrid dealer/limit order market (such as the NYSE) provides better liquidity to small traders and large institutional traders, while a pure limit order market (like Euronext) offers better liquidity for mid-size orders.

In an empirical analysis, Kavajecz and Odders-White (2001) find that dealers on the NYSE respond strongly to changes in the limit order book when setting the prices and depths at which they are willing to trade. In fact, economically, this turns out to be the most important factor for revising price schedules.

C. Two Limit Order Markets

Foucault and Menkveld (2006) study competition between two limit order books. More specifically, they analyze the entry of a new market (i.e. EuroSETS from the London Stock Exchange), next to an incumbent market (Euronext) in the Netherlands. They first develop a theoretical model and subsequently perform an empirical analysis. In the model, they allow for differences in order submission fees, and introduce two types of brokers:

(i) smart routers, who route orders across markets to obtain the best execution price and
(ii) non smart routers, who ignore quotes in the entrant market and always trade in the incumbent market.
The latter traders generate so-called trade-troughs\textsuperscript{18}. Their model generates two main empirical predictions. First, other things equal, consolidated depth at a certain price (i.e. the sum of all shares available at that price or better in both markets) should be larger after EuroSETS entry. This result is driven by the absence of time priority across markets. The reasoning is that it allows traders to jump ahead of the queue of limit orders in one market by submitting a limit order in the competing market. Second, the model predicts that an increase in the proportion of smart routers coincides with an increase of liquidity supply by the entrant market EuroSETS. The intuition is that more smart routers increase the execution probability of limit orders submitted to the entrant market.

Their empirical analysis confirms these predictions. More specifically, they find an increase in consolidated depth after the entry of EuroSETS. Secondly, also depth on the incumbent market Euronext increases after the entry. The reasoning is that Euronext reduced its fees around the entry of EuroSETS. The resulting increase in depth more than compensates the loss of order flow to the entrant market. Finally, EuroSETS has lower spreads and a larger share in consolidated depth for stocks with a larger proportion of smart routers.

D. Alternative Trading Systems

Recently, a number of alternative trading systems emerged. Broadly speaking, these can be divided in two main categories: Electronic Communication Networks (ECNs) and Crossing Networks (CNs).

1. ECNs

ECNs are trading systems that essentially operate as fully anonymous, public limit order books. They are mainly successful in attracting order flow from NASDAQ (a dealer market). ECNs obtain 42% of share volume in NASDAQ listed stocks, compared to only 3% for NYSE listed stocks (see Stoll (2006)). One explanation is that the NASDAQ reform of 1997 increased the competition between dealers and ECNs by providing traders with a better access to ECNs. In this way, traders can e.g. save the spread, charged by dealers. Moreover, as NYSE is a hybrid market, there is already “internal” competition between NYSE market makers and the NYSE limit order book.
Weston (2002) investigates the impact of ECNs on NASDAQ dealers. He finds that the growth of ECNs is associated with higher liquidity on NASDAQ as quoted, effective, and relative bid-ask spreads are lower and depths are larger. Moreover, markets are less concentrated. Fink, Fink and Weston (2006) study in addition the competitive impact of ECNs on NASDAQ. Their results show that an increase in ECN trading may have caused some traditional market makers to exit the market for market making. Overall, ECNs seem indeed to provide a source of competition to traditional NASDAQ dealers.

Huang (2002) compares the quality of the quotes in ECNs and those posted by NASDAQ market makers. He shows that ECNs not only post informative quotes, but also, compared to market makers, ECNs post quotes rapidly and are more often at the inside. Additionally, ECN quoted spreads are smaller than dealer quoted spreads. These results may suggest that the benefits from intermarket competition may outweigh possible costs of market fragmentation. Barclay, Hendershott and McCormick (2003) find that ECN trades are smaller than NASDAQ trades. However, ECN trading explains from 60% to 100% more of the efficient stock price variance than NASDAQ trades. In other words ECN trades are much more informative and ECNs attract more informed trades. However, ECN trades have higher expected ex ante trading costs because market makers on NASDAQ are able to differentiate the least informed (and hence most profitable) trades and offer them better execution. These trades have lower effective spreads. However, because trades on ECNs are more informed, they tend to have lower ex-post trading costs as measured by the realized spread.

Conrad, Johnson and Wahal (2003) investigate proprietary data on order submissions to CNs, ECNs and brokers. They find that in general realized execution costs are lower on the ATS. The cost difference between day crosses and broker-filled order is 30 basis points, the difference between ECN and broker trades is 66 basis points. After the reduction in tick size on the main markets, the cost difference for CNs was not affected while the difference for ECNs declined.

2. Crossing Networks

The other form of ATS are CNs. These are defined by the SEC as “systems that allow participants to enter unpriced orders to buy and sell securities, these orders are crossed at a specified time at a price
derived from another market”. Note that CNs are thus periodic trading systems. If executed, orders obtain a better price (midquote), but traders face the risk of non-execution\(^9\). As CNs clearly do not contribute to price discovery, they need an efficient base market. Moreover, compared to other trading systems, CNs are very opaque, since they do not disseminate any information about order imbalances, the identity of traders, .... A detailed discussion of CNs can be found in Degryse, Van Achter and Wuyts (2007).

The interaction between a CN and a dealer market (DM) is modeled in a static model in Hendershott and Mendelson (2000). The dynamics of the interaction between a CN and a dealer market are studied in Degryse, Van Achter and Wuyts (2006). Hendershott and Mendelson (2000) consider a random number of informed and uninformed traders, who simultaneously submit single-unit orders to either the DM or the CN. When orders are not executed on the CN, they can be resubmitted to the dealer. A trader’s choice depends on its own characteristics (e.g. his patience to trade and his valuation of the stock), as well as other parameters (such as execution probabilities on the CN, the spread in the DM, ...). Their model shows that each market caters for a specific type of trader. Moreover, DMs are influenced in two opposite ways by competition from the CN. On the one hand, there is risk sharing as dealers’ inventory and adverse selection costs are lowered by exclusive CN traders, resulting in narrower spreads. On the other hand, opportunistic CN trading (i.e. using the DM as “market of last resort”) may widen the DM-spread. The reasoning is that, in this case, the CN is skimming off part of the uninformed traders. Consequently, this fraction of uninformed traders cannot be “used” anymore by dealers to compensate their losses to informed traders. Within the CN, also two opposite forces are at work. First, a positive liquidity externality exists, as an increase in CN trading volume benefits all CN traders and attracts additional liquidity. Second, when the CN becomes sufficiently liquid, this liquidity externality may be dominated by a negative crowding externality: low-liquidity preference traders compete with the higher-liquidity value traders on the same market side. Combined with the competition effect, the resulting overall impact remains ambiguous. The emergence of the additional CN trading venue benefits some traders, while harming others.

Degryse, Van Achter and Wuyts (2006) investigate the interaction of a CN and a continuous (one-tick) DM. More specifically, they analyze the impact on the composition and dynamics of the order flow on
both systems. They develop the analysis for three possible market designs:

(i) transparency,
(ii) complete opaqueness, and
(iii) partial opaqueness.

The benchmark transparency case reflects that traders are fully informed about past order flow and hence observe the prevailing state of the CN’s order book before determining their strategy. This results in pre- and post-trade transparency. However, in reality CNs are rather opaque. This is incorporated by analyzing two different degrees of opaqueness: partial and complete. While partial opaqueness implies that traders observe previous trades at the DM but not submissions to the CN, complete opaqueness entails that traders are uninformed on both past CN and DM order flow. Their model leads to several results. First, in common to the three informational settings, they find that an increase in the DM’s relative spread augments the CN’s order flow. Therefore, it can be expected that CNs will be more successful in markets where spreads are substantial. At the same time price discovery should be sufficiently informative as the CN “free rides” on information about prices from the DM. Second, a CN and a DM cater to different types of traders. Investors with a high willingness to trade are more likely to opt for immediacy and prefer to trade at a DM. The existence of a CN results in “order creation”: investors with a low willingness to trade submit orders to a CN whereas they would never trade at a DM. Third, the transparency and partial opaqueness settings produce systematic patterns in order flow. In particular, for the transparency case, they find that the probability of observing a CN order at the same side of the market is smaller after such an order than if it was not. Also, the probability of observing a sell at the DM decreases and the probability of a buyer trading on the DM increases when the previous order was a CN buy. Fourth, their results highlight that it is important to take into account the interaction between trading systems when measuring “normal” order flow. For example, when looking at an individual trading system, some order or trade flow sequences could wrongly be interpreted as being driven by information events, whereas they are caused by the interaction of trading systems.

Empirically, Gresse (2006) shows that risk-sharing benefits from CN trading outweigh fragmentation and cream-skimming costs. Moreover, dealers’ spreads are negatively correlated with CN trading.
She also documents that execution probabilities on CNs are rather low (2-4%). Fong, Madhavan and Swan (2004) compare the price impact of block trades across the upstairs market, a CN and the limit order book. The ATS are found to be beneficial to market participants as the upstairs market and the CN offers a lower price impact. Moreover, they do not harm liquidity in the main market.

VII. CONCLUSION

This paper discussed the determinants and implications of a key element of well-functioning financial markets: liquidity. This concept was defined as how easy traders can buy or sell large numbers of shares without large price effects. I argued that liquidity comprises several interacting dimensions: spread, depth, resiliency and immediacy. These dimensions were discussed for a quote driven and an order driven market. Also other design elements of trading system, such as transparency, anonymity and the presence of a trading floor, were investigated. Subsequently, the focus shifted from one market to the impact of intermarket competition on liquidity. In bringing together some main conclusions from the analysis, I focus below on four beneficiaries of liquidity: traders, stock exchanges (or alternative trading systems), listed firms and the stability of the financial system.

First, liquidity is obviously relevant for traders as it directly determines their cost of trading. Order handling costs, inventory and asymmetric information were shown to drive trading costs. Moreover, also market design is relevant. However, results are not always uniform. While transparency in general tends to benefit traders, certain types of traders may prefer less transparent markets, e.g. when they need to carry out a large transaction and do not want to reveal their trading intention. Similarly, lower tick sizes will lower spreads, but a too low tick size may actually harm traders as e.g. limit order traders may not receive appropriate compensation for the risks associated with a limit order. Liquidity also determines the return traders require to invest in a stock. The reasoning is that liquidity and liquidity risk are factors that are priced in the market.

Secondly, liquidity is important for stock exchanges and trading systems in general. In competition with each other and with alternative trading systems, liquidity tends to be an important argument to attract order flow and listings. Moreover, liquidity exhibits an
externality: liquid markets tend to attract even more liquidity. An optimal design of the trading system can help trading venues to attract traders. Moreover, trading venues can tailor their trading rules such that they cater to specific types of traders. The emergence of ATSs is an illustration of this point, e.g. a CN can allow large traders to execute their trades without large price impacts. But also the traditional exchanges differentiate across traders. Both NYSE and NASDAQ acquired an ECN and have set-up (or announced) a CN.

Thirdly, listed firms benefit from more liquidity. Liquidity is beneficial for the cost of capital of a firm through its impact on expected returns required by investors. Obviously, firms then have an incentive to increase the liquidity of their equity such that they can benefit from a lower cost of capital. Although a number of elements that impact liquidity are outside the firm’s control (i.e. those related to the setup of the trading system, e.g. anonymity, transparency, the presence of dealers, ...), the firm is able to influence some factors. First, it can set-up an effective and open communication with the press and analysts. Disclosing more and better information should narrow the gap between informed and uninformed traders. This reduces asymmetric information in the market, leading to a reduction in the spread. Secondly, firms can request a listing or cross-listing on a liquid exchange. Pagano, Randl, Roell and Zechner (2001) indeed find that firms are more likely to cross-list in liquid markets.

Finally, liquidity is also important for the stability of the financial system as a whole. The above discussed “bright view” in liquidity argues that liquid markets invite a larger number of traders to participate in the market. This larger base can then increase the ability of the system to cushion large shocks.

While the survey in this paper provided some evidence on liquidity, there are still ample areas for future research. First, the sometimes conflicting empirical evidence calls for more research on the optimal design of markets. Secondly, traditional exchanges are to an increasing extend confronted with competition from alternative trading systems. An analysis of the (dynamic) impact of these ATSs and of their functioning is still in its infancy. Thirdly, a clear tendency exists towards globalization of trading: recently NYSE and Euronext merged, while plans are announced for an alliance with the Tokyo Stock Exchange. This would create an almost global market on equity trading. The impact of this globalization on liquidity has not yet been extensively studied. Fourth, while the literature has reached agreement
on the fact that liquidity and liquidity risk are priced in the market, little is known about which elements of market design affect this pricing of liquidity (risk). Fifth, to what extent are portfolio decisions of traders influenced by liquidity risk and execution risk? All these questions provide an intriguing and challenging path for future research.

NOTES

1. For a definition of the concept “liquidity”, we refer to Section III.
2. Tick size (or minimum price variation) refers to the minimum amount in which the price of a security can change.
3. A limitation of the current paper is that I restrict the discussion to stock markets. However, a wide number of insights can be translated to other markets as well. Lyons (2001) presents an overview of the market microstructure approach to FOREX markets. An analysis of bond market liquidity can be found in Chordia, Sarkar and Subrahmanyam (2005), while Cheung, de Jong and Rindi (2005) discuss the microstructure of MTS, the trading system for government bonds. Hartmann, Manna and Manzanares (2001) analyze the microstructure of the money market. In these contributions, the insights from the literature on stock markets are applied. Therefore, the methodologies, insights and intuition presented in the current paper, also extend to a range of other financial markets.
4. See e.g. Domowitz and Steil (1999) or Jain (2002) for recent surveys of stock markets.
5. ECNs basically function as an anonymous limit order book. CNs are systems that allow participants to enter unpriced orders to buy and sell securities; orders are crossed at a specified time at a price derived from another market. CNs are thus periodic systems.
6. As we will explain below, in limit order markets, unexecuted limit orders form the supply of liquidity while market orders are demanding liquidity.
7. For excellent surveys, we refer to Madhavan (2000) and Biais, Glosten and Spatt (2005). An overview of the early (mainly theoretical) literature can be found in O’Hara (1995).
8. The main assumptions are:
   1) Buy and sell orders arrive randomly;
   2) Expected returns are constant over time;
   3) The market exhibits strong-form efficiency (i.e. an incoming buy or sell order does not lead to a revision of the value of the asset, as is the case e.g. in models of asymmetric information, see further).
9. Recall that on the NYSE, dealers compete with a limit order book.
10. To be precise: the current best quote is above the expected value of the asset rounded to the nearest tick.
11. We focus on dynamic models. For a static model of limit order markets, see e.g. Glosten (1994).
12. Picking off risk refers to the possibility that, after a limit order has been submitted, new public information might arrive. This may create a winner’s curse problem for limit order traders since their orders are more likely to be executed (i.e. picked off) at a loss when their orders become mispriced.
13. A best limit update is recorded when either the best ask or bid price, and/or depth at these prices change.
14. Note that in part because of this paper, the SEC imposed reforms on Nasdaq in 1997, which alleviated the issue, see Barclay et al. (1999).
15. A notable exception is in-house matching or internalization.
16. This is the so-called OpenBook service, introduced on January 24, 2002. It shows the aggregate limit order volume available in the book at each price. The information about depth is updated every ten seconds. It reflects only the depth in the limit order book, and e.g. not those offered by specialists. Note that this information is not for free since subscribers pay a fee for the service.
17. An agent can be a dealer or a limit order trader.
18. These are trades where there is a violation of price priority. In other words, such trades are not executed at the best possible price. The reasoning in Menkveld and Foucault (2006) that some traders only consider one market, and not the entrant market (which potentially may offer a better price).
19. If at the time of the cross, there are more buy than sell orders, part of the buy orders will not execute. The exact choice of which buy orders are then executed depends on the matching algorithm implemented by a CN.
20. Cream-skimming refers to the fact that some uninformed traders prefer the CN and cannot be used anymore by the dealer to compensate his losses on informed trades.

REFERENCES


