DISCUSSION PAPER SERIES

No. 3288

THE MICROSTRUCTURE OF STOCK MARKETS

Bruno Biais, Larry Glosten
and Chester S Spatt

FINANCIAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP3288.asp

ISSN 0265-8003

THE MICROSTRUCTURE OF STOCK MARKETS

Bruno Biais, Université des Sciences Sociales de Toulouse and CEPR
Larry Glosten, Columbia University
Chester S Spatt, Carnegie Mellon University

Discussion Paper No. 3288
March 2002

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 3288

March 2002

ABSTRACT

The Microstructure of Stock Markets*

We survey the literature analysing the price formation and trading process, and the consequences of market organization for price discovery and welfare. We develop a united perspective on theoretical, empirical and experimental approaches. We discuss the evidence on transaction costs and the price impact of trades and its analyses in terms of adverse selection, inventory costs and market power. We review the extent to which the associated frictions can be mitigated by such features of market design as the degree of transparency, the use of call auctions, the discreteness of the pricing grid and the regulation of competition between liquidity suppliers or exchanges.

JEL Classification: D82, G10, G14 and G20
Keywords: bid-ask spread, liquidity, market design, market microstructure and transactions costs

Bruno Biais
Directeur de Recherche
Université des Sciences Sociales
Place Anatole-France
31000 Toulouse Cédex
FRANCE
Tel: (33 5) 6112 8598
Fax: (33 5) 6112 8637
Email: biais@cict.fr

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=111911

Larry Glosten
Columbia University
418A Uris Hall
Columbia Business School
New York
New York 10027
USA
Tel: (1 212) 854 2476
Fax: (1 212) 662 8474
Email: lrg2@columbia.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=116442

Chester S Spatt
Graduate School of Industrial
Administration
Carnegie Mellon University
Posner Hall Room 253A
Pittsburgh 15213
USA
Tel: (1 412) 268 2234
Fax: (1 412) 268 8896
Email: cspatt@andrew.cmu.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=118944

* We are thankful for helpful guidance and comments from the editors (Ravi Jagannathan and Maureen O'Hara) and the referee, as well as Catherine Casamatta, Thierry Foucault, and Christine Parlour.

Submitted 11 March 2002

The Microstructure of Stock Markets

Mark Garman (1976) quite aptly coined the phrase "market microstructure" as the title of an article about market making and inventory costs. The phrase became a descriptive title for the investigation of the economic forces affecting trades, quotes and prices. Our review will cover not only what research has had to say about the nature of transaction prices, but also the broader literature on the interrelation between institutional structure, strategic behavior, prices and welfare.

In perfect markets, Walrasian equilibrium prices reflect the demand curves of all potential investors. While the determination of these fundamental equilibrium valuations is the focus of (most of) asset pricing, market microstructure studies how, in the short term, transaction prices converge to (or deviate from) long–term equilibrium values. Walras himself was concerned about the convergence to equilibrium prices, through a tâtonnement process. One of the first descriptions of the microstructure of a financial market can be found in the Elements d'Economie Politique Pure (1874), where he describes the workings of the Paris Bourse. Walras's field observations contributed to the genesis of his formalization of how supply and demand are expressed and markets clear.[2]

Short–term deviations between transaction prices and long–term fundamental values arise because of frictions reflecting order–handling costs, as well as asymmetric information or strategic behavior. Another potential source of market power stems from the delegation of trade execution to financial intermediaries. Delegation arises because most potential investors cannot possibly spend their time monitoring the market and placing and revising supply and demand curves for financial assets. Only a small subset of all economic agents become full–time traders, and stand ready to accommodate the trading needs of the rest of the population. This raises the possibility that these key liquidity suppliers could behave strategically. The organization of financial markets defines the rules of the game played by investors and liquidity suppliers. These rules affect the way in which prices are formed and trades determined, as well as the scope for asymmetric information or strategic behavior, and thus the frictions and transactions costs arising in the trading process.

The importance of market microstructure is illustrated both by the resources devoted to the trading process

[2] Walker (1987) offers a historical perspective on this aspect of the genesis of general equilibrium theory.

and by the magnitude of transactions costs incurred by investors. While the cost of transacting could seem small, the volume of transactions makes the overall economic effect non-trivial. For example, in 1999, roughly 200 billion shares traded on the NYSE alone. A transaction cost charge of only two and one–half cents implies a flow of 5 billion dollars. Indeed, the volume of activity is very sensitive to the level of transactions costs, as illustrated by the dramatic increase in turnover during the last 25 years. While this increase is partly due to phenomena which are somewhat outside the scope of market microstructure, such as the development of derivative trading, it also reflects the decline in trading costs that have resulted from the deregulation of commissions, improvements in trading technology, and the increase in the competitiveness and openness of exchanges.

Market microstructure analysis has influenced the regulation and development of financial markets. The discussions of a number of security market issues have been markedly informed by the microstructure literature. Much of the discussion of SOES ("small order execution system") on NASDAQ and the so-called "SOES bandits" revolved around the informational issues raised by this literature. The NASDAQ collusion case resulted from an empirical market microstructure study and its resolution involved very substantial changes in the structure of the market. This outcome resulted from a number of microstructure analyses performed on behalf of both sides of the debate. The effects of decimalization, payment for order flow and transparency (price and quote reporting) are additional examples of issues engaging both regulators and microstructure researchers.

To survey the insights offered by the microstructure literature we develop a synthetic framework offering a unifying perspective on the theoretical analyses. We also confront the theoretical and empirical studies. We discuss which theoretical predictions have been tested, and to what extent they have been rejected or found consistent with the data. We also rely on the theoretical analyses to offer an interpretation for empirical findings. We thus show how the market microstructure literature, building upon first economic principles such as adverse selection and strategic behavior, provides a tool to analyze traders behavior and market design, and offers a rationale for a large array of stylized facts and empirical findings. Our endeavor to integrate the theoretical and empirical sides of the literature differs from O'Hara (1997), whose book provides a rich survey of the theoretical literature. Madhavan (2000) provides an interesting unifying perspective on the microstructure literature, building on an empirical specification in the line of Hasbrouck (1988). While we also discuss this important line of research, our emphasis on theoretical microeconomic foundations and on the strategic behavior of market participants differs from Madhavan (2000). We concentrate on the portion of the literature that addresses price formation and market design, while not addressing other important issues such as the interactions between market microstructure and corporate finance or asset pricing.

Section 1 surveys the first generation of the market microstructure literature, analyzing the price impact of trades and the spread, assuming competitive suppliers of liquidity. Under this assumption, the revenues of the agents supplying liquidity, corresponding to the spread, simply reflect the costs they incur: order-handling costs (Roll, 1984), adverse–selection costs (Kyle, 1985, Glosten and Milgrom, 1985, Glosten, 1994) and inventory costs (Stoll, 1978). While this literature identified these costs theoretically, it also developed empirical methodologies to analyze data on transaction prices and quantities and estimate trading costs, through the relation between trades and prices and the bid-ask spread (Roll, 1984, Glosten and Harris, 1988, and Hasbrouck, 1988). This literature has shown that trades have both a transitory and a permanent impact on prices. While the former can be traced back to order-handling and inventory costs, the latter reflects information, although not necessarily private information. Furthermore, as data on inventories became available, empirical studies of specialists' or traders' inventories examined the relevance of the inventory paradigm. While this literature has shown that inventory considerations have an impact on the trades of liquidity suppliers, the empirical significance of the impact of inventories on the positioning of their quotes is less clear.[3]

In Section 2 the competitive assumption is relaxed to discuss the case where the supply of liquidity is provided by strategic agents bidding actively in the market. Their market power can lead to a relative lack of liquidity, as shown theoretically by Glosten (1989), Bernhardt and Hughson (1997) and Biais, Martimort and Rochet (2000), and empirically by Christie and Schultz (1994) and Christie, Harris and Schultz (1994). As the focus of the market microstructure literature was shifting from competitive to strategic liquidity suppliers, empirical studies went beyond the analyses of transactions prices and quantities. We survey the insights offered by the literature on quotes and order placement strategies.

Building on the concepts and insights presented in the previous sections, as well as on recent theoretical, empirical and experimental studies, Section 3 discusses the consequences of such aspects of the organization of the markets as the timing of trades, transparency, the thinness of the pricing grid, or the different forms of competition.

- The literature suggests that call auctions can enhance liquidity by concentrating trades at one point in

[3]We also discuss how the first generation of the market microstructure literature conceptualized liquidity in financial markets as reflecting the incentives of the traders to cluster to benefit from the additional liquidity they provide to one another (Admati and Pfleiderer, 1988, and Pagano, 1989).

time and can facilitate price discovery; however, for large trades both empirical and theoretical analyses suggest that the continuous market also offers a useful trading venue.

- The literature points to the benefits of allowing investors to compete to supply liquidity by placing limit orders, to the adverse–selection problems generated by asymmetric access to the marketplace (e.g., Rock, 1990), and to the usefulness of repeated trading relationships to mitigate adverse–selection.

- While the theoretical literature emphasizes the benefits of transparency to mitigate frictions due to asymmetry and strategic behavior, empirical and experimental results are much more ambiguous.

- While very coarse pricing grids can artificially increase spreads, theoretical and empirical analyses suggest that very thin pricing grids can magnify the market power of liquidity suppliers with better access to the marketplace.

- Empirical market microstructure studies also show that while market fragmentation can reduce competition within each of the market centers, it can enhance competition across exchanges.

Section 4 offers a brief conclusion and sketches some avenues of further research.

1 Competitive market makers and the cost of trades

In the first part of this section, to provide a unified perspective (and when it is possible without excessive complications) we survey the theoretical literature within the framework of a simple synthetic model of the market for a risky asset with expected final cash flow, or fundamental value, $E(v) = \pi$, and with N competitive market makers.[4] In the second part of the section, we survey empirical analyses testing and estimating these models.

1.1 Theoretical analyses

In this subsection, we first survey pricing, trades and welfare with order-handling costs, inventory costs, and adverse selection costs in a uniform price auction. Second, we consider alternative modelling frameworks: We discuss the case where limit–orders are placed and then the market order is executed against these schedules.

[4] For the sake of brevity we only mention the assumptions and results, omitting the proofs. The latter are available upon request for the interested reader.

Since, in that market structure, different limit orders, placed at various prices, are executed at those prices, as in a discriminatory auction. We also discuss the case where adverse selection is generated by asymmetries in the timing of the moves, rather than by private signals about the value of the asset.

1.1.1 Order–handling costs and the bid–ask bounce

In the line of Roll (1984) suppose the N market makers are risk neutral and incur an identical cost $\frac{c}{2}q^2$ to trade q shares. This reflects order–handling costs (but not other components of the spread, reflecting inventory costs, adverse selection, or market power, analyzed below). Suppose a market order to buy Q shares has been placed by an uninformed liquidity trader. The market makers each sell $\frac{Q}{N}$ shares at the ask price:

$$A = \pi + (\frac{c}{N})Q, \tag{1}$$

reflecting their marginal cost. Similarly, if the liquidity trader had placed a sell order, the bid price would have been:

$$B = \pi - (\frac{c}{N})Q. \tag{2}$$

Correspondingly the spread is: $2\frac{c}{N}Q$. Generalizing this simple model i) to allow the fundamental value to follow a random walk, and ii) assuming the market orders are i.i.d., there is negative serial autocorrelation in transaction price changes (or returns), due to the bouncing of transaction prices between the bid and the ask quote.

1.1.2 Inventory

Now suppose the market makers are risk averse, as first analyzed by Stoll (1978) and by Ho and Stoll (1981 and 1983). To simplify the analysis we will hereafter focus on CARA utility functions and jointly normally distributed random variables. Denote the constant absolute risk aversion index of the market makers κ, σ^2 the variance of the final cash flow of the asset $(V(v))$, and I the average inventory position of the market makers. When the liquidity trader submits a market order to buy Q shares, the ask price is set as the marginal valuation of the shares by the competitive market makers:

$$A = [\pi - \kappa\sigma^2 I] + (\frac{c + \kappa\sigma^2}{N})Q. \tag{3}$$

Symmetrically the bid price is:

$$B = [\pi - \kappa\sigma^2 I] - (\frac{c + \kappa\sigma^2}{N})Q. \tag{4}$$

The midpoint of the spread (m) is equal to the fundamental value of the asset (π) minus a risk premium compensating the market makers for the risk of holding their initial inventory ($\kappa\sigma^2 I$). Market makers with very long positions are reluctant to add additional inventory and relatively inclined towards selling. Consequently, their ask *and* bid prices will be relatively low. Similarly, market makers with very short inventory positions will tend to post relatively higher quotes. Because of the central role of inventory considerations in this analysis, it is often referred to as the inventory paradigm.

The spread reflects the risk–bearing cost incurred by market makers building up positions to accommodate the public order flow. The price impact of trades is increasing in trade size, as well as the risk aversion of the market makers κ and the variance of the value σ^2.

While the above analysis, in the line of the work of Stoll (1978), is cast in a mean-variance framework in which the link between prices and inventory is linear, under alternative parameterizations inventory effects can be nonlinear. For example, the impact of inventory on prices could be relatively strong for extreme inventory positions.

It is delicate to extend the above simple one–period case to a dynamic context, yet the analysis suggests the following implications of the inventory paradigm for the dynamics of prices and trades: Increases in inventory lead to decreases in prices, which in turn attract buy orders and lead to decreases in inventory. Correspondingly, the midquote is mean-reverting as a reflection of the mean reversion in inventories.

Amihud and Mendelson (1980) analyze an alternative model, where dealers are risk neutral, and yet set prices to manage their inventory positions, because they face constraints on the maximum inventory they can hold. In this dynamic model mean reversion in inventories also arises, along with a nonlinear impact of inventory on pricing.

While individuals are indeed likely to exhibit risk aversion, it is less obvious why the banks, securities houses and other financial institutions employing dealers would be averse to diversifiable risk. To speak to this issue it could be fruitful to analyze theoretically the internal organization of these financial institutions. For example, suppose the dealers need to exert costly but unobservable effort to be efficient and take profitable inventory positions. To incentivize them to exert effort, it is necessary to compensate them based on the profits they make. In this context, even if diversifiable risk does not enter the objective function of the financial institution, it plays a role in the objective function of an individual dealer quoting bid and ask prices.

1.1.3 Adverse selection generated by private information about the final value of the asset

Now consider the case where the market order is placed by an investor trading both for liquidity and informational motives. Considering informed investors is in the line of Bagehot (1971), Grossman and Stiglitz (1980), Kyle (1985), Glosten and Milgrom (1985), and Easley and O'Hara (1987). To study the consequences of adverse selection, while avoiding the unpalatable assumption of exogenous noise traders,[5] and still building on the insights of the inventory paradigm, we now extend the simple model introduced above to the asymmetric information case.[6]

Suppose the market order is placed by a strategic, risk–averse agent, with CARA utility. Denote her risk aversion parameter γ. She is endowed with L shares of the risky asset, and has observed a signal s on the final value v. Specifically assume that: $v = \pi + s + \epsilon$, where π is a constant, $E(s) = 0, E(\epsilon) = 0$, and σ^2 now denotes the variance of ϵ. The market makers do not know exactly the inventory shock of the informed trader. From their viewpoint L is a random variable.[7] Consequently, as the informed order reflects both her signal (s) and her risk–sharing need (L), it provides the market makers with a noisy estimate of the private signal. Denote,

$$\theta = s - \gamma\sigma^2 L,$$

and:

$$\delta = \frac{V(s)}{V(s) + (\gamma\sigma^2)^2 V(L)}.$$

δ is a measure of the information content of the order flow, as it is increasing in the informativeness of the insider signal, and decreasing in the non–informational component of the trade. δ also measures the magnitude of the adverse selection problem. For example, $\delta = 0$ corresponds to the case where there is no private information. θ reflects the valuation of the strategic informed trader for the asset: it is increasing in the private signal, and decreasing in the inventory of the trader.

[5] The exogneous noise trading assumption raises the issue why there exists noise traders willing to lose money. It also makes it impossible to conduct any welfare analysis or to compare different market structures, since it prevents accounting for the impact of the market structure on noise trading. Glosten (1989) and Spiegel and Subrahmanyam (1992) endogenize liquidity trading resulting from rational risk–sharing motives.

[6] Subrahmanyam (1991) extends the analysis of Kyle (1985) to the case of risk–averse market makers posting reservation quotes.

[7] We maintain the assumption, which greatly facilitates the algebraic calculations, that s, ϵ, and L are jointly normal and independent.

If $\delta < \frac{1}{2}$, i.e., the magnitude of the adverse selection problem is limited, then there exists a perfect Bayesian equilibrium where the trade of the informed agent (Q) is affine and increasing in θ and the equilibrium price (P) and the updated conditional expectation of the asset value are affine in the informed trade (Q). More precisely,

$$E(v|Q) = [\delta m + (1-\delta)\pi] + \delta(2\lambda + \gamma\sigma^2)Q, \tag{5}$$

$$Q = \frac{(\pi - m) + \theta}{2\lambda + \gamma\sigma^2}, \tag{6}$$

and:

$$P = m + \lambda Q, \tag{7}$$

where,

$$m = \pi - \frac{\kappa V(v|\theta)}{1-\delta} I, \tag{8}$$

and:

$$\lambda = \frac{\frac{c}{N} + \frac{\kappa V(v|\theta)}{N} + \gamma\sigma^2\delta}{1 - 2\delta}. \tag{9}$$

When $\delta = 0$, i.e., there is no private information, this simplifies to the above presented Roll/Ho and Stoll model. Symmetrically, in the case where market makers are risk neutral ($\kappa = 0$), and there is no order–handling cost ($c = 0$), we obtain a specification similar to Kyle, where prices are equal to updated expectations of the value of the asset, conditional on the order flow. Buy orders convey good news and drive ask prices up, while sell orders convey bad news and push bid prices down. In the general case where $\delta > 0, \gamma > 0$ and $c > 0$, the informational component of the spread is added to those reflecting risk aversion and order-handling costs. The larger the size of the order, the larger its impact on prices. The strategic insider is aware of this effect, and limits the size of the trade to limit its impact.[8] This provides a theoretical framework within which to analyze liquidity: when information asymmetries are severe, market makers have limited risk–bearing ability or order–handling costs are large, trades have a strong impact on prices, which can be interpreted as a form of illiquidity.

[8]An additional way for informed agents to limit part of their price impact is to sell information to other investors, as analyzed by Biais and Germain (2001).

Welfare The profits of the informed agent ($Q(v-p)$) are the mirror image of the losses of the uninformed agents. From a utilitarian perspective, and with CARA utilities, this transfer has no direct impact on social welfare. Nevertheless, information asymmetries do reduce social welfare, because they reduce the risk–sharing gains from trades which can be achieved in the marketplace. This is just another form of the lemon's problem, and is conceptually very similar to the consequences of adverse selection in insurance markets analyzed by Rothschild and Stiglitz (1976). In the remainder of this subsection we illustrate these points in the context of our simple model.[9] For simplicity assume liquidity suppliers are risk neutral (i.e., $\kappa = 0$) and there are no order–handling costs (i.e., $c = 0$). In this case, to maximize gains from trade, the risk–averse agent should entirely trade out of his endowment shock. Denote this first–best trade $Q^* : Q^* = -L$. Because of information asymmetries and strategic behavior, however, the equilibrium trade (Q) is less responsive to inventory shocks than the first–best trade. Indeed, simple manipulations yield:

$$|\frac{\partial Q}{\partial L}| = 1 - 2\delta < |\frac{\partial Q^*}{\partial L}| = 1.$$

This lower responsiveness of trades to endowment shocks reduces the risk–sharing gains from trade and correspondingly, social welfare. The greater the magnitude of the adverse–selection problem, measured by δ, the lower the second–best welfare.

Note that the informed agent scales back his trade because he is aware of its impact on prices. That impact, which can be measured by the slope of the price function, λ, which in this case simplifies to $\gamma\sigma^2\delta/(1-2\delta)$, is also increasing in the magnitude of the adverse–selection problem, δ. Hence, the reduction in social welfare and gains from trade goes hand in hand with the reduction in market liquidity caused by information asymmetries.

Equilibrium multiplicity and endogenous liquidity Restricting the focus to linear strategies leads to uniqueness of the equilibrium. Yet, as is general in signalling games, equilibrium multiplicity can arise. With normal distributions Bhattacharya and Spiegel (1991) examine non–linear strategies.[10]

Building upon the modelling framework and insights of Glosten and Milgrom (1985), Dow (1998) shows that with rational expected utility maximizing liquidity traders, multiple equilibria can arise, corresponding to

[9] Our simple model is amenable to welfare analysis, since there are no noise traders, and all agents are expected utility maximizers.

[10] Biais and Rochet (1997) analyze the class of (non–linear) perfect Bayesian equilibria arising for arbitrary (bounded support) distributions, in trading games where, as in the specification above, the informed agent is risk averse and trades both to exploit his signal and to share risk. Bagnoli, Viswanathan and Holden (2001) and Noldeke and Troeger (1998) study the links between the linearity of the equilibrium and the normality of the distributions.

different levels of endogenous liquidity. Indeed, if it is anticipated that liquidity will be large, liquidity traders trade intensively, and the spread is tight. On the contrary, if low liquidity is expected, uninformed trading is reduced, the proportion of informed trades is large and the spread wide. Hence there are different equilibrium of endogenous liquidity, with different levels of risk sharing and welfare. This contrasts with the analysis of Rochet and Vila (1994), which establishes equilibrium uniqueness in a variant of the Kyle (1985) model, without parametric assumptions on the distributions of the random variables, and under the assumption that the informed trader can place limit orders. The reason why uniqueness obtains in this context is because of the zero–sum property of the game played by the informed agents and the noise traders. This is in contrast with Dow (1998) where the different equilibria correspond to different levels of risk–sharing gains from trade.

Equilibrium multiplicity, and coordination on endogenous liquidity also arise in Admati and Pfleiderer (1988), Pagano (1989) and Foster and Viswanathan (1990). In these models, investors choose to concentrate their trades on a single market or at a single point in time, to benefit from the liquidity externalities generated by other traders.[11] This theory of clustering in trades offers an interpretation for the observed intraday patterns in volume, whereby trading tends to be concentrated at certain points within the trading day. Yet, while empirically clustering occurs at the opening and the closing of the market, this does not follow directly from these theoretical analyses. Hong and Wang (2000) complement them by studying the case where, while informational and non–informational shocks occur continuously over time, the market is periodically closed. This model is able to generate several stylized facts well documented by empirical studies, such as U shapes in trading volume (Jain and Joh, 1988) or in stock returns (Harris, 1986, Smirlock and Starks, 1986, and Wood, McInish and Ord, 1985).

1.1.4 Adverse selection in a discriminatory pricing auction/limit order book

We now turn to a different market structure in which the competitive liquidity suppliers first place limit–order schedules and then the market order is placed and executed against these schedules. As shown by Rock (1990) and Glosten (1994), the nature of adverse selection is somewhat different in that context. Liquidity suppliers cannot condition on the total quantity of the market order (Q) when they place their orders. Rather, they know that if a limit order to sell placed at a given price p is hit, then the total size of the market order is at least as large as the cumulated depth of the book ($q(p)$) up to that price. Consequently, the expectation of the value of the security given that this order has been hit is the following "upper–tail expectation":

[11] Note that in Pagano (1989) there is no adverse selection.

$$E[v|Q > q(p)]. \quad (10)$$

In this context, if the liquidity suppliers are risk neutral and competitive, ask prices are equal to such "upper–tail expectations" while, symmetrically, bid prices are lower–tail expectations.

An important feature of prices set in this discriminatory–pricing context is that there is a "small–trade spread", i.e., infinitesimal trades have a discrete impact on prices. This contrasts with the uniform–price mechanism analyzed above where the price impact is commensurate with the size of the trade. This small–trade spread arises because the ask price for an infinitesimal buy order (denote it ι) impounds non–infinitesimal information content. Indeed, the conditioning set, in the upper–tail expectation: $E[v|Q > \iota]$, includes the case where the total quantity is small (Q close to ι), as well as the cases where it is much greater ($Q >> \iota$).

1.1.5 Adverse selection generated by asymmetries in the timing of the moves

Copeland and Galai (1983) offer an alternative perspective on adverse selection in financial markets. They point out that limit orders or market makers' quotes are exposed to the risk of being "picked off" when the market valuation is changing, resulting in unprofitable execution. This adverse selection problem is of a different nature than that studied by Kyle (1985) and Glosten and Milgrom (1985). While Kyle (1985) and Glosten and Milgrom (1985) emphasize asymmetric information about the value of the asset, such that adverse selection would arise even if the market makers and informed traders moved simultaneously, in Copeland and Galai (1983) the asymmetry in the timing of the moves is at the origin of the adverse selection.

In the line of Copeland and Galai (1983), Foucault, Roëll and Sandas (1999) analyze the case where market makers must decide how often to review market conditions and refresh their quotes. There is a tradeoff between the cost of frequent reviews and the benefits of being picked off less frequently. In addition, there is an externality between market makers, since the frequency with which one market maker reviews his quotes has an impact on the magnitude of the adverse selection problem faced by his competitors. In this context, the frequency of quote revisions, the size of the bid–ask spread, and the magnitude of the adverse selection problem are jointly determined in equilibrium. When one market maker revises his quotes, if the others are informed of this (for example, by a special signal on their trading screens), they rapidly change their own quotes. This offers a theoretical interpretation for the empirical finding in Biais, Hillion and Spatt (1995) that after the best ask or bid has been cancelled, possibly because it had become out of line with the valuation of the stock, it is often the case that another cancellation takes place very rapidly on the same side of the book. While this interpretation corresponds to trading on the Paris Bourse, Foucault, Roëll and Sandas (1999) emphasize the consequence of quote staleness in the presence of SOES "bandits" in the NASDAQ.

1.2 Empirical analyses

In this subsection, we first propose an empirical counterpart to the synthetic theoretical model presented above.[12] Then, we survey empirical approaches and results in light of this synthetic framework.

1.2.1 A simple synthetic framework

The empirical counterpart of the price equation (7) is:

$$P_t = m_t + \lambda_t \, Q_t, \tag{11}$$

where P_t is the transaction price at time t, m_t is the midpoint, and λ_t can be interpreted as the effective half–spread at the time of the transaction Q_t. Typically, the index t is taken to be discrete, and represents a clock measured in number of trades.

In the theoretical analysis above, the impact of the inventory of the market makers on the mid–quote is reflected in equation (8). Its empirical counterpart is:

$$m_t = \pi_t - bI_t, \tag{12}$$

where, reflecting the time series nature of the data, the variables m_t, π_t, and I_t are indexed by time.

To study time series data, we need to specify the dynamics of I_t. A natural candidate would be:

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

A more general formulation is:

$$I_{t+1} = aI_t - Q_t + u_{t+1} \tag{13}$$

[12]The reader should bear in mind the following caveats however. For simplicity, we ignore the effects of discrete prices. Furthermore, and maybe more importantly, a fully–fledged model of the dynamics of trades and quotes in the presence of inventory and information effects can potentially give rise to rather intractable nonlinearities and non–stationarity. For simplicity in the empirical specification presented here we ignore these difficulties and treat the time series of observations as generated by the repetition of one–period models.

Technically, $a < 1$ ensures that the impact of trades on inventories is not permanent. Economically this may reflect several effects: First, the set of agents supplying liquidity is not constant as agents can exit or enter the pool of market makers. Second, liquidity suppliers can unwind their trades in other markets, or hedge them in other securities or markets. u_{t+1} can be thought of as a random exogenous shock on the inventories of the market makers.

In the theoretical analysis in the previous subsection, the trade is given by equation (6). Its empirical counterpart is:

$$Q_t = A_t - d(m_t - \pi_t) + \eta_t, \tag{14}$$

where η_t is the unpredictable component of the trade, conveying noisy information about the insider signal, $d(m_t - \pi_t)$ reflects the impact of market makers inventories on trades, and A_t reflects potential additional predictability in demand at time t. Substituting $m_t - \pi_t = -bI_t$, from equation (12), the trade equation is:

$$Q_t = A_t + bdI_t + \eta_t, \tag{15}$$

i.e., the greater the inventory of the liquidity suppliers, the more they are expected to sell, and correspondingly the more the liquidity trader is expected to buy.

In the theoretical analysis in the previous subsection, the update of the value of the asset conditional on the trade is given by equation (5). To specify its empirical counterpart note that, since changes in expectations must not be forecastable, changes in the true price in response to trade must be a function only of the unanticipated trade. Generalizing slightly the linear equation (5) consider a quadratic polynomial in the signed trade. Also the empirical result in Jones, Kaul and Lipson (1994b) that it is the occurrence of trades rather than the size of trades which conveys information, suggests including in the regressors the discrete variable, η_{0t}, taking the value 1 for purchases and -1 for sales (as in Glosten and Harris (1988)). This leads to the following specification for the permanent response to trades:

$$\pi_{t+1} = \pi_t + z_0\eta_{0t} + z_1\eta_t + z_2\eta_t^2 + v_{t+1}, \tag{16}$$

where the error term (v_{t+1}) is typically assumed i.i.d.[13]

[13] George, Kaul and Nimalendran (1991) generalize equation (16) to account for serial correlation. They show that studies failing to account for serial correlation present in the data overestimate the adverse selection effect.

1.2.2 Surveying several empirical analyses in light of our simple synthetic framework

Order handling costs Roll's (1984) model corresponds to the case where in (11) there is a constant spread, in (12) $b = 0$, in (14) $d = A_t = 0$, η_t are i.i.d and take the value 1 or -1 with equal probability, and in (16) $z_0 = z_1 = z_2 = 0$.

As shown, by Roll (1984), in this model the spread is equal to $-cov(P_{t+1} - P_t, P_t - P_{t+1})$. Hence computing the covariance between consecutive price changes provides an estimate of the spread, even when data on bid and ask quotes or trade sizes is not available. Because of the relatively strong assumptions in this model and because the bid–ask bounce does not play a large role in the variance of returns measured at low frequency, the Roll estimator is not very well adapted for low frequency data. On the other hand, with daily or higher frequency data, the Roll estimator provides a good estimate of the spread. For example, using daily data, Stoll (2000) finds that the covariance between consecutive price changes is negative for 99% of the stocks on the NYSE and 100% of the stocks on Nasdaq, and that the estimates of the relative spread using the Roll (1984) approach are very close to the observed quoted spreads.

Inventory costs In the context of the synthetic specification outlined in equations (11) to (16), the case where inventory (and order handling) costs influence quotes and trades but there no adverse selection corresponds to the case where $z_0 = z_1 = z_2 = 0$. In this context, Ho and Macris (1984) offer an empirical analysis of price and trades dynamics in options markets.

Equation (15), which specifies the dynamics of signed trades, implies they should reflect the inventories of the liquidity suppliers. Consistent with this equation, several papers have provided empirical evidence consistent with the view that market makers tend to sell (resp. buy) when they hold long (resp. short) positions. Hasbrouck and Sofianos (1993) and Madhavan and Smidt (1993) find that there is reversion of specialist inventories towards their mean, though at a slow rate.[14] The order of magnitude of the readjustment lag is between a day and a month. Madhavan and Sofianos (1998) find that specialists participate more actively as sellers (resp. buyers) when they hold long (resp. short) positions. Lyons (1995) provides evidence consistent with inventory effects in the foreign exchange market. Manaster and Mann (1996) find that Chicago Mercantile Exchange (CME) market makers with relatively long (resp. short) positions tend to sell (resp. buy). Also in line with the inventory

[14] Hasbrouck and Sofianos (1993) show that inventory dynamics vary across stocks. They also find little evidence that specialists are hedging their positions across stocks or with options.

control theory, Reiss and Werner (1998) and Hansch, Naik and Viswanathan (1998) find that, on the London Stock Exchange, dealers with long positions tend to sell to dealers with short positions. This collection of findings suggests that the reversion of market maker inventories is a robust feature of many diverse trading mechanisms, consistent with $bd > 0$ in equation (15).

Equation (12) implies that the midquote should be decreasing in the inventory of the liquidity suppliers. The empirical evidence on this impact of inventories on prices and quotes is ambiguous, however. On the one hand, consistent with equation (12), Madhavan and Smidt (1993) find that increases in the inventory of the specialist lead to decreases in quotes. On the other hand, Madhavan and Sofianos (1998) find that the specialists control their inventories through the timing of the direction of their trades rather than by adjusting quotes. Furthermore, Manaster and Mann (1996) find that, contrary to the implications of the pure inventory theory, market makers with long (resp. short) positions tend to sell at relatively large (resp. buy at small) prices.[15] This suggests that theories of pricing by market makers need to reflect additional features besides the pure theory of inventory control.

Adverse selection Glosten and Harris (1988) offers one of the first empirical specifications in line with the adverse selection paradigm. Their model corresponds to the case where there are no inventory effects so that in (12) $b = 0$ and in (14) $d = 0$ and $A_t = 0$.[16] They estimated this market microstructure model using intraday data and found that significant amounts of NYSE common stock spreads were due to asymmetric information. Several more recent studies offer empirical results consistent with the adverse selection model:

λ in equation (11) is a measure of the depth of the market (as λ goes up, depth is reduced).As the informational motivation of trades becomes relatively more important, λ goes up. Consistent with this prediction, Lee, Mucklow and Ready (1993) find that around earnings announcements (when adverse selection is likely to be more severe) depth is reduced and spreads widen on the NYSE. This is more pronounced for announcements with larger subsequent price changes.[17] In addition, Sofianos (1995) finds that specialists on average incur positioning losses on their inventory, which are compensated by their revenues from spreads.[18] Furthermore, the adverse

[15] Of course, the structure of the market studied by Manaster and Mann (1996) on the CME is very different than the NYSE specialist system.

[16] In addition, Glosten and Harris (1988) developed a methodology to take into account discreteness of the price grid (a feature of the data not taken into account in the empirical specification (11) to (16)).

[17] Similarly, Kavajecz (1999) finds that both specialists and limit order traders reduce depth around information events.

[18] See also the results of Naik and Yadav (1999).

selection model predicts that the informational price impact of trades should be commensurate with the private signal underlying the informed trade. Consistent with this, Seppi (1992) finds positive correlation between price changes associated with block trades and subsequent innovations in earnings announcements. Also consistent with equations (11) to (16), Huang and Stoll (1994) find that, after a large purchase, occurring at a price significantly above the midquote, the midquote is expected to go up, reflecting the impact of the trade on the bid and ask prices.

Both inventory and adverse selection theories predict that trades impact prices, but the former predicts that this impact should be transient, while the latter predicts that this impact should be permanent. This permanent impact goes through the impact of unexpected trades on expectations, modeled in equation (16) above. In the context of the pure inventory/order–handling cost paradigm, z_0, z_1 and z_2 should be 0, in contrast to the prediction of the adverse selection paradigm. Hasbrouck (1991) analyzes the joint process of trade and quote revisions using a VAR approach.[19] In fact, manipulating equations (11) through (16) and taking $a < 1$ in equation (13), we obtain the VAR specification in Hasbrouck (1991). He finds that trades do have a permanent impact, inconsistent with the hypothesis that there is no information content in trades.[20] In a similar spirit, and in the line of the seminal work of Kraus and Stoll (1972) and Holthausen, Leftwich and Mayers (1990), a body of empirical literature has studied the permanent price impact of block trades, reflecting adverse selection, and their transient impact, likely to reflect inventory and liquidity considerations.

The econometric approach taken by Easley, Kiefer and O'Hara (1997) to study adverse selection is quite different from that outlined above (in equations (11) to (16)). They consider a theoretical model in the line of Glosten and Milgrom (1985) and Easley and O'Hara (1992). In particular, as in Easley and O'Hara (1992), informed agents can be trading in the marketplace with some probability. Relying on this theoretical model, Easley, Kiefer and O'Hara (1997) write down the likelihood that, on a given day, a given number of purchases, sales and time intervals without trades are observed and estimate the deep parameters of the model. They find that the probability that there is private information in the marketplace on a given day is 75%, and that, when there is private information, the probability that a given trader is informed is equal to 17%.[21] One methodological

[19] This VAR specification is richer than the specification in Hasbrouck (1988), which i) analyzed the univariate process of signed trades and ii) regressed quotes changes onto trades.

[20] The impact of *public* information upon price changes is analyzed in Jones, Kaul and Lipson (1994a).

[21] Bernhardt and Hughson (1995) offer an alternative approach to estimating the probablility that a given trader is informed. Their estimate varies between 0.43% and 9%, depending on the specification.

difference between their approach and the specification in equations (11) to (16) is that Easley, Kiefer and O'Hara (1997) base their estimation on the likelihood of trades, rather than the joint likelihood of trades and prices.

1.3 Summary of the results and avenues of further research

Table 1 offers a summary of the theoretical and empirical results surveyed in this section. The perfect market hypothesis, under which trades have no impact on prices, is strongly rejected. Furthermore, the market microstructure literature provides insights as to the causes of this rejection. The hypothesis that market makers face no inventory constraints is rejected. In addition, trades have a permanent impact on prices. That this impact is *permanent* (as shown by the work of Hasbrouck) is important because it points at information effects, while analyses restricted to the short–term impact of trades on prices could not disentangle inventory effects (as studied by Ho and Macris, 1984) from adverse selection (as studied in Glosten and Harris, 1988). Another interesting piece of evidence on the long–term impact of trades on prices stems from the foreign exchange market. While macroeconomic variables, such as interest rates, fail to explain variations in exchange rates (e.g., Meese and Rogoff, 1983), Evans and Lyons (1999) find that signed order flow has significant explanatory power. It would be interesting to find out if this impact of order flow on asset prices also arises in low–frequency stock price data.

While the finding that trades have a permanent impact on prices is consistent with the adverse selection theory (and therefore does not lead to rejecting it), further work is needed to test that paradigm. Other phenomena besides adverse selection, such as the reaction of traders and investors to public information, could lead to positive correlation between the direction of trades and that of price changes. An important avenue for further research is to find out the extent to which the permanent impact of trades on prices reflects private as opposed to public information. Neal and Wheatley (1998) offer intriguing results regarding this issue. They estimate for closed-end funds a market microstructure econometric model similar to that described in the above section. While for these assets there is very little scope for asymmetric information about the liquidation value, the estimates of the adverse selection component are large and significant. They conclude that this suggests that either adverse selection arises primarily from factors other than the current liquidation value or that the empirical models are misspecified.

The empirical analyses surveyed above offer interesting characterizations of high–frequency data and insights on models, relying for example on VAR analysis or structural estimation of deep parameters by maximum likelihood analysis. Sandas (2001) offers a structural econometric analysis of Glosten's (1994) model of competitive liquidity supply in an electronic limit order book. Using the GMM overidentifying restrictions approach enables him to both estimate the deep parameters of the model, and test the null that the model is consistent with the data. In its richest parametrization, the model is rejected for 50% of the stocks. The directions in which the model is rejected point at interesting avenues of further research: the slope of the limit order book appears to be steeper than predicted by the theory. This could reflect market power as discussed in the next section. Also, the estimates of order–handling costs tend to be negative, which could be the result of limit orders being placed by patient investors desiring to trade. A promising avenue of further research is to develop additional tests of the restrictions imposed by theoretical market microstructure models.[22]

2 Active bidding, strategic liquidity suppliers and endogenous liquidity demand

Instead of focusing on competitive market makers, the second generation of the market microstructure literature considers strategic agents, bidding proactively to exploit market conditions and possibly private information, while supplying liquidity. Competition between liquidity suppliers is similar to competition between bidders in an auction. Models based on inventory effects are similar to analyses of private–value auctions, while adverse selection–based models parallel common–value auctions. When the number of liquidity suppliers is limited, because inventory–holding and adverse–selection costs reduce their willingness or ability to supply liquidity, strategic market makers can earn rents, reflecting their market power.

2.1 Strategic liquidity supply without adverse selection

2.1.1 Market power

To the extent that the overall supply of liquidity is limited, as a consequence of risk aversion, limits on the maximum position of market makers, or restricted entry in the market–making business, traders posting quotes strategically can earn rents.

To illustrate the consequences of liquidity suppliers' risk aversion under imperfect competition, consider the model presented in the previous section. For simplicity assume there is no information asymmetry ($\delta = 0$) and

[22] Foster and Viswanathan (1995) offer a structural estimation of a Kyle (1985) model. They find that the model is rejected because it cannot explain simultaneously the level of volume and the covariation between price changes and trades.

no order–handling costs ($c = 0$), but the N liquidity suppliers are risk averse ($\kappa > 0$) and strategic. Suppose the liquidity suppliers compete in price schedules and the investor responds by determining her optimal trade. Along with strategic considerations, the price schedules arising in equilibrium reflect the costs faced by the liquidity suppliers. As shown in Subsection 1.2, the marginal cost of supplying the Q^{th} share is: $\pi - \kappa\sigma^2 I + \kappa\sigma^2 Q/N$ (see equation (3), above). Note that this marginal cost is increasing in Q. Consistent with results obtained in industrial organization, this induces deviations from competitive pricing as shown by Biais, Foucault and Salanié (1998), Roëll (1999), and Viswanathan and Wang (1999). Taking a different approach, Dutta and Madhavan (1997) use a repeated prisoner's dilemma model to show that dealers can implicitly collude on wide spreads.

These theoretical analyses are consistent with several empirical studies suggestive of strategic behavior by liquidty suppliers. Christie and Schultz (1994) and Christie, Harris and Schultz (1994) document the use of a wide pricing grid to sustain large spreads on the NASDAQ.[23] As a consequence of the resulting controversy, the SEC required that public investors be allowed to supply liquidity by placing limit orders, thereby competing with the NASDAQ dealers. Barclay, Christie, Harris, Kandel and Schultz (1999), study the consequences of this reform implemented in 1997. They find that quoted and effective spreads after the implementation of the reform fell substantially from their pre-reform level. Additionally, they find that an even larger decline in the spread occurred from 1994 to 1996 (i.e., before the reform) as a consequence of the adverse publicity and investigations. The impact of the controversy in reducing spreads is analogous to the reaction to the Christie and Schultz (1994) paper that is documented in Christie, Harris and Schultz (1994). Naik and Yadav (1999) analyze empirically the consequences of the reform which took place in 1997 in the London Stock Exchange, allowing the public to compete with dealers through the submission of limit orders. They find that the effective spread decreased significantly, and that this decrease is larger than that documented by Barclay et al. (1999) for NASDAQ. They also find that the cross–subsidization among trade sizes has disappeared, leading to a decline in trading costs for small trades and an increase in these costs for large trades. These results suggest that allowing all investors to place limit orders leads to a reduction in the market power of the dealers. While non–anonymity is key in supporting such collusion on large spreads in a repeated interaction environment, the anonymity prevailing in ECNs makes it less likely to emerge. This is consistent with the finding by Simaan, Weaver and Whitcomb (1999) that odd–tick avoidance is less prevalent in ECNs.

While the empirical and theoretical papers surveyed above suggest that, when the number of dealers is finite,

[23] See also Huang and Stoll (1996).

liquidity supply is imperfect, dealers' entry could be expected to mitigate, or eliminate this imperfection. Indeed, Wahal (1997) shows that on the Nasdaq entry and exit of market makers is a pervasive phenomenon, and entry leads to declines in spreads. However, the empirical results of Ellis, Michaely and O'Hara (2002) show that the competitive pressure exerted by such entry is limited. They find that entering market makers fail to capture a meaningful share of the market. Correspondingly, one dealer tends to dominate trading in a stock, which tends to increase spreads.

2.1.2 Dynamic order placement strategies

For tractability, the synthetic model presented in the first section of this survey, as well as a large part of the microstructure literature, is based on a one period analysis. This approach does not capture the dynamic nature of liquidity provision and order placement strategies in the market place. Several empirical and theoretical papers offer insights into this dynamic process, however.

While in many theoretical analyses (including those surveyed above) some traders are exogenously assumed to use limit orders and others market orders, in practice, investors can choose between limit and market orders. Foucault (1999) endogenizes this choice in a stationary dynamic model. He analyzes theoretically the investor's decision to hit the current quote or place a limit order as a function of the state of the order book, imposing rational expectations about the endogenous probability of execution of limit orders. Consistent with intuition, it is optimal to place limit orders when the spread is large, while it is optimal to place market orders when the spread is tight. Parallel to this theoretical analysis, Harris and Hasbrouck (1996, see Table 5) find that for stocks with a 1/4 spread (in their sample period the tick size was 1/8) the execution performance of orders placed within the quotes dominates that of market orders.[24] This is also consistent with the empirical analysis of the conditional frequencies of different strategies in the electronic limit order book in Paris by Biais, Hillion and Spatt (1995). They find that when the spread is relatively wide liquidity is often supplied (limit order suppliers beat the existing quote), while when the spread is relatively narrow, investors are more willing to accept the prevailing liquidity, which is being offered on relatively favorable terms, and place market orders.[25] This gives

[24] Harris and Hasbrouck (1996, Table 3) also include some summary statistics on order frequencies and observe that in their dataset the most commonly used limit order tends to be the best performing order.

[25] The finding, by Madhavan and Sofianos (1998) that the specialist tends to participate more in trades when the spread is large is consistent with the specialist following a similar type of liquidity supply strategy. Further insights on the role of the specialist and that of limit orders in the supply of liquidity are offered in the theoretical analysis of Seppi (1997).

rise to mean reversion in the bid–ask spread and negative serial autocorrelation in ask (or bid) price changes, as, when the ask price has moved above its equilibrium level, it is undercut by a more favorable limit order to sell.[26] Note that this reversion to the mean is not instantaneous, as it takes some time for the liquidity providers to identify these order placement opportunities. Yet, Biais, Hillion and Spatt (1995) find that order placement occurs more rapidly when the spread is large than when it's tight. This relatively fast reaction reflects the speed with which investors monitoring the market seize the opportunity to supply liquidity, when the latter is scarce and well compensated. While Harris and Hasbrouck (1996) and Biais, Hillion and Spatt (1995) use descriptive statistics to focus on economic interpretations, Bisière and Kamionka (1998), Lo, MacKinlay and Zhang (2001) and Engle (2000) analyze the order and trade dynamics using more sophisticated econometric techniques.[27]

Parlour's (1998) theoretical analysis of dynamic order placement studies how investors trade off less attractive pricing against the improved price priority obtained by jumping ahead of the queue of limit orders and undercutting the current best quotes. Consistent with this theoretical analysis, Biais, Hillion and Spatt (1995) find that investors are more likely to place limit orders within the quotes when the depth at the quotes is large.

An interesting avenue of further research is to integrate the insights generated by these dynamic analyses of order placement strategies with the results obtained in a one period adverse selection context, summarized throughout the present survey. Ultimately, this research program could yield a characterization of dynamic order placement strategies, reflecting liquidity, adverse selection and market power considerations.

2.2 Strategic liquidity supply with adverse selection

2.2.1 Uninformed liquidity suppliers

While in the previous sub–section we considered the case where liquidity suppliers are risk averse but there is no adverse selection (similarly to private value auctions), we now turn to the case where the liquidity suppliers are risk neutral but face an informed agent (as in common value auctions). As in the previous subsection, and as in Subsection 1.4, we consider the case where first liquidity suppliers compete in price schedules and then the informed agent determines her optimal trade: Q. To determine their optimal price schedules, the strategic market makers must evaluate the cost of supplying liquidity. The cost of supplying the q^{th} share is equal to:

[26] This is related to, but different from the negative autocorrelation in transaction price changes generated by the bid–ask bounce, analyzed in Roll (1984).

[27] Griffiths, Smith, Turnbull and White (2000) offer an interesting empirical approach to the characterization of order placement strategies and their performance.

$E(v|Q > q)$, as in equation (10). As in the previous section, this marginal cost is increasing in trade size, but unlike in the previous section this reflects adverse selection rather than risk aversion.

First consider the case where there is only one, monopolistic, liquidity supplier, as in Glosten (1989). Very much in the line of textbook monopoly theory, the selling prices quoted by the monopolist are equal to the sum of his cost and a monopolistic markup, $m_1(q)$, where the subscript refers to the fact that there is only one liquidity supplier:

$$A(q) = E(v|Q > q) + m_1(q).$$

Similar to analyses of price discrimination, the markup of the monopolist reflects the elasticity of the demand he faces, which in turn reflects the distribution of the different types of agents:

$$m_1(q(\theta)) = \frac{1 - F(\theta)}{f(\theta)},$$

where θ is the valuation of the informed agent for the asset (as explained in Subsection 1.1.3), $q(\theta)$ is the optimal trade size for the agent whose type is θ, and $F(\theta)$ is the c.d.f. of agents' types, while $f(\theta)$ is the corresponding density. One can draw an analogy between the results obtained by Glosten (1989) and those obtained by the analysis of monopoly pricing in the presence of information asymmetries on private values by Goldman, Leland and Sibley (1984). What is much less standard in Glosten (1989), relative to that literature, is that in the common value environment he analyzes, the marginal cost for the monopolist of supplying shares is endogenous.

Second, consider the case where there are $N > 1$ strategic liquidity suppliers competing in price schedules. Bernhardt and Hughson (1997) show that in this case there exists no equilibrium where oligopolists earn zero expected profits, very much in contrast with the zero profits earned by competitive market makers in the Kyle (1985) signalling paradigm. Biais, Martimort and Rochet (2000), characterize the equilibrium price schedules arising in this context. As in the monopoly case, they are the sum of a cost component and a mark–up, $m_N(q)$. The latter is decreasing in the number of liquidity suppliers and goes to 0 as N goes to infinity. In that limiting case the oligopolistic equilibrium converges to the competitive equilibrium analyzed by Glosten (1994). The contrast between the oligopolistic rents arising in the competition in schedules game for finite N, and the zero–profit result arising in the Kyle (1985) signalling game can be interpreted as follows:

In the signalling case, having all observed the size of the market order Q , all the market makers conduct the same Bayesian inference and identically update their expectation of the value of the security: $E(v|Q)$. Price competition between the market makers then amounts to standard Bertrand competition, with identical and constant marginal cost, equal to the conditional expectation of the value of the security. In this context, if all the market makers but one quoted ask prices strictly above this cost, it would be optimal for the last market maker to undercut them.

In the screening game, the market makers post their schedules before observing the size of the market order. The latter is then split across the market makers, to minimize (resp. maximize) the total price paid for the purchase (resp. sale). Now consider the problem faced by one of the market makers, confronted with the equilibrium schedules of his competitors. Combining these schedules with the best response of the informed agent, one obtains the residual demand curve faced by the market maker. Confronted with this residual demand, the market maker behaves similarly to a monopolist.[28] When setting his ask price schedule, he has to trade off increases in sales against reductions in margins. This tradeoff reflects the (finite) elasticity of the residual demand curve faced by each market maker. Note that this finite elasticity, and the resulting oligopoly rents would not arise if there were no information asymmetry about the common value of the asset.[29] Indeed, this information asymmetry implies that the marginal cost of supplying the shares for a given market maker (equal to the expectation of the value of the security conditional on the amount he sells) increases with the quantity he sells. This increase in marginal costs, in turn implies that the price of each market maker is increasing in the quantity he sells. It is this increase which implies that the elasticity of the residual demand faced by each market maker is finite. Note also that this does not arise either in the signalling case, where, since the size of the market order is publicly announced to all the market makers before they quote their prices, their marginal cost (which only reflects this announcement) is independent of the quantity they sell. Finally note that the link between the valuation of the asset by each market maker and the amount he trades is similar to the winner's curse arising in the common value auction analyzed by Milgrom and Weber (1982).

[28] In the limiting case where the number of liquidity suppliers goes to infinity, the same results as in Glosten (1994) obtain.

[29] With private values, $m_N(q) = 0$, so that the equilibrium outcome of the screening game is the same as that of the signalling game. This equivalence arising for private values, as well as the difference between the two games with common values, is reminiscent of the results obtained in auction theory. As shown by Milgrom and Weber (1982), the revenue–equivalence theorem obtained for private values does not hold with common values.

2.2.2 Informed liquidity suppliers

Liquidity suppliers can directly observe signals about the asset payoff: The first analyses of strategic liquidity supply by informed agents were offered by Kyle (1984 and 1989). Kyle (1989) analyzes the linear equilibria of Cournot competition between risk–averse informed oligopolists in a one–shot, uniform–price call auction. As in the Cournot model, the oligopolists, who trade on the basis of their private signals, limit the size of their trades to limit their market impact, and earn market power rents. Note however, that liquidity demand is simply assumed to be an exogenous noise term. Also, in such market setups, there are generally a multiplicity of equilibria, as shown for example by Klemperer and Meyer (1989). Hence focusing on the linear equilibrium is not without consequence.

In the spirit of the auction–theoretic work of Milgrom, Engelbrecht and Weber (1982), Calcagno and Lovo (1998) offer an extension of Kyle's (1985) dynamic analysis to the case where the informed agent is not the market order trader, but one of two risk–neutral market makers. As in Kyle (1985), the market makers compete in prices for market orders. The latter stem from exogenous noise traders. At each point in time there is an auction, whereby each of the market makers places one bid price and one ask price (conditional on his observation of past prices and trades), and the exogenous noise trader hits the best bid or ask. The uninformed market maker understands he faces a winner's curse, and factors it into his bidding strategy. In equilibrium he earns zero expected profits. In order to preserve his rent, the informed agent follows mixed strategies, such that his quotes are only partially revealing, except at the last round of the game. The informed agent faces a tradeoff, between larger quantities (and thus larger immediate profits) and more information revelation. This is similar to the tradeoff arising in Kyle (1985). The difference is that, in Calcagno and Lovo (1998), it is the quotes of the market makers, rather than the market order flow, which partially reveal information, and thus lead the price discovery process. It could be interesting, in future research, to empirically test the extent to which market makers, rather than their customers, possess private information.

Manaster and Mann (1996) provide empirical evidence which speaks to this issue. They find that CME market makers tend to sell at relatively high prices and to buy at relatively low prices. This is consistent with market makers taking positions based on superior information about the likely evolution of prices.[30] Such information

[30] Interestingly, the empirical result obtained by Manaster and Mann (1996), and its theoretical counterpart based on informed market makers, go in the opposite direction than the empirical results obtained for the NYSE and London Stock Exchange which corresponded to the situation where the market makers are uninformed and face superiorly informed traders (see Table 1)

could be gathered based on market information, such as orders or the observation of other market participants on the floor. While this private information could be the source of the profitability of market makers, there may be an alternative interpretation, emphasizing the market power of market makers, which enables them to buy at relatively cheap (bid) prices and sell at relatively expensive (ask) prices.

Liquidity suppliers can also observe pieces of market information: Vayanos (1999) offers a dynamic extension of Kyle (1989), where strategic risk–averse agents have private information about their endowments, while information about the dividend flow is public. To share risk, agents with long positions are inclined to sell, while agents with short positions are inclined to buy. This is similar to the case analyzed in Subsection 1.3, where the valuation of the strategic trader for the asset (θ) was shown to be decreasing in her endowment in the stock. In this context, the equilibrium aggregate valuation is decreasing in aggregate holdings. Hence information about endowments is not only informative about private values but also about common values, since endowments affect market prices. Hence, since they reflect endowments, trades convey signals relevant for pricing the asset. Consequently, they have an impact on prices. To reduce this impact the agents reduce the aggressiveness of their trades. This limits their ability to share risk and is therefore detrimental to social welfare. This distortion would not arise if endowments were public information. Similarly, Cao and Lyons (1999) analyze how dealers can extract from their order flow information about aggregate holdings and therefore, market pricing, and Viswanathan and Wang (1999) analyze the case where traders, informed about the asset payoff, transmit orders to dealers, who use the information content of these orders in the interdealer market.

2.3 Conclusion and implications

The findings of the second generation of market microstructure research surveyed in this section are summarized in Table 2, Panel A and B. Overall they suggest that the assumption that liquidity providers are competitive, although convenient to simplify theoretical analyses, does not arise out of the formal treatment of realistic institutional arrangements for trade. In a variety of market structures, a very large number of liquidity suppliers is needed for the equilibrium to be approximated by a zero–profit condition. Hence, oligopolistic rents must be taken into account, along with inventory, adverse selection and order–handling costs, to understand the sources of transactions costs. From a policy perspective, this suggests that exchange regulators and organizers must foster entry and competition for the supply of liquidity, in order to reduce market power, and consequently transactions costs.

In light of the analyses surveyed above, an interesting avenue for further research could be to disentangle, econometrically, the consequences of market power on spreads from the competitive components of spreads. Another important issue is to identify the nature of private information, by disentangling fundamental information about individual firms from signals inferred from the observation of the trading process. While the former can be obtained by investors and asset managers, market makers and traders have special access to the latter. This can contribute to their market power.

3 Market design

The organization of the market can be seen as the extensive form of the game played by investors and traders. It determines the way in which the private information and strategic behavior of the traders affect the market outcome. Similar to auction design or mechanism design, market microstructure analyzes how these rules of the game can be designed to minimize frictions and thus optimize the efficiency of the market outcome. In this section we review the body of empirical and theoretical results comparing the determination of prices and allocations within several particular market structures. At a higher level these results can be used to provide important insights about the efficient design of trading systems.

3.1 Call versus Continuous

3.1.1 Concentrating trades at one point in time can be efficient

In practice, relatively inactive stocks are traded on a call market, as the level of activity increases, this is combined with a continuous market (in which the specialist adds liquidity); very actively traded assets are exchanged on a continuous basis. This suggests that call markets can be used to concentrate liquidity when the latter is not plentiful. This raises the question why economic agents need to be instructed to concentrate their trades, by means of such a trading mechanism.

This may be related to the public good nature of liquidity. Admati and Pfleiderer (1988) and Pagano (1989) show that clustering of trades naturally arises, even if it is not mandated by the structure of the market. Yet, in such a context, multiple equilibria can arise, due to the strategic complementarities among liquidity supplies. Hence mandating concentration of trades and orders, by using a call market, can be seen as a device to help traders coordinate on an equilibrium, in order to minimize trading costs.

The theoretical analysis of Vayanos (1999), mentioned in the previous section, offers another reason why mandating agents to concentrate their trades, as in a call market, can be welfare improving. As discussed above, he shows that, to reduce their price impact, the strategic agents split their trades. Since this reduces their ability to trade out of their endowment shocks, this reduces the gains from trade achieved in the marketplace. Gains from trade could be improved if the trader could credibly commit to engage in a single trade. In this case, a liquidity supplier could accommodate his risk–sharing demand, at the price corresponding to his trade size, without fearing that additional trades would take place in the future, altering further the value of the stock.[31] The smaller the time interval between trading opportunities (h), the greater the ability of traders to strategically split their trades and the greater the welfare loss. This loss is maximized when h goes to 0. Batching orders at discrete points in time, as in a call market, may enhance welfare, by enhancing the ability of strategic traders to commit to a single trade.

The theoretical analysis of Copeland and Galai (1983) presented in Section 1, suggests yet another reason why call trading can be efficient. In their analysis, adverse selection stems from asymmetry in the timing of the moves—providers of liquidity must quote based on current information, while a future market order that hits that quote is based on the information available at that future time. In a call auction, providers of liquidity must quote, but they can deliver the quote just before the known time of the call. Thus, if the arrival rate of information is high relative to the arrival rate of orders, as would be the case in a thinly–traded security, the call auction can minimize informational differences at the time of trade and lead to greater risk sharing. On the other hand, if the arrival rate of orders is high relative to the arrival rate of new information, the gains from periodic calls are small, and offset by the gains to traders of being able to rebalance their portfolios when they choose.

Trading halts can be viewed, at least in part, as an institutional response to these economic forces. A trading halt occurs when the arrival rate of information is high. The halt itself sends a signal to the traders who monitor their quotes relatively infrequently, giving them the opportunity to revise their limit orders. Consistent with these remarks, Corwin and Lipson (2000) find that cancellations and the placement of new limit orders are particularly frequent during trading halts and that a large proportion of the order book at the resumption of trading is composed of orders placed during the halt.

[31] This is in the same spirit as in the upstairs block market analyzed by Seppi (1990).

3.1.2 The informational efficiency of call auctions

Amihud and Mendelson (1987) found that on the NYSE the opening price was noisier than the closing price.[32] One possible interpretation was that the market mechanism used at the opening, similar to a uniform–price call auction, was less efficient than the mechanism used at the close, i.e., continuous trading. An alternative interpretation was that the opening price was more difficult to find than the closing price, reflecting the contrast between the uncertainty following the overnight trading period, and the price discovery achieved at the end of the trading day. To differentiate across these two alternative interpretations, Amihud and Mendelson (1991) and Amihud, Mendelson and Murgia (1990) analyzed markets where call auctions were held at other points in time than the opening. They found that prices set in such call auctions were not less efficient than comparable continuous market prices. This leads to the conclusion that the relative inefficiency of the opening call auction price does not reflect the trading mechanism but the fact that overnight the market is closed and the price discovery associated with the price discovery process halted.

To cope with the difficulty of the discovery of opening prices many stock exchanges have introduced tâtonnement procedures. For example during the preopening period, in the Paris Bourse, agents can place, revise or cancel orders and indicative prices reflecting aggregate supply and demand are displayed. Biais, Hillion and Spatt (1999) find empirically that while in the early phase of the preopening period the hypothesis that indicative prices are pure noise cannot be rejected, towards the end of the preopening the hypothesis that prices are rational expectations of the value of the asset cannot be rejected.

These empirical findings can be interpreted in light of recent experimental studies. Schnitzlein (1996) finds that the informational efficiency of prices is not significantly different in a one-shot, uniform price, call auction, and in a continuous market. Biais and Pouget (2000) find that, while the mere presence of an opening call auction is not sufficient to improve drastically the informational efficiency of prices, the combined effect of a preopening period and a call auction does improve the informational efficiency of the price discovery process.

[32] Ronen (1992) notes that measures of informational efficiency computed for several stocks over the same period of time are likely not to be independent, since the dynamics of the prices of these stocks are correlated. She proposes a GMM–based method, to deal with this correlation.

3.1.3 Uniform pricing in call auctions

Another difference between call auctions and continuous trading is that in the former all trades are executed at a single uniform price, while in the latter, as orders walk up or down the book, and as the latter evolves, trades are filled at different prices. In the previous sections we reviewed the difference between equilibrium outcomes arising under uniform pricing (Kyle, 1985, 1989) and those arising with discriminatory pricing (Glosten 1994, Bernhardt and Hughson, 1997, Biais, Martimort and Rochet, 2000). These analyses show that there is a small trade spread in the discriminatory–price auction but not in the uniform–price auction. On the other hand, for large trades, transactions costs are lower in the discriminatory–price auction than in the uniform–price auction. This is consistent with the empirical results of Kehr, Krahnen and Theissen (1998) who find that, on the Frankfurt Stock Exchange, for small trades transactions costs are lower in the call market, while for large trades they are lower in the continuous market.

3.1.4 Conclusion

The literature surveyed in this subsection is summarized in Table 3, Panel A. Overall it suggests that, call auctions can enhance welfare, and possibly the informational efficiency of the market. Continuous markets, however, can offer a useful complement to opening call auctions. Studying the complementary, and possibly the competition, between these two market structures can offer an interesting avenue for further research.

3.2 Who should supply liquidity?

Liquidity can be supplied by a variety of agents including limit order traders, dealers, floor brokers and specialists. These parties can be subject to different priority rules, enjoy market power to different degrees and possess differing amounts of information. For example, the specialist on the NYSE enjoys less priority than all other market participants, but has access to better information about orders in the book and pricing in related markets due to his last mover advantage. Another example relates to NASDAQ and the London Stock Exchange where until recently only dealers had the opportunity to post quotes. In pure limit order markets, such as the Paris Bourse, differences among liquidity suppliers are less important, though only brokers can observe the identification code of the other brokers transmitting orders. A fundamental issue in the design of trading systems concerns the determination of the different rules applying to liquidity suppliers and the information to which they have access. The NYSE specialist example illustrates some important adverse selection issues that arise as a consequence of the asymmetry in the timing of trading opportunities of different liquidity suppliers.

When a marketable orders arrives on the trading floor, the specialist can decide to allow the order to be executed against the outstanding limit orders, or to fill the order himself. He can achieve that by undercutting the book or by "stopping" market orders and guaranteeing execution at the posted quote or price improvement. Such price improvements are frequent since, as documented by Sofianos (1995), specialists trade extensively inside the quoted spread. As the specialist possesses information about the potential information content of the order, based for example upon the pricing in related markets, the size of the incoming order or the identity of the potential counterparty, he can condition his decisions upon information not available to the investors when they placed limit orders in the book.[33] As shown by Rock (1990), this creates an adverse selection problem for these investors and discourages them from providing liquidity. As stated in equation (10) in Section 1.4: $E(v|Q > q)$ is the cost of the q^{th} unit sold by the limit order traders, where Q is the size of the market order. The opportunistic intervention of the specialist modifies the distribution of Q, and raises the cost of the limit order traders. This reduces the extent to which they provide liquidity for the market. Rock (1990) shows that, when the specialist is risk neutral (so that there are no risk–sharing benefits from splitting the trade between him and the limit order traders), the adverse selection problem is so extreme that the limit order book entirely dries out, and the specialist is the only liquidity provider.[34] Consistent with this analysis, Ready (1999) finds that orders that the specialist stops are more profitable to the liquidity supplier than orders which are allowed to transact against the limit order book.

On the NYSE the specialist observes the orders that are in the book immediately prior to the opening, and can use this information to choose his own supply or demand. This raises essentially the same adverse selection problem as in Rock (1990). In fact, the magnitude of this problem may be especially large at the opening relative to the trading day due to the large uncertainty about the valuation of the stock and the considerable private information obtained by the specialist through the observation of the supply and demand stemming from many orders. Stoll and Whaley (1990) relate empirically the monopoly power of the specialist at the opening to the statistical properties of opening prices, namely that the open–to–open volatility is larger than the close–to–close volatility, and that the overnight innovations in returns are partially reversed during the day. Madhavan and

[33] This is consistent with the finding in Madhavan and Sofianos (1998) that specialists participate more in smaller trades.

[34] Seppi (1997) extends the analysis of Rock (1990) to compare the performance of a pure limit order market to that of a hybrid market with a specialist and competing limit orders. He finds that a hybrid market provides better liquidity to small retail and institutional trades, while a pure limit order market may offer better liquidity on mid-size orders.

Panchapagesan (2000) enrich our understanding of these issues by analyzing data about the limit orders in the book at the opening and the specialist's opening trade. They find that this trade tends to bring the opening price closer to the fundamental value of the asset. While they interpret this result as suggesting that the specialist enhances price discovery at the opening, we offer the alternative interpretation that the specialist buys (resp. sells) when the price reflecting the orders in the book is undervalued (resp. overvalued) as suggestive of adverse selection.

Benveniste, Marcus and Wilhelm (1992) offer an interesting counter–argument to the view that the status of the specialist enhances adverse selection. They argue that the repeated and non–anonymous interaction between the specialist and floor brokers can help to cope efficiently with information asymmetries. Consider the opposite situation, whereby investors would infrequently and anonymously interact in the marketplace. In that setting there would not be significant reputational costs to being opportunistic in the trading process. In contrast, because the brokers non–anonymously and repeatedly interact with the specialist, they would bear large reputational costs if they were to misrepresent their trading intentions to him.[35] Consistently with this argument, Venkataram (2001) finds empirically that, other things equal, the NYSE is more liquid than the Paris Bourse.

Table 3, Panel B, summarizes the theoretical and empirical analyses of the role of the specialist surveyed in this subsection. Overall these analyses suggest that, to reduce market power, and consequently transactions costs, all investors should be granted the ability to supply liquidity on equal conditions (level playing field). Indeed, most major markets (including the NYSE, the NASDAQ, the London Stock Exchange, the Tokyo Stock Exchange, XETRA and the Paris Bourse) now allow for the placement of limit orders by all investors.

3.3 Transparency

In transparent markets abundant information is available to investors and traders about orders and quotes (ex–ante transparency) and about transactions (ex–post transparency). As this tends to equalize information across market participants, transparency reduces the magnitude of adverse selection problems. Since these problems, as shown in Subsection 1.3, reduce the gains from trade, transparency can be anticipated to increase welfare. Indeed, within the context of an adverse selection–based model of the spread (in the same spirit as the synthetic

[35] Leach and Madhavan (1993) put forward another reason why the specialist system might induce more efficient markets than an open limit order book. They show that the specialist can use its monopolistic situation to experiment with prices and thus induce more informative order flow, thereby expediting price discovery.

model outlined in Subsection 1.3), Pagano and Roëll (1996) show theoretically that transparency reduces the transaction costs incurred by uninformed investors. Consistent with that analysis, Flood, Huisman, Koedjick and Mahieu (1999) find that pre–trade transparency narrows spreads in experimental financial markets.

One could argue, however, that trade disclosure can make it harder to supply liquidity to large traders. After large trades, in a transparent markets, the market maker can be in a difficult bargaining position to unwind his inventory. Naik, Neuberger, and Viswanathan (1999) offer an interesting counterargument. After the risk–averse dealer has bought a block from a potentially informed trader, he seeks to unload his position. Yet to mitigate his price impact, he reduces the size of his trade, thus reducing his ability to share risk.[36] This does not arise with trade disclosure. In that case, since the market has already taken into account the information content of the trade, the dealer can unwind his inventory with little incremental price impact. Consequently, trade disclosure enhances risk sharing. The empirical evidence in Gemmill (1996) is consistent with the view that transparency at least does not reduce liquidity. Gemmill (1996) analyzes liquidity in the London Stock Exchange under three publication regimes: from 1987 to 1988 dealers had to immediately report their trades, from 1991 to 1992 they had to do so within 90 minutes, while from 1989 to 1990 they had 24 hours to do so. He finds that there is no gain in liquidity from delayed publication of block trades, as the spreads and the speed of price adjustment are not affected by the disclosure regime.

Yet, in a dynamic trading environment, transparency can have ambiguous consequences, as shown by the experimental and theoretical analyses of Bloomfield and O'Hara (1999 and 2000). Consider a two–period model where informed agents can submit market orders to one of N liquidity suppliers. If the market is opaque, only the liquidity supplier who accommodated the order at the first period is informed about its occurrence, sign and size. Consequently, at the first period, liquidity suppliers quote relatively tight spreads, to attract order flow, and thus acquire private information.[37] At the second period, however, the liquidity suppliers who did not participate in the initial trade face a double winner's curse problem: i) with respect to the informed agent, and ii) with respect to the informed liquidity supplier. This relatively severe adverse selection problem widens their spreads. The market spread is wide also, as the informed liquidity supplier finds it optimal to undercut his competitors by just one tick. Thus, different temporal patterns emerge in the transparent and the opaque market. While in the latter, spreads may be initially relatively large, they decrease fast, as information is revealed through time.

[36] This is similar to the effect analyzed in Vayanos (1999).

[37] This could provide a interpretation for the findings by Madhavan, Porter and Weaver (2000) that an increase in ex–ante transparency on the Toronto Stock Exchange in 1990 led to an increase in spreads.

In the former, in contrast, while initial spreads are relatively tight, later spreads tend to remain relatively high. It could be interesting to test this result by comparing U shape intraday patterns in spreads across markets with different levels of transparency.

The above results are summarized in Table 3, Panel C. The contrasting conclusions reached by the different studies reflect the facets of market liquidity on which they focus. It could be interesting to integrate these different perspectives, to identify and quantify the tradeoffs among the different aspects of transparency. This could be useful to evaluate the overall impact of transparency on welfare.

3.4 Tick Size

Early studies of the discreteness of transaction prices documented the pervasiveness of clustering on round prices (Harris, 1991) and developed econometric methodologies to bridge the gap between theoretical models with continuous prices and discrete transactions price data (Glosten and Harris, 1988, Hausman, Lo and MacKinlay, 1992). More recently, the emphasis has shifted towards analyzing the consequences of price discreteness on trading strategies and market outcomes.

Coarse pricing grids can mechanically constrain liquidity suppliers and result in excessively large spreads.[38] Consider for example a discrete–price version of the Glosten (1994) model, as in Sandas (2001). Assuming that time precedence holds, and that there are a large number of potential traders, equilibrium requires that at each price the last share offered just breaks even. In line with the discussion of limit order books above, this implies that the lowest offer, A_1, is the smallest allowable price greater than or equal to the expectation of the value of the asset, conditional on the market order being strictly positive: $E[V|Q > 0]$. The quantity offered at this price is Q_1 such that: $A_1 = E[V|Q > Q_1]$. If A_i is the i^{th} offer price, the cumulative quantity offered at A_i or lower, Q_i, is given by $A_i = E[V|Q > Q_i]$. The equilibrium limit order schedule with discrete pricing is a step function with points of upward jumps (in the case of the offer) lying on the continuous price equilibrium schedule. This model predicts that a decrease in the tick size will generally reduce the quoted spread, reduce the amount offered at each price, but leave cumulative quantity at the original set of prices unchanged.

While the conclusion that finer grids lead to tighter spreads is in line with the stylized facts, there is also empirical evidence that tightening the pricing grid can reduce the overall depth of the order book, see e.g., Goldstein and Kavajecz (2000) and Jones and Lipson (2001). This reduction in liquidity could reflect an increase

[38] Harris (1994) develops an econometric methodology to assess the consequences on the spread of a reduction in the tick size.

in the magnitude of the Rock (1990) adverse selection problem, since reducing the tick size makes it less costly for the specialist to undercut the book.[39] Making it less costly to undercut also undermines the value of the time priority enjoyed by limit orders, and thus discourages their placement, as noted by Harris (1994).[40] Indeed, Cordella and Foucault (1999) show that, with relatively coarse prices, liquidity suppliers find it advantageous to rapidly quote the narrowest possible spread, to benefit from time priority at this relatively advantageous price. By making time priority less valuable, fine ticks reduce the cost of hiding orders – as the main cost of the latter is that they do not benefit from time priority. Consistently with this point, Harris (1998) finds that on the Toronto Stock Exchange and the Paris Bourse the fraction of orders that is hidden is relatively larger when the tick size is relatively finer.

The results presented in this subsection are summarized in Table 3, Panel D. These analyses suggest that, while tick size may a priori seem a relatively trivial issue, it can have significant consequences in the market by emphasizing the consequences of other imperfections, such as for example the Rock (1990) adverse selection problem, or the non–competitive behavior of liquidity suppliers, as illustrated by Christie and Schultz (1994).

3.5 Intermarket Competition

3.5.1 The costs of fragmentation

As discussed in Subsection 1.3, since orders provide liquidity to one another, there is a natural tendency for trades to concentrate on one market (see Pagano (1989), Admati and Pfleiderer (1988) and Chowdhry and Nanda (1991)). While these analyses suggest that market fragmentation should not arise in equilibrium, they are developed under the assumption that liquidity suppliers are competitive. Strategic liquidity suppliers can find it optimal to provide liquidity outside the primary market, thus inducing market fragmentation.[41] For example, they can offer "quote matching," i.e., promise to execute a maximum number of shares at the market quote determined in the primary market. This "cream skimming" is possible when time priority is not enforced across exchanges.[42] As "quote matchers" capture relatively less informed orders, the adverse selection problem faced

[39] This suggests that the reduction in the tick size should lead to an increase in the specialist's participation rate. Consistent with this analysis, the specialist's participation rate was larger in 1998 than in 1996 (see NYSE 1998 Fact Book).

[40] See also Spatt and Srivastava (1994).

[41] Blume and Goldstein (1997) and Bessembinder and Kaufman (1997) provide empirical evidence of effective fragmentation.

[42] Deviation from time priority can arise in the U.S., as specialists on one exchange can match the National Best Bid and Offer, and thus execute orders even if they did not previously post the best bid or offer. Our analysis of the negative consequences of this

limit orders traders in the primary market is increased, similar to the effect of the specialist in Rock (1990). Consistent with this analysis, Easley, Kiefer and O'Hara (1996) find that there is a significant difference in the information content of orders executed in New York and Cincinnati.

To illustrate these points, suppose that a quote matching exchange promises to transact Q_0, and suppose for simplicity that all small orders go to the quote matching exchange. Then the market ask will be the smallest allowable price greater than $E[V|Q > Q_0]$, which exceeds $E[V|Q > 0]$. The remainder of the equilibrium limit order schedule will be unaffected. With a relatively large tick size, quote matching will have no effect on the nature of the quotes. However, as the tick size gets smaller, or the adverse selection problem gets larger, quote matching is predicted to have more of an effect, and correspondingly market fragmentation widens the spread.[43] It should also be noted that according to this model, quote matching will be profitable at any tick size, no matter how small. Thus, if the model is correct, payment for order flow should be unaffected by a reduction in the tick size.

3.5.2 The benefits of intermarket competition

While the above arguments imply that fragmentation reduces market quality when the liquidity suppliers in the central marketplace are competitive, this result is not necessarily upheld when these liquidity suppliers enjoy market power. In that case, the presence of a second market can exert a beneficial competitive pressure on the central market. Several empirical studies actually point in that direction. Battalio, Greene and Jennings (1997) study the impact of a reform which allowed brokers to execute their customer orders themselves on the Boston and Cincinnati Stock Exchanges without respecting the time priority of other dealers on these exchanges. They find that the ability of brokers to preference their own specialist units led to a substantial diversion of executions from the NYSE to these regional markets. As this took place, the NYSE spread actually declined. Similarly, Battalio (1997) finds that NYSE spreads are reduced after Madoff Securities begins purchasing order flow to attract order flow away from the NYSE. In the same spirit, the results of Lightfoot, Martin, Peterson and Sirri (1999) do not support the hypothesis that preferencing arrangements reduce the quality of financial markets. Neal (1987) finds that options which can be traded in several exchanges have relatively lower spreads.

feature of the National Market System is consistent with the finding in the industrial organization literature that price matching is anti–competitive.

[43] In contrast, in a private value environment, fragmentation does not generate adverse selection, and thus does not widens the average spread, as shown by Biais (1993).

Furthermore, while fragmentation reduces the incentives to supply liquidity in the primary in market, it need not imply reduced aggregate depth, as shown in Glosten (1998). Consider two competing pure limit order books, I and II, each honoring time precedence among its own quoters, but not across markets. Market order users randomly send their order to one or the other of the exchanges. However, order–handling rules require that if an order exhausts the quantity on one exchange the remainder is sent to the other exchange for execution. Let μ be the probability that a market order is sent to exchange I. The last share at the lowest offer A, on exchange I will execute if 1) the market order is sent to exchange I and it is larger than Q_I or 2) the market order is sent to exchange II and it is larger than $Q_{II} + Q_I$. Thus, the quantities Q_I and Q_{II} must satisfy:

$$\mu(A - E[V|Q > Q_I]) \Pr(Q > Q_I) + (1 - \mu)(A - E[V|Q > Q_{II} + Q_I]) \Pr(Q > Q_{II} + Q_I) = 0,$$

and:

$$(1 - \mu)(A - E[V|Q > Q_{II}]) \Pr(Q > Q_{II}) + \mu(A - E[V|Q > Q_I + Q_{II}]) \Pr(Q > Q_I + Q_{II}) = 0.$$

Thus, the ask price must be greater than $E[V|Q > Q_I]$, and lower than $E[V|Q > Q_{II} + Q_I]$. Thus there is a reduced incentive to quote quantity on each exchange. However, the aggregate quantity, $Q_I + Q_{II}$ will be larger when there are two exchanges. In effect, competition between the exchanges forces the quoters to compete on the average share rather than the marginal share, thus reducing the profitability of the infra-marginal shares. As the tick size decreases, the magnitude of this effect decreases, and in the limit disappears.

3.5.3 The organization of intermarket competition affects its efficiency

While the above discussed results lead to a somewhat ambiguous conclusion, they may reflect some specific features of the architecture of U.S. markets that would not arise in other contexts. First, in a setting where time priority would be enforced across markets the above discussed negative effects of intermarket competition would not arise. Second, the profitability of attracting orders away from the NYSE may reflect some privileges enjoyed by those present on the trading floor. Indeed the specialist does not have to fully disclose the order book to the public, and floor brokers have the option to share trades with the book. This can create rents for those present on the floor, and correspondingly transactions costs for the other players. In this context, it can be profitable to attract orders away from the exchange to side step these rents. This suggests that in a context where i) time priority would be enforced across markets, and ii) no one would benefit from a privileged status, the competition between markets would not have negative effects. Note that conditions i) and ii) would hold in the case of competition between electronic limit order books. The consolidation of all sources of liquidity that would arise in this context is reminiscent of the analysis of "the inevitability of an electronic limit order book" analyzed in Glosten (1994).

In addition, the coexistence of markets could be useful to reap the benefits from competition among exchanges, especially with respect to the dynamics of the market structure and the incentives to innovate in developing new trading mechanisms and technologies. For example, the modernization of European stock markets since the mid–eighties, including the switch to continuous trading and electronic markets, was spurred by the competitive pressure of the London based International SEAQ. Note however that competition between exchanges may not always lead to optimal market structures, as shown by the theoretical analysis of Foucault and Parlour (1999). In their model, stock exchanges choose listing fees and trading costs, which determine their attractiveness for firms interested in listing and for investors. As firms differ in the extent to which they value decreases in trading costs, and as different combinations of fees and costs can be viewed as differentiated products, two competing exchange can find it optimal (in order to reduce competitive pressure) to design different fees and costs structures, and serve two different market niches. The corresponding duopolistic equilibrium can fail to implement the welfare maximizing market structure, and even lead to lower welfare than a monopolistic situation.

4 Conclusion

The market microstructure literature has generated a rich body of theoretical and empirical analyses and established a number of facts. Inventory control considerations are important for liquidity suppliers and explain part of the impact of trades. The evidence is also consistent with the view that another part of this impact stems from the information content of orders and the associated adverse selection. In several market contexts, liquidity suppliers, such as dealers or specialists, have been found to exercise market power. This also explains part of the impact of trades.

Several policy implications emerge from these theoretical and empirical analyses. Because market power is at the root of a significant part of the transactions costs incurred by investors, open access to the supply of liquidity should be encouraged. To minimize adverse selection constraints, the different suppliers of liquidity should be allowed to intervene on a level playing field, in terms of market information, priority, or order–handling procedures. Along with the competition among liquidity suppliers within one market, the competition across marketplaces plays an important role in curbing market power and intermediation rents. The analyses also yield several insights into different aspects of the design of markets: the use of call auctions, preceded by preopening sessions, is likely to facilitate information aggregation, price discovery and risk sharing; while coarse pricing grid are likely to generate transactions costs, in particular by facilitating the exercise of market power, infinitesimal ticks may not be optimal; while the theoretical literature generally advocates transparency, in practice strategic considerations might undermine the benefits of transparency.

These policy implications, and the underlying theoretical and empirical microstructure literature, have played a key role in the rapid evolution of financial markets over the last decades. The Christie and Schultz (1994) study has greatly contributed to opening the access to the supply of liquidity to many investors. Also because of technological advances, there has been a general move around the globe towards electronic limit order books, with an opening call auction followed by a continuous trading session (Paris, Tokyo, Frankfurt, London, and to some extent NASDAQ). This move has been facilitated and in part motivated by numerous empirical and theoretical studies documenting the advantages of this market architecture.

Perhaps the next challenge for the application of market microstructure analysis is to translate it into a body of applicable methods, to be used by investors and traders in the design of their order placement strategies.[44] Quite promisingly in this respect, major financial players are actually currently developing tools to measure liquidity and design trading robots. One of the difficulties, in this venture, is to bridge the gap between the qualitative implications from stylized theoretical models and the precise quantitative rules required in practice.

Market microstructure still offers numerous exciting avenues of further research.

- Obviously, analyzing the interactions between market microstructure and asset pricing or corporate finance is an important task – although these topics are beyond the scope of the present survey. It would also be fruitful to analyze the interplay between market microstructure and psychology considerations – not unlike those discussed by behavioral finance (see e.g., Biais, Hilton, Mazurier and Pouget, 2001).

- Another fascinating topic is the analysis of the consequences of technological advances, such as the internet, for the architecture and workings of financial markets. Because it lowers entry costs, the internet has intensified competition between market centers, and erodes the rents earned by financial intermediaries. It

[44]This line of research could build on the insights on institutional trading behavior offered by Chan and Lakonishok (1995), Keim and Madhavan (1995), and Cheng and Madhavan (1997), the econometric approach developed by Lo, MacKinlay, and Zhang (2001), and the analysis of order placement strategies by Biais, Hillion and Spatt (1995) and Harris and Hasbrouck (1996).

will be important and interesting to study the industrial organization of these phenomena.[45]

- Yet another key element in the dynamics of exchanges and markets is their ownership and governance structure.[46] Several exchanges have recently gone public (e.g., the Paris, London and Frankfurt Bourses). It will be useful to analyze the implications of such changes.

- New progresses in market microstructure analysis will also stem from the use of new methodologies. Several recent studies have relied on experimental economics to study aspects of market design.[47] This approach is very promising, for at least three reasons. First it enables one to vary the institutional context and the structure of the market, an option which is rather infrequently available in field research.[48] Second, it enables the researchers to observe important elements which are difficult to disentangle for field research: information sets, potential gains from trade, equilibrium behavior.[49] Third, it enables to measure the extent to which agents converge to, or deviate from, equilibrium behavior, and how this is related to the organization of the market.[50] Another tool which might be fruitfully used by microstructure analysis is mechanism design theory, which has already been very useful in other (related) fields such as the analysis of spectrum auctions or IPOs.[51] This approach might be helpful for analyzing the welfare consequences of market design.

- Like a large fraction of the market microstructure literature, the present survey is devoted to the analysis of stock markets.[52] The analysis of other markets (e.g., derivatives, foreign exchange, or energy markets), and their comparison with stock markets is an important avenue of research.[53] Among the many issues which arise, one can point at the following: How does the optimal organization of the market reflect the

[45] Some preliminary investigations in these topics include Simaan, Weaver, and Whitcomb (1999), Huang (2000), and Biais, Bisière and Spatt (2001).

[46] Davis (1999) offers interesting historical insights in the evolution of these institutions.

[47] Building on the seminal papers of Plottt and Sunder (1982, 1988), early experimental studies of financial markets include Copeland and Friedman, (1987, 1991).

[48] See, e.g., Bloomfield and O'Hara (1998, 1999, 2000), Schnitzlein (1996), and Flood et al, (1999).

[49] See, e.g., Pouget (2001).

[50] See, e.g., Biais and Pouget (2000).

[51] See, e.g., Benveniste and Spindt (1989), Benveniste and Wilhelm (1990), Spatt and Srivastava (1991), and Biais, Bossaerts and Rochet (2001).

[52] Except for the interesting paper by Evans and Lyons (1999), see also Lyons (1995).

[53] Biais and Hillion (1994) and Easley, O'Hara and Srinivas (1998), analyze the joint workings of stock and options markets.

characteristics of the asset traded in this market? Is it likely that all financial markets will converge towards a standard model, such as electronic limit order books? If such convergence takes place, is it going to rely on an integrated internet architecture?

References

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

Amihud, Y., and H. Mendelson, 1987, "Trading Mechanisms and Stock Returns: An Empirical Investigation," *Journal of Finance*, 42, 533-555.

Amihud, Y., and H. Mendelson, 1991, "Volatility, Efficiency and Trading: Evidence from the Japanese Stock Market," *Journal of Finance*, 46, 1765-1789.

Amihud, Y., and H. Mendelson, 1980, "Dealership Market," *Journal of Financial Economics*, 8, 31-53.

Amihud, Y., H. Mendelson, and M. Murgia 1990, "Stock Market Microstructure and Return Volatility: Evidence from Italy," *Journal of Banking and Finance*, 14, 423-440.

Bagehot, W., 1971, "The Only Game in Town," *Financial Analysts Journal*, 22, 12-14.

Bagnoli, M., S. Viswanathan, and C. Holden, 2001, "On the Existence of Linear Equilibria in Models of Market Making," *Mathematical Finance*, 11, 1-33.

Barclay, M., W. Christie, J. Harris, E. Kandel, and P. Schultz, 1999, "The Effects of Market Reform on the Trading Costs and Depths of Nasdaq Stocks," *Journal of Finance*, 54, 1-34.

Battalio, R., 1997, "Third Market Broker-Dealers: Cost Competitors or Cream Skimmers?" *Journal of Finance*, 52, 341-352.

Battalio, R., J. Greene, and R. Jennings, 1997, "Do Competing Specialists and Preferencing Dealers Affect Market Quality?" *Review of Financial Studies*, 10, 969-993.

Benveniste, L., A. Marcus, and W. Wilhelm, 1992, "What's Special About the Specialist?," *Journal of Financial Economics*, 32, 61-86.

Benveniste, L., and P. Spindt, 1989, "How Investment Bankers Determine the Offer Price and Allocation of New Issues," *Journal of Financial Economics*, 24, 343-361.

Benveniste L., and W. Wilhelm, 1990, "A Comparative Analysis of IPO Proceeds Under Alternative Regulatory Environments," *Journal of Financial Economics*, 28, 173-207

Bernhardt, D., and E. Hughson, 1997, "Splitting Orders," *Review of Financial Studies*, 10, 69-101.

Bernhardt, D., and E. Hughson, 1995, "Intraday Trade in Dealership Markets," working paper.

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, 293-319.

Bhattacharya, U., and M. Spiegel, 1991, "Insiders, Outsiders, and Market Breakdowns," *Review of Financial Studies*, 4, 255-282.

Biais, B., 1993, "Price Formation and Equilibrium Liquidity in Fragmented and Centralized Markets," *Journal of Finance*, 48, 157-185.

Biais, B., P. Bossaerts, and J. Rochet, 2001, "An Optimal IPO Mechanism," forthcoming, *Review of Economic Studies.*

Biais, B., T. Foucault, and F. Salanie, 1998, "Floors, Dealer Markets and Limit Order Markets," *Journal of Financial Markets*, 1, 253-284.

Biais, B., and L. Germain, 2001, "Incentive Compatible Contracts for the Sale of Information," forthcoming, *Review of Financial Studies.*

Biais, B., and P. Hillion, 1994, "Insider and Liquidity Trading in Stock and Options Markets," *Review of Financial Studies*, 7, 743-780.

Biais, B., P. Hillion, and C. Spatt, 1995, "An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse," *Journal of Finance*, 50, 1655-1689.

Biais, B., P. Hillion, and C. Spatt, 1999, "Price Discovery and Learning during the Preopening Period in the Paris Bourse," *Journal of Political Economy*, 107, 1218-1248.

Biais, B., D. Martimort, and J. Rochet, 2000, "Competing Mechanisms in a Common Value Environment," *Econometrica*, 68, 799-838.

Biais, B., and S. Pouget, 2000, "Microstructure, Incentives and Equilibrium Discovery in Experimental Financial Markets," working paper, Toulouse University.

Biais, B., and J. Rochet, 1997, "Lecture Notes in Mathematics," 1656, Springer Verlag, *Financial Mathematics.*

Biais, B., D. Hilton, K. Mazurier, and S. Pouget, 2001, "Psychological Traits and Trading Strategies," working paper, Toulouse University.

Biais, B., C. Bisiere and C. Spatt, Imperfect competition in financial markets: Island versus Nasdaq, Working paper, Toulouse University.

Bisiere, C., and T. Kamionka, 1998, "Timing of Orders, Order Aggressiveness and the Order Book in the Paris Bourse," forthcoming, *Annales d'Economie et de Statistique.*

Bloomfield, R., and M. O'Hara, 1998, "Does Order Preferencing Matter?" *Journal of Financial Economics*, 50, 3-37.

Bloomfield, R., and M. O'Hara, 1999, "Market Transparency: Who Wins and Who Loses?" *Review of Financial Studies*, 12, 5-35.

Bloomfield, R., and M. O'Hara, 2000, "Can Transparent Markets Survive?" *Journal of Financial Economics*, 55, 425-459.

Blume, M., and M. Goldstein, 1997, "Quotes, Order Flow, and Price Discovery," *Journal of Finance*, 52, 221-244.

Calcagno, R., and S. Lovo, 1998, "Bid-Ask Price Competition with Asymmetric Information between Market Makers," working paper, CORE.

Cao, C., and R. Lyons, 1999, "Inventory Information," working paper, University of California at Berkeley.

Chan, L., and J. Lakonishok, 1995, "The Behavior of Stock Prices Around Institutional Trades," *Journal of Finance*, 50, 1147-1174.

Cheng, M., and A. Madhavan, 1997, "In Search of Liquidity: Block Trades in the Upstairs and Downstairs Markets," *Review of Financial Studies*, 10, 175-203.

Chowdhry, B., and V. Nanda, 1991, "Multimarket Trading and Market Liquidity," *Review of Financial Studies*, 4, 483-511.

Christie, W., J. Harris, and P. Schultz, 1994, "Why Did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes?" *Journal of Finance*, 49, 1841-1860.

Christie, W., and P. Schultz, 1994, "Why do NASDAQ Market Makers Avoid Odd-Eighth Quotes?" *Journal of Finance*, 49, 1813-1840.

Copeland, T., and D. Friedman, 1987, "The Effect of Sequential Information Arrival on Asset Prices: An Experimental Study," *Journal of Finance*, 42, 763-797.

Copeland, T., and D. Friedman, 1991, "Partial Revelation of Information in Experimental Asset Markets," *Journal of Finance*, 46, 265-295.

Copeland, T., and D. Friedman, 1992, "The Market Value of Information: Some Experimental Results," *Journal of Business*, 65, 241-266.

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

Cordella, T., and T. Foucault, 1999, "Minimum Price Variations, Time Priority and Quote Dynamics," *Journal of Financial Intermediation*, 8, 141-173.

Corwin, S., and M. Lipson, 2000, "Order Flow and Liquidity around NYSE Trading Halts," *Journal of Finance*, 55, 1771-1805.

Davis, L., 1999, "Membership Rules and the Long Run Performance of Stock Exchanges: London, New York, and Paris - Some Lessons from History," working paper, California Institute of Technology.

Declerck, F., 2001, "Why markets should not necessarily reduce tick size", Working paper, Toulouse University.

Dow, J., 1998, "Self-Sustaining Liquidity in an Asset Market with Asymmetric Information," working paper, European University Institute.

Dutta, P., and A. Madhavan, 1997, "Competition and Collusion in Dealer Markets," *Journal of Finance*, 52, 245-276.

Easley, D., N. Kiefer, and M. O'Hara, 1996, "Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow," *Journal of Finance*, 51, 811-833.

Easley, D., N. Kiefer, and M. O'Hara, 1997, "One Day in the Life of a Very Common Stock," *Review of Financial Studies*, 10, 805-835.

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

Easley, D., and M. O'Hara, 1992, "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 577-605.

Easley, D., M. O'Hara, and P. Srinivas, 1998, "Option Volume and Stock Prices: Evidence on Where Informed Traders Trade," *Journal of Finance*, 53, 431-465.

Ellis, K., R. Michaely, and M. O'Hara, 2002, "The making of a dealer market: From entry to equilibrium in the trading of Nasdaq stocks", forthcoming *Journal of Finance*.

Engle, R., 2000, "The Econometrics of Ultra-High Frequency Data," *Econometrica*, 68, 1-12.

Evans, M., and R. Lyons, 1999, "Order Flow and Exchange Rate Dynamics," NBER Working Paper No. W7317.

Flood, M., R. Huisman, K. Koedijk and R. Mahieu, 1999, "Quote Disclosure and Price Discovery in Multiple-Dealer Financial Markets," *Review of Financial Studies*, 12, 37-59.

Foster, D., and S. Viswanathan, 1990, "A Theory of the Interday Variations in Volume, Variance, and Trading

Costs in Securities Markets," *Review of Financial Studies*, 3, 593-624.

Foster, D., and S. Viswanathan, 1995, "Can Speculative Trading Explain the Volume–Volatility Relation?" *Journal of Business and Economic Statistics*, 13, 379–396.

Foucault, T., 1999, "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market," *Journal of Financial Markets*, 2, 99-134.

Foucault, T., and C. Parlour, 1999, "Competition for Listings," working paper.

Foucault, T., A. Roëll, and P. Sandas, 1999, "Imperfect Market Monitoring and SOES Trading," forthcoming *Review of Financial Studies.*

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

Gemmill, G., 1996, "Transparency and Liquidity: A Study of Block Trades on the London Stock Exchange Under Different Publication Rules," *Journal of Finance*, 51, 1765-1790.

George, T., G. Kaul, and M. Nimalendran, 1991, "Estimation of the Bid-Ask Spread and Its Components: A New Approach," *Review of Financial Studies*, 4, 623-656.

Glosten, L., 1989, "Insider Trading, Liquidity, and the Role of the Monopolist Specialist," *Journal of Business*, 62, 211-235.

Glosten, L., 1994, "Is the Electronic Open Limit Order Book Inevitable?" *Journal of Finance*, 49, 1127-1161.

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

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

Glosten, L., 1998, "Precedence Rules," working paper, Columbia University.

Goldman, M., H. Leland, and D. Sibley, 1984, "Optimal Nonuniform Prices," *Review of Economic Studies*, 51, 305-319.

Goldstein, M., and A. Kavajecz, 2000, "Eighths, Sixteenths and Market Depth: Changes in Tick Size and Liquidity Provision on the NYSE," *Journal of Financial Economics*, 56, 125-149.

Griffiths, M., B. Smith, S. Turnbull, and R. White, 2000, "The Costs and Determinants of Order Aggressiveness," *Journal of Financial Economics*, 56, 65-88.

Grossman, S., and J. Stiglitz, 1980, "On the impossibility of informationally efficient markets", *American Economic Review*, 393–408.

Hansch, O., N. Naik, and S. Viswanathan, 1998, "Do Inventories Matter in Dealership Markets? Evidence from the London Stock Exchange," *Journal of Finance*, 53, 1623-1656.

Harris, L., 1998, "Does a Large Minimum Price Variation Encourage Order Exposure?" working paper, University of Southern California.

Harris, L., 1986, "A Transactions Data Study of Weekly and Intradaily Patterns in Stock Prices," *Journal of Financial Economics*, 16, 99-117.

Harris, L., 1991, "Stock Price Clustering and Discreteness," *Review of Financial Studies*, 4, 389-415.

Harris, L., 1994, "Minimum Price Variations, Discrete Bid-Ask Spreads, and Quotation Sizes," *Review of Financial Studies*, 7, 149-178.

Harris, L., and J. Hasbrouck, 1996, "Market vs. Limit Orders: The SuperDOT Evidence on Order Submission Strategy," *Journal of Financial and Quantitative Analysis*, 31, 213-231.

Hasbrouck, J., 1991, "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46, 179-207.

Hasbrouck, J., 1988, "Trades, Quotes, Inventories, and Information," *Journal of Financial Economics*, 22, 229-252.

Hasbrouck, J., and G. Sofianos, 1993, "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists," Journal of Finance, 48, 1565-1593.

Hausman, J., A. Lo, and C. MacKinlay, 1992, "An Ordered Probit Analysis of Transaction Stock Prices," *Journal of Financial Economics*, 31, 319-379.

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

Ho, T., and H. Stoll, 1983, "The Dynamics of Dealer Markets Under Competition," *Journal of Finance*, 38, 1053-1074.

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

Holthausen, R., R. Leftwich, and D. Mayers, 1990, "Large-Block Transactions, The Speed of Response, and Temporary and Permanent Stock-Price Effects," *Journal of Financial Economics*, 26, 71-95.

Hong, H., and J. Wang, 2000, "Trading and Returns Under Periodic Market Closures," *Journal of Finance*, 55, 297-354.

Huang, R., 2000, "Price Discovery by ECNs and Nasdaq Market Makers," working paper, Vanderbilt University.

Huang, R., and H. Stoll, 1994, "Market Microstructure and Stock Return Predictions," *Review of Financial Studies*, 7, 179-213.

Huang, R., and H. Stoll, 1996, "Dealer Versus Auction Markets: A Paired Comparison of Execution Costs on NASDAQ and the NYSE," *Journal of Financial Economics*, 41, 313-357.

Jain, P., and G. Joh, 1988, "The Dependence Between Hourly Prices and Trading Volume," *Journal of Financial and Quantitative Analysis*, 23, 269-283.

Jones, C., G. Kaul, and M. Lipson, 1994a, "Information, Trading, and Volatility," *Journal of Financial Economics*, 36, 127-154.

Jones, C., G. Kaul, and M. Lipson, 1994b, "Transactions, Volume, and Volatility," *Review of Financial Studies*, 7, 631-651.

Jones, C., and M. Lipson, 2001, "Sixteenths: Direct Evidence on Institutional Execution Costs," *Journal of Financial Economics*, 59, 253-278.

Kehr, C., J. Krahnen, and E. Theissen, 1998, "The Anatomy of a Call Market: Evidence from Germany," working paper.

Keim, D., and A. Madhavan, 1995, "Anatomy of the Trading Process: Empirical Evidence on the Motivation for and Execution of Institutional Equity Trades," *Journal of Financial Economics*, 37, 371-398.

Klemperer, P., and M. Meyer, 1989, "Supply Function Equilibria in Oligopoly under Uncertainty," *Econometrica*, 57, 1243-1277.

Kraus, A., and H. Stoll, 1972, "Price Impacts of Block Trading on the New York Stock Exchange," *Journal of Finance*, 27, 569-587.

Kyle, A., 1984, "Market Structure, Information, Futures Markets and Price Formation," in *International Agricultural Trade: Advanced Readings in Price Formation, Market Structure and Price Instability*, edited by Gary Storey, Andre Schmitz and Alexander H. Sarris, Boulder, Colorado, Westview Press.

Kyle, A., 1989, "Informed Speculation with Imperfect Competition," *Review of Economic Studies*, 56, 317-355.

Kyle, A., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315-1335.

Leach, C., and A. Madhavan, 1993, "Price experimentation and security market structure", *Review of Financial Studies*, Vol 6, No 2, 375–404.

Lee, C., B. Mucklow, and M. J. Ready, 1993, "Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis," *Review of Financial Studies*, 6, 345-374.

Lightfoot, L., P. Martin, M. Peterson, and E. Sirri, 1999, "Order Preferencing and Market Quality on United States Equity Exchanges," working paper.

Lo, A., C. MacKinlay, and J. Zhang, 2001, "Econometric Models of Limit-Order Executions," *Journal of Financial Economics*, forthcoming.

Lyons, R., 1995, "Tests of Microstructural Hypotheses in the Foreign Exchange Market," *Journal of Financial Economics*, 39, 321-351.

Madhavan, A., 2000, "Market Microstructure," *Journal of Financial Markets*, 3, XXXX.

Madhavan, A., and V. Panchapagesan, 2000, "Price Discovery in Auction Markets: A Look Inside the Black Box," *Review of Financial Studies*, 13, 627-658.

Madhavan, A., D. Porter, and D. Weaver, 2000, "Should Securities Markets be Transparent?" working paper, University of Southern California.

Madhavan, A., and S. Smidt, 1993, "An Analysis of Changes in Specialist Inventories and Quotations," *Journal of Finance*, 48, 1595-1628.

Madhavan, A., and G. Sofianos, 1998, "An Empirical Analysis of NYSE Specialist Trading," *Journal of Financial Economics*, 48, 189-210.

Manaster, S., and S. Mann, 1996, "Life in the Pits: Competitive Market Making and Inventory Control," *Review of Financial Studies*, 9, 953-975.

Meese, R. and K. Rogoff, 1983, "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?" *Journal of International Economics*, 14, 3-25.

Milgrom, P., and R. Weber, 1982, "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089-1122.

Naik, N., A. Neuberger, and S. Viswanathan, 1999, "Trade Disclosure Regulation in Markets with Negotiated Trades," *Review of Financial Studies*, 12, 873-900.

Naik, N., and P. Yadav, 1999, "The Effects of Market Reform on Trading Costs of Public Investors: Evidence from the London Stock Exchange," working paper.

Neal, R., 1987, "Potential Competition and Actual Competition in Equity Options," *Journal of Finance*, 42, 511–531.

Neal, R., and S. Wheatley, 1998, "Adverse Selection and Bid-Ask Spreads: Evidence from Closed-End Funds," *Journal of Financial Markets*, 1, 121-149.

Noldeke, G., and T. Troeger, 1998, "On Linear Equilibria in the Kyle Model," unpublished working paper.

Pagano, M., 1989, "Trading Volume and Asset Liquidity," *Quarterly Journal of Economics*, 104, 255-276.

Pagano, M., and A. Röell, 1996, "Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading," *Journal of Finance*, 51, 579-611.

Parlour, C., 1998, "Price Dynamics in Limit Order Markets," *Review of Economic Studies*, 11, 789-816.

Petersen, M., and D. Fialkowski, 1994, "Posted Versus Effective Spreads," *Journal of Financial Economics*, 35, 269-292.

Plott, C., and S. Sunder, 1982, "Efficiency of Experimental Security Markets with Insider Information: An Application of Rational-Expectations Models," *Journal of Political Economy*, 90, 663-698.

Plott, C., and S. Sunder, 1988, "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56, 1085-1118.

Pouget, S., 2001, "The Walrasian Tâtonnement to Economize on Cognitive Transactions Costs," working paper, Toulouse University.

Ready, M., 1999, "The Specialist's Discretion: Stopped Orders and Price Improvement," *Review of Financial Studies*, 12, 1075-1112.

Reiss, P., and I. Werner, 1998, "Does Risk-Sharing Motivate Interdealer Trading?," *Journal of Finance*, 53, 1657-1703.

Roëll, A., 1999, "Liquidity in Limit Order Book Markets and Single Price Auctions with Imperfect Competition," working paper, Princeton University.

Rochet, J.C., and J. Vila, 1994, "Insider Trading without Normality," *Review of Economic Studies*, 61, 131-152.

Rock, K., 1990, "The Specialist's Order Book and Price Anomalies," working paper, Harvard University.

Roll, R., 1984, "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market," *Journal of Finance*, 39, 1127-1139.

Ronen, T., 1992, "Tests and Properties of Variance Ratios in Microstructure Studies," working paper.

Rothschild, M., and J. Stiglitz, 1976, "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics, 90, 629–649.*

Sandas, P., 2001, "Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market," *Review of Financial Studies*, 14, 705-734.

Schnitzlein, C., 1996, "Call and Continuous Trading Mechanisms Under Asymmetric Information: An Experimental Investigation," *Journal of Finance*, 51, 613-636.

Seppi, D., 1992, "Block Trading and Information Revelation Around Quarterly Earnings Announcements," *Review of Financial Studies*, 5, 281-305.

Seppi, D., 1997, "Liquidity Provision with Limit Orders and a Strategic Specialist," *Review of Financial Studies*, 10, 103-150.

Seppi, D., 1990, "Equilibrium Block Trading and Asymmetric Information," *Journal of Finance*, 45, 73-94.

Simaan, Y., D. Weaver, and D. Whitcomb, 1999, "The Quotation Behavior of ECNs and NASDAQ Market Makers," WFA.

Smirlock, M., and L. Starks, 1986, "Day-of-the-Week and Intraday Effects in Stock Returns," *Journal of Finance*, 17, 197-210.

Sofianos, G., 1995, "Specialist Gross Trading Revenues at the New York Stock Exchange," working paper, New York Stock Exchange.

Spatt, C., and S. Srivastava, 1994, "Notes on Private Information and the Organization of Securities Markets," in Robert A. Schwartz (ed.), *Global Equity Markets: Technological, Competitive, and Regulatory Challenges*, Irwin Professional Publishing, Chicago, IL.

Spatt, C., and S. Srivastava, 1991, "Preplay Communication, Participation Restrictions, and Efficiency in Initial Public Offerings," *Review of Financial Studies*, 4, 709-726.

Spiegel, M., and A. Subrahmanyam, 1992, "Informed Speculation and Hedging in a Noncompetitive Securities Market," *Review of Financial Studies*, 5, 307-329.

Stoll, H., 1978, "The Supply of Dealer Services in Securities Markets," *Journal of Finance*, 33, 1133-1151.

Stoll, H., and R. Whaley, 1990, "Stock Market Structure and Volatility," *Review of Financial Studies*, 3, 37-71.

Stoll, H., 2000, "Friction", *Journal of Finance*, 1479–1518.

Subrahmanyam, A., 1991, "Risk Aversion, Market Liquidity, and Price Efficiency," *Review of Financial Studies*, 4, 417-441.

Vayanos, D., 1999, "Strategic Trading and Welfare in a Dynamic Market," *Review of Economic Studies*, 66, 219-254.

Venkataram,, K., 2001, Automated versus floor trading: an analysis of execution costs on the Paris and New York Stock Exchanges, *Journal of Finance*, 145-1485.

Viswanathan, S., and J. Wang, 1999, "Why is Interdealer Trading so Pervasive in Financial Markets," working paper, Duke University.

Wahal, S., 1997, "Entry, exit,, market makers, and the bid–ask spread", *Review of Financial Studies*, Vol 10, No 3, 871–901.

Walker, D., 1987, "Walras's Theories of Tatonnement," *Journal of Political Economy*, 95, 758-775.

Walras, L., *Eléments D'économie Politique Pure, ou Théorie de la Richesse Sociale*. Lausanne: Corbaz, 1874. 2d rev. ed. Lausanne: Rouge, 1889.

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

TABLE 1: COMPETITIVE MARKET MAKERS AND THE COST OF TRADES:

	Theoretical implications	Empirical results
Inventory paradigm	As market makers buy (sell) and their inventory increases (decreases), they seek to sell (buy) back. Consequently they lower (raise) their quotes, thus controlling the order flow they attract, and bringing their inventory back to their preferred position. [Stoll, 1978, Amihud and Mendelson, 1980, Ho and Stoll, 1981 and 1983,]	Market makers with long (resp. short) positions tend to sell (resp. buy) back (Hasbrouck and Sofianos (1993), Madhavan and Smidt (1993), Manaster and Mann (1996), Reiss and Werner (1998), Hansch et al (1998)). Empirical results on the impact of inventory positions on prices are less clearcut: On the NYSE, increases (resp. decreases) in the inventory of the specialist lead to decreases (resp. increases) in his quotes (Madhavan and Smidt, 1993), and after price rises (resp. decreases), where he is likely to have sold (resp. bought), the specialist is more likely to be on the bid side of the book (Kavajecz, 1999). But, on the CME, floor brokers tend to sell at high prices and buy at low prices (Manaster and Mann, 1996).
Adverse selection paradigm	Trading with privately informed investors leads to losses for market makers. They set spreads to generate revenues compensating for these losses. Correspondingly, spreads are larger when adverse selection is more severe. [Glosten and Milgrom, 1985, Kyle, 1985, Easley and O'Hara, 1987]	Trades have a permanent impact on prices (Hasbrouck (1991), Holthausen, Leftwich and Mayers (1990)), consistent with transactions reflecting private *or public* information (Neal and Wheatley, 1998). Block trades predict subsequent innovations in earnings announcements (Seppi, 1992). Market makers incur positioning losses on their inventory (Sofianos, 1995, Naik and Yadav, 1999). Spreads increase and depth decreases before earnings announcements (Lee et al 1993, Kajavecz 1999).

TABLE 2: STRATEGIC LIQUIDITY SUPPLY

Panel A: Analyses which do not rely on adverse selection.

Type of market	Theoretical analyses	Empirical results
Order driven markets	It is optimal to place limit orders within the spread when the latter is large, and to place market orders when the spread is tight (Foucault, 1999). Investors trade-off time prio rity and price when deciding where to place orders in the book (Parlour, 1998).	On the NYSE, for stocks with ¼ spread, limit orders within the quotes dominates outperform market orders (Harris and Hasbrouck, 1996). On the Paris Bourse, limit orders within the quotes are more frequent when the spread is large and when the depth at the quotes is large (Biais, Hillion and Spatt, 1995).
Dealer markets	When the number of suppliers is limited noncompetitive behavior is to be expected. This is reinforced by repeated interaction between dealers (Dutta and Madhavan, 1997), or when dealers have increasing marginal costs of supplying liquidity (Biais, Foucault & Salanié, 1998, Roëll, 1999, Viswnathan and Wang, 1999).	For comparable stocks, Huang and Stoll (1996) find larger spreads on NASDAQ than on the NYSE. NASDAQ quotes before 1994 were consistent with collusion between dealers (Christie and Schultz, 1994 and Christie, Harris and Schultz, 1994). After 1997, as limit order traders were authorized to compete with dealers to supply liquidity, spreads were significantly reduced on NASDAQ (Barclay et al, 1999) and on the London Stock Exchange (Naik and Yadav, 1999).

TABLE 2: STRATEGIC LIQUIDITY SUPPLY

Panel B: Analyses relying on adverse selection

Class of models	Theoretical analyses	Empirical results
Uninformed market makers	Because of adverse selection, the marginal cost of supplying liquidity is increasing (Glosten, 1994). Consequently, with a finite number of market makers, there are oligopolistic rents, increasing in the degree of adverse selection (Bernhardt and Hughson, 1997, Biais, Martimort and Rochet, 2000).	The reduction in spreads following the 1997 NASDAQ market reform was for more pronounced for small volume firms (Barclay et al, 1999), which are those for which adverse selection problems are likely to be the most severe. Sandas (2001) finds steeper limit order schedules than predicted by the Glosten (1994) competitive model.
Informed market makers	Informed market makers inject noise in their quotes to avoid immediate full revelation and preserve their informational edge, yet their quotes and trades reveal some information to the market (Calcagno and Lovo, 1999).	CME market makers' trades tend to be in the right direction (Manaster and Mann, 1996). .

TABLE 3: MARKET DESIGN.

Panel A: Call versus continuous market.

Issues	Theoretical analyses	Empirical analyses
Efficient pricing		NYSE opening call auction prices are noisier than closing prices, but the Tokyo Stock Exchange midday call auction price is not (Amihud and Mendelson, 1987, 1991). The preopening period conveys information in the Paris Bourse (Biais, Hillion and Spatt, 1999), and eriod enhances the effiency of the opening call in experimental markets (Biais and Pouget, 2001).
Depth and spread	When liquidity suppliers have convex costs, market impact is greater (resp. lower) for large (resp. small) quantities in discriminatory than in uniform price auctions (Roell, 1999, Viswanathan and Wang, 1999, Biais, Martimort and Rochet, 2000).	On the Frankfurt Stock Exchange, for small trades transactions costs are lower in the opening call, while for large trades market impact is lower in the continuous trading phase (Kehr, Krahnen and Theissen, 1998)
Risk Sharing	Strategic agents with private information about their risk sharing needs limit their trades to reduce market impact. This is stronger with continuous trading than when there is a larger time interval between trades (Vayanos, 1999)	

TABLE 3: MARKET DESIGN

Panel B: The specialist

Due to the presence of the specialist, adverse selection can be :	Theoretical analyses	Empirical analyses
worsened …	When an order arrives on the floor the specialist can choose to undercut the book, to stop the order or to let the it hit the book. This creates an adverse selection problem (Rock, 1990). To the extent that small trades have a lower informational content it can be advantageous for the specialist to step up and execute these (Seppi, 1997).	Petersen and Falkowski (1994) and Sofianos (1995) document order stopping. Consistent with the analysis of Rock (1990), Ready (1999) finds that stopped orders are more profitable for liquidity suppliers than orders allowed to trade with the book.
or mitigated.	The specialist, interacting repeatedly with brokers, can extract private information from them, thus reducing adverse selection (Benveniste et al, 1992).	

TABLE 3: MARKET DESIGN.

Panel C: Transparency

Form of transparency	Theoretical analysis	Empirical analysis	Experimental analysis
Ex ante	Better dealer information reduces adverse selection and hence spreads (Pagano and Roell, 1996).	An increase in ex-ante transparency on the Toronto Stock Exchange led to an increase in spreads (Madhavan, Porter and Weaver , 2000).	Pre-trade transparency narrows spreads (Flood et al., 1999).
Ex post	Trade disclosure enhances risk sharing (Naik et al., 1999).	In the London Stock Exchange, dealer spreads were not affected by changes in the trade disclosure regime (Gemmill, 1996).	Opening spreads are larger, but subsequent spreads tighter, when ex-post transparency is enhanced (Bloomfield and O'Hara, 1999, 2001).

TABLE 3: MARKET DESIGN.

Panel D: Tick size

	Theoretical arguments	Empirical results
Minimum tick size constrains market makers	Coarse price grid can constrain spreads to be excessively wide (Harris, 1994). Non-infinitesimal tick size can enhance the incentives of the dealers to be the first one to post the narrowest possible spread, and thus reduce average spread size (Cordella and Foucault, 1999).	Harris (1994) develops an econometric methodology to estimate the reduction in spreads that would be generated by a reduction in tick size.
Asymmetry in the timing of the moves	When market orders are transmitted to the NYSE, the floor traders and the specialist have the option to undercut the limit orders in the book. This creates a winner's curse problem (Rock, 1990), especially with thin ticks (Seppi, 1997). In electronic limit order markets, such as Xetra or the Paris Bourse, this problem should not arise.	There was a reduction in the cumulated depth of the NYSE book after the reduction of tick size from one eighth to one sixteenth (Goldstein and Kavajecz, 2000). No such reduction is observed when the price grid becomes thinner on the Paris Bourse (Demarchi, 1999, Declerck, 2001).