Chapter 10
Security Market Microstructure: The Analysis of a Non-frictionless Market

Reto Francioni, Sonali Hazarika, Martin Reck, and Robert A. Schwartz

10.1 Introduction

Security market microstructure addresses issues that involve the implementation of portfolio (investment) decisions in a marketplace. Implementation entails the placement and handling of orders in a securities market, and their translation into trades and transaction prices. The process links fundamental information concerning equity valuation (which is of primary concern to portfolio managers) to prices and transaction volumes that are realized in the marketplace. The quality of the link depends on the rules, procedures, and facilities of a securities market, and on the broader regulatory and competitive environment within which the market operates.

Widespread interest on the part of the securities industry, government, and academia is testimony to the importance of market microstructure analysis. The subject addresses issues that concern investors, broker/dealer intermediaries, market regulators, exchanges, and other trading venues as well as the broad economy. Interest in microstructure has increased sharply over the past three and a half decades, spurred in particular by three events: the US Securities and Exchange Commission's (SEC) Institutional Investor Report (1971), the passage by the US Congress of the Securities Acts Amendments of 1975, and the sharp stock market drop on October 19, 1987. Further, the advent of computer-driven trading in recent years has enabled researchers to capture electronically the full record of all trades and quotes, and this has provided empirical researchers with far richer data (referred to as "high-frequency data") for analyzing trading and price setting.

This chapter includes material from [1] and from Schwartz which was reprinted in [2].

R. Francioni (✉) • M. Reck
University of Basel/Deutsche Börse Group, Basel, Switzerland
e-mail: reto.francioni@unibas.ch

S. Hazarika • R.A. Schwartz
Zicklin School of Business, Baruch College, CUNY, New York, NY, USA

R. Francioni, R.A. Schwartz (eds.), *Equity Markets in Transition*,
DOI 10.1007/978-3-319-45848-9_10

Over the years, microstructure analysis has expanded and, concomitantly, exchange structure has strengthened. We consider both of these developments in this chapter. First, we set forth the major challenges that the microstructure literature addresses. Second, we consider the properties of a frictionless trading environment. Third, we present a broad view of the direction in which microstructure analysis has been and is evolving. Fourth, we turn to one application—the design of an actual marketplace: Deutsche Börse's electronic trading system, Xetra. The German market was the last of the major European bourses to introduce an electronic trading platform, and it is state of the arts, which makes Deutsche Börse a particularly interesting case in point. Fifth, in the concluding section, we consider the bumpy and hazardous road that takes us from theory to the development of an actual marketplace.

10.2 Microstructure's Challenge

Microstructure analysis has four broad applications. First (and this is a key focus of the chapter), it gives guidance to market structure development. The link with market structure is straightforward: the critical factor that drives microstructure analysis is friction in the marketplace (i.e., the explicit and implicit costs of implementing portfolio decisions), and trading costs depend on the architecture of the marketplace which determines how orders are handled and turned into trades. The flipside of friction is illiquidity, and a primary function of a market center is to amass liquidity.

Microstructure's second application is to facilitate the development of trading strategies and algorithms for asset managers and broker/dealer intermediaries. The importance of this application is evident in the current development of computer-driven algorithmic trading. Algorithms can be fine-tuned to take account of, for example, the probability of a limit order executing, time-of-day effects such as market openings and closings, search for liquidity in a fragmented environment, and choice of a trading modality (e.g., a continuous limit order book market, a quote-driven dealer market, a periodic call auction, a block trading facility, or hybrid combinations of the above).

The third application of microstructure analysis concerns tests of market efficiency. In the 1970s, at a time when the subject was first emerging, the efficient market hypothesis (EMH) was widely accepted by financial economists as a cornerstone of modern portfolio theory, and it continues to receive broad academic support today. The hypothesis addresses informational as distinct from operational efficiency (the latter refers to the containment of transaction costs by superior market design). According to the EMH, a market is informationally efficient if no participant is able to achieve excess risk-adjusted returns by trading on currently available information. Many of the EMH tests have considered one major part of the information set—market information (e.g., recent quotes, trading volume, and transaction prices). If prices properly reflect all known information, then (in a frictionless market at least) they must change randomly over time, hence the term "random walk." Earlier studies, based on daily data, generally supported the random walk hypothesis. However, with the advent of high-frequency data, the footprints of complex correlation patterns have been detected. This observation, along with superior knowledge of the impact of trading costs on returns behavior, is casting a new light on market efficiency. Whether inefficiency is thought of in operational or informational terms, the EMH is not as stellar as it once was.

In its fourth application, microstructure analysis sheds light on how new information is incorporated into security prices. In a zero-cost, frictionless environment, share values would be continuously and instantaneously updated with the release of new information. In actual markets, however, information must be received and assessed, traders' orders must be placed and processed, and executions must be delivered and accounts cleared and settled. Costs, both explicit (e.g., commissions) and implicit (e.g., market impact), are incurred throughout this chain of events. Highlighted in much microstructure literature are the costs that some participants incur when, in an asymmetric information environment, other participants receive information first and trade on it to the disadvantage of the uninformed.

Asymmetric information is not the only reality, however. In light of the size, complexity, and imprecision of much publicly available information, one might expect that investors in possession of the same (large) information set will form different expectations about future risk and return configurations. This situation is referred to as "divergent expectations."[1] Asymmetric information and divergent expectations together reflect a rich set of forces that impact the dynamic behavior of security prices.

This overview of microstructure's four broad applications underscores that trading frictions are the subject's raison d'être. Participant orders cannot be translated into trades at zero cost (markets are not perfectly liquid), and trades typically are not made at market clearing (i.e., equilibrium) prices. Trading decision rules (algorithms) are needed because the costs of implementing portfolio decisions can sharply lower portfolio performance. In fact, much algorithmic trading is designed to control trading costs, rather than to exploit profitable trading opportunities. Today, trading is recognized as an activity that is both distinct from investing and equivalently professional. Market structure is of concern to the buy-side desks precisely because markets are not perfectly liquid, and neither are they perfectly efficient, either informationally or operationally. Consequently, better market structure can deliver superior portfolio performance for participants.

What is the economic service, one might ask, that an equities market provides? The fuzzy link that connects information and prices in the non-frictionless environment underscores two major market functions—price discovery and quantity discovery. Price discovery refers to participants collectively searching for equilibrium prices. Quantity discovery refers to the difficulty that participants who would be willing to trade with each other actually have finding each other and trading when markets are fragmented. This difficulty is accentuated because some participants (primarily institutional investors) do not immediately reveal the total size of their orders (doing so would unduly drive up their market impact costs).

[1] For a recent discussion, see Davis et al. [72].

Market structure affects both the accuracy of price discovery and the completeness of quantity discovery. The link between market structure and price discovery depends on the environment within which participants are operating. At one end of the spectrum, investors can be equally informed and form homogeneous expectations based on the information. At the other end, they can be differentially informed and form divergent expectations with regard to commonly shared information. When investors who share common information all agree on share values (i.e., have homogeneous expectations), prices can be "discovered" in the upstairs offices of research analysts. When investors are not equally informed, and when they form different expectations based on common information, prices must be discovered in the marketplace. In this second environment, the economic service provided by an exchange is clear—it "produces the price."

Regarding quantity discovery, handling the orders of large institutional customers is a challenge. It is not at all uncommon for an institution to want to buy or to sell, for instance, 500,000 shares of a company that has an average daily trading volume of 300,000 shares. Executing an order of this size can easily drive prices away from the trader before the job has been completed. The adverse price move is a *market impact cost*. Institutions attempt to control their market impact costs by trading patiently and, as much as possible, invisibly. Good market structure can help. To this end, a number of alternative trading systems (ATSs) have been formed in recent years, and dark (i.e., non-transparent) liquidity pools have emerged.

With prices discovered in the marketplace, participants employ trading strategies when they come to the market to implement their portfolio decisions. Participants with differential information that will soon become public determine how best to meter their orders into the market so as to move prices to new levels with *minimal* speed. Additional questions that any trader might ask include the following: "If I trade now, at the current moment, how will the price that I will receive compare with the average price that shares are trading at today?" "Is price currently at a sustainable, validated level, or is it likely to move higher or lower in the coming hours, minutes, or even seconds?" "Would I do better to be patient and place a limit order, or submit a market order and get the job done right away?" "Should I attempt to trade now in the continuous market, or wait for a closing call?" The orders that a set of participants reveal to the market depend on how questions such as these are answered, and prices that are set and trading volumes that are realized depend on the orders that are revealed.

The categories of trading costs that receive the most attention on the part of exchanges, regulators, and academicians are generally those that are the most straightforward to measure: commissions and bid-ask spreads. Increasingly precise measures of market impact are also becoming available, and this cost too is being widely taken into account. On the other hand, the opportunity cost of a missed trade, being far more difficult to quantify, is often overlooked. Also more challenging is quantifying a cost that has received little formal attention: realizing executions at poorly discovered prices. The problem, of course, is that equilibrium values are not observable and appropriate benchmark values are not easily defined.

10.3 The Perfectly Liquid Environment of CAPM

Peter Bernstein's [3] piece in the *Journal of Portfolio Management* has the intriguing title, "The Surprising Bond Between CAPM and the Meaning of Liquidity." In it he wrote, "The more liquid an asset, the greater the dominance of systematic risk over stock specific risk." We build on this insight in this section. In so doing, we formalize the fact that the capital asset pricing model describes an extreme case, a totally frictionless world where liquidity is infinite and systematic risk has complete dominance over stock-specific risk. The analysis provides a good platform from which to launch a discussion of market microstructure, the study of a non-frictionless environment.

CAPM models the price of the individual equity shares that, in aggregate, comprise the market portfolio. Following standard methodology, we start our analysis of the frictionless environment by taking the market portfolio to be one single asset (e.g., an all-encompassing exchange traded fund). We consider the demand of an agent to hold shares of this one risky asset when the only alternative is the riskless asset. We show that an individual agent's demand curve to hold shares of the risky asset is downward sloping, and then use this curve to re-derive certain key CAPM equations to show that the associated demand to hold shares of each individual equity issue in that portfolio is infinitely elastic, and that therefore the market for the individual shares is infinitely liquid.

In the CAPM world, each individual equity issue in the market portfolio has an intrinsic value that is given by the parameter that locates the height (on the price vector) of that infinitely elastic demand. In the section that follows, we turn to the non-frictionless environment of microstructure analysis where individual stock demand curves are downward sloping, the liquidity of individual shares is, therefore, finite, and individual shares do not have intrinsic values.

To obtain the representative investor's demand curve to hold shares of the risky market portfolio, first we state the agent's utility (of wealth) function. The demand curve to hold shares of the market portfolio may then be obtained directly from the utility function. The derivation follows Ho et al. [4].

We make the following assumptions:

- The investor's portfolio comprises a risk-free asset and one risky asset (shares of the market portfolio).
- Share price and share holdings are continuous variables.
- Short selling is unrestricted.
- The existence of a brief trading period, T_0 to T_1, which is followed by a single investment period, T_1 to T_2.
- All transactions made during the trading period are settled at point in time T_1.
- The investor seeks a portfolio at the beginning of the investment period (at time T_1) that will maximize the expected utility of wealth to be realized at the end of the investment period (at time T_2).

- Investor expectations with respect to the share price at the end of the investment period (at time T_2) are exogenously determined (expectations are independent of the current price of shares).
- Investors are risk averse.

The following variables are used:

C_0 = holdings of the risk-free asset at the beginning of the trading period (T_0).
C_1 = holdings of the risk-free asset at the beginning of the investment period (T_1).
N_0 = number of shares of the market portfolio held at the beginning of the trading period (T_0).
N_1 = number of shares of the market portfolio held at the beginning of the investment period (T_1).
$R_0 - 1$ = risk-free rate of interest over the trading period.
$R_1 - 1$ = risk-free rate of interest over the investment period.
P_1 = price at which shares of the market portfolio are purchased or sold during the trading period.
P_2 = price at which shares of the market portfolio can be sold at the end of the investment period (T_2).
$r_m = P_2/P_1 - 1$ = return on the market portfolio.
Q = number of shares traded by the investor at the beginning of the investment period (T_1); $Q > 0$ indicates a purchase; $Q < 0$ indicates a sale.

10.3.1 The Expected Utility of End-of-Period Wealth

The participant starts the *investment period* with C_1 dollars of the risk-free asset and N_1 shares of the market portfolio (the risky asset). Therefore, wealth at T_2 is given by $C_1R_1 + N_1P_2$. As of T_1, this wealth is uncertain because P_2 is uncertain. As of T_1, the expected utility of end of period wealth can be written as

$$EU\left(C_1R_1 + N_1P_2\right) \tag{10.1}$$

The investor starts the *trading period* with C_0 dollars of the risk-free asset and N_0 shares of the risky asset. If during the trading period the decision maker were to exchange holdings of the risk-free asset for Q shares of the risky asset at a price of P_1, the *expected utility of end-of-period wealth*, written as a function of P and Q, given N_0 and C_0, would be

$$h\left(P_1, Q \mid N_0, C_0\right) = EU\left[\left(C_0R_0 - QP_1\right)R_1 + \left(N_0 + Q\right)P_2\right] \tag{10.2}$$

where $C_0R_0 - QP_1 = C_1$ and $N_0 + Q = N_1$. Equation (10.2) can be rewritten as

$$h\left(P_1, Q \mid N_0, C_0\right) = c + gQ\left(a - bQ - P_1\right)$$

where

$$\begin{aligned}
c &= U(W) - \pi N_0^2 U'(W) / R_1 \\
g &= U'(W) R_1 \\
a &= \left[E(P_2) - 2\pi N_0\right] / R_1 \\
b &= \pi / R_1 \\
\pi &= -\frac{1}{2}\left[U''(W) / U'(W)\right] \mathrm{Var}(P)
\end{aligned} \tag{10.3}$$

The step from (10.2) to (10.3) involves a Taylor expansion of the investor's utility around the expected value of wealth if the investor does not trade.[2] The procedure is a convenient way of introducing the variance term into the utility function.[3]

10.3.2 The Reservation Demand Curve

Equation (10.3) can be further assessed with the use of risk aversion and risk premium measures that are defined in Appendix 1. Specifically, using (10.3), we now obtain both a *reservation price demand curve* and an *ordinary demand curve* to hold shares of the risky asset. We consider the reservation demand curve first.

The reservation price for a purchase or a sale is the maximum price the decision maker would be willing to pay to buy a given number of shares ($Q>0$), or the minimum price the decision maker would be willing to receive to sell a given number of shares ($Q<0$) when the only alternative is not to trade at all. Equation (10.3) shows that, if no trade is made (that is, if $Q=0$), the decision maker's expected utility is equal to c. The reservation price for any value of Q is the price that equates the expected utility [$h(P_1, Q|N_0, C_0)$] if the trade is made, with the expected utility (c) if no trade is made. Thus the reservation price for any value of Q is given by

$$h\left(P^R, Q \mid N_0, C_0\right) = c \tag{10.4}$$

where P^R is the reservation price associated with the trade of Q shares. Given (10.3), for (10.4) to be satisfied, we must have $a - bQ - P_1 = 0$. Hence the reservation price demand curve is

$$P^R = a - bQ \tag{10.5}$$

[2] For a discussion of the Taylor procedure see, for example, R. G. D. Allen, *Mathematical Analysis of Economists,* London, England: Macmillan, 1960.

[3] Two further assumptions are required to obtain (10.3): (1) the third derivative of utility with respect to wealth is small enough to ignore; and (2) the squared deviation of the expected rate of return on the risky asset from the risk-free rate is small enough to ignore.

10.3.3 The Ordinary Demand Curve

Using (10.3), we can also obtain the ordinary demand curve. At any value of P_1, the decision maker selects the value of Q that maximizes expected utility. Hence, the ordinary price demand curve is given by

$$\frac{\partial h}{\partial Q}\left(P^0, Q \mid N_0, C_0\right) = 0 \tag{10.6}$$

where P^0 is the "ordinary" price associated with the trade of Q shares. Differentiating h in (10.3) with respect to Q, setting the derivative equal to zero, and rearranging gives

$$P^0 = a - 2bQ \tag{10.7}$$

10.3.4 The Risk Premium and the Market Price of Risk

When the investor has traded the optimal number of shares of the market portfolio at the market determined price per share, his or her risk premium can be related to the market price of risk. Assessing, the ordinary demand curve at $P^0 = P_1$ gives

$$P_1 = \frac{E\left(P_2\right)}{R_1} - \frac{2\pi N_1}{R_1} \tag{10.8}$$

Multiplying by R_1/P_1, rearranging, and recognizing that $[E(P_2)/P_1] - 1 = E(r_m)$ and $R_1 - 1 = r_f$, we get

$$\frac{2\pi N_1}{P_1} = E\left(r_m\right) - r_f \tag{10.9}$$

Therefore, we have

$$\pi_{M\%} = E\left(r_m\right) - r_f \tag{10.10}$$

where $\pi_{M\%}$ is the marginal risk premium (see Appendix 1). Note that the right-hand side is the price of risk. We thus see that the investor achieves an optimal holding of the risky asset by obtaining the number of shares that equates the marginal risk premium with the market price of risk.

10.3.5 The Investor's Optimal Point on the Capital Market Line

The demand model can be used to assess the investor's optimal point on the capital market line. Let r_p be the return on the combined portfolio (N_1 shares of the market portfolio and C_1 dollars of the risk-free asset). From Appendix equation (10.28) we have

$$\pi = \pi_{rp}\left(\frac{W}{N_1}\right)^2$$

which, using $R_A = -U''(W)/U'(W)$, the measure of absolute risk aversion, can be written as

$$\pi = \frac{1}{2} R_A Var\left(r_p\right)\left(\frac{W}{N_1}\right)^2 \tag{10.11}$$

Because $\sigma_p = (NP/W)\sigma_m$, we have $\mathrm{Var}(r_p) = \sigma_p(NP/W)\sigma_m$ and can write (10.11) as

$$\pi = \frac{1}{2} R_A \sigma_p \left(\frac{PW}{N_1}\right)\sigma_m \tag{10.12}$$

Substituting (10.12) into (10.9) and simplifying give

$$R_R \sigma_P = \frac{E\left(r_m\right) - r_f}{\sigma_m} \tag{10.13}$$

where $R_R (= WR_A)$ is the measure of relative risk aversion.

Equation (10.13) shows that for the investor to hold an optimal combined portfolio, the market price of risk per standard deviation of the market portfolio must be equal to the investor's coefficient of relative risk aversion times the standard deviation of the combined portfolio's return.

Letting $w = N_1 P_1 / W$, substituting $w\sigma_m, = \sigma_p$ into (10.13), and rearranging give

$$w = \frac{E\left(r_m\right) - r_f}{Var\left(r_m\right) R_R} \tag{10.14}$$

Equation (10.14) shows that the percentage of wealth that the risk-averse participant invests in the market portfolio is positively related to the expected return $E(r_m)$, and negatively related to r_f, $\mathrm{Var}(r_m)$, and R_R. Investors all face the same values of $E(r_m)$, r_f, and $\mathrm{Var}(r_m)$, but differ according to their degree of risk aversion. More risk-averse investors (larger R_R) have smaller optimal values of w and hence are more apt to lend at the risk-free rate (which implies $w<1$); less risk-averse investors (smaller R_R) have larger optimal values of w and hence are more likely to borrow at the risk-free rate (which implies $w>1$).

The right-hand side of (10.13) is the market price of risk per standard deviation of the market portfolio. The total compensation for risk taking is the price of risk times the standard deviation that the investor accepts (here, the standard deviation of the combined portfolio). Multiplying both sides of (10.13) by σ_p, we obtain

$$R_R\,Var\left(r_p\right)=\left[\frac{E\left(r_m\right)-r_f}{\sigma_m}\right]\sigma_p \tag{10.15}$$

Adding r_f to both sides of (10.15) gives the investor's total compensation for waiting and for risk taking:

$$E\left(r_p\right)=r_f+R_R\,Var\left(r_p\right)=r_f+\left[\frac{E\left(r_m\right)-r_f}{\sigma_m}\right]\sigma_p \tag{10.16}$$

Equation (10.16) shows that the location of the investor's optimal point on CAPM's capital market line depends on his or her measure of relative risk aversion (R_R).

10.3.6 The ith Risky Asset's Point on the Security Market Line

We now assess the demand model to show the location of an *i*th risky asset on the security market line. In so doing, we establish that the demand for the *i*th risky asset is infinitely elastic. Equation (10.10) shows that the marginal risk premium for each investor, as a percentage of P_1, will equal $E(r_m)-r_f$. Therefore, for each investor,

$$\frac{R_A\,Var\left(P_2\right)N_1}{P_1}=E\left(r_m\right)-r_f \tag{10.17}$$

It follows from Equation (10.17) that investors with lower values of R_A hold a larger number of shares, such that the product R_AN_1 is the same for all investors. Because $r_m=(P_2/P_1)-1$, $Var\left(r_m\right)=Var\left(P_2\right)/P_1^2$. Substituting $Var\left(r_m\right)P_1^2=Var\left(P_2\right)$ into (10.17) and simplifying give

$$R_A\,Var\left(r_m\right)P_1N_1=E\left(r_m\right)-r_f \tag{10.18}$$

Using $P_1N_1=wW$ we obtain

$$wR_R\,Var\left(r_m\right)=E\left(r_m\right)-r_f \tag{10.19}$$

Equation (10.19) can be interpreted as an equilibrium condition for each investor. Because $wR_R = R_A N_1 P_1$, and given that the product $R_A N_1$ is constant across investors, $R_R w$ is constant across all investors. [It is also clear from (10.19) that the product wR_R must be constant across all investors, because $E(r_m)$, r_f, and $\mathrm{Var}(r_m)$ are the same for all.]

The equilibrium condition for each investor with respect to the market portfolio implies an equilibrium condition for each investor with respect to any ith risky asset in the market portfolio. The CAPM shows that the relevant measure of risk for the ith risky asset is $\beta_i = \sigma_{im}/\mathrm{Var}(r_m)$. Therefore, writing $\mathrm{Var}(r_m) = \sigma_{im}/\beta_i$, substituting into (10.19), and multiplying both sides by β_i we get

$$wR_R\sigma_m = \beta_i\left[E(r_m) - r_f\right] \tag{10.20}$$

Adding r_f to both sides of (10.20) gives CAPM's security market line,

$$r_f + wR_R\sigma_{im} = r_f + \beta_i\left[E(r_m) - r_f\right] = E(r_i) \tag{10.21}$$

where $E(r_i)$ is the expected return on the ith stock in the market portfolio. Equation (10.21), assessed at $w = 1$, shows that the expected return for the ith risky asset depends on its covariance with the market return, and on the measure of relative risk aversion for an investor whose optimal combined portfolio contains the market portfolio only. The equation also shows that the ith risky asset's specific location on the security market line depends on the covariance of the asset's return with the return on the market portfolio, and hence that its expected return depends only on β_i, its systematic risk.

It follows from the above discussion that the demand to hold shares of the market portfolio is downward sloping, while the demand for each individual stock in the market portfolio is infinitely elastic. The reason is that perfect substitutes do not exist for the aggregate portfolio, but they do exist for the individual stocks. Only one factor characterizes any ith stock—β_i, its covariance with the market. But the covariance for any stock can be duplicated exactly by an appropriate combination of two or more other stocks, and all holdings that have the same covariance must yield the same expected return. If they were to yield different expected returns, an unlimited number of shares of the higher yielding position would be bought, and an unlimited number of shares of the lower yielding position would be sold short until, with costless trading, the buying and selling pressures bring the two prices into exact equality. Unlimited buying (selling) at any price lower (higher) than the beta appropriate, CAPM price manifests an infinitely elastic demand to hold shares. That is, at an infinitesimally higher price no shares will be held, and at an infinitesimally lower price demand will be unlimited.

Bernstein's [3] two insights immediately follow: a stock's systematic risk totally dominates its specific risk, and the market for each ith stock is infinitely liquid at the price which translates into $E(r_i)$, its systematic risk-appropriate return.

As we turn to the non-frictionless market, the infinitely liquid, infinitely elastic property of CAPM is a good point of departure from the frictionless world. A common denominator in many microstructure analyses is that the demand to hold shares of individual stocks is downward sloping (which means that shares do not have intrinsic values). Market makers post bid and ask quotes which, when raised, result in more public sales to the market maker and which, when lowered, result in more public purchases from the market maker. Bid and ask quotes can be distributed over multiple price points in competitive dealer markets as well as on public limit order books. Trading is not costless. Both explicit costs (e.g., commissions and taxes) and implicit costs (e.g., market impact costs) are incurred. Information is complex and imprecise, and thus investors commonly disagree about its interpretation. Arbitrage is not costless, and perfect substitutes for individual issues do not exist. Share values depend not only on the calculations of systematic risk in the upstairs markets, but also on how orders interact in the marketplace. As a consequence of all of this, trades that are made and the transaction prices that they are made at also depend on the structure of a marketplace. Microstructure analyses address these realities that CAPM does not comprehend.

10.4 What Microstructure Analysis Has to Offer: Personal Reflections

In this section we review the development of microstructure analysis. Our objective is not to provide a comprehensive survey of the literature, but to highlight some of the important themes that have given guidance to market structure development. More detailed information can be obtained from Cohen et al. [5] who have provided an early survey of the field; from O'Hara [6] who discusses important theoretical microstructure models; from Madhavan and Ananth. Market Microstructure. Journal of Financial Markets [7]; Biais et al. [8]; and Parlour and Seppi [9] who have provided more recent surveys; and from Hasbrouck, Joel. Empirical Market Microstructure. Oxford University Press [10] who deals with empirical microstructure research and research methodology. We first focus on the early literature, next turn to more recent developments, and lastly present our thoughts concerning an important direction in which future microstructure research ought to head.

10.4.1 The Early Focus

The first contributions to the new field in financial economics that came to be called "microstructure" were made by a couple of people who participated in the SEC's Institutional Investor Report (1971). A handful of others independently started to focus on microstructure topics in the early 1970s. Eventually a few of the early researchers came to recognize the commonality of their interests and, applying the title of Garman's [11] well-known paper, "Market Microstructure," they gave the field its name.

Much of the early literature focused on dealers and exchange specialists. These market makers were viewed as the suppliers of immediacy to investors, and the spread was considered the price they charge for providing this service in an environment where order arrival is nonsynchronous. Of key importance was the relationship between spreads and the costs of market making.

The earlier market maker studies were in large part motivated by a desire to determine whether or not these intermediaries were realizing monopoly profits and, if so, whether or not their profits were attributable to market making being a natural monopoly. Spreads that are greater than the costs of market making would be taken as an indication of monopoly power on the part of the dealers, and spreads that were negatively related to trading volumes would indicate economies of scale in market making, which could imply a natural monopoly [12]. Spreads were indeed found to decrease with transaction volume, but reasons other than market making being a natural monopoly were advanced [13, 14].

The general picture which emerged was that the trading costs incurred by investors could be lowered by strengthening competition between market maker intermediaries. In particular, competition in the NYSE market was deemed inadequate, as specialists and the exchange itself were viewed as having monopoly positions: each stock was assigned to just one specialist; the NYSE's order consolidation rule (Rule 390) precluded in-house executions by requiring that exchange members send their orders for NYSE-listed securities to an exchange; and commissions were fixed and unjustifiably high [15].[4]

Not surprisingly, the focus on the market maker firms led several researchers to model market maker pricing decisions (i.e., the setting of their bid and ask quotes). These included Bagehot et al. [16], Stoll et al. [17], Amihud et al. [18], Ho et al. [19–21], and Mildenstein et al. [22]. With one exception [16], the early formulations dealt with inventory considerations. A market maker firm holding an undesirably long position would lower the quotes (i.e., lower the offer so as to sell more shares, and reduce the bid so as to discourage others from selling shares to it). Reciprocally, a market maker who was short would raise the quotes. This response on the part of the public (buy more shares when the market maker's offer is lower, and sell more share when the market maker's bid is higher) is evidence that the public's demand to hold shares of any specific stock was taken to be downward sloping. A variety of mathematical tools were used to solve for optimal market maker quotes. These models also gave further insight into the cost components of the market maker's spread [23].

While insightful, the early inventory-based pricing models suffered from some shortcomings. First, the early formulations for the most part assumed monopoly market makers, even though some of these models were applied to markets such

[4] Another major issue addressed by the microstructure literature at that time was the impact of information on trading volume and price ([74–76]).

as the New York Stock Exchange where exchange specialists were in fact competing with other floor brokers and customer limit orders [24]. The application of theory further suffered from the reality that the price of immediacy for an investor is not the spread of an individual market maker, or even an average market maker spread, but the inside spread (i.e., the lowest ask across all market makers minus the highest bid).[5] It is important to note that dealer spreads could individually remain relatively invariant with respect to transaction volume while the inside spread fell appreciably.

A further shortcoming of most of these earlier models is that they did not take account of a major cost incurred by market makers: the losses generated by trading with better informed investors. Recognition of this reality (which is also outside the scope of the frictionless world of CAPM) led to a development that did much to establish microstructure as an important new field in financial economics—the introduction of market maker models that were based, not on inventory management, but on controlling the cost incurred when some investors are in possession of information that the market maker and other investors have not yet received. Bagehot et al. [16] was the first to embark on this line of thought. He was later followed by, among others, Glosten et al. [25] and Kyle, Albert. "Continuous Auctions and Insider Trading." Econometrica, 53 [26].

With information asymmetries, the market maker always looses when trading with a better informed participant. For microstructure theorists at the time, this meant that, for the dealer market not to fail, some investors must trade for reasons that are not related to information.[6] Liquidity considerations (i.e., an investor's personal cash flow needs) were one such motive for public buying and selling. A third participant type was also introduced along with the liquidity traders—noise traders (participants who trade on price moves as if they contain information when in fact they do not). This trio of informed traders, liquidity traders, and noise traders was used to show how markets could function and, in so doing, enable new information to be incorporated into security prices (Grossman et al. [27], Milgrom, Paul and Nancy Stokey. "Information et al. [28]," Kyle, Albert. "Continuous Auctions and Insider Trading." Econometrica, 53 [26], Glosten et al. [25], Copeland et al. [29], and Easley and Maureen O'Hara. "Order Form and Information in Securities Markets." Journal of Finance 46 [30–32]).

At this stage in its early development, the microstructure pricing models were predominantly market maker models. One exception should be noted, however: a National Book System proposed by Mendelson et al. [33] contained a comprehensive description of an order-driven automated trading system that provided guidance for designing the first exchange-based electronic trading systems. For a more recent discussion of automated trading systems, see Domowitz, Ian and Benn Steil. "Automation et al. [34]." Most equity markets around the globe are now order-driven, limit order

[5] For further discussion, see Cohen et al. [5].

[6] A market supported by informational trading only can indeed function if agents trade with each other because their expectations are divergent. When the information that triggers trading is common knowledge, the condition may be thought of as one where agents are agreeing to disagree.

book markets that might include market makers in a hybrid structure (as does the NYSE), but are not basically quote-driven (i.e., dealer) markets (as was the old Nasdaq and London Stock Exchange). The limit order book markets are driven by the orders placed by the investors themselves, not by market maker intermediaries.

10.4.2 The Current Focus

Over the years, microstructure analysis has grown extensively on both the theoretical and empirical fronts. Concomitantly, the securities markets themselves have evolved, becoming evermore technologically developed, more global in outreach, but also more fragmented between different trading facilities. One important new direction microstructure research has taken is to further model the order-driven market, an environment where natural buyers and sellers provide immediacy to each other because some, who are patient, are willing to post limit orders while others, who demand immediacy, choose to submit market orders that execute against the posted limit orders. Understanding the costs of, and motives for, placing limit orders as distinct from market orders was called for.

With limit orders, the very existence of the bid-ask spread has to be explained. That is, with a sufficiently large number of participants placing priced orders, one might expect that orders would be posted at virtually every available price point in the neighborhood of equilibrium, and that the spread would disappear. Cohen, Maier, Schwartz, and Whitcomb (CMSW) made this point in their review paper [5], and they analyzed the existence of the spread in Cohen et al. [35].[7] They further write, "With regard to modeling the market spread, we suggest that a straightforward aggregation from individual spreads is not possible in a system where there is no clear distinction between demanders and suppliers of immediacy, and where traders meet in a dynamic, interactive environment that incorporates the impact of investor order placement strategies." Strategic order placement clearly required further analysis.

The task, however, was not simple. Some of the first papers in this area assumed, as is true for a dealer market, that limit order and market order participants are two separate, exogenously fixed groups that are separated by a firewall [36]. This assumption, while simplifying mathematical modeling, unfortunately distills out much of the richness of an order-driven market. More recent models have eliminated the firewall (Handa et al. [37]; Foucault et al. [38]; Parlour et al. [39]; Handa et al. [40]; Foucault et al. [41]; and Goettler et al. [42]). With the choice between limit order and market order endogenous, for any market to function, participants must divide naturally into four groups which reflect two dichotomies (one between buyers and sellers and the other between limit order and market order placers), not the standard two (buyers and sellers).

[7] Cohen et al. [35] describe the trade-off between execution probability and price improvement in the optimal choice between limit and market orders.

With order type selection endogenous in the order-driven market, the balance between immediacy demanders and immediacy suppliers becomes a second equilibrium that must be understood. That is, one needs to recognize the conditions under which some participants will choose to be liquidity demanders (place market orders) while others choose to be liquidity suppliers (place limit orders). If a reasonable balance is not achieved between these two groups, the order-driven market will fail (as indeed it does for thinner, small cap stocks). Increasingly, these issues have been handled, and some sophisticated limit order models have been developed.[8]

Microstructure analysis of trading systems has expanded to include periodic call auctions.[9] The economics of a call auction are quite different from those of continuous trading and, consequently, so too are the order placement strategies that participants should employ when they approach a call market. Call auctions do not, by their very nature, supply immediacy. Rather, orders that are entered during a call's book-building phase are held for a periodic crossing at a single clearing price at the (generally predetermined) time of the market call. Consequently, buy and sell orders submitted to a call do not execute when they arrive even if they match or cross in price (matching and crossing orders execute immediately in a continuous trading environment). This being the case, limit and market orders have a different meaning in a call: limit orders do not supply immediacy to market orders, and market orders are simply extremely aggressively priced limit orders (i.e., a market order to sell in a call effectively has a limit price of zero, and a market order to buy effectively has a limit price of infinity).

Today, virtually all modern, electronic exchanges open and close their continuous markets with call auctions. Consequently, participants face further decisions when operating in a call plus continuous, hybrid market: how to submit an order to a call auction which is followed by continuous trading (e.g., an opening call), and how to submit an order to a continuous trading environment that is followed by a call auction (e.g., a closing call). Taking these tactical decisions into account is part of the complexity of microstructure analysis.

Technological developments have simultaneously enabled new trading venues to emerge (which can fragment markets) while providing connectivity between them (which can consolidate markets). Concurrently, regulatory initiatives have been motivated by the desire to intensify inter-market competition. Questions can be raised, however, concerning fragmentation of the order flow. The conventional wisdom has been that the consolidation of order flow improves liquidity, and exposing each order to all other displayed orders gives investors the best prices for their trades. Consolidating trading in a single market provides incentives to liquidity suppliers to compete aggressively for market orders by revealing their trading interest, and by being the first to establish a more favorable price (if time is used as a secondary priority rule).

On the other hand, arguments in favor of trading on multiple markets include the benefits of inter-market competition, and the fact that traders with disparate motives for trading may want different marketplaces to trade in (i.e., the "one-size-does-not-fit-all" argument). And so, different markets develop to serve diverse investor needs

[8] See Back et al. [77] for a recent discussion and further references.

[9] See Economides et al. [78] for a description of alternative call market structures.

(such as achieving a faster execution vs. obtaining a better price). One growing need among large institutional investors, the ability to trade large orders with minimal market impact, has led to the advent of dark pool, block trading facilities such as Liquidnet, Pipeline, and ITG's Posit that aid in quantity discovery. This development in the industry has spawned a related line of research on off-exchange and upstairs trading ([43–46]).

A spectrum of market quality issues have been of long and continuing importance to microstructure researchers. These include market transparency,[10] both pre- and post-trade [47], the accentuation of intraday price volatility, and correlation patterns which have been observed in high-frequency data [48]. Other important issues include price clustering and tick sizes [49–51]. Applications such as transaction cost analysis (TCA) and algorithmic trading have received increasing attention [52]. The relative performance of floor-based vs. electronic trading is another important issue [34].

A major line of empirical research was pioneered by Hasbrouck et al. [53, 54] who decomposes transaction prices into two components: a random walk component and a stationary component. The random walk component is identified with an efficient price that the market is trying to discover. The stationary component is viewed as microstructure noise. Microstructure noise is commonly explained by features such as the bid-ask spread, market impact, and discreteness of the pricing grid. The noise component has also been attributable to price discovery itself being a dynamic process (Menkveld et al. [55], and Paroush et al. [56]).[11]

Numerous empirical studies have focused on two of the world's premier markets, the New York Stock Exchange and Nasdaq [54, 57–65], among others. Many other studies have considered European markets, Asian markets, and other markets around the world (e.g., [66–68]).[12] Across all of these markets, structural and performance differences have been noted, but also major similarities have been observed. It is apparent that, despite the influence of historic and cultural considerations, trader behavior and market performance around the globe depend largely on microstructure realities. Alternatively stated, trading rooms and markets around the world bear striking resemblances to each another.

Another recent line of research has considered how search costs affect bid-ask spreads in financial markets. To this end, Duffie, Pedersen, and Garleanu [69] present a dynamic model of market makers under the assumption of no inventory risk and information that is symmetrically distributed. They show that sophisticated investors who have better search and bargaining abilities face tighter bid-ask spreads. This is in contrast to traditional information-based models which imply that spreads are wider for more sophisticated (i.e., better informed) investors.

---

[10] Trading systems differ in their degree of transparency Pagano et al. [79] investigate whether greater transparency enhances market liquidity by reducing the opportunities for taking advantage of uninformed participants.

[11] Also see Hasbrouck, Joel. "One security et al. [65], Harvey et al. [80], and Jones et al. [81]." Further references are provided by Menkveld et al. [55].

[12] Also see Bessler, Wolfgang. Editor, Bösen et al. [82] for discussion and further references.

As we have noted, unlike in the frictionless market arena of CAPM, amassing liquidity is a primary function of a marketplace and market structure features are generally designed with liquidity implications in mind. Asset managers also take liquidity into account, along with the two other standard variables of modern portfolio theory, risk, and return. Difficulties in defining, measuring, and modeling liquidity are formidable, however, and the literature that deals with it directly is relatively sparse [70]. Nevertheless, liquidity considerations have permeated the microstructure literature, both explicitly and implicitly.[13]

Looking back over the development of microstructure analysis, two observations stand out. First, microstructure studies have in multiple ways given direction to market structure development. Second, to a remarkable extent, the various theoretical microstructure models that are center stage today, and many empirical analyses that are based upon them, share a common structural framework—the asymmetric information paradigm. This consistency is desirable in that it implies that the field has grown by accretion rather than by replacement. Consequently, new insights are more apt to refine than to contradict old conclusions.

Consistency, however, is not desirable if the common structural framework becomes overly rigid and restrictive, and if it yields incomplete and/or misleading answers to questions involving trader behavior, market structure, and regulatory policy. At times, a literature starts to advance along new fronts. We consider this possibility next for the microstructure literature.

10.4.3 Future Directions

As we have noted, the current focus in the literature is on asymmetric information-based models, which are characterized as follows. Trading is driven by informational change, liquidity needs, and noise trading. The information motive for trading is the first mover of the three (liquidity and noise trading are required so that a market will not fail). Further, order arrival in the continuous environment is generally taken to be asynchronous. For a continuous trading regime to function with asynchronous order arrival, the presence of a limit order book and/or a market maker intermediary is required.

Information trading is of keen interest because it represents the process by which new information is reflected in share values. In the standard asymmetric information models, it is assumed that all participants in possession of the same information form equivalent expectations concerning future risk and return configurations. When information changes, however, participants may not all receive the news at the same time; some receive it before others, a reality that, at any point in time, can divide traders into two groups—the informed and the uninformed.

[13] For further discussion and references regarding liquidity see Amihud et al. [83]; Chordia et al. [84, 85]; Hasbrouck, Joel and Duane Seppi. "Common Factors in Prices et al. [86]; Amihud et al. [87]; and Pástor et al. [88]."

Informed participants will never trade with each other; consequently, liquidity and noise traders must be present for a market to function. As noted, asymmetry of information, for the most part, lies at the heart of the standard microstructure models of today.

The homogeneous expectation assumption has been tempered of late. As a further departure from the infinitely liquid, zero-cost environment of CAPM, it is being recognized that some participants produce “private information” (namely, that they further process information so as to gain insights that are not immediately available to others). Whether participant expectations differ because of the actual production of private information, or simply because different people interpret the same information or news announcement differently, the expectations of a group of investors can be divergent.

Also at the heart of the asymmetric information models is the presumption that a stock has a fundamental value that bears a unique relationship, not to trader activity in the marketplace, but to the fundamental information that informed traders possess. The process of information being fully reflected in prices under asymmetric information is the act of informed and uninformed agents trading with each other until any discrepancy between a market price and a fundamental value is eliminated. The process can be viewed as arbitrage. In the earlier dealer models, the market maker was assumed to know a stock’s fundamental value. In later models, informed traders but not the market maker know the fundamental values [26]. Especially in the later models, price discovery is not instantaneous; rather, it is a protracted process that depends on the individual strategies employed by the informed and uninformed agents.

In recent years, an alternative paradigm has been emerging: a divergent expectations environment [71]. While institutionally realistic, this paradigm has met with considerable academic resistance. For one thing, homogeneous expectations environments are far easier to deal with mathematically and homogeneity has, in many applications, proven to be a useful modeling assumption. The assumption has also been retained for another reason. As an attribute of individual rationality, it is presumed that intelligent agents facing the same information and applying the same (correct) analytic techniques will reach the same conclusions and, therefore, will have homogeneous expectations.

Fundamental information, however, is enormous in scope. It is complex and imprecise, and our tools for analyzing it are relatively crude. In the presence of fuzzy information, expectations can be divergent. Allowing for divergent expectations opens another path for microstructure analysis, and it introduces new questions concerning agent behavior, market structure, and regulatory policy. Moreover, a further element can enter the analysis in a divergent expectations environment: along with forming their own opinions, agents may also respond to the opinions of others, i.e., exhibit adaptive valuation behavior [56, 72].[14] Just how agents commu-

[14] Adaptive valuation behavior refers to individual agents becoming more bullish (bearish) when learning of the relatively bullish (bearish) attitudes of others.

nicate with each other and respond to each others' opinions is a subject for ongoing research. The topic also opens another interface with behavioral finance.

Price discovery acquires a different meaning in a divergent expectations environment, and this has important implications for market structure. When asymmetric information characterizes a community of investors, the strategic behavior of informed agents can affect the path that price takes when news moves a share value from one equilibrium to another, but the new equilibrium is path *independent*. With divergent expectations, the new equilibrium is path *dependent*—it depends on how the opinions of a diverse set of agents are integrated [56]. Alternatively stated, with divergent expectations, price discovery is a coordination process and, as such, is directly effected by market structure.

In the standard asymmetric information environment, the key dichotomy is between informed and uninformed participants. But a second dichotomy also exists—one that separates large institutional customers from small retail customers. One might expect that the informed investor set would largely comprise the institutional customers. After all, the institutions are professional, they can afford to continuously monitor information and respond to news, and their very size (all else constant) reduces their per share cost of doing so. With divergent expectations, however, there is no presumption that institutional customers can, because of their size, consistently evaluate shares more accurately. On the contrary, institutions commonly disagree with each other and, as a consequence, commonly trade with each other.

In the divergent expectations environment, institutional investors do not necessarily have an advantage over retail customers as fundamental analysts. In fact, their size makes trading more difficult and they incur higher transaction costs. So what accounts for their popularity? The value added by the mutual funds, pension funds, etc. comes largely from their ability to facilitate diversification. Further, they can bring a systematic, professional, and disciplined approach to portfolio management [72].

10.5 From Theory to Application

Microstructure analysis is inherently involved with analyzing the detailed functioning of a marketplace. The literature has a strong theoretical component and, to a large extent, is structured to yield insights into the effect of market design (structure and regulation) on market performance. Hopefully, theory can provide a broad roadmap for real-world market architects to follow. In this section we provide a broad overview of major technology and regulatory changes that have taken place in the USA and Europe.[15]

[15] Further discussion of market structure development is provided by Harris and Larry. Trading and Exchanges: Market Microstructure for Practitioners [89].

10.5.1 Technological Developments

Two exogenous forces have driven market structure change: technology and regulation. Regarding technology, the first big step was taken in 1971 in the USA when the National Association of Securities Dealers (NASD) introduced an electronic automated quotation (AQ) display system called NASDAQ. The Toronto Stock Exchange was the first exchange to introduce an electronic order-driven platform, its Computer Assisted Trading System (CATS); the year was 1977. Following in Toronto's footsteps, London instituted SEAQ in 1986, Paris rolled out its Cotation Assistée en Continu (CAC) in 1986, and Deutsche Börse's Xetra came to life in 1997. Also in 1997 the London Stock Exchange introduced its Stock Exchange Trading System (SETS) limit order platform. By the end of the twentieth century most of the exchanges in Europe had converted to electronic limit order book platforms.

Change came more slowly in the USA. Instinet introduced an electronic platform in 1969. Nearly 30 years later, Instinet became known as an Electronic Communications Network (ECN). In short order, a slew of other ECNs emerged, led most prominently by Archipelago and Island. In 2002, Nasdaq implemented its own electronic platform which, at the time, was called "SuperMontage." Most recently, in the Spring of 2006, the newly privatized NYSE Group initiated its Hybrid Market, a facility that has transformed the Big Board from a floor-based "slow" market into a hybrid that includes a "fast market" electronic venue. As of this writing, the floor-based component of the NYSE's hybrid has been markedly reduced in importance. Several specialist firms have ceased operations, other floor brokers have departed, and the trading room areas have collapsed from five to two.

10.5.2 Regulatory Initiatives

Major regulatory initiatives have played an important role in jump-starting these market structure changes. The 1975 Congressional Securities Acts Amendments was the first sizable regulatory foray into market structure development. The Amendments precluded the fixing of commission rates and mandated the development of a National Market System (NMS). In 1997, the US Securities and Exchange Commission instituted its new Order Handling Rules (OHRs), which require that market makers holding customer limit orders display those orders in their quotes, and that dealers at least match any quotes that they themselves display on an ECN (either by bettering the quotes that they offer customers or by posting their superior quotes in Nasdaq's SuperMontage). Following the OHRs, three other regulatory initiatives were introduced in the USA in relatively fast succession. In 2000, the NYSE, under pressure from the SEC, rescinded its order consolidation rule (Rule 390). In 2001, the US markets completed the transition from fractional to decimal pricing, which resulted in the minimum tick size decreasing from 1/16 of a dollar or 6.25

cents (it had earlier been 1/8 of a dollar or 12.5 cents) to one cent. In 2005, the SEC adopted Regulation NMS, the key provision of which is that better priced limit orders cannot be traded through (the trade-through rule was fully implemented in 2007).

On the eastern side of the Atlantic, the first major regulatory initiative was taken in 1993 when the Investment Services Directive opened the door for cross-border trading by introducing the single European passport. As discussed in Schwartz, Robert and Reto Francioni. Equity Markets in Action and Sons [2], "*Passporting* defines a system of mutual acceptance of other EU countries' rules without truly harmonizing all of the details of the various rules." Major regulatory change is currently coming again to the European arena in the form of the Markets in Financial Instruments Directive (MiFID). Key provisions in MiFID include a best execution requirement (echoes of the 1975 US Securities Acts Amendments), a quote disclosure requirement for upstairs broker/dealers (echoes of the US Order Handling Rules), and the disallowance of order focusing rules (echoes of the US SEC pressuring the withdrawal of NYSE Rule 390). A major regulatory difference is that no trade-through rule has been imposed on the European markets (unlike under the US SEC's Reg NMS).

10.6 Deutsche Börse: The Emergence of a Modern, Electronic Market

We turn in this section to the designing of an actual marketplace. Our focus is on Deutsche Börse: it is the dominant stock exchange in Germany, the last of the major European bourses to go electronic, and its technology is state of the art.

Important insights were gained from the microstructure literature during Xetra's planning period and the system's implementation has marked a huge step forward for Germany's equity markets. But our roadmap, which is undoubtedly incomplete today, was even more limited in the 1994–1997 years when Xetra was being designed. And, there is always the danger that the cartographer whose map is being used has some misconceptions (e.g., believes in the existence of the Northwest Passage).

10.6.1 The German Equities Market in the Mid-1990s

As recently as the mid-1990s, the German market had major structural defects that would undermine its competitiveness in the European arena. In recognition of this, Deutsche Börse, the newly founded exchange operator of the Frankfurter Wertpapierbörse (FWB), became the leading force for change.[16]

[16] FWB also owned the futures and options exchange Deutsche Termine Börse. After the 1997 merger with SOFFEX, DTB became Eurex.

In the mid-1990s, Frankfurt's trading floor was the major marketplace for German stocks, but the German market was badly fragmented. Kursmaklers, the equivalent of specialists, concentrated much of the liquidity in their order books. A primitive (by today's standards) electronic trading system, IBIS (which was owned by FWB), operated in parallel with the floor trading. IBIS's central component was an open limit order book that had hit and take functionality, but did not match orders automatically. The electronic system captured about 40 % of the trading volume in the 30 large-cap DAX stocks, but no link existed between IBIS and the floor. Seven other floor-based regional exchanges were also operating in Germany with technical infrastructures that were similar to those in Frankfurt. In total, the regionals at that time were attracting roughly 10 % of German exchange-based trading volume. Moreover, off-board trading has been (and still is) prevalent in Germany [73].

Transparency for floor trading (pre-trade transparency in particular) was low. Quotes were not distributed publicly (they were available on the floor only). Price priority between different trading venues was not enforced and orders executed in one market commonly traded through orders waiting to be executed in another market. Market manipulation and other abuses of power and position were believed to be rife on the old Frankfurt floor. Given the appreciable market fragmentation, poor transparency, imperfect inter-market linkages, and dubious floor behavior, transaction costs were high. Changes, both structural and regulatory, were called for. The result was the development of Xetra, an electronic order-driven trading system that comprises two principal modalities—a continuous order book platform and periodic single-price call auctions.[17]

10.6.2 Designing a New Trading System

Xetra's development started in 1994, and the system was launched in 1997.[18] Strong external forces also motivated this reengineering of Deutsche Börse's market structure: regulatory reform, soaring trading volumes, pan-European harmonization of the exchange industry, vibrant cross-border competition for order flow, and rising concerns of market participants about the future performance of Germany's financial markets.

Through Xetra's design stage, microstructure theory, even as it existed at the time, was an indispensable guide. This new field in financial economics, with its origin in issues concerning the competitive and architectural structure of an equity market, should have been able to give guidance to the development of an actual marketplace such as Xetra. To an extent, it has fulfilled its promise. The literature gave Deutsche Börse a broad roadmap, and it has highlighted underlying relationships and other important considerations that a market architect should be aware of.

[17] For further discussion and descriptions, see Francioni et al. [1].

[18] Appendix 2 provides details of Xetra's design.

Building the Xetra model involved specifying principles that the new market should implement, and the system's functionality also had to be defined. Most importantly, the new market system was to provide equal and decentralized access to all of its participants. Further, the system's functionality and the market information delivered to users (both pre- and post-trade) were to be the same for all traders. A trader's location should not matter. With this in mind, Deutsche Börse's fundamental architectural decision was to structure a hybrid market that included two major modalities—a continuous electronic order-driven platform, and periodic call auctions that were used primarily for market openings and closings.[19]

An absolutely critical attribute of an order-driven trading system is its ability, vis-à-vis its competitors, to win the battle for liquidity. Regarding this matter, the earlier microstructure literature has given some guidance, but liquidity is a complex attribute to deal with. As it is not easy to define and measure, liquidity has been very difficult to model and assess. However, as noted above, the measurement and analysis of liquidity are currently attracting considerably more attention in the microstructure literature.

Price discovery and transparency are two other issues for which the microstructure literature has provided valuable guidance. The architects at Deutsche Börse recognized that price discovery is a primary function of a market center, and their major reason for introducing the call auctions was to sharpen its accuracy, particularly at market openings and closings. Understanding that transparency is important while recognizing that it should not be excessive, the decision was made to disclose only the indicative clearing price (not the full book of orders) in the pre-call, book-building period.

Microstructure literature has given insights into the operations of the public limit order book for continuous trading. At the time, recognition was also emerging of periodic call auctions, a modality that was clearly differentiated from, but could effectively be used with, the continuous market. With regard to continuous trading, microstructure analyses of the use of limit and market orders and of the interaction between these two order types proved to be most valuable. However, a deeper understanding of the economics of an order-driven market now exists than was the case in the 1994–1997 period when Xetra was being designed.

Another important contribution of microstructure theory has been the classification of traders according to their needs for immediacy and their propensities to be either givers or takers of liquidity. The differentiation between informed and uniformed traders also proved to be valuable, particularly with respect to the market maker role that has been incorporated into Xetra. Specifically, market makers, referred to as "designated sponsors," were included to bolster liquidity provision for smaller cap stocks. A balance had to be achieved between the obligations imposed on the designated sponsors and the privileges granted to them. To accomplish this, information had to be assessed concerning the role of dealers in general (e.g., NASDAQ-type market makers) and specialists in particular (e.g., NYSE-type

[19] Interestingly, the microstructure literature on call auctions was relatively sparse at that time. For an early discussion, see Handa et al. [90].

specialists). That balance defined the designated sponsors' role in Xetra, and secured their willingness to accept it. Market microstructure insights also yielded the understanding needed to transform the specialist role into the newly designed designated sponsor role.

But designing an automated trading systems is indeed a complex task, and the gap between theory and implementation is both large and intricate. Trading decisions can be made in a large variety of ways that run the gamut from humans interacting directly with humans without computers to humans trading via electronic order handling and execution systems and to computers making trading decisions that are sent electronically to a computerized market (e.g., computer-driven algorithmic trading). Since the mid-1990s, market structure development has involved mainly the design of an electronic trading facility.

Deutsche Börse took account of the fact that automation impacts both the way in which trading decisions are made and the process by which prices are determined and trades executed in a market center. An electronic market requires the specification of an array of critical features (e.g., the trading modalities employed, rules of price and quantity determination, and basic features such as order types and trading parameters). With an electronic market, the software that implements a desired market structure must be specified on a level of detail that far exceeds what is required for human intermediated trading.

For instance, a human agent (specialist) has historically handled price determination at NYSE openings. This function is performed with reference to various rules, but the specialist is also free to exercise reasonable judgment. Further, human-to-human interactions can evolve naturally as problems, opportunities, and new competitive pressures arise. In contrast, with a fully electronic opening, every possible condition that can occur must be recognized and a rule for dealing with it specified, and electronic interaction can be changed only by rewriting the code that specifies with step-by-step precision just how orders are handled and turned into trades and transaction prices.

How does one achieve the precise specifications that a computerized trading system must have? In 1994, the market architects at Deutsche Börse could study the operations of other electronic platforms (e.g., CATS in Toronto and CAC in Paris). Doing so was helpful but of limited value given that Deutsche Börse was looking to develop a distinctive system.

When moving into new territory, market structure development is a venture. How does one know in advance whether or not it will work? How can one determine whether or not the new system will be viable from a business perspective? Nevertheless, design decisions have to be made, technical requirements must be specified, and the system must be built. The decisions involved represent huge financial bets on whether or not a new market structure will attract sufficient liquidity. Prototyping a new market in the design phase helps the assessment process, but doing so was considerably more difficult in 1994 than it is today with the advent of superior information technology and testing capabilities. In 1994, the architects were forced to rely more on their own educated judgment and on any insights they might gain from microstructure research.

Those who are involved in the design of an actual market realize that the devil is in the details. Consider, for instance, the specification of a call auction. A call has excellent theoretical properties, but how should an actual auction be designed? It is straightforward to say that the market clearing price in a call auction should be the value that maximizes the number of shares that trade. But what should the specific rule be for selecting the clearing price if two prices both result in the same maximum trade size? Additionally, how transparent should the book be in the pre-call, order entry period? Are further design features needed to counter the possibility of gaming? And so on.

Other considerations that for the most part are outside the scope of the microstructure literature also came into play during the design of Xetra. Information technology issues such as scalability, open architecture, and system reliability are of critical importance. So too are procedures for post-trade clearing and settlement. One of the final steps in the structural design of the new German market was the introduction in 2003 of a central counterparty (with a CCP, counterparty risk management was centralized and trading became fully anonymous, both pre- and post-trade). Electronic trading is also a prerequisite for highly efficient straight-through processing (STP involves all stages of a trade's life cycle). Information technology has further facilitated the timely capture of market data (all trades, quotes, market index values, etc.) and has expedited its delivery to users. With regard to these diverse applications, Deutsche Börse has achieved a closer integration between trading on Xetra and the broader market infrastructure.

10.7 Conclusion: The Roadmap and the Road

A market architect must have a roadmap that, broadly speaking, says where one ought to head and roughly how to get there. To this end, the microstructure literature has added clarity, articulation, and intellectual support. Briefly stated, the objective is to reduce trading frictions (costs), sharpen price discovery, and facilitate quantity discovery. The means of achieving this broad objective involve the amassing of liquidity. This is done through the appropriate use of limit order books for both continuous and call auction trading and, where appropriate, the inclusion of broker/dealer intermediaries. Further insights are gained from microstructure's in-depth analyses of trading motives (new information, liquidity needs, and technical trading signals). The literature has also provided guidance with regard to issues such as transparency and the consolidation (fragmentation) of order flow.

But theory, even if it does provide a good roadmap, can take one only so far. The closer one gets to the design of an actual system, the more apparent the complexities of trading and trading systems become. The road actually traveled is indeed bumpy and hazardous. System designers know that "the devil is in the details." They have to grapple with issues ranging from scalability, reliability, and other IT requirements to business considerations concerning the ultimate profitability of a trading venue. The market architects at Deutsche Börse recognized these issues and their new system, Xetra, has marked a huge step forward for the German equity market.

Today, important problems persist with regard to market design in Germany (and in all other markets around the world). Two fundamental questions concerning market architecture that have yet to be adequately answered are the following: (1) What is the best way to deal with large, institutional orders? (2) How is liquidity creation best handled for mid-cap and small-cap stock? At the same time, important microstructure topics continue to emerge at the academic research desks. Are there limits beyond which microstructure theory cannot provide guidance? Are there limits to the level of efficiency that a real-world market can ever achieve? Undoubtedly, both answers are "yes" but, without question, neither of these limits has as of yet been reached. Quite clearly, microstructure research and the design of an actual marketplace remain works in progress.

10.8 Appendix 1: Risk Aversion and Risk Premium Measures

Our analysis of the perfectly liquid CAPM environment makes reference to two measures of risk aversion and to several dimensions of a risk premium. We provide details concerning both of these in this appendix.

10.8.1 Risk Aversion

We use two risk aversion measures: (1) $R_A = -U''(W)/U'(W)$ is a measure of absolute risk aversion, and (2) $R_R = WR_A$ is a measure of relative risk aversion. Because $U'' < 0$ for a risk averse decision maker, R_A, $R_R > 0$ for risk aversion. Larger values of R_A and R_R indicate higher degrees of risk aversion. R_A is a measure of absolute risk aversion because it reflects the decision maker's reaction to uncertainty in relation to the *absolute* (dollar) gains/losses in an uncertain situation. R_R is a measure of relative risk aversion because it reflects the decision maker's reaction to uncertainty in relation to the *percentage* gains/losses in an uncertain situation.[20]

10.8.2 Risk Premiums

A *risk premium* is the minimum dollar compensation a decision maker requires to hold a risky asset in place of an alternative that involves no risk. Specifically, a decision maker would be indifferent between a riskless investment with a certain return of D dollars and a risky investment with an expected dollar return of $E(Z)$ equal to

[20] For further discussion, see J. Pratt, "Risk Aversion in the Small and the Large," *Econometrica,* January 1964.

D plus the investor's risk premium. In general, the investor's risk premium depends upon his or her utility function and initial wealth, and upon the distribution of Z.

Pi (π) in (10.3) is a risk premium: π equals one-half of R_A (the measure of the investor's absolute risk aversion) times Var(P_2), which measures the absolute (dollar) risk attributable to holding one share of the market portfolio. The uncertainty associated with holding N shares of the risky asset is Var(NP_2) = N^2 Var(P_2); thus the total risk premium for holding N shares is

$$\pi_T = \pi N_1^2 \tag{10.22}$$

Dividing (10.22) by N_1 (=N_0+Q) gives the risk premium per share (the average risk premium):

$$\pi_A = \pi N_1 \tag{10.23}$$

Differentiating (10.22) with respect to N_1 gives the risk premium for a marginal share (the marginal risk premium):

$$\pi_m = 2\pi N_1 \tag{10.24}$$

Dividing (10.24) by P_1 expresses the marginal risk premium as a percentage of current price:

$$\pi_{M\%} = \frac{\pi_M}{P_1} = \frac{2\pi N_1}{P_1} \tag{10.25}$$

The return on the combined portfolio of N_1 shares of the market portfolio and C_1 dollars of the risk-free asset is

$$r_P = \left(\frac{P_2}{P_1} - 1\right)\left(\frac{P_1 N_1}{W}\right) + \left(1 - \frac{P_1 N_1}{W}\right) r_f \tag{10.26}$$

and the variance of the return on the combined portfolio is

$$Var\left[\left(\frac{P_2}{P_1}\right)\left(\frac{P_1 N_1}{W}\right)\right] = \left(\frac{N_1}{W}\right)^2 Var\left(P_2\right) \tag{10.27}$$

Thus the investor's risk premium associated with the uncertain return realized from the combined portfolio is

$$\pi_{rp} = \left(\frac{N_1}{W} \right)^2 \pi \tag{10.28}$$

10.9 Appendix 2: Designing Xetra

This appendix provides further detail on the development and design of Deutsche Börse's electronic trading platform, Xetra. The first steps in designing Xetra involved specifying principles that the new market should implement, and defining the system's functionality. This was done by Deutsche Börse working together with key market participants. Most importantly, the new market system was to provide equal and decentralized access to all its participants. Further, the system's functionality and the market information delivered to users (whether pre- or post-trade) were to be the same for all traders. A trader's location should not matter.

Equity trading in the German market has been and continues to be order driven. This was true both for IBIS and for floor trading that was managed by a Kursmakler acting in the capacity of auctioneer, broker, and dealer. It was clear from the beginning that Xetra should run an open limit order book (open in the sense that aggregated order volume is displayed at all price points in the order book). Additionally, order matching was automated and trader anonymity ensured.

Core features of an electronic trading system are determined by the market structure that it implements. The structure defines how orders are handled and translated into trades and transaction prices. Xetra's market model comprises diverse submodels, each with a single trading modality, or a combination of multiple modalities (i.e., it is a hybrid). Most importantly, Xetra implements both continuous trading and periodic call auction trading. This differentiation is required to cope with liquidity differences among stocks, and different liquidity needs among users depending on the size of their orders and motives for trading. The market for all stocks opens and closes with a call auction, while less liquid stocks trade in multiple call auctions per day.

Once the building blocks were defined (i.e., continuous trading and call auctions), and their combinations specified, the next design step was to detail the specific features of each of the modalities. Those features are either static (i.e., represent basic structures such as the order book) or dynamic (i.e., define processes and behavior such as order matching). The next two sections of this appendix consider the systems design in more detail for continuous trading and periodic call auction trading, respectively.

10.9.1 Continuous Trading

By the mid-1990s, order books for continuous trading with price and time priorities had been implemented around the globe. In designing Xetra, Deutsche Börse's market architects could refer to a wide range of existing examples, and to a broad microstructure literature. Once the eligible order types were identified, the center piece of the development was the definition of the detailed rules of price-time matching. The complexity of this definition was broken down into a finite set of individual cases that involved various order book situations combined with various incoming orders, for which the trading outcome was to be defined by a rule. All rules collectively described the dynamics of order matching.

A major challenge in designing continuous trading involves the measures that should be taken to provide an orderly market in periods of sharply elevated price volatility. To deal with this, the concept of a "price corridor" was formulated. Diverse corridors around historical prices were defined that set the benchmark for an "orderly" price for the next trade. If a price occurred that lay outside its corridor, trading was to be halted (briefly) with the entire order book transported into a call auction. The purpose of the call was to allow the market to consolidate in both space and time. Trading in the continuous market was resumed upon completion of the call.

Lastly, all trading parameters for the continuous platform had to be determined. This included specifying tick sizes, breadth of the price corridors, durations, and timings. Together, this provided a comprehensive overview of the "steering wheels" for the newly designed market.

10.9.2 Call Auction Trading

The purpose of Xetra's call auctions is threefold: (1) to open and close continuous trading, (2) to trade less liquid stocks in multiple calls per day with no continuous trading offered, and (3) to stabilize the market in times of large price moves. Despite those multiple purposes, a single design was defined for the auctions. Additionally, certain key consistencies between continuous trading and the call had to be achieved. For example, both limit and market orders that could be submitted to continuous trading were allowed entry into the call order book. This seemingly simple requirement was complicated to implement because it expanded the universe of possible order book configurations (and therefore necessitated more complex matching rules). Additional procedures for setting the clearing price were also required to guard against erroneous pricing that could be caused by market orders overpowering an insufficient number of limit orders. As with continuous trading, price and time priority execution rules were stipulated.

Most crucial was the degree of transparency that the calls would offer. Sufficient information about the order book had to be delivered for market participants to have relevant price and quantity information concerning actual market situations, but detailed information was suppressed to inhibit excessive information leakage and gaming. The pre-call information now available in Xetra is the highest bid and the lowest offer posted in the call when these orders do not cross, or the indicative call auction price that is calculated when the order book is crossed. In other words, the full order book content is not visible—pre-call, the Xetra screen displays only the potential outcome of the call at each point in time.

When Xetra was under development, call auction trading at prespecified times was managed on the floor by specialists who were responsible for price determination, timing, and provision of dealer liquidity. The challenge was to reengineer the call so that it could be run by a computer, not by a human intermediary. The issue that Deutsche Börse was facing was also grappled with by market microstructure academicians and other market architects. Substantial external guidance was received in the planning process. In particular, important inputs were obtained concerning the optimal degree of transparency for the call's anti-gaming measures. The availability at the time of a variety of different call auction designs (both used and proposed) enabled Xetra's calls to be designed relatively quickly.

10.9.3 Electronic Trading for Less Liquid Stocks

Kursmaklers (specialists) on the Frankfurt floor (both today and in the past) provide immediate liquidity at times when external liquidity is insufficient. The desire was strongly expressed, with two provisos, for a market maker to be incorporated into Xetra's order-driven model for less liquid stocks. The two provisos were that (1) market participants must all have equal access to information, and (2) equal access to functionality must be maintained at a maximum level. Consequently, any changes that would favor the dealers were kept to a minimum.

The dealers were referred to as "designated sponsors." Like market makers in general, the designated sponsors were given both privileges and obligations. The primary obligation is that, on request of other market participants, the designated sponsor must provide quotes for a minimum volume and maximum spread in a stock during continuous trading. Additionally, multiple designated sponsors were included, so that they might compete with each other. Concurrently, the fulfillment of each sponsor's obligation is measured, and the results are published.

The designated sponsors' primary privilege is that they can see the identity of the quote requesters in an environment that otherwise ensures complete anonymity. Further, a sponsor balances the order book in all call auctions for the stocks that it is registered in. This gives the designated sponsors a last mover advantage (the freedom to trade against any imbalance that might exist at the market clearing price). With this privilege, a designated sponsor can influence the clearing price so as to execute orders that otherwise would not have transacted in that call. Lastly, the designated sponsors, depending on their measured performance, receive fee reductions.

10.9.4 Xetra's Implementation and the Migration of Liquidity to Xetra Since 1997

Xetra went operational in Fall 1997. At the beginning, the new system attracted roughly 60 % of trading in the most liquid segment of the market, the 30 DAX stocks. Trading on Xetra for mid-cap stocks was not as successful—market share for this segment of the market was about 20 %, as the less liquid stocks largely continued at that time to trade on the floor. But the 1997 launch was just the start of a sequence of releases that have continued through the current time.

One more recent innovation was the "continuous call auction." With this facility, calls are not held at prespecified times but are triggered by the occurrence of a "critical" liquidity situation. The continuous call comprises a dealer-auctioneer who is responsible for providing a base level of liquidity in each call, as well as controlling its timing. Additionally, Xetra allows internalization of trading by member firms. Consequently, Xetra, which originally started as an exchange trading system, now also serves as the technical platform for OTC trading.

Major innovations have benefited a broad range of cap sizes and, across the board, floor trading has continued to decline. Xetra has now been rolled out to 260 member firms in Europe, and its market share currently stands at 95 % of all on-exchange trading in Germany today.

References

1. Francioni, Reto, Sonali Hazarika, Martin Reck and Robert A. Schwartz. "Equity Market Microstructure: Taking Stock of What We Know." *Journal of Portfolio Management*, Fall, 2008, forthcoming.
2. Schwartz, Robert and Reto Francioni. *Equity Markets in Action*, John Wiley and Sons, 2004.
3. Bernstein, Peter. "The Surprising Bond between CAPM and the Meaning of Liquidity." *Journal of Portfolio Management,* Fall 2007.
4. Ho, Thomas, Robert Schwartz, and D. Whitcomb. "The Trading Decision and Market Clearing Under Transaction Price Uncertainty." *Journal of Finance,* March 1985.
5. Cohen, Kalman, Steven Maier, Robert Schwartz, and David Whitcomb. "Market Makers and the Market Spread: A Review of Recent Literature." *The Journal of Financial and Quantitative Analysis*, 1979, 14, pp. 813–835.
6. O'Hara, Maureen. Market Microstructure Theory. Basil Blackwell, 1997, Cambridge, MA.
7. Madhavan, Ananth. Market Microstructure. *Journal of Financial Markets*, 2000, 3, pp. 205–258.
8. Biais, Bruno, Bruno, Larry Glosten, and Chester Spatt. "Market microstructure: A Survey of Microfoundations, Empirical Results, and Policy Implications." *Journal of Financial Markets*, 2005, 8, pp. 217–264.
9. Parlour, Christine and Duane Seppi. "Limit Order Markets: A Survey." Forthcoming in Handbook of Financial Intermediation & Banking edited by A.W.A. Boot and A. V. Thakor, 2008.
10. Hasbrouck, Joel. Empirical Market Microstructure. Oxford University Press, 2007.
11. Garman, Mark. "Market Microstructure." *Journal of Financial Economics*, June 1976, pp. 33–53.

12. Stigler, George. "Public Regulation of the Securities Markets." *Journal of Business*, April 1964, pp. 117–142.
13. Smidt, Seymour. "Which Road to an Efficient Stock Market: Free Competition or Regulated Monopoly?" *Financial Analyst Journal,* 27, 1971, pp. 18–20, 64–69.
14. Tinic, Sneha. "The Economics of Liquidity Services." *The Quarterly Journal of Economics*, 86, 1, 1972, pp. 79–93.
15. Tinic, Sneha, and Richard West. "The Securities Industry under Negotiated Brokerage Commissions: Changes in the Structure and Performance of New York Stock Exchange Member Firms" Bell Journal of Economics, 11, (Spring 1980), pp. 29–41.
16. Bagehot, Walter (pseudonym). "The Only Game in Town." *Financial Analysts Journal,* (Mar/April 1971), pp. 12–14, 22.
17. Stoll, Hans. "The Supply of Dealer Services in Securities Markets." *Journal of Finance*, 33, 1978, pp. 1133–1151.
18. Amihud, Yakov, and Haim Mendelson. "Dealership market: Market-Making with Inventory." *Journal of Financial Economics*, 8, 1980, pp. 31–53.
19. Ho, Thomas, and Hans Stoll. "On Dealer Markets under Competition." *Journal of Finance*, 35, 1980, pp. 259–267.
20. Ho, Thomas, and Hans Stoll. "Optimal Dealer Pricing under Transactions and Return Uncertainty." *Journal of Financial Economics*, 9, 1981, pp. 47–73.
21. Ho, Thomas, and Hans Stoll. "The Dynamics of Dealer Markets Under Competition." *Journal of Finance*, 38, 1983, pp. 1053–1074.
22. Mildenstein, Eckart, and Harold Schleef. "The Optimal Pricing Policy of a Monopolistic Marketmaker in the Equity Market." *Journal of Finance*, 38, 1983, pp. 218–231.
23. Stoll, Hans. "Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests," *Journal of Finance*, 44, 1989, pp. 115–134.
24. Demsetz, Harold. "The Cost of Transacting." *The Quarterly Journal of Economics*, 82, 1, 1968, pp. 33–53.
25. Glosten, Lawrence, and Paul Milgrom. "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents." *Journal of Financial Economics*, 14, 1985, pp. 71–100.
26. Kyle, Albert. "Continuous Auctions and Insider Trading." *Econometrica*, 53, 1985, pp. 1315–1335.
27. Grossman, Sandford, and Joseph Stiglitz. On the Impossibility of Informationally Efficient Markets. *American Economic Review*, 70, 3, 1980, pp. 393–408.
28. Milgrom, Paul and Nancy Stokey. "Information, Trade and Common Knowledge." *Journal of Economic Theory*, 26, 1, 1982, pp. 17–27.
29. Copeland, Thomas, and Dan Galai. "Information Effects on the Bid-Ask Spread." *Journal of Finance*, 38, 1983, pp. 1457–1469.
30. Easley, David, and Maureen O'Hara. "Price, Trade Size, and Information in Securities Markets." *Journal of Financial Economics*, 19, 1987, pp. 69–90.
31. Easley, David, and Maureen O'Hara. "Order Form and Information in Securities Markets." *Journal of Finance* 46, 1991, pp. 905–928.
32. Easley, David, and Maureen O'Hara. "Time and the Process of Security Price Adjustment." *Journal of Finance*, 47, 1992, 577–606.
33. Mendelson, Morris, Junius Peake and T. Williams, "Toward a Modern Exchange: The Peake-Mendelson-Williams Proposal for an Electronically Assisted Auction Market," in E. Bloch and R. Schwartz, eds. *Impending Changes for Securities Markets: What Role for the Exchange?* JAI Press, 1979.
34. Domowitz, Ian and Benn Steil. "Automation, Trading Costs, and the Structure of the Securities Trading Industry." Brookings-Wharton Papers on Financial Services, 1999, 33–92.
35. Cohen, Kalman, Steven Maier, Robert Schwartz, and David Whitcomb. "Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread." *The Journal of Political Economy*, April 1981, pp. 287–305.

36. Glosten, Lawrence. "Is the Electronic Open Limit Order Book Inevitable?" *Journal of Finance*, 49, 1994, pp. 1127–1161.
37. Handa, Puneet, and Robert Schwartz. "Limit Order Trading." *Journal of Finance*, December 1996a, pp. 1835–1861.
38. Foucault, Thierry. "Order Flow Composition and Trading Costs in a Dynamic Order Driven Market." *Journal of Financial Markets,* 2, 1999, pp. 99–134.
39. Parlour, Christine. "Price Dynamics in Limit Order Markets." *Review of Financial Studies,* 11, 1998, pp. 789–816.
40. Handa, Puneet, Robert Schwartz, and Ashish Tiwari. "Quote Setting and Price Formation in an Order Driven Market." *Journal of Financial Markets*, 2003, 6, pp 461–489.
41. Foucault, Thierry, Ohad Kaden and Eugene Kandel. "The Limit Order Book as a Market for Liquidity." *Review of Financial Studies,* 18, 2005, pp. 1171–1217.
42. Goettler, Ronald, Christine Parlour, and Uday Rajan. "Equilibrium in a Dynamic Limit Order Market." *Journal of Finance,* 60, 2005, pp. 2149–2192.
43. Grossman, Sandford. "The Information Role of Upstairs and Downstairs Markets." *Journal of Business*, 1992, 65, pp. 509–529.
44. Keim, Donald, and Ananth Madhavan. "The Upstairs Market for Large-Block Transactions: Analysis and Measurement of Price Effects." *Review of Financial Studies,* 9, 1996, pp. 1–36.
45. Madhavan, Ananth, and Minder Cheng. "In Search of Liquidity: An Analysis of Upstairs and Downstairs Trades." *Review of Financial Studies*, 10, 1997, pp. 175–204.
46. Seppi, Duane. "Equilibrium Block Trading and Asymmetric Information." *Journal of Finance*, 45, 1990, pp. 73–94.
47. Porter, David and Daniel Weaver. "Post-Trade Transparency on Nasdaq's National Market System." *Journal of Financial Economics*, 50, 2, 1998, pp. 231–252.
48. Engle, Robert, and Clive Granger. "Co-integration and Error Correction: Representation, Estimation and Testing." *Econometrica*, 55, 1987, pp. 251–276.
49. Angel, James. "Tick Size, Share Prices, and Stock Splits." *Journal of Finance*, 52, 1997, pp. 655–681.
50. Harris, Larry. "Stock Price Clustering and Discreteness." *Review of Financial Studies*, 4, 1991, pp. 389–415.
51. Harris, Larry. "Minimum Price Variations, Discrete Bid-Ask Spreads, and Quotation Sizes." *Review of Financial Studies*, 7, 1994, pp. 149–178.
52. Domowitz, Ian, Jack Glen and Ananth Madhavan. "Liquidity, Volatility, and Equity Trading Costs across Countries and Over Time." *International Finance* 4, 221–256, 2001.
53. Hasbrouck, Joel. "Assessing the Quality of a Security Market: A New Approach to Transaction-Cost Measurement." *Review of Financial Studies,*, 6 (1), 1993, pp. 191–212.
54. Hasbrouck, Joel, and George Sofianos. "The Trades of Market Makers: An Empirical Analysis of NYSE Specialists." *Journal of Finance*, 48, 1993, pp. 1565–1593.
55. Menkveld, Albert, Siem Koopman and André Lucas. "Modelling Around-the-Clock Price Discovery for Cross-Listed Stocks Using State Space Methods." *Journal of Business and Economic Statistics,* 25 (2), 2007, pp. 213–225.
56. Paroush, Jacob, Robert Schwartz, and Avner Wolf. "The Dynamic Process of Price Discovery in an Equity Market." Working paper, Baruch College, CUNY, 2008.
57. Barclay, Michael and Terrence Hendershott and Timothy McCormick. "Competition Among Trading Venues: Information and Trading on Electronic Communications Networks." *Journal of Finance*, 58, 2003, pp. 2637–2666.
58. Bessembinder, Hendrik, and Herbert Kaufman. "A Comparison of Trade Execution Costs for NYSE and Nasdaq-listed Stocks." *The Journal of Financial and Quantitative Analysis*, 32, 1997a, pp. 287–310.
59. Bessembinder, Hendrik, and Herbert Kaufman. "A cross-exchange comparison of execution costs and information flow for NYSE-listed stocks." *Journal of Financial Economics*, 46, 1997b, pp. 293–319.
60. Bessembinder, Hendrik. "Trade Execution Costs on Nasdaq and the NYSE: A Post-Reform Comparison." *The Journal of Financial and Quantitative Analysis*, 34, 1999, pp. 387–407.

61. Bessembinder, Hendrik. “Quote-Based Competition and Trade Execution Costs in NYSE-Listed Stocks”, *Journal of Financial Economics*, 70, 2003, pp. 385–422.
62. Christie, William, Jeffrey Harris, and Paul Schultz. Why Did NASDAQ Market Makers Stop Avoiding Odd-Eighth Quotes? *Journal of Finance*, 49, 1994, pp. 1841–1860.
63. Christie, William, and Paul Schultz. Why do NASDAQ Market Makers Avoid Odd-Eighth Quotes? *Journal of Finance*, 49, 1994, pp. 1813–1840.
64. Hasbrouck, Joel. “Measuring the Information Content of Stock Trades.” *Journal of Finance*, 46, 1991, pp. 179–207.
65. Hasbrouck, Joel. “One security, many markets: Determining the Contribution to Price Discovery.” *Journal of Finance*, 50, 1995, pp. 1175–1199.
66. Biais, Bruno., Pierre Hillion, and Chester Spatt. “An Empirical Analysis of the Limit Order Book and the Order Flow in the Paris Bourse.” *Journal of Finance*, 50, 1995, pp. 1655–1689.
67. Ozenbas, Deniz, Robert Schwartz, and Robert Wood. “Volatility in U.S. and European Equity Markets: An Assessment of Market Quality,” *International Finance,* Volume 5 Number 3, Winter 2002, pp. 437–461.
68. Sandas, Patrick. “Adverse Selection and Competitive Market Making: Empirical Evidence from a Limit Order Market.” *Review of Financial Studies*, 14, 2001, pp. 705–734.
69. Duffie, Darrell, Lasse Pedersen and Nicolae Garleanu. “Valuation in Over-the-Counter Markets.” *Review of Financial Studies*, 20, 2007, pp. 1865–1900.
70. Bernstein, Peter. “Liquidity, Stock Markets and Market Makers.” *Financial Management,* Summer 1987, pp. 54–62.
71. Miller, Edward, “Risk, Uncertainty and Divergence of Opinion,” *Journal of Finance,* 1977, 32, 1151–1168.
72. Davis, Paul, Michael Pagano, and Robert Schwartz. “Divergent Expectation.” *Journal of Portfolio Management*, Fall 2007, forthcoming.
73. Davis, Paul, Michael Pagano, and Robert Schwartz. “Life After the Big Board Goes Electronic.” *Financial Analysts Journal*, Volume 62, Number 5, September/October 2006, pp. 14–20.
74. Beja, Avraham and Nils H. Hakansson. “Dynamic Market Processes and the Rewards to Up-to-Date Information.” *Journal of Finance*, 32, 1977, pp. 291–304.
75. Beja, Avraham and M. Barry Goldman. “On the Dynamic Behavior of Prices in Disequilibrium.” *Journal of Finance*, 35, 1980, pp. 235–248.
76. Copeland, Thomas. “A Model of Asset Trading under the Assumption of Sequential Information Arrival.” *Journal of Finance*, 31, 1976, pp. 1149–1168.
77. Back, Kerry and Shmuel Baruch. “Working Orders in Limit-Order Markets and Floor Exchanges.” *Journal of Finance*, 62, 2007, pp. 1589–1621.
78. Economides, Nicholas, and Robert Schwartz. “Electronic Call Market Trading.” *Journal of Portfolio Management*, Spring 1995, pp. 10–18.
79. Pagano, Marco and Ailsa Röell. “Transparency and Liquidity: a Comparison of Auction and Dealer Markets with Informed Trading”, *Journal of Finance*, 51, 1996, pp. 579–611.
80. Harvey, Campbell and Roger Huang. “Volatility in the Foreign Currency Futures Market.” *Review of Financial Studies,* 4, 1991, pp. 543–569.
81. Jones, Charles, Gautam Kaul and Marc Lipson. “Information, Trading and Volatility.” *Journal of Financial Economics*, 36, 1994, pp. 127–154.
82. Bessler, Wolfgang. Editor, Bösen, Banken und Kapitalmärkte, Duncker & Humblot, 2006.
83. Amihud, Yakov and Haim Mendelson. “Asset Pricing and the Bid-Ask Spread.” *Journal of Financial Economics*, 17, 1986, pp. 223–249.
84. Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam. “Commonality in Liquidity.” *Journal of Financial Economics*, 56, 2000, pp. 3–28.
85. Chordia, Tarun, Richard Roll and Avanidhar Subrahmanyam. “Liquidity and Market Efficiency.” *Journal of Financial Economics*, 87, 2008, pp. 249–268.
86. Hasbrouck, Joel and Duane Seppi. “Common Factors in Prices, Order Flows and Liquidity.” *Journal of Financial Economics*, 59, 2001, pp. 383–411.

87. Amihud, Yakov. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects." *Journal of Financial Markets*, 5, 2002, pp. 31–56.
88. Pástor, Luboš and Robert Stambaugh. "Liquidity Risk and Expected Stock Returns." *Journal of Political Economy*, 113, 2003, pp. 642–685.
89. Harris, Larry. Trading and Exchanges: Market Microstructure for Practitioners. 2003, Oxford University Press.
90. Handa, Puneet and Robert Schwartz. "How Best to Supply Liquidity to a Securities Market." *Journal of Portfolio Management*, Winter 1996b, pp. 44–51.
91. Schwartz, Robert. *Reshaping the Equity Markets: A Guide for the 1990s,* Harper Business, 1991.