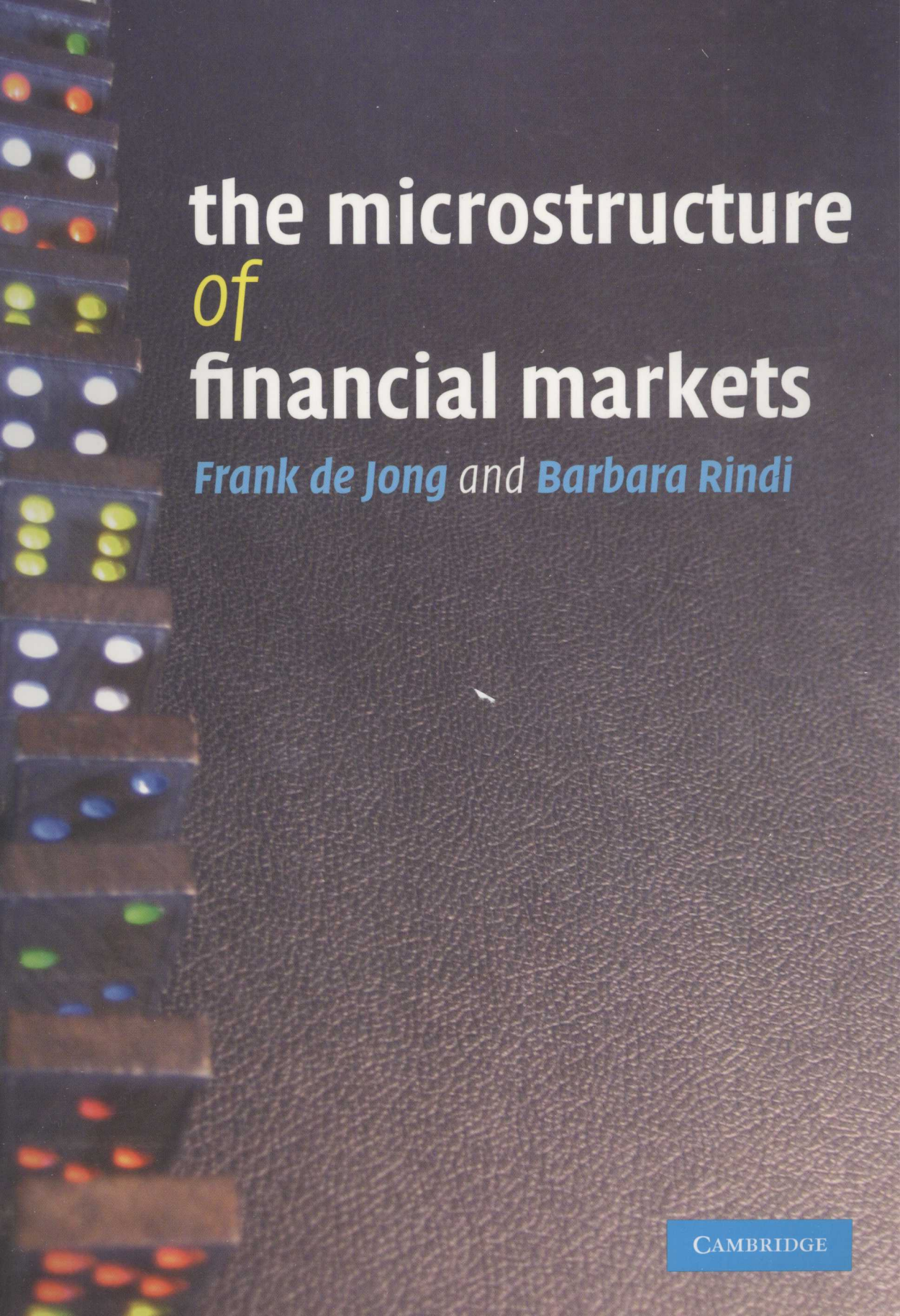
the microstructure
of
financial markets
Frank de Jong and Barbara Rindi
CAMBRIDGE

The Microstructure of Financial Markets

Frank de Jong

and

Barbara Rindi

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521687270

First published 2009

Printed in the United Kingdom at the University Press, Cambridge

A catalogue record for this publication is available from the British Library

ISBN 978-0-521-86784-9 hardback
ISBN 978-0-521-68727-0 paperback

Preface

This is a textbook on economic models of financial markets (market microstructure). The material is intended for PhD and Master's (MSc or MPhil) students in economics or finance. Readers are expected to have some background in microeconomic theory, basic finance and some statistics. The aim is to provide the student with the tools to be able to read and appreciate academic papers on market microstructure. The book can be used for a full semester course on financial markets and market microstructure.

The authors have been using draft versions of the book chapters in their PhD and Master's courses at Bocconi University, the University of Amsterdam and Tilburg University. We thank the students of these courses for their questions and suggestions. We would also like to thank several colleagues who provided us with invaluable feedback and comments. We thank our referees for their suggestions, and in particular Paolo Vitale for a thorough review of the manuscript. We are grateful to Roberto Battalio, Bruno Biais, Sabrina Buti, Paolo Colla, Gene Kandel, Jeremie Lefebvre, Marco LiCalzi, Angelo Ranaldo and in particular Ohad Kadan, Giovanna Nicodano and Ingrid Werner for their detailed comments. Our thanks also to Luisella Bosetti, Fabio Deotto, Luca Filippa, Martino Ghezzi, Alan Hodson, Larry Leibowitz, Enrico Mandelli, Clare McQuitty and other practitioners and regulators who gave us the opportunity to learn about the working of real financial markets.

Contents

Figures

Tables

Introduction

In the traditional approach to financial economics the price formation process is a 'black box' in which there is no explicit role for financial market structure. Aspects such as dealers, spread and the organizational structure of financial markets are ignored. In Stoll's (2000) words, 'thirty years ago friction in financial markets was largely ignored in the theory of finance'. In the last two decades a substantial literature on the process of financial price formation has developed and is generally known under the name *market microstructure*. This literature relaxes the assumptions of the traditional asset pricing models, such as the absence of transaction costs, homogeneous and symmetric information, and looks inside the 'black box' to get a better understanding of intraday price dynamics. A related objective is to formulate recommendations for the optimal design and regulation of financial markets. Market microstructure also provides new techniques for estimating transaction costs and market liquidity. This is of practical relevance for professional investors, who are interested in efficient order execution and the effect of liquidity on asset prices.

Amihud and Mendelson (1987) was one of the first papers to draw researchers' attention to the relevance of market structure. They found that the variance of the open-to-open returns was higher than that of the close-to-close returns for the hundred most liquid stocks on the New York Stock Exchange (NYSE). This suggested that, due to the different market organization of the opening and the closing, agents trading at the opening were exposed to greater volatility than traders at the close.

In the following decade a substantial empirical literature focused on the relevance of financial market microstructure to price formation in stocks, bonds, derivatives and foreign exchange order flows. One of the purposes of microstructure models is to understand the impact of the organizational structure and the design of financial markets on trading costs and asset prices. These models are then used to construct indicators of market quality, which can serve to assess the utility of regulatory interventions.

The traditional representation of price formation in financial markets is an invisible hand or a Walrasian auctioneer equating supply and demand at the equilibrium price. However, real-world examples of financial markets show that price formation is more complex since (i) traders do not arrive simultaneously at the market-place and (ii) information is asymmetric. These imperfections generate trading costs and make

the market protocol in which trading takes place a relevant factor. Building on these frictions, two main strands of standard theoretical microstructure literature developed, namely inventory-based and information-based models. The former assigns a primary role to market-makers as liquidity providers (professionals who undertake to supply liquidity to the market) and shows how the bid–ask spread compensates them for price risk on inventory; the latter focuses on asymmetric information among market participants and shows how market-makers set the bid–ask spread to compensate for adverse selection costs. Applications of microstructure models (for an overview, see e.g. Biais, Glosten and Spatt, 2005 and Madhavan, 2000) include estimations of transaction costs and liquidity as well as the study of the price discovery process. This illustrates how and how quickly private information is reflected in market prices.

Microstructure models can also be used to explain apparently abnormal behaviour of prices and volumes. For instance, it has been noticed that the intraday pattern of average volumes and stock returns is U-shaped (Jain and Joh, 1988; Wood, McInish and Ord, 1985): at the end and at the beginning of the trading day an increase in both volumes and volatility is observed. Microstructure models (Admati and Pfleiderer, 1988; Hong and Wang, 2000) were constructed to explain why volumes and volatility are concentrated at certain times of day and how the relationship between the two variables can be explained. Another well-known example of an abnormal price pattern, termed the 'NASDAQ controversy', is described in the adjoining box.

The NASDAQ controversy

A well-known example of abnormal price behaviour is the finding of Christie and Schultz (1994) on the distribution of price quotes for the hundred most traded stocks on the National Association of Security Dealer Automated Quotation (NASDAQ) exchange. They found a virtually total absence of odd-eighth quotes. By contrast, for a sample of a hundred stocks with similar characteristics on the NYSE, they observed that the quotes were distributed evenly over the full spectrum of the eighths. Christie, Harris and Schultz (1994) showed that the absence of odd-eighth quotes on NASDAQ produced a wider spread, suggesting collusive behaviour on the part of market-makers. Another well-known early empirical finding that questioned the effective degree of competition within and between markets was that of Huang and Stoll (1996), who showed that execution costs were twice as high for a sample of NASDAQ stocks as for a matched sample of NYSE stocks. They proposed two explanations for the lower degree of competition on the NASDAQ: the common use by NASDAQ dealers and brokers of practices such as internalization, preferencing and crossing orders (see Chapter 10); and the availability of alternative trading systems where dealers could quote non-competitive prices. The Securities and Exchange Commission (SEC) reacted to these findings in 1997 by introducing its 'order handling rules', which significantly enhanced competition, both within and between markets.

Microstructure models generally begin by observing that markets can differ in a number of respects. The differences may involve the typology of participants, their attitude towards risk, the existence of fixed entry costs, the organizational structure of trades and the degree of transparency. In the standard models, market participants are usually classified by the type of information they hold and their motive for trading.

Informed traders have access to private information about the liquidation value of the asset traded and/or the identity of the other agents in the market-place. Private information may derive from insider information on the future value of the asset, from (costly) research on the asset's value, or from knowledge of the order flow; costly information can also be obtained through analysts. In the latter cases, information can be acquired only at a cost and it is not illegal. For instance, most investment banks allocate resources to obtaining information on the flow of orders or to procuring the greatest possible number of institutional clients. This kind of privileged information is typical of secondary markets in treasury bonds and foreign exchange. Furthermore, the information available to the insiders may differ in its precision. In market microstructure models the precision of the information is defined as the inverse of the variance of the asset's liquidation value, conditional on the information available to the insiders.

Uninformed traders have no private information; usually the models make distinctions among noise traders, who are purely liquidity-motivated, hedgers, who trade to cover possible adverse fluctuations and are therefore modelled as risk-averse, and market-makers or dealers, who provide liquidity.

As far as preferences are concerned, in standard models of market microstructure, agents are generally assumed to be either risk-neutral or risk-averse. Risk-neutral agents are interested in the average (or expected) return to the asset, whether this is certain or uncertain. For example, take two assets, A and B, with A yielding \$8 or \$12 (with probability $\frac{1}{2}$) depending on the state of the world, and B yielding a certain return of \$10. Risk-neutral agents will be indifferent between these two asset returns. Risk-averse agents, however, demand a premium on the return (or a discount on the price) to hold the risky asset.

At times, the cost of acquiring information is modelled in order to make the number of agents with privileged information endogenous. It may be useful, in fact, to evaluate the possible effect of changes in regulations or market organizational structure, bearing in mind that these changes can modify the insiders' incentive to trade.

Outline of the book

A road-map for this book is suggested in Figure Intro.1. Chapter 1 provides a detailed overview of the organizational structure of financial markets around the world, plus an in-depth description of the working of the limit order book, which is an increasingly important market structure. It also describes the trading protocol on the New York Stock Exchange.

Chapters 2 to 5 present the basic models of market microstructure, based on asymmetric information and inventory control. Chapter 2 starts with rational expectations

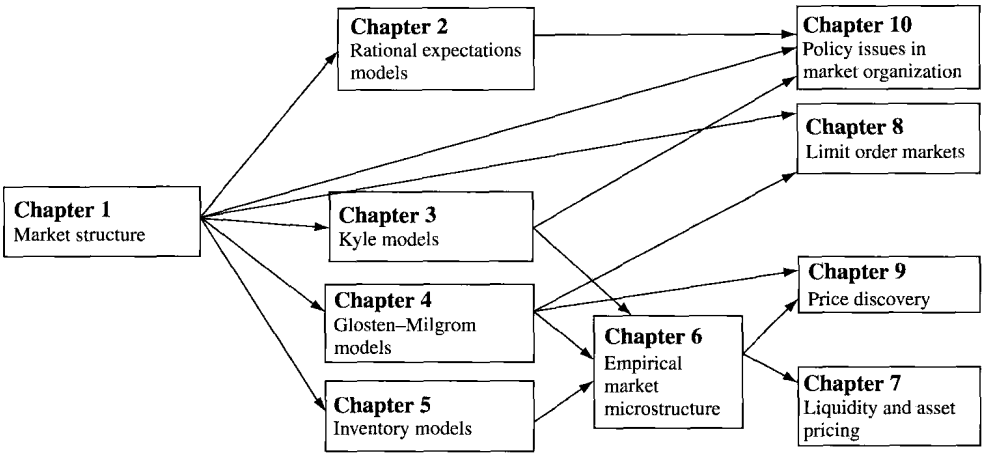

Figure Intro.1. Road-map of the book.

models for financial market equilibrium. Strictly speaking, these are not models of the workings of a financial market but illustrations of the equilibrium outcome that the trading process may produce. This chapter also introduces several useful tools, such as Bayesian updating and the rational expectations equilibrium. Chapter 3 discusses the Kyle (1985) model with an informed trader who maximizes trading profits by strategically choosing the size of his order. This model is the workhorse of theoretical market microstructure, and also indeed of many models in corporate finance. The model implies that order flow moves prices: buy orders drive prices up, sell orders down. Chapter 4 presents Glosten and Milgrom's (1985) model, which instead of Kyle's order-driven market considers a dealership market, showing that differences in information between traders create an adverse selection problem for the dealer. The dealers respond by driving a wedge between buying and selling prices, the bid–ask spread. Chapter 5 provides an alternative explanation for the bid–ask spread. Even if all traders have the same information, a spread may arise endogenously in the market. If traders come to the market randomly, a counterparty cannot be found immediately. The dealer absorbs the temporary imbalances in supply and demand, but if he is risk-averse, he has to be compensated for the risk of price fluctuations on the inventory: this compensation is provided by the bid–ask spread.

These chapters form the core of the market microstructure theory. The subsequent chapters build on these theories and develop empirical methods (Chapters 6 and 7) and offer more advanced, recent applications of market microstructure (Chapters 8 to 10).

Chapter 6 deals with empirical models and the estimation of transaction costs. First it discusses Roll's (1984) estimator; this model shows that transaction costs induce negative serial correlation in returns and provides a method to estimate bid–ask spreads using return data alone. The chapter then presents the Glosten and Harris (1988) model for estimating transaction costs from intraday price and trade data. The chapter also discusses several sources of intraday data and methods for estimating liquidity in

the absence of high-frequency data. Chapter 7 establishes the relevance of liquidity and transaction costs for asset pricing. This chapter builds a bridge between market microstructure and asset pricing.

The last three chapters contain more advanced and recent material. Chapter 8 focuses on developments of models that describe the working of the open limit order book (OLOB), the most common type of financial trading system. Chapter 9 is devoted to time-series models for price discovery, i.e. the empirical modelling of the convergence to the equilibrium price. It presents the econometric tools for the analysis of price and trade data, from one or several markets. Finally, Chapter 10 applies the theoretical models to the analysis of various policy issues in the organization of financial markets, with sections on transparency, dual capacity trading, and the debate on consolidation or fragmentation of trading.

The book can be used for a full semester course on financial markets and market microstructure. For a shorter course, we suggest taking Chapter 1 (market organization), Chapters 2–5 (theory), and Chapters 6–7 (empirical models and asset pricing).

Several other textbooks on market microstructure are available. O'Hara (1995) was the first, and it is chiefly theoretical in nature. More recent, and compared to the present book more specialized texts, are the following. Lyons (2001) focuses on the microstructure of the foreign exchange market. Harris (2003) gives a very detailed introduction to markets and trading structures. Hasbrouck (2007) gives a comprehensive overview of empirical methods in microstructure.

References

Admati, A. and P. Pfleiderer, 1988, 'A theory of intraday patterns: volume and price variability', *Review of Financial Studies,* 1, 3–40.

Amihud, Y. and H. Mendelson, 1987, 'Trading mechanism and stock returns: an empirical investigation', *Journal of Finance*, 42, 533–55.

Biais, B., L. Glosten and C. Spatt, 2005, 'Market microstructure: a survey of microfoundations, empirical results, and policy implications', *Journal of Financial Markets,* 8, 217–64.

Christie, W. and P. Schultz, 1994, 'Why do NASDAQ market makers avoid odd-eighth quotes?', *Journal of Finance*, 49, 1813–40.

Christie, W., J. Harris and P. Schultz, 1994, 'Why did NASDAQ market makers stop avoiding odd-eighth quotes?', *Journal of Finance*, 49, 1841–60.

Glosten L. and E. Harris, 1988, 'Estimating the component of the bid/ask spread', *Journal of Financial Economics*, 21, 123–42.

Glosten L. and J. Milgrom, 1985, 'Bid, ask and transaction prices in a specialist market with heterogeneously informed traders', *Journal of Financial Economics*, 14, 71–100.

Harris, L. 2003, *Trading and Exchange*, Oxford University Press.

Hasbrouck, J. 2007, *Empirical Market Microstructure*, Oxford University Press.

Hong, H. and J. Wang, 2000, 'Trading and returns under periodic market closures', *Journal of Finance,* 55, 297–354.

Huang, R, and H. Stoll, 1996, 'Dealer versus auction markets: a paired comparison of execution costs on NASDAQ and the NYSE', *Journal of Financial Economics,* 41, 313–57.

Jain, O and G. Joh, 1988, 'The dependence between hourly prices and trading volume', *Journal of Financial and Quantitative Analysis*, 23, 269–83.

Kyle, A. 1985, 'Continuous auctions and insider trading', *Econometrica*, 53, 1315–35.

Lyons, R. 2001, *The Microstructure Approach to Exchange Rates*, Cambridge, Mass.: MIT Press.

Madhavan, A. 2000, 'Market microstructure: a survey', *Journal of Financial Markets*, 3, 205–58.

O'Hara, M. 1995, *Market Microstructure Theory*, Oxford: Blackwell.

Roll, R. 1984, 'A simple implicit measure of the bid–ask spread in an efficient market', *Journal of Finance*, 39, 1127–39.

Stoll, H. 2000, 'Friction', *Journal of Finance*, 55, 1479–1514.

Wood, R., T. McInish and J. Ord, 1985, 'An investigation of transactions data for NYSE Stocks', *Journal of Finance*, 40, 723–39.

1 Institutions and market structure

This chapter sets out the basics of institutions and market organization, concentrating on those elements which are most useful in discussing the theory. The book by Harris (2003) provides an excellent reference manual and where further clarification is required or if more in-depth examination is desired, we refer the student to that text.

1.1 Trading protocol

The organizational structure of a financial market comprises the rules that regulate trading and the trading procedures. The market structure indicates who can access the trading venue, which instruments can be traded, the location of the trading sessions and the order routing systems. The market structure also spells out the rules governing the access of market participants to public and private information and so affects the degree of pre- and post-trade transparency decisively. It follows that the market structure has an impact on agents' trading strategies and therefore on market quality.

This chapter discusses how different market structures can be classified, according to both the order execution systems and to the type of trading sessions. As far as execution systems are concerned, markets can be classified as either order-driven or quote-driven. Markets based on the direct interaction of agents' orders are commonly called order-driven, whilst those where contracts must be fulfilled through intermediaries are called quote-driven; markets where both systems are in use are called hybrid. As will be discussed later in this chapter, most financial markets today have a hybrid protocol, as it is quite flexible in accommodating the liquidity needs of different instruments.

1.1.1 Order-driven markets

In order-driven markets the prices at which contracts are executed may be determined either at the same time the orders are transmitted to the market-place or afterwards. Investors' buy and sell orders are matched directly, without intermediaries. Liquidity is guaranteed by a constant flow of orders from market participants. There are no designated market-makers. The only intermediary on the market is the broker, who transmits clients' orders, but does not take own positions in the asset traded. All order-driven

Order-driven markets

Auction markets

Order matching rules: order precedence rules and trade pricing rules

- ➢ *Call (or batch) markets*
- ○ *Oral auction (open-outcry in floors or pits)*
- ○ *Electronic auction*
- ➢ *Continuous market (limit order book)*

Crossing networks

Order matching rules: order precedence rules and derivative pricing rules

Quote-driven markets

- ➢ *Screen-based markets – dealer markets*
- ➢ *Continuous auction markets*
- ➢ *Brokered markets*

Hybrid markets

Figure 1.1. Market structures.

markets are based on order precedence rules that rank and match orders for execution, but they do not all use the same pricing rule. As shown in Figure 1.1, these markets can be organized either as auctions or as crossing networks. Auction markets are based on specific trade pricing rules, while crossing networks use prices determined in other markets.

Auction markets The prevailing type of market in the leading financial centres at present is the auction. Auction markets can take two forms: call markets or continuous auctions. In call (or batch) auctions, orders are submitted simultaneously, whereas in continuous auctions agents can submit orders at any time during the *trading phase*.

In auction markets, there are two basic types of order – *limit orders* and *market orders*. A limit order specifies both a quantity and a maximum price or minimum price for execution depending on whether it is a buy or a sell order. A market order only specifies a quantity and will be executed at the best price available. A market order will always be executed as long as there is sufficient supply or demand, but the price may be unfavourable. Limit orders have a guaranteed price, but immediate execution is not guaranteed unless there are matching orders on the opposite side of the market. If immediate execution is not possible, the limit order will be placed in the *limit order book* and remain there until it is either executed against a new incoming order or cancelled. Limit orders are valuable for patient traders who wait for favourable prices, but they carry the risk of non-execution. Market orders are more suitable for impatient traders who want to execute the order immediately and are less sensitive to price.

In *call markets,* orders entered for the call can be publicized either orally or electronically. Call markets are characterized by a high degree of pre-trade transparency since traders can observe each other's orders during the auction.

In *oral call auctions*, also termed open-outcry auctions, agents cry their offers to buy or sell face-to-face on a trading floor. The rules governing oral auctions are quite simple. Firstly, it is compulsory for traders to communicate both prices and executions publicly. Secondly, traders with the highest bid and lowest ask gain precedence for order execution (*price priority*). Thirdly, traders who bid the best prices earlier gain precedence over those who bid the same prices later (*time priority*). Order matching rules are hierarchical, with price priority primary and time priority secondary. There exist other secondary precedence rules that discriminate in favour of certain classes of agents, for example, the public versus the specialists on the floor of the NYSE (*class priority* or *public order precedence*). Equilibrium trade prices in oral auctions are determined either by an auctioneer who minimizes net demand imbalances, or by brokers on the floor who match orders for different investors. The trade price rule in oral auctions is therefore quite simple, with orders executed at the best quoted prices. The most prominent example of an open-outcry auction is the futures market organized by the Chicago Board of Trade.

In *electronic call auctions,* orders are submitted to a computerized system during a predetermined period of time, and all trades in a stock take place at the same time and at the same equilibrium price under the uniform pricing rule (see Figure 1.2).

The price rule governing *electronic continuous auctions* is, instead, discriminatory. Prices are formed over time as traders observe the order flows, and orders are executed one by one as submitted, at the available prices. Market participants observe past transactions before submitting an order. This kind of market is generally automatic and is

Oral auction (open-outcry auction)
- Rules of pre-trade transparency
- Order precedence rules (time and price priority, public precedence)
- Trade pricing rules governed by the best quoted prices

Electronic call auction, continuous auction and crossing networks
- *Order precedence rules*
  - Time priority
  - Price priority
  - Visibility and size restriction
- *Trade pricing rules*
  - Uniform price (electronic call)
  - Discriminatory price (continuous call)
  - Derivative price (crossing network)

Figure 1.2. Trading rules.

the most frequent mode of trading derivatives and stocks.[1] This is currently the most common form of order-driven market, and is structured as an open limit order book (OLOB). The first characteristic of a platform organized as an OLOB is the accumulation of orders in an electronic book. A limit order is registered in the book and executed when an order of opposite sign and identical or better price is entered. An 'at best' order is carried out at the best price available in the order book. Price and time priority rules govern the order book. The limit orders executed first are those with the best price (price priority). If there are several orders for a given price, the first one to be executed can be selected in a number of ways. The most common rule is time priority: first submitted, first executed. There are also other rules. Size priority considers the number of shares in the order. Another type of priority rule is based on origin; for instance, at the NYSE orders from the public are filled first, those from specialists only afterwards.

Continuous and batch markets each have their own specific advantages. Continuous markets provide immediacy, enhance intraday price discovery and allow for easy enforcement of priority rules. Batch markets should reduce execution costs and settlement noise and thus provide better price stabilization, especially for thinly traded stocks.

Electronic communication networks (ECN) also work as open limit order books. These are forms of the Alternative Trading System (ATS) and describe trading systems that bring buyers and sellers together for the electronic execution of trades. They are registered under the Securities and Exchange Commission regulation which governs special purpose trading facilities that are not exchanges. The Commission has defined an ECN as *any electronic system that widely disseminates orders entered into it by the subscribers to third parties, and permits such orders to be executed in whole or in part.* Subscribers to ECNs can be retail investors, institutional investors, market-makers and other broker-dealers. The definition specifically excludes internal broker-dealer order-routing systems and crossing systems. Examples of ECNs are INET, Archipelago, BRUT and ATTN.

Opening and closing auctions Batch auctions can be used at the opening, at the closing and during intraday trading halts. The way in which the opening and closing auctions function in the European exchanges is almost identical, whereas some differences characterize the American markets. The US markets are discussed in section 1.1.3. Here we outline the functioning of the European auctions.

The opening auction has three phases. The first is the pre-opening phase, where all agents submit proposals to a centralized body that acts as the auctioneer. During the pre-opening there are no trades, only order submissions. Investors can observe the orders submitted and the tentative clearing price, which the algorithm computes continuously. This process of price formation follows order matching and trade pricing

[1] Examples of derivative markets are MATIF and MONEP for France, EUREX for Switzerland and Germany and IDEM for Italy; examples of stock markets are Euronext, Deutsche Börse (XETRA) for Germany, MTA for Italy, LSE (SETS) for the UK and TSE (STP) for Japan.