



ELSEVIER

Journal of Financial Markets 5 (2002) 127–167

www.elsevier.com/locate/econbase

Journal of
FINANCIAL
MARKETS

Market architecture: limit-order books versus dealership markets[☆]

S. Viswanathan^{a,*}, James J.D. Wang^b

^a *Fuqua School of Business, Duke University, Box 90120 Durham, NC 27708-0120, USA*

^b *Thunderbird, The American Graduate School of International Management,
Glendale, AZ 85306, USA*

Abstract

We analyze the customer's choice with respect to a limit-order book, a dealership market, and a hybrid market structure that combines the two. The customer's sell order is competed for and divided among a finite number of risk-averse market makers. We present a general characterization of equilibrium in the limit-order book. We show that when the order flow has a linear hazard ratio, the limit order book is preferred by risk neutral customers. However, a risk averse customer will prefer to trade in a dealership market when the number of market makers is large. Further, for risk averse customers, the hybrid market structure can dominate the dealership market and the limit-order book. The results are driven by a tradeoff between two features of the equilibrium demand schedules: a bid-shading effect that operates differently in a limit-order book compared with a dealership market, and a zero-quantity bid-ask spread that is present in the limit-order book only. © 2002 Published by Elsevier Science B.V.

JEL classification: D43; D44; D82; G19

Keywords: Market microstructure; Dealership; Limit-order book

[☆] For many helpful comments, we thank Bruno Biais, Francesca Cornelli, Ananth Madhavan, Christine Parlour, Ailsa Röell, Duane Seppi, participants of the May 1998 NBER Microstructure Conference, the 1998 WFA Meetings in Monterey, California, the 1998 EFA Meetings in Fontainebleau, France, and the 1999 Econometric Society Meetings in New York City.

*Corresponding author. Tel.: +1-919-660-7784; fax: +1-919-660-7971.

E-mail address: viswanat@mail.duke.edu (S. Viswanathan).

1. Introduction

An interesting regularity of stock markets is that small orders get executed via a limit-order book and large orders get executed via a dealership market. For example, on the New York Stock Exchange, small orders are filled in the downstairs market (a limit-order book with a specialist), while large orders are filled in the upstairs market (a dealership market). De Jong et al. (1995) document a similar phenomenon in Europe where small orders in French stocks have lower trading costs in the Paris Bourse, which is an electronic limit-order market, while large orders in the same stocks receive better execution in the London Stock Exchange, which is a nonanonymous dealership market.¹

Markets that are exclusively dealership (e.g., NASDAQ and London) have faced continued criticism in their execution of small orders, indicating that there is a perception that an exclusive dealership market is not the efficient mechanism for executing small orders. For example, work by Christie and Schultz (1994) and by Huang and Stoll (1996) suggests that the costs of trading are much higher for NASDAQ-traded stocks than for NYSE-traded stocks. The controversy over the poor quality of executions on NASDAQ led the SEC to require NASDAQ market makers to display customer limit orders that better market maker quotes (this rule was effective on January 20, 1997).² While London has lower execution costs than NASDAQ (see Hansch et al., 1998), it has also moved to electronic order book trading for the FT100 stocks as of October 20, 1997.³ Inter-market competition, especially the poorer executions offered in London for small orders (relative to other European exchanges) led to these changes. London now uses an electronic order book for smaller orders, while large orders are still routed through a dealership mechanism. These changes indicate that at least for smaller customer orders there is a preference for the limit-order book over the dealership market.

Markets that involve primarily institutional trading such as the foreign exchange market, the bond market, and the when-issued market in U.S. Treasury securities are still organized as dealership markets. Thus large-sized order flow is primarily executed in dealership markets across many securities.

We believe that this regularity is not an accident and that the division of order flow between these two different market structures poses an important challenge for the theory of market microstructure. As we argue below, the extant market microstructure theories have (with few exceptions) focused on

¹Section 4 contains more discussion on empirical evidence in this regard.

²Market makers have the choice of displaying customer quotes as the best bid or executing against the limit order at the customer's quoted price or shipping the order to an electronic crossing network (ECN) to be crossed at the customer's quoted price.

³See London Stock Exchange press release at the web site <http://www.londonstockex.co.uk>.

understanding the nature of trading in a given market and not on the choice of market structure itself.

The fundamental questions that we seek to answer in this paper are the following: Under what circumstances would a customer prefer a limit-order book to a dealership market and vice versa? Under what circumstances would a hybrid market that allows for a limit-order book for small orders and a dealership structure for large orders improve upon the exclusive dealership or limit-order book structure? Thus our emphasis is on the comparison of customer welfare across the three market structures being considered: limit-order books, dealership markets and hybrid markets. Such comparative results may shed new light on important issues such as why the pure dealership structure is perceived as being a poor market for small orders, and why the limit-order book is inefficient in handling large order sizes.

We consider a model where a supplier of order flow (to fix the discussion, we will focus on the bid side and consider the “customer” as a seller) has orders that vary in size from a minimum to a maximum. A finite number of risk-averse market makers (the “dealers”) compete for this order flow by submitting demand-supply curves or bid-ask schedules, i.e., bids and asks for every quantity. In equilibrium, the market makers’ trading strategies account for the pricing and order-routing rules adopted by the exchange. Focusing on the bid side, the equilibrium bids for various quantities are at prices less than the marginal valuation of the underlying asset, which would be the price in a competitive market. We term the difference between the marginal value of the asset (which is the competitive price) and the bid (at any given quantity) as bid-shading. This bid-shading occurs because of the strategic market power of the market makers (there are only a finite number of them). While this bid-shading is present in both the dealership market and the limit-order book, it operates differently in the two market structures. Specifically, the amount of bid-shading is increasing in the quantity obtained in a dealership market and is decreasing with quantity in a limit-order book.

In a dealership market, a customer’s order is filled at a single price — the market clearing price. The price that a market maker bids for a given quantity does not affect the price bid at other quantities. In equilibrium, market makers engage in bid-shading that is increasing in quantity and independent of the distribution of order flow. In a limit-order book, the market makers pay bid prices starting from the highest price on their books and move down their bid schedules until they reach the market clearing price. The price bid in a limit-order book for a given quantity affects the total payment when the market maker is allocated an equilibrium quantity higher than that given quantity. This effect tends to make market makers bid more aggressively in the limit-order book and implies that the equilibrium bid schedules are flatter (and thus the demand function, the inverse of the bid schedule, is more price-elastic) in the limit-order book than in the dealership market. Since the bid for the first

marginal unit in the limit-order book affects the total payment when any equilibrium quantity is allocated to the dealer, the market maker bids below his marginal valuation on this first marginal unit, leading to zero-quantity price discount (and thus a zero quantity bid–ask spread). In contrast, the bid–shading in the dealership market is increasing in quantity and at zero quantity, there is no bid–ask spread.

We focus on the ex ante comparison of market architecture, i.e., we compare expected customer welfare across different market structures. We find that the limit-order book is preferred by all risk-neutral customers. While the book involves a zero-quantity spread, it provides better prices at large orders and this suffices to ensure its dominance. However, when customers are risk-averse, the dominance of the limit-order book over a dealership market ceases to be true. In fact, the dealership market becomes the customer's preferred choice when the number of market makers is large. As the number of market makers increases, the demand curve (the inverse of the bid schedule) in a limit-order book becomes steeper while the demand curve in a dealership market becomes flatter. This exacerbates the level of price variation in the limit-order book and thus expected customer welfare is higher in the dealership market.

When the dealership market dominates the limit-order book, we find that the hybrid limit-order/dealership structure improves on the dealership market from the customer's perspective. In a hybrid market, limit orders are only accepted for quantities that are smaller than some exogenously fixed level. Orders greater than this cutoff level cannot be executed in the book and have to be routed to the dealership market. Restricting the size of trades on the limit-order book reduces the zero-quantity spread and, as a result, the price paid for the small orders is higher as the limit-order schedules shift upward. Due to the small variation in size for orders filled in the book, price variation is also limited. Consequently, the hybrid limit-order book/dealership market improves the expected price received by the customer without increasing his price execution risk and hence generates higher customer utility than the pure dealership market. Similarly, we show that as long as the number of market makers is not too small, the hybrid market improves upon the limit-order book. Our results provide a justification for the recent move to a hybrid market structure (small orders go to a limit order book, large orders go to a dealership market) on the NASDAQ and the London equity markets.

Related to the issues discussed here, Glosten (1994) presents a model of the limit-order book where the exact order size is unknown and cannot be contracted upon. Hence a provider of order flow can divide orders if she so desires (this is what he refers to as an anonymous market). Consequently, market makers have to condition on the event that the order they are executing is of larger size than their position in the limit-order book. This tends to induce market makers to flatten their submitted bid schedules. Importantly, Glosten shows that this anonymity of markets (and the consequent ability to divide

order flow) is a problem for dealership markets but not for the limit-order book.⁴ Thus, he argues that limit-order books possess a robustness property in anonymous markets that dealership markets do not possess. Seppi (1997) analyzes a limit-order book with discreteness where a monopolist specialist can step ahead of the limit orders.

We differ from Glosten (1994) and the literature that follows⁵ in several important ways. First, we consider a world where all market makers compete strategically. In contrast, Glosten allows only for monopolistically competitive limit-order submitters while Seppi (1997) only allows the specialist to play strategically (the limit-order submitters are competitive). Second, we focus on the welfare of the supplier of order flows under these market settings. In addressing the question of which market structure provides better execution for a given provider of order flow, we draw on a parallel literature on divisible good auctions due to Wilson (1979), Kyle (1989), Back and Zender (1993), Wang and Zender and Viswanathan and Wang (1996).⁶ When a customer sell order of sufficiently large size reaches the limit-order book market, it will “march down” the bid schedule contained in the book and transactions will take place at multiple prices. Therefore, a discriminatory auction is used to capture this multi-price feature of the book. On the other hand, in a dealership market, all orders are executed at the market clearing price. Thus, the dealership market is well approximated by a uniform-price auction.⁷ Applying the theory of divisible good auctions, we present a systematic approach to solving for equilibria in dealership markets and in limit-order books.

While the auctions literature serves as our starting point, the emphasis of this paper is different. First, the existing auction literature typically analyzes the incentives of a risk neutral seller who sells one unit of the good; in contrast we

⁴ Bernhardt and Hughson (1997) analyze the multiplicity of equilibria that arises when order flow is divisible. See also Rock (1996).

⁵ In a paper that is closer to ours in its focus, Parlour and Seppi (1997) compare the limit-order/specialist market structure in Seppi (1997) with the pure limit-order book in Glosten (1994). In contrast, our paper involves a comparison between a limit-order book, a dealership market, and a hybrid limit-order/dealership market structure. Pirrong (1999) considers a model that focuses on the competition between exchanges.

⁶ Bernheim and Whinston's (1986) work on menu auctions is related to this literature. Maskin and Riley (1989) present an optimal mechanism designs viewpoint towards multi-unit auctions with private values.

⁷ We recognize that dealership markets and limit-order books differ along many dimensions (tick size, anonymity, visibility of order flows, post-trading reporting, and retrading opportunity for the liquidity providers). In this paper we choose to highlight their different pricing rules. In many dealership markets, the incoming customer order is filled by a single market maker who then retrades with other market makers using an inter-dealer trading system. In another paper (Viswanathan and Wang, 2001), we analyze the purpose and efficiency of such two-stage trading mechanisms and compare them with the limit-order book.

focus on the case of a risk-averse supplier of order flow who has a random supply.⁸ Second, we study market structures that involve limit orders up to a certain critical size and a dealership market for larger sizes. As we have argued above, equity markets seem to be converging to this hybrid market structure, and one aim of this paper is to explore why such hybrid market structures are viewed as superior to a pure dealership market.

Three papers, Biais et al. (1998), Biais et al. (2000), and Röell (1998), are related to our paper.⁹ The first paper considers several market structures when the order size is known. In such an environment, the limit-order book has a unique efficient outcome, while a dealership market has multiple inefficient equilibria. Our paper and the other two cited papers consider market structures where the exact order size is unknown to the market makers when they submit their bidding schedules. The paper by Biais et al. (2000) is closer to our paper in that it considers a limit-order book where the marginal valuation is downward-sloping due to adverse selection. In contrast, we follow an approach where the marginal valuation is downward-sloping because of the risk-aversion of the market makers (inventories). However, the characterization they obtain for bid-shading is similar. In fact, we prove in our paper that in the linear marginal valuations case, the “information” and “inventory” approaches are observationally equivalent. Röell (1998) also considers the comparison between limit-order books and dealership markets using the linear equilibrium under the exponential distribution (unbounded support) and finds that the limit-order book dominates the dealership market (for both the “information” and “inventory” approach). Our paper differs from hers in three ways. First, our characterization of the necessary and sufficient conditions for equilibrium in the limit-order book holds for arbitrary distributions. Second, for specific comparisons between limit-order books and dealership markets, we focus on distributions that admit linear hazard ratios over bounded support; in contrast, Röell (1998) focuses on the exponential distribution, which has constant hazard ratios and unbounded support. Third, we consider hybrid market architectures.

The rest of the paper is organized as follows. In Section 2, we characterize the equilibrium strategies in a dealership market and in a limit-order book. The choice of market structure from the customers’ perspective is analyzed in Section 3. We start with a comparison of an exclusive dealership market versus

⁸Quantity uncertainty from the bidders’ perspective is a feature of several divisible good auction models that allow for noncompetitive (or quantity) bids. However, in these models the seller is only concerned with the distribution of prices and not the distribution of both prices and quantities as the total quantity sold is fixed. That is not true when the seller has a noisy supply, and hence we obtain different results.

⁹The work of Ho and Stoll (1983) and of Biais (1993) is based on unit-demand auctions and thus cannot be used to address the issues considered in this paper.

an exclusive limit-order book. We then consider a hybrid market where small orders get executed in a limit-order book and larger orders go to dealership market. In Section 4 we study several important extensions of the model. We also relate the model's findings to the empirical evidence on transaction costs in different market settings. Section 5 concludes.

2. The model

For exogenous reasons, a customer comes to the exchange floor and submits a sell order¹⁰ of size $\tilde{z}$.¹¹ We assume that the customer has a mean-variance preference:

$$U(\tilde{w}) = E[\tilde{w}] - A \text{Var}[\tilde{w}], \quad (1)$$

where $\tilde{w}$ is the proceeds from the sale, and A parameterizes his disutility from uncertainty.

The market makers act to maximize a mean-variance derived utility of profit with the risk aversion parameter, ρ .¹² There are $N > 2$ risk-averse market makers, each submitting demand curves competing for the customer's order. That is, these demand curves, $x_i(p)$, $i = 1, 2, \dots, N$ (or, equivalently, their inverse functions are the bid schedules, $p_i(x)$) represent the market makers' trading strategies. Since we assume that the underlying security is perfectly divisible, we will only consider piecewise continuously differentiable demand functions in this paper.

The market clearing price is the highest price at which there is no excess demand:

$$p = \begin{cases} \max \left\{ p \mid \sum_{i=1}^N x_i(p) \geq z; p \geq 0 \right\} & \text{if } \left\{ p \mid \sum_{i=1}^N x_i(p) \geq z; p \geq 0 \right\} \neq \emptyset, \\ 0 & \text{if } \left\{ p \mid \sum_{i=1}^N x_i(p) \geq z; p \geq 0 \right\} = \emptyset. \end{cases}$$

The prices at which the customer's order will be filled depend on the market structure in place. In a dealership market, a customer's order is filled at a single

¹⁰Buy orders can be analyzed analogously. For this paper, however, we will maintain the assumption that $N\bar{v} \gg \rho\tau_v^{-1}$, in order to restrict the discussion to just one side of the market, i.e., the customer sells and the market makers buy.

¹¹In Section 4.3 we allow for the customer to choose the selling quantity endogenously and show that the qualitative results remain the same.

¹²For the dealership market, this quadratic objective function can be derived from the standard combination of constant absolute risk aversion utility functions and normally distributed asset values. For the limit-order book market, these assumptions do not translate into a quadratic objective function in the dynamic optimization problem. In Section 4.2, we state and prove a characterization of the equilibrium strategies for the limit-order book when each market maker maximizes a CARA utility function rather than a mean-variance objective function.

price – the market clearing price. In a limit-order book, the market makers pay bid prices starting from the highest price on their books and move down their bid schedules until they reach the market clearing price.

The above description of the pricing rules is reminiscent of the nature of auction rules adopted in the sales of divisible goods. In fact, the analogy between a uniform-price auction and a dealership market, and that between a discriminatory auction and a limit-order market, will be exploited in our subsequent analysis.

For the random asset value, we use $E[\tilde{v}] = \bar{v}$ and $\text{Var}[\tilde{v}] = \tau_v^{-1}$ to denote its mean and variance, respectively. The quadratic form of the derived utility implies that the market maker's marginal valuation is linear in his quantity allocation. The other random variable of the model, $\tilde{z}$, is assumed to be distributed independently of $\tilde{v}$. As we will show later, the precise distributional properties of $\tilde{z}$ are not important for equilibrium in a dealership market. For a limit-order book, however, the distribution of $\tilde{z}$ matters and we present a general characterization of the solution.

Whenever explicit calculation is required, we assume that $\tilde{z} \in [0, 1]$ has a density function and a cumulative density function of the following forms:¹³

$$g(z) = \frac{1}{\theta}(1-z)^{(1-\theta)/\theta}, \quad (2)$$

$$G(z) = 1 - (1-z)^{1/\theta}, \quad (3)$$

where the parameter $\theta \in (0, \infty)$ is related to the mean and variance of the distribution as follows:

$$E[\tilde{z}] \equiv \bar{z} = \frac{\theta}{1+\theta},$$

$$\text{Var}[\tilde{z}] \equiv \sigma_z^2 = \frac{\theta^2}{(1+\theta)(1+3\theta+2\theta^2)}.$$

We note that the case of $\theta = 1$ corresponds to the uniform distribution.¹⁴ It can be shown that the above class of densities is the only class of bounded distributions for which the hazard ratio is a linear function of the underlying random variable. Specifically, the hazard ratio for the above distribution is:

$$H(z) = \frac{1-G(z)}{g(z)} = \theta(1-z). \quad (4)$$

¹³Note that the choice of $[0, 1]$ as support is without loss of generality. We could choose $[0, s]$ as the support and the corresponding cdf would be $G(z) = 1 - (1 - z/s)^{1/\theta}$.

¹⁴To facilitate welfare comparison of the hybrid market with the dealership and the limit-order book markets, we specialize to the uniform distribution for $\tilde{z}$ in Section 3.2.

This property will be exploited when we search for linear solutions in the limit-order book.

We assume that all market makers are similar in all regards, and thus, only symmetric strategy equilibrium is sought.

2.1. Characterizing the equilibrium strategies

To study the dealership market and limit-order book within the same analytical framework, we introduce the price discrimination parameter $\alpha \in \{0, 1\}$. Specifically, market maker i 's uncertain bidding profit, using strategy $x_i(p)$, is:

$$\tilde{\pi}_i = \tilde{v}x_i(p) - (1 - \alpha)px_i(p) - \alpha \int_0^{x_i(p)} p(q) dq.$$

The case of $\alpha = 1$ corresponds to the limit-order book, whereas the case of $\alpha = 0$ corresponds to the dealership market. For the model above between N market makers and the customer, the equilibrium strategies used by the market makers are the solutions to an ordinary differential equation as described below.

Proposition 2.1. When the market makers act to maximize a mean-variance derived utility, the following is a necessary condition for any symmetric strategy Nash equilibrium in demand functions that are piecewise continuously differentiable:

$$\sum_{j \neq i}^N x_j'(p) = -\frac{(1 - \alpha)x_i(p) + \alpha H(z)}{\bar{v} - p - \rho\tau_v^{-1}x_i(p)}, \quad (5)$$

where $H(z)$ is the hazard ratio for $\tilde{z}$.

Eq. (5) is also sufficient for a symmetric strategy equilibrium when the following condition holds for all $z > \hat{z}$ and $\hat{p}(z) = p(\hat{z})$:

$$\alpha H(z) > \alpha H(\hat{z}) - \frac{\rho\tau_v^{-1}(z - \hat{z})[(1 - \alpha)x_i(\hat{p}(z)) + \alpha H(\hat{z})]}{\bar{v} - \hat{p}(z) - \rho\tau_v^{-1}x_i(\hat{p}(z))}. \quad (6)$$

Further, when $\alpha = 1$ (limit-order book), there is an unique equilibrium with the property that $p(z) \rightarrow \bar{v} - (\rho\tau_v^{-1}/N)$ as $z \rightarrow 1$.¹⁵

Proof. See Appendix A. $\square$

Eq. (6) is always true for $\alpha = 0$. Therefore, solutions to Eq. (5) for the uniform-price auctions are indeed equilibrium strategies. In addition, we can

¹⁵All other downward sloping equilibria of the limit-order book have the property that the price at the highest customer order, $z = 1$, is $-\infty$.

also show that all linear solutions to Eq. (5), for both uniform-price and discriminatory auctions, satisfy the sufficient condition in Eq. (6).¹⁶ More generally, Eq. (6) holds if the hazard ratio of the random variable $\tilde{z}$ does not decrease too rapidly.¹⁷

Equation (5) can be rewritten to express the marginal price bid for quantity x as follows:

$$p = \bar{v} - \rho \tau_v^{-1} x_i(p) + \frac{(1 - \alpha)x_i(p) + \alpha H(z)}{\sum_{j \neq i}^N x'_j(p)}. \quad (7)$$

The first two terms are just the market makers' marginal valuation of the underlying risky asset. Note that we focus on customer sales; thus Eq. (7) represents the market makers' bid function. In a perfectly competitive market, the market makers bid their marginal values. Since we focus on downward-sloping demand curves, the last term is always negative and thus represents the amount of *bid-shading* relative to the marginal values.

The extent of bid-shading is different depending on whether trading takes place in a dealership market or in a limit-order book. From Eq. (7), the bid-shading is $x_i(p)/[\sum_{j \neq i}^N x'_j(p)]$ in a dealership market ($\alpha = 0$). Thus, there is always bid-shading when $x_i > 0$ and there is no bid-shading at $x_i = 0$, i.e., there is no zero-quantity price discount. In a limit-order book market, the bid-shading is measured by $H(z)/\sum_{j \neq i}^N x'_j(p)$. Therefore, it is zero for the largest quantity that can be obtained in equilibrium (because $H(z) = 0$ at $z = 1$). However, in a limit-order book, bid-shading always exists at zero quantity (because $H(z) > 0$ at $z < 1$). We refer to this price discount as the (bid side of) bid–ask spread at zero quantity.¹⁸

With linear bidding schedules (i.e., when $x'(p)$ is a constant across all market makers in equilibrium), the above discussion leads to the conclusion that the extent of bid-shading is increasing along the quantity axis for a dealership market and decreasing for a limit-order book. Therefore, the equilibrium demand function (the inverse of the bid schedule) should be more price-elastic in a limit-order book than in a dealership market. At the same time, bidding in the limit-order book is characterized by a price discount at zero quantity.

The preceding discussion demonstrates the existence of a fundamental tradeoff between the zero-quantity spread and the degree of price competition in comparing a limit-order book with a dealership market (see Fig. 1). As we will show, the main results of the paper can be understood in view of this

¹⁶ See Appendix A for a direct check on the second-order condition.

¹⁷ Hence an increasing hazard ratio is sufficient but not necessary for existence.

¹⁸ Symmetrically, a similar effect exists on the ask side. This is what Glosten (1989) refers to as zero-quantity bid–ask spread. The results in Glosten (1989) are similar to those for our limit-order book except that the response of the other limit order submitters does not exist in his monopoly specialist model.

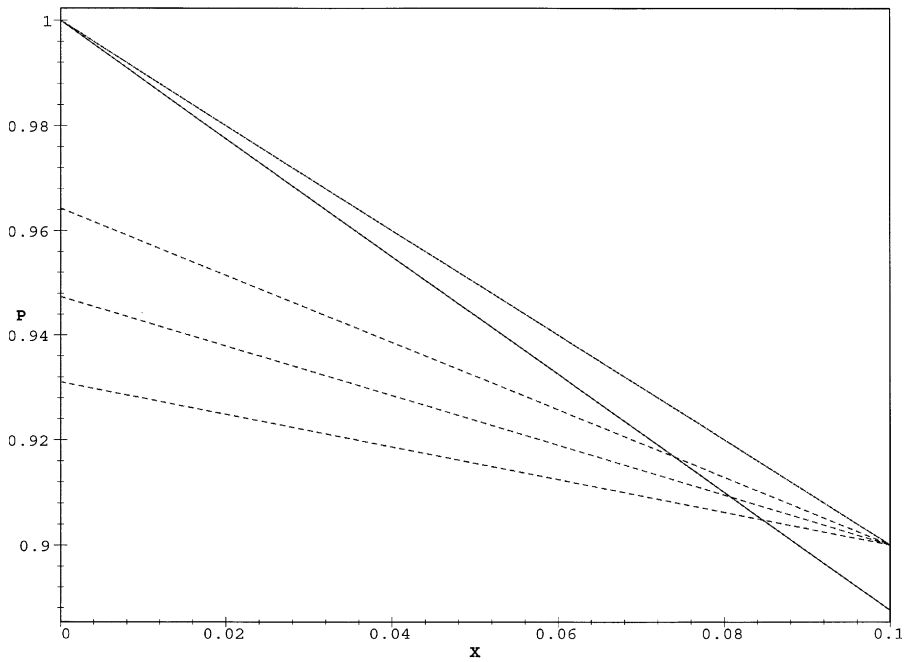


Fig. 1. Price versus quantity. The higher solid line is the dealers marginal valuation curve. The lower solid line is his inverse demand curve in the dealership market. The three dashed lines are his bid schedules (inverse demand curves) in the limit-order book market for different mean values of customer order size (from top to bottom, $\theta = 1/2, 1, 2$, respectively).

tradeoff which characterizes the general solutions of the model. To facilitate a welfare comparison of the different market settings, explicit linear solutions will be derived as special cases of Proposition 2.1, and they serve as the basis for addressing some important market architecture design issues.

2.2. Dealership market

The dealership market can be described as a special case of Proposition 2.1 when $\alpha = 0$.

Corollary 2.1. In a dealership market consisting of N market makers with quadratic derived utility, there is a symmetric, linear strategy equilibrium in which all market makers use the bid schedule:

$$p_i(x) = \bar{v} - \frac{\rho\tau_v^{-1}(N-1)}{N-2}x. \quad (8)$$

This equilibrium is unique within the class of equilibria with linear strategies.

The equilibrium demand curve for any individual market maker is independent of the precise distributional assumptions about the customer order. In fact, the equilibrium identified in Proposition 2.1 is unique among all possible equilibria characterized by symmetric, linear demand functions. Furthermore, this equilibrium possesses a “no-regret” property — no market maker would want to change his pricing behavior unilaterally even if he had the opportunity to do so after observing the actual customer order size. Both of these features speak to the robustness of the pricing equilibrium in a dealership environment.

Given the equilibrium strategy, it is straightforward to derive the expressions for the market clearing price and the quantity allocation for market makers:

$$\tilde{p} = \bar{v} - \frac{\rho\tau_v^{-1}(N-1)}{N-2} \frac{\tilde{z}}{N}, \quad (9)$$

$$\tilde{x}_i = \frac{\tilde{z}}{N}. \quad (10)$$

The customer’s total payoff from selling in a *dealership* market is:

$$\begin{aligned} \tilde{w}_d &= \sum_{i=1}^N \tilde{x}_i \tilde{p} = \tilde{z} \tilde{p} \\ &= \bar{v} \tilde{z} - \frac{\rho\tau_v^{-1}(N-1)}{N(N-2)} \tilde{z}^2 \\ &= b_d \tilde{z} - c_d \tilde{z}^2. \end{aligned} \quad (11)$$

As expected, $\tilde{w}_d \rightarrow \bar{v} \tilde{z}$ as $N \rightarrow \infty$. That is, as the number of market makers becomes large, the market becomes more competitive from the customer’s perspective. The difference between $\tilde{w}_d$ and $\bar{v} \tilde{z}$ corresponds to the rent extracted by the market making sector as a whole.

2.3. Limit-order book

With a limit-order book, the market maker’s profit is a function not only of the market-clearing price but also of other points on her bidding curves. In particular, the price bid for customer quantity x is also relevant for all customer quantities higher than x . As a result, the analysis of the limit-order book is more complicated than that of the dealership market.

Corollary 2.2. Assume the customer order, $\tilde{z}$, has a linear hazard ratio over the interval $[0, 1]$. In a limit-order book consisting of N market makers with quadratic derived utility, there is a symmetric, linear strategy equilibrium in which all

market makers use the bid schedule:

$$p_i(x) = \bar{v} - \frac{\rho\tau_v^{-1}[\theta + (N-1)x]}{N(1+\theta) - 1}. \quad (12)$$

This equilibrium is unique within the class of linear strategy equilibria. Further, this is the unique equilibrium with the property that $p(z) \rightarrow \bar{v} - \rho\tau_v^{-1}1/N$ as $z \rightarrow 1$.

Proof. See Appendix B. $\square$

Comparing Eq. (12) with Eq. (8), we note that the equilibrium bid schedule is flatter in the limit-order book than in a dealership setting. That is, the presence of price discrimination in the book intensifies the price competition among the liquidity providers. From the customer's perspective, there is also a cost associated with trading against the book since the book entails a zero-quantity discount, the size of which is directly proportional to the maximum size of customer orders.¹⁹

Using the above equilibrium strategy, we have the following expressions for the market-clearing price and the quantity allocation for market makers:

$$\tilde{p} = \bar{v} - \frac{\rho\tau_v^{-1}\theta}{N(1+\theta) - 1} - \frac{(N-1)\rho\tau_v^{-1}}{N(1+\theta) - 1} \frac{\tilde{z}}{N}, \quad (13)$$

$$\tilde{x}_i = \frac{\tilde{z}}{N}. \quad (14)$$

The customer's total payoff from selling against a limit-order book is equal to the integration of Eq. (13) from 0 to z :

$$\begin{aligned} \tilde{w}_b &= \left[\bar{v} - \frac{\rho\tau_v^{-1}\theta}{N(1+\theta) - 1} \right] \tilde{z} - \frac{\rho\tau_v^{-1}(N-1)}{2N[N(1+\theta) - 1]} \tilde{z}^2 \\ &\equiv b_b \tilde{z} - c_b \tilde{z}^2. \end{aligned} \quad (15)$$

Again, the customer's payoff from his sell order behaves intuitively: $\tilde{w}_b \rightarrow \bar{v}\tilde{z}$ as $N \rightarrow \infty$.

Notice that the behavior of the linear equilibrium solution as the number of market makers, N , changes, is different depending on the market structure. In the case of a dealership market, the equilibrium demand curves submitted by the market makers become flatter as N increases. This is quite intuitive since we expect the presence of more market makers to cause more intensified inter-dealer competition, which results in more elastic trading strategies. In contrast, the linear solution for the limit-order book becomes steeper as N becomes greater.

¹⁹ Suppose the support for $\tilde{z}$ is $[0, r]$ instead of $[0, 1]$; then it is straightforward to show that the solution changes to $\bar{v} - \rho\tau_v^{-1}[\theta r + (N-1)x]/[N(1+\theta) - 1]$.

To understand the comparative statistic results in N , we fix the market makers' total risk aversion by fixing ρ/N and then let N approach infinity.²⁰ This is equivalent to fixing the total market making capacity and examining the effect of having fewer or more market makers. As N is increased, in a dealership market, the equilibrium demand curve pivots around the highest price level, $p(\tilde{z} = 0)$, to become flatter. This is in accordance with the usual intuition that the presence of more market makers enhances competition and improves the terms of trade at all trade sizes. For a limit-order book, the effect of increasing N is to pivot the demand curve around the lowest price level, $p(\tilde{z} = 1)$, to become steeper. Again, more market makers provide for better prices for the customer at all trade sizes. In this case, even though the demand curve becomes more inelastic as N increases, the intercept of the demand curve moves higher to offset the effect of a steeper curve.

3. Comparing market architectures

We will focus on the comparison of market architectures under the premise that the customers' choice of an exchange floor is made prior to knowing their order size. Since firms list on exchanges and exchanges choose market architectures, we take the view that firms pick the market architecture that maximizes customer welfare.²¹ Hence, we analyze the customer's choice from an *ex ante* standpoint.

We first consider a customer who has to choose between a dealership market and a limit-order book. Based on the intuitions that are developed in this setting, we then explore the consequence of introducing a hybrid structure of the two mentioned markets.

3.1. Limit-order books versus dealership markets

Proposition 3.1. On an ex ante basis, a risk-neutral customer prefers to trade in a limit-order book instead of in a dealership market.

Proof. See Appendix C. $\square$

Fig. 2 plots the customer's revenue from his sell order as a function of the actual order sizes. Two aspects of Fig. 2 are worth noting. First, the limit-order book dominates the dealership market for large order sizes, and the reverse is

²⁰ We are grateful to Ailsa Röell for suggesting the following intuitive explanation.

²¹ We take the number of market makers as given. Pirrong (1999) presents a model where the number of members in an exchange is endogenous.

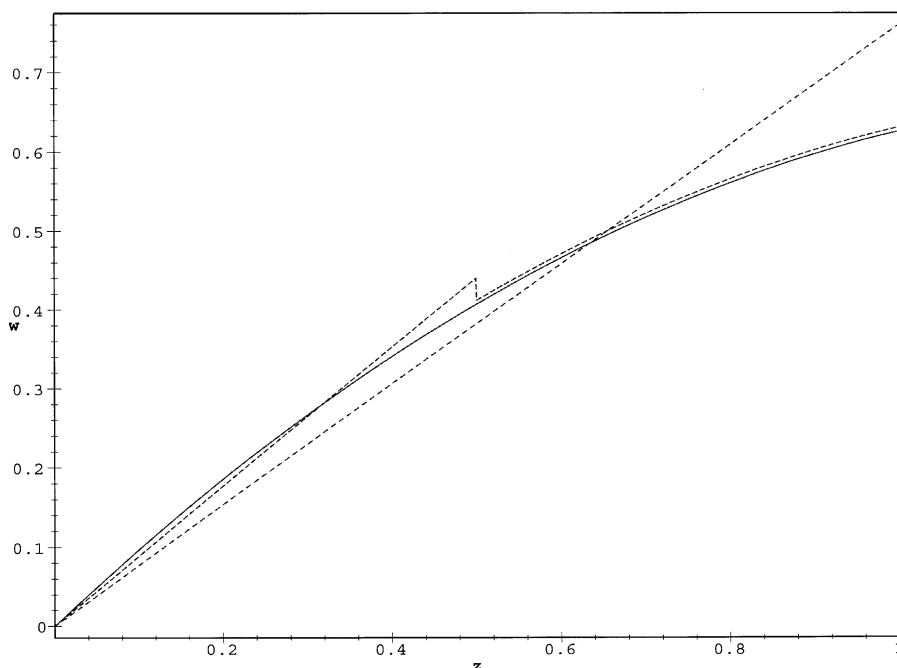


Fig. 2. Customers revenue, w , versus customer order size, z . The solid line is for dealership market. The smooth dashed line is for the limit-order book market. The kinked dashed line is for the hybrid market ($r = 0.5$).

true for small order sizes. Second, the customer's revenue as a function of $\tilde{z}$ has a larger curvature in a dealership market than in a limit-order book.²²

In a limit-order book, the market makers submit demand curves that are more price-elastic (i.e., the bid schedules are flatter) than the demand curves used in a dealership market. This leads to more heightened price competition among submitters of limit orders — a factor that tends to favor the book from the customer's viewpoint. At the same time, bid schedules in limit-order books contain a price discount which is proportional to the average order size. This second factor is what makes the limit order book less desirable for a given customer order size that is small.²³

²² This is because $c_d > c_b$ from Eqs. (8) and (12). It is related to the fact that the equilibrium bid schedule is flatter in a limit-order book, a point we discuss below.

²³ For very small order sizes, customers are selling at almost no discount in a dealership market, whereas customers must sell with a zero-quantity discount in a limit-order book market. At small order sizes, the issue of price elasticity is not important.

The steeper bid schedule in a dealership market causes the revenue function to be more concave, which in turn implies that at large order sizes the customers obtain less revenue in a dealership market than in a limit-order book. The customer's revenue in a limit-order book is initially (i.e., at small order sizes) smaller than the revenue in a dealership market. Because of the difference in the curvature of the revenue function, however, the revenue in a limit-order book actually becomes greater than that in a dealership market at larger order sizes. Proposition 3.1 states that the revenue function, when integrated over all customer order sizes, is *always* greater in a limit-order book market than in a dealership market.

Obviously, Proposition 3.1 does not preclude the possibility that a risk-averse customer could find the dealership market to be the preferred trading environment. The next result identifies the conditions under which a dealership market is preferred by the customer over a limit-order book market.

Proposition 3.2. There exist nonempty sets of parameter values such that, on an ex ante basis, a risk-averse customer prefers to trade in a dealership market instead of a limit-order market.

Proof. See Appendix D. $\square$

As the number of market makers increases, the equilibrium demand schedule in a limit-order book becomes steeper (less competitive) and the level of the price discount declines. Thus, these two effects tend to offset one another. In contrast, increasing the number of market makers always makes the demand curve in a dealership market flatter and thus intensifies bidder competition. Consequently, as compared to the book, the dealership market becomes more attractive to the customer as the number of market makers increases. We verified this observation in numerical computation and illustrated it in Fig. 3. (Fig. 4 provides a similar comparison with endogenous customer orders; see Section 4.3 for more details.)

3.2. Hybrid markets

As we have discussed in the introduction, equity markets have converged in recent years to a hybrid structure that has a limit-order book for small orders and a dealership market for large orders. This has always been the case with the NYSE, which has both the “downstairs” market (a specialist/limit-order book structure) and the “upstairs” market (a dealership market). On markets that were exclusively dealerships like the London Stock Exchange and NASDAQ, market competition (in the case of London) and regulatory pressure (in the case of NASDAQ) have resulted in the establishment of execution systems for smaller orders that resemble limit-order books.

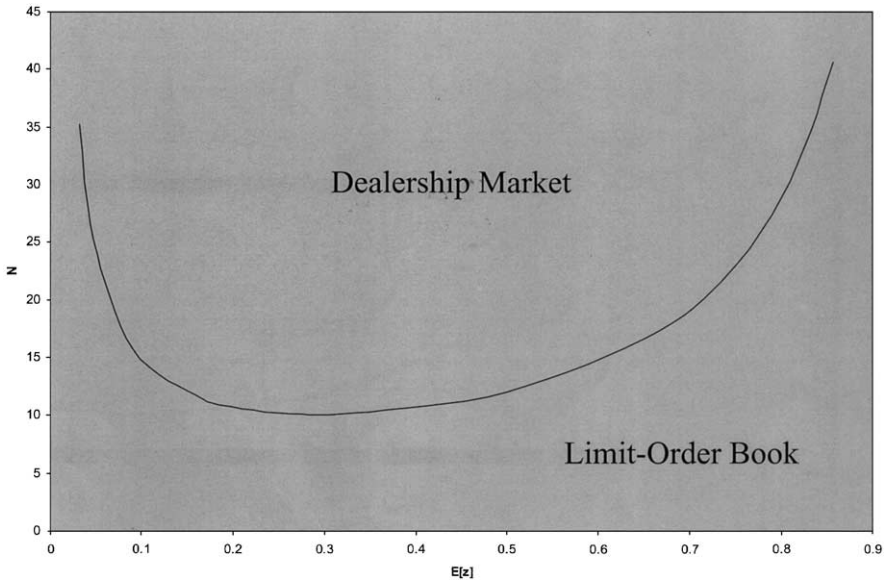


Fig. 3. The number of dealers, N , versus the mean customer order size, $E[z] = \theta/(1 + \theta)$. The area above the curve is where the customer prefers the dealership market over the limit-order book market. In this graph, the quantity the customer sells is determined exogenously.

Here we consider a hybrid market in which small orders (order sizes that are below a certain cutoff point) are cleared against a limit-order book and the remaining orders of greater sizes are routed to a dealership market. If such a cutoff point is announced and agreed upon by all market participants, it has the effect of changing the distribution of orders. An interesting issue to explore is whether such a hybrid market structure can improve on a pure limit-order book and a pure dealership market from the customer's perspective.

In the following analysis, we set $\theta = 1$, i.e., we consider the uniform distribution only. In a hybrid market where all customer orders of size less than $r \in [0, 1]$ are directed to a limit-order book, and orders greater than or equal to r are directed to a dealership market, the equilibrium demand curves are as follows:

$$p(x) = \begin{cases} \bar{v} - \frac{\rho\tau_v^{-1}}{2N-1}[r + (N-1)x] & \text{if } 0 \leq x < \frac{r}{N}, \\ \bar{v} - \frac{\rho\tau_v^{-1}(N-1)}{N-2}x, & \text{if } \frac{r}{N} \leq x. \end{cases} \quad (16)$$

The customer's payoff from trading in a *hybrid* market is:

$$\tilde{w}_h = \begin{cases} b'_b \tilde{z} - c'_b \tilde{z}^2 & \text{if } \tilde{z} \in [0, r), \\ b'_d \tilde{z} - c'_d \tilde{z}^2 & \text{if } \tilde{z} \in [r, 1], \end{cases} \quad (17)$$

where b_d, c_d are given in Eq. (11), and

$$b'_b = \bar{v} - \frac{\rho\tau_v^{-1}r}{2N-1}, \quad (18)$$

$$c'_b = \frac{\rho\tau_v^{-1}(N-1)}{2N(2N-1)}. \quad (19)$$

Therefore, the comparison of market structures from the customer's perspective reduces to a comparison of

$$E[U(\tilde{w}_h)] = \int_0^r U(b'_bz - c'_bz^2) dz + \int_r^1 U(b_dz - c_dz^2) dz \quad (20)$$

for a hybrid market, versus $E[U(\tilde{w}_d)]$ for an exclusive dealership market, and versus $E[U(\tilde{w}_b)]$ for an exclusive limit-order book market, where:

$$E[U(\tilde{w}_d)] = \int_0^1 U(b_dz - c_dz^2) dz, \quad (21)$$

$$E[U(\tilde{w}_b)] = \int_0^1 U(b_bz - c_bz^2) dz. \quad (22)$$

Proposition 3.3. On an ex ante basis, a risk neutral customer prefers to trade in a limit-order book instead of in a hybrid market, which in turn is preferred over a dealership market.

The above result can be shown straightforwardly by computing:

$$E[\tilde{w}_b] - E[\tilde{w}_h] = \frac{\rho\tau_v^{-1}(1-r^3)}{2(N-2)(2N-1)} > 0,$$

$$E[\tilde{w}_h] - E[\tilde{w}_d] = \frac{\rho\tau_v^{-1}r^3}{2(N-2)(2N-1)} > 0.$$

We observe that Proposition 3.1 can be viewed as a corollary of Proposition 3.3.

These results suggests that, from a risk neutral customer's viewpoint, the hybrid market is a monotonic interpolation between a dealership market ($r=0$) and a limit-order book ($r=1$) in the sense that $\forall 0 \leq r < r' \leq 1$, $E[\tilde{w}_h(r)] < E[\tilde{w}_h(r')]$.

For a risk-averse customer, the comparison of a hybrid market structure with a pure dealership market or a pure limit-order book is more complicated.

Proposition 3.4. There exist nonempty sets of parameter values such that, on an ex ante basis, a risk-averse customer prefers to trade in a hybrid market over both the dealership market and the limit-order book. In particular, the hybrid

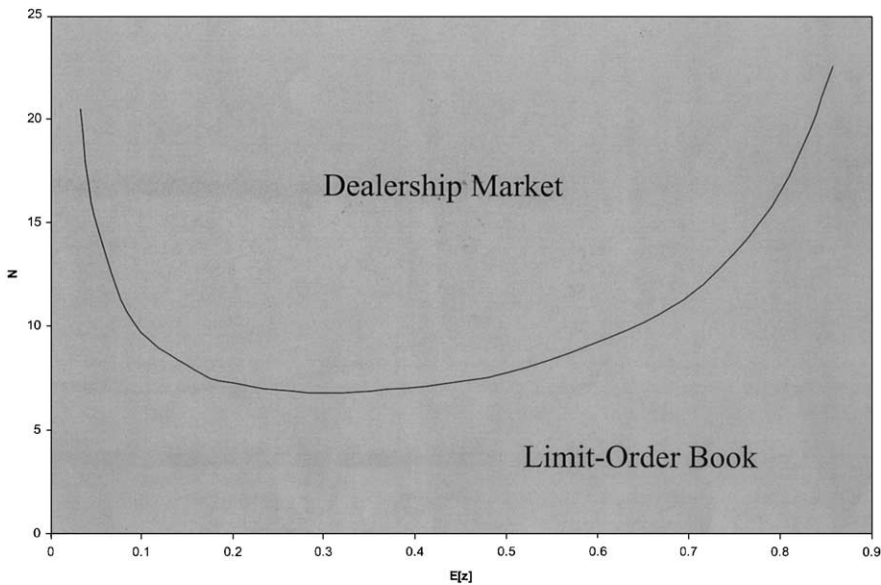


Fig. 4. The number of dealers, N , versus the mean customer order size, $E[z] = \theta/(1 + \theta)$. The area above the curve is where the customer prefers the dealership market over the limit-order book market. In this graph, the quantity the customer sells is determined endogenously.

market always dominates the dealership market for small values of the cutoff customer order size, r .

Proof. See Appendix E. $\square$

Proposition 3.4 is illustrated in Figs. 5 and 6. In Fig. 5, we show that a hybrid market with a relatively small cutoff level r always improves upon a pure dealership market. The reason is that a hybrid market and a dealership market are different only for customer orders that are smaller than r . If r is very small, the price discount effect associated with the limit-order book is negligible. Since the limit-order book always dominates the dealership market for these small orders, replacing a pure dealership market by a hybrid market with a small r must be welfare improving.

If the cutoff size, r , is sufficiently large, however, the hybrid market can dominate the dealership market for small values of N . According to Proposition 3.2, the limit-order book is a preferred choice when N is small. Thus, the hybrid market, which has the limit-order book features at below-cutoff order sizes, will be preferred by the customer over a pure dealership market.

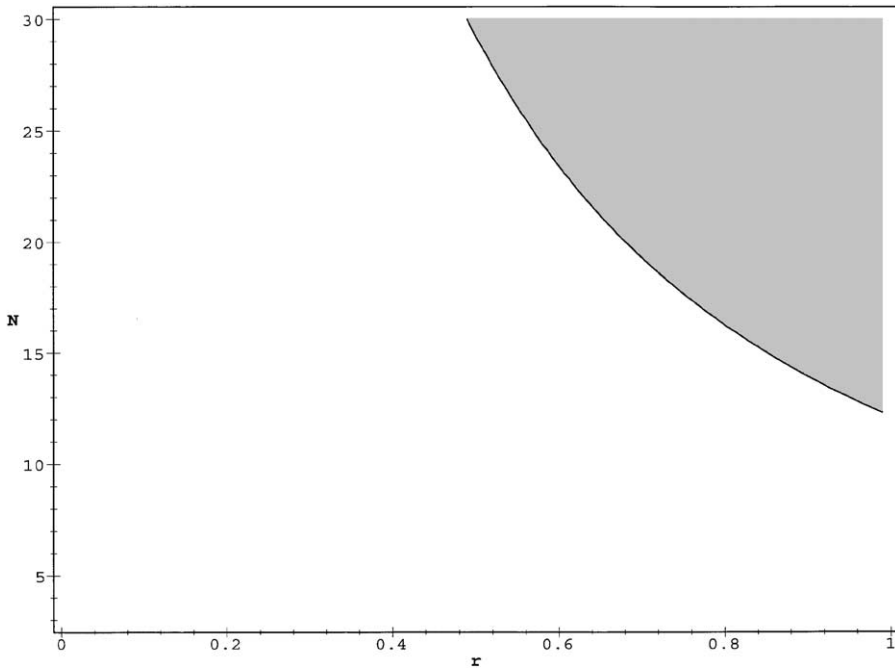


Fig. 5. The number of dealers, N , versus r , the cutoff customer order size in a hybrid market. The unshaded area is where the customer prefers the hybrid market over the pure dealership market.

The comparison of the hybrid market and the limit-order book is illustrated in Fig. 6. There, we show that the hybrid market can improve upon a pure limit-order book when the number of dealers, N , is large. Again, the intuition can be seen from the result in Proposition 3.2. The difference between a hybrid market and a pure limit-order book is similar to that between a dealership market and a limit-order book. Since the dealership market is expected to dominate a limit-order book when N is large, the hybrid market also is preferred over a pure limit-order book market under the same condition.

Proposition 3.4 demonstrates that, when structured appropriately, the hybrid limit-order book/dealership market generates higher trading revenue for the customer than the pure dealership market. Further, if the number of market makers is not too small, hybrid markets also dominate the limit-order book. We believe that these insights provide a rationale as to why many securities markets around the world are evolving toward a hybrid market structure.

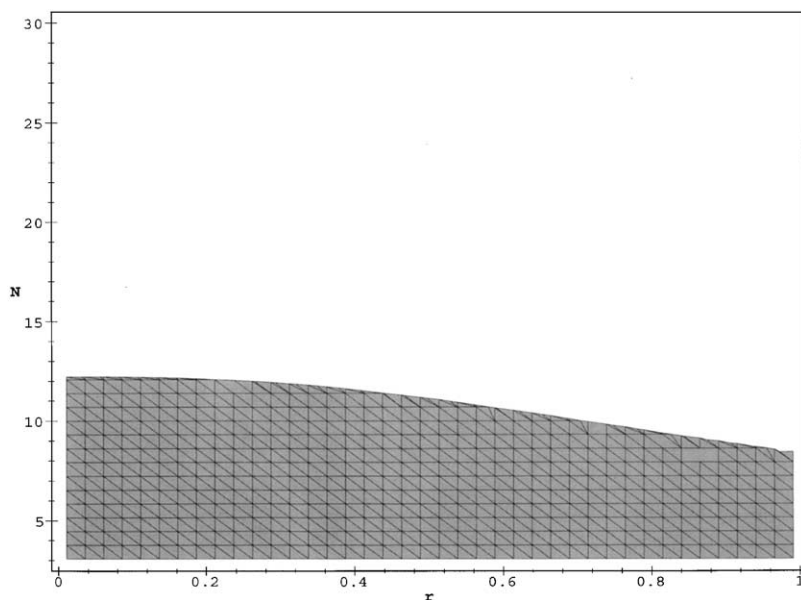


Fig. 6. The number of dealers, N , versus r , the cutoff customer order size in a hybrid market. The unshaded area is where the customer prefers the hybrid market over the pure limit-order book market.

4. Extensions

4.1. Extension I: customer information

Our results thus far belong to the “inventory paradigm”. An easy way to see this is by introducing an ex ante inventory (call it I) for all dealers. Then all previous results hold with the transformation $\bar{v} \rightarrow \bar{v} - \rho\tau_v^{-1}I$.

Informational advantage on the part of some customers is another often-cited source of “market imperfection” which prevents market makers from pricing in a fully competitive fashion. This approach is used in Glosten (1994) and in Biais et al. (2000). We show the observational equivalence of the linear information model and our inventory approach.²⁴

²⁴Despite the similarity in analytics, the underlying intuitions for the two models are not identical. For example, in the inventory model, an individual market maker's valuation depends on the share of quantity that he obtains. In the asymmetric information model, market maker's marginal valuation depends on the amount of total demand in the market.

Instead of positing risk-aversion, let us suppose that the (risk neutral) market makers use the following form of updating of the underlying asset value based on the actual customer order size, z :

$$E[\tilde{v}|\tilde{z} = z] = \bar{v} - \xi z \quad (\xi > 0). \quad (23)$$

Then, by tracing through the proof as outlined in Appendix A, it can be shown that the resulting equilibrium characterization is given by:

$$p = \bar{v} - \xi z + \frac{(1 - \alpha)x_i(p) + \alpha H(z)}{\sum_{j \neq i}^N x'_j(p)} \quad (24)$$

and that the sufficient condition is identical to Eq. (6).²⁵ Comparing the above relation with Eq. (7) reveals that this particular representation of private customer information is observationally indistinguishable from the inventory case. One need only make the transformation $\rho\tau_v^{-1} \leftrightarrow N\xi$ to go from one setup to the other. This is not surprising, as the main role of inventories or information is to make the marginal valuation curve downward sloping.²⁶ Hence comparisons that we make between limit-order books and dealership markets directly extends to the information model with a linear updating rule.

4.2. Extension II: CARA utility

In the preceding analysis we assumed that the market makers' objective function is of the mean-variance type. For the uniform-price auction, it can be shown that this is equivalent to the primitive assumptions of CARA utility and normally distributed asset values. For the discriminatory auction case, however, this equivalence does not hold and we provide a result that holds for CARA utility.

Proposition 4.1. When the market makers in a limit-order book all have the same CARA utility and the asset values are normally distributed, the following is a necessary condition for any symmetric strategy Nash equilibrium in demand functions that are piecewise continuously differentiable:

$$\sum_{j \neq i}^N x'_j(p) = -\frac{H(z)}{\bar{v} - p - \rho\tau_v^{-1}x_i(p)}, \quad (25)$$

²⁵ Eq. (24) is the first order condition obtained by Biais et al. (2000).

²⁶ The limits of these two models as the number of bidders goes to infinity is very different as the inventory model converges to perfect competition while the information model converges to the equilibrium in Glosten (1994). However, for any given N and $\rho\tau_v^{-1}$, there is a value of ξ (depending on N) that yields observational equivalence.

where

$$H(z) = \frac{\Lambda(z)}{g(z)}, \quad (26)$$

$$\frac{\partial \Lambda(z)}{\partial z} = \rho \Lambda(z) \left\{ \left[\bar{v} - p(z) - \rho \tau_v^{-1} \frac{1}{N} \right] \frac{1}{N} - p(z) \right\} - g(z). \quad (27)$$

Eq. (25) is also sufficient for a symmetric strategy equilibrium when the following condition holds for all $z > \hat{z}$ and $\hat{p}(z) = p(\hat{z})$:

$$H(z) > H(\hat{z}) - \frac{\rho \tau_v^{-1} (z - \hat{z}) H(\hat{z})}{\bar{v} - \hat{p}(z) - \rho \tau_v^{-1} x_i(\hat{p}(z))}. \quad (28)$$

Also, there is a unique equilibrium with the property that as $p(z) \rightarrow \bar{v} - \rho \tau_v^{-1} 1/N$ as $z \rightarrow 1$.²⁷

Further we compare the equilibrium price under CARA with the equilibrium price under mean-variance utility (considering the equilibrium where $p(z) \rightarrow \bar{v} - \rho \tau_v^{-1} 1/N$) and we find that

$$p(z)|_{\text{CARA}} > p(z)|_{\text{MV}} \quad (29)$$

Proof. See Appendix F. $\square$

Market makers who have CARA utility are averse to price-execution risk (in contrast to market makers with mean variance utility who are only averse to payoff risk) and hence bid more aggressively in an attempt to reduce this price-execution risk. A similar result is shown by Röell (1998) in the context of the exponential distribution. An implication of Proposition 3.4 is that the welfare comparison in Proposition 3.2 continues to hold with market makers who have CARA utility.

4.3. Extension III: endogenous customer order sizes

We have analyzed a model in which the customer order size is exogenously drawn from some distribution, $\tilde{z} \in [0, 1]$. We now allow the customer to choose the quantity he sells, $q(z) \in [0, z]$, given an exogenous endowment shock $\tilde{z}$.

In this subsection, we set the customer's risk-aversion coefficient to be the same as the market makers', i.e., $A = \rho/2$, to simplify the computation. This does not change the results in any qualitative way.

²⁷ Again, all other downward-sloping equilibria have the property that the price at the highest customer order, $z = 1$, is $-\infty$.

In a dealership market, the strategy used by the market makers does not depend on the distribution of the customer orders. As a result, the equilibrium price is the following (see Eq. (9)):

$$p(q) = \bar{v} - \frac{(N-1)\rho\tau_v^{-1}}{N-2} \frac{q(z)}{N}. \quad (30)$$

Thus, the customer maximizes the following objective function by selecting the quantity to sell, $q(z)$:

$$\begin{aligned} U_d(\tilde{w}) &= E[\tilde{w}] - \frac{\rho}{2} \text{Var}[\tilde{w}] \\ &= (z - q)\bar{v} + qp(q) - \frac{\rho\tau_v^{-1}}{2}(z - q)^2, \end{aligned} \quad (31)$$

where

$$\tilde{w} = (z - q)\bar{v} + qp(q). \quad (32)$$

It is easy to check that the optimal choice of quantity is:

$$q^*(z) = \frac{z}{1 + 2(N-1)/N(N-2)}, \quad (33)$$

and the maximized objective function is:

$$U_d^* = (z - q^*)\bar{v} + q^*p^*(q^*) - \frac{\rho\tau_v^{-1}}{2}(z - q^*)^2 = zp^*(q^*). \quad (34)$$

In a limit-order book, the equilibrium strategy of the market makers will depend on the distribution of the customer orders, which is endogenous. Therefore, we need to determine the customer's and the market makers' equilibrium strategies jointly.²⁸

However when the customer sells a positive quantity to the dealers, the optimal quantity he sells is $q^* = z - (\bar{v} - p)/\rho\tau_v^{-1}$. For very small customer order size, the customer will not trade at all because of the presence of zero-quantity spread in the limit-order book. Thus, we conjecture the following optimal trading strategy for the customer:

$$q^*(p, z) = \begin{cases} z - \frac{\bar{v} - p}{\rho\tau_v^{-1}} & \text{if } z \in [\underline{s}, 1], \\ 0 & \text{if } z \in [0, \underline{s}), \end{cases} \quad (35)$$

where the cutoff size $\underline{s}$ will be determined by the bidding behavior of the dealers.

²⁸ See Appendix G for derivation of the results here.

Given this conjecture, we consider the truncated distribution of customer orders over $[\underline{s}, 1]$. The market makers' equilibrium strategy is:²⁹

$$x_i(p) = \gamma \left[\bar{v} - \frac{2\rho\tau_v^{-1}\theta}{N(1+\theta) + 2\theta + \sqrt{N^2(1+\theta)^2 - 4\theta}} - p \right],$$

$$i = 1, 2, \dots, N, \quad (36)$$

where

$$\gamma = \frac{N(1+\theta) - 2 + \sqrt{N^2(1+\theta)^2 - 4\theta}}{2(N-1)\rho\tau_v^{-1}}. \quad (37)$$

Furthermore, the threshold size of the order below which the customer does not sell is:

$$\underline{s} = \frac{2\theta}{N(1+\theta) + 2\theta + \sqrt{N^2(1+\theta)^2 - 4\theta}} \in (0, 1). \quad (38)$$

Using the above results, it is straightforward to show that the customer's expected utility from trading in a limit-order book is:

$$U_b^* = \begin{cases} \bar{v}(z - q^*(p, z)) - \frac{\rho\tau_v^{-1}}{2}(z - q^*(p, z))^2 + p^*q^* + \frac{(q^*(p, z))^2}{2N\gamma} & \text{if } z \in [\underline{s}, 1], \\ \bar{v}z - \frac{\rho\tau_v^{-1}}{2}z^2, & \text{if } z \in [0, \underline{s}], \end{cases} \quad (39)$$

where the equilibrium price, p^* , is given by:

$$p^*(z) = \bar{v} - \rho\tau_v^{-1} \frac{2N\theta/[N(1-\theta) + \sqrt{N^2(1+\theta)^2 - 4\theta}] + z}{1 + N\rho\tau_v^{-1}\gamma}. \quad (40)$$

With endogenous customer order size, the range of quantities that the customer sells is different under the two market structures. In the dealership market, it is

$$q^* \in \left[0, \frac{1}{1 + 2(N-1)/N(N-2)} \right], \quad (41)$$

²⁹ The linear hazard ratio distributions that we consider (see Eq. (2)) preserve the linear hazard ratio property over the truncated support $[\underline{s}, 1]$.

whereas for the limit-order book it is:

$$q^* \in \left[\underline{s} - \frac{\bar{v} - p^*(\underline{s})}{\rho \tau_v^{-1}}, 1 - \frac{\bar{v} - p^*(1)}{\rho \tau_v^{-1}} \right]. \quad (42)$$

Through numerical computation, we can establish that the comparison of customer welfare across these two market structures is similar to that in Proposition 3.2. Specifically, a risk-averse customer prefers to trade in a dealership market rather than a limit-order book when the number of market makers is large. In other words, our main insight about the customer's preference between a dealership market and a limit-order book is robust even when the customer chooses his trading quantities optimally.

Comparing Fig. 4 with Fig. 3, we find that allowing the customer to choose his trading quantities endogenously expands the parameter region where a dealership market is preferred to a limit-order book. This is primarily due to the fact that, when the customer is not forced to sell, there is an “no-trade” zone in the limit-order book.

4.4. Discussion

We next relate our theoretical results to the existing evidence on trading costs across different market arenas. De Jong et al. (1995) compare trading costs for French equities listed on both the Paris Bourse (a limit-order book market) and London Stock Exchange (a dealership market). They find that the Paris Bourse has lower transaction costs for small sizes. For very large sizes, the Paris limit-order book is often too thin and causes the average bid–ask spread to rise steeply. While trading costs (measured by the bid–ask spread as a percentage of the transaction price) are higher in London for small transactions (roughly up to what is called the “normal market size”), London attracts the bulk of transaction volumes in large order sizes. In fact, the average value of transactions in London is about 10 times the average value of regular transactions in Paris. This apparent bifurcation of smaller orders being executed in Paris and larger orders in London is consistent with the results of our equilibrium analysis.

While our findings are generally consistent with existing empirical evidence on trading costs for stocks listed on both the dealership market and limit-order book, the model also yields additional predictions that are testable.³⁰ For

³⁰ Researchers also studied the transaction costs for equity issues from Belgium (Anderson and Tychon, 1993; Degryse, 1999), Italy (Impenna et al., 1995), and Germany (Schmidt and Iverson, 1992; Brown, 1994) that are simultaneously listed on their domestic market (resembling electronic limit-order book trading) and on the London Stock Exchange. Booth et al. (1999) provides evidence on a matched sample of NASDAQ stocks and German stocks traded on IBIS, which is an automated market. By and large, the conclusions are similar to the work of Röell (1992) and De Jong et al. (1995) on French stocks.

example, the model predicts that, *ceteris paribus*, the liquidity providers are more price competitive in the limit-order book than in the dealership market. That is, the equilibrium bid schedules used by market makers are flatter in the limit-order book than in the dealership. In addition, the difference in the price elasticity of demand schedules across the two market structures is more pronounced for stocks that are relatively more volatile. Finally, the impact of changing the size of the market making sector on equilibrium strategy is different across the two markets. As the number of market makers increases, market maker competition intensifies in a dealership market as measured by a flatter demand curve. With a limit-order book, the presence of more market makers can make the equilibrium demand steeper while reducing the zero-quantity spread.

5. Conclusion

As discussed in the introduction, there has been a convergence in recent years in equity markets towards a hybrid limit-order book/dealership market structure, where smaller orders are executed on the limit-order book, while large orders are executed in the dealership market. In this paper, we analyze the customer's choice among a limit-order book, a dealership market, and a hybrid market structure of the two. The customer's order is competed for and divided among risk-averse market makers (limit-order providers). Drawing on the analogy between a limit-order book and a discriminatory auction and the analogy between a dealership market and a uniform-price auction, we present a general approach to solving for equilibria in dealership markets and in limit-order books.

Our main conclusions are as follows: (1) a risk-neutral customer prefers to trade in a limit-order market instead of in any hybrid or dealership markets; (2) a risk-averse customer prefers to trade in a dealership market over a limit-order book market when the number of market makers is large and when the average order size is large; and (3) for risk-averse customers, the hybrid market structure, when properly structured, dominates both the dealership market and the limit-order book. Findings of the paper can be understood in terms of the tradeoff between two features of the equilibrium demand schedules: a bid-shading effect that operates differently in a limit-order book compared to a dealership market, and a zero-quantity spread that is present in the limit-order book only.

Appendix A. Proof of Proposition 2.1

In the following we characterize market maker i 's optimal trading strategy, $x_i(p)$, in response to the strategy used by the other dealers, $x_j(p)$ ($\forall j \neq i$). The

total amount of supply is $\bar{z} \in [0, 1]$. We use $g(z)$ and $G(z)$ to denote the pdf and cdf for $\bar{z}$. We also define bidder i 's (upward-sloping) residual supply curve given z as $h(p, z)$ where

$$h(p, z) = z - \sum_{j \neq i} x_j(p). \quad (\text{A.1})$$

In what follows, we will suppress the dependence on z unless necessary, i.e., we will write $h(p)$ instead of $h(p, z)$.

With a price discrimination parameter $\alpha \in \{0, 1\}$, bidder i 's uncertain bidding profit is the following:

$$\tilde{\pi}_i = \tilde{v}x_i(p) - TP_i, \quad (\text{A.2})$$

where his total payment is:

$$\begin{aligned} TP_i &= (1 - \alpha)px_i(p) + \alpha \int_0^{x_i(p)} p(q) dq \\ &= (1 - \alpha)px_i(p) + \alpha \left[\sum_{j=1}^N \int_0^{x_j(p)} p(q) dq - \sum_{j \neq i} \int_0^{x_j(p)} p(q) dq \right] \\ &\equiv (1 - \alpha)px_i(p) + \alpha[A - B]. \end{aligned} \quad (\text{A.3})$$

We note the following relations for use in subsequent calculation:

$$\frac{dA}{dz} = \sum_{j=1}^N p \frac{dx_j}{dz} = p \frac{dz}{dz} = p, \quad (\text{A.4})$$

and

$$\frac{\partial B}{\partial p} = \sum_{j \neq i}^N p \frac{\partial x_j(p)}{\partial p} = -p \frac{\partial h(p)}{\partial p}. \quad (\text{A.5})$$

We assume that market maker i chooses his optimal trading strategy by maximizing the following derived mean-variance utility function:³¹

$$E_{\bar{z}} \left[\bar{v}h(p) - \frac{\rho\tau_v^{-1}}{2}(h(p))^2 - TP_i \right]. \quad (\text{A.6})$$

Defining $A(z)$ as the state variable and $p(z)$ as the control variable, we can analyze the problem using the following Lagrangian:

$$L = g(z) \left[\bar{v}h(p) - \frac{\rho\tau_v^{-1}}{2}h^2(p) - (1 - \alpha)ph(p) - \alpha A(z) + \alpha B(p) \right] + \lambda p, \quad (\text{A.7})$$

³¹ See Appendix F for an alternative formulation with CARA utility.

with the constraint that the state variable $A(z) \geq 0$. We use instead the constraint $A(0) \geq 0$ (with downward sloping demand curves as a solution this will imply $A(z) \geq 0$). The optimality condition is:

$$0 = \frac{\partial L}{\partial p} = g(z)[\bar{v} - p - \rho\tau_v^{-1}h(p)]\frac{\partial h(p)}{\partial p} - (1 - \alpha)h(p)g(z) + \lambda. \quad (\text{A.8})$$

The adjoint equation is:

$$-\frac{\partial \lambda}{\partial z} = \frac{\partial L}{\partial A} = -\alpha g(z), \quad (\text{A.9})$$

and the transversality conditions are

$$\lambda(1) = 0, \quad A(0)\lambda(0) = 0. \quad (\text{A.10})$$

Using the transversality condition that $\lambda(1) = 0$, the adjoint equation can be integrated to obtain

$$\lambda(z) = -\alpha[1 - G(z)]. \quad (\text{A.11})$$

Since this implies that $\lambda(0) = -\alpha$, we must have $A(0) = 0$ (this will be satisfied by our candidate solution because when $z = 0$, the symmetric equilibrium gives zero quantity to each bidder).

Combining Eqs. (A.8) and (A.11), we have:

$$[\bar{v} - p - \rho\tau_v^{-1}h(p)]\frac{\partial h(p)}{\partial p} = (1 - \alpha)h(p) + \alpha\frac{1 - G(z)}{g(z)}, \quad (\text{A.12})$$

which is equivalent to:

$$\sum_{j \neq i}^N x'_j(p) = -\frac{(1 - \alpha)x_i(p) + \alpha[1 - G(z)]/g(z)}{\bar{v} - p - \rho\tau_v^{-1}x_i(p)}. \quad (\text{A.13})$$

Because the Lagrangian is linear in the state variable, A , sufficiency of the solutions to the above necessary condition to form an equilibrium is guaranteed if the Lagrangian is (quasi) concave in the control variable, p .

Taking the derivative of Eq. (50) with respect to p , we have the following second-order condition:

$$g(z)[\bar{v} - p - \rho\tau_v^{-1}h(p)]\frac{\partial^2 h(p)}{\partial p^2} - g(z)\left[(2 - \alpha) + \rho\tau_v^{-1}\frac{\partial h(p)}{\partial p}\right]\frac{\partial h(p)}{\partial p}, \quad (\text{A.14})$$

which is clearly negative for all downward sloping, linear strategies (which implies upward-sloping residual supply curves) as $\partial^2 h(p)/\partial p^2 = 0$ and $\partial h(p)/\partial p > 0$.

A more general approach is as follows. We examine Eq. (A.8) at two price levels, $p(z)$ and $\hat{p}(z) > p(z)$. Quasi-concavity requires that:

$$\begin{aligned}
 & g(z)\{[\bar{v} - p - \rho\tau_v^{-1}h(p)]h'(p) - (1 - \alpha)h(p)\} + \lambda(z) \\
 & > g(z)\{[\bar{v} - \hat{p} - \rho\tau_v^{-1}h(\hat{p})]h'(\hat{p}) - (1 - \alpha)h(\hat{p})\} + \lambda(z) \\
 & = g(z)\{[\bar{v} - \hat{p} - \rho\tau_v^{-1}(z - Nx_j(\hat{p}))]h'(\hat{p}) - (1 - \alpha)h(\hat{p})\} + \lambda(z) \\
 & = g(z)\{[\bar{v} - \hat{p} - \rho\tau_v^{-1}(\hat{z} - Nx_j(\hat{p}))]h'(\hat{p}) - (1 - \alpha)h(\hat{p}) \\
 & \quad + \rho\tau_v^{-1}(\hat{z} - z)h'(\hat{p})\} + \lambda(z) \\
 & = g(z)\{[\bar{v} - p(\hat{z}) - \rho\tau_v^{-1}h(p(\hat{z}))]h'(p(\hat{z})) - (1 - \alpha)h(p(\hat{z})) \\
 & \quad + \rho\tau_v^{-1}(\hat{z} - z)h'(p(\hat{z}))\} + \lambda(z),
 \end{aligned}$$

where, in the last step, we used the fact that $\hat{p}(z) = p(\hat{z})$ for some $\hat{z} < z$.

Noting Equation (54), the above inequality can be rewritten as:

$$\begin{aligned}
 & \alpha H(z) > \alpha H(\hat{z}) + \rho\tau_v^{-1}(\hat{z} - z)h'(p(\hat{z})), \\
 & = \alpha H(\hat{z}) - \frac{\rho\tau_v^{-1}(z - \hat{z})[(1 - \alpha)x_i(\hat{p}) + \alpha H(\hat{z})]}{\bar{v} - \hat{p} - \rho\tau_v^{-1}x_i(\hat{p})}. \tag{A.15}
 \end{aligned}$$

The above sufficient condition needs to be true for all $z > \hat{z}$ and suffices for prices observed in equilibrium. For a price such that $p > p(0)$, one has that $x_j(p) = 0$ and $x'_j(p) = 0$. Hence $[\bar{v} - p - \rho\tau_v^{-1}h(p)]h'(p) = 0$ and the above inequality still holds. For $p < p(z = 1)$, we assume that other limit-order submitters submit a price $p = p(z = 1)$ for all quantities greater than s/N (this is the off-equilibrium extension). We need this extension only if $p(z = 1)$ is not $-\infty$. Hence, we obtain that $h'(p) = -\infty$. Hence, $[\bar{v} - p - \rho\tau_v^{-1}h(p)]h'(p) = -\infty$ and the above inequality still holds. $\square$

Appendix B. Proof of Corollary 2.2

Here we specialize to the case of $\alpha = 1$ in Proposition 2.1. Furthermore, we assume $H(z) = \theta(1 - z)$ for $z \in [0, 1]$.

Substituting the conjectured equilibrium demand schedule $x_i(p) = \mu - \gamma p$, $i = 1, 2, \dots, N$ into the market clearing condition, we have $z = N(\mu - \gamma p)$. Plugging into Eq. (5), it becomes:

$$-(N - 1)\gamma = -\frac{\theta[1 - N(\mu - \gamma p)]}{\bar{v} - p - \rho\tau_v^{-1}x_i},$$

or

$$x_i = \frac{\bar{v} - p}{\rho\tau_v^{-1}} - \frac{\theta[1 - N(\mu - \gamma p)]}{(N - 1)\gamma\rho\tau_v^{-1}}.$$

From the above relations, we can solve for:

$$\gamma = \frac{N(1 + \theta) - 1}{(N - 1)\rho\tau_v^{-1}},$$

$$\mu = \gamma\bar{v} - \frac{\theta}{N - 1}.$$

The solution in inverse demand form is given in Corollary 2.2. $\square$

Appendix C. Proof of Proposition 3.1

The customer's random wealth is of the form $\tilde{w} = b\tilde{z} - c\tilde{z}^2$. His utility function is:

$$U(\tilde{w}) = E[\tilde{w}] - A \text{Var}[\tilde{w}],$$

where

$$E[\tilde{w}] = b\frac{\theta}{1 + \theta} - c\frac{2\theta^2}{1 + 3\theta + 2\theta^2}, \quad (\text{C.1})$$

and

$$\begin{aligned} E[\tilde{w}^2] = & b^2 \frac{2\theta^2}{1 + 3\theta + 2\theta^2} + c^2 \frac{24\theta^4}{1 + 10\theta + 35\theta^2 + 50\theta^3 + 24\theta^4} \\ & - 2bc \frac{6\theta^3}{1 + 6\theta + 11\theta^2 + 6\theta^3}, \end{aligned}$$

and

$$\begin{aligned} \text{Var}[\tilde{w}] = & E[\tilde{w}^2] - E^2[\tilde{w}] \\ = & b^2 \frac{\theta^2}{(1 + 2\theta)(1 + \theta)^2} + c^2 \frac{4\theta^4(5 + 11\theta)}{(1 + 7\theta + \theta^2)(1 + 3\theta + 2\theta^2)^2} \\ & - 2bc \frac{4\theta^3}{(1 + \theta)^2(1 + 5\theta + 6\theta^2)}. \end{aligned} \quad (\text{C.2})$$

For risk-neutral customers $A = 0$, and therefore we need to show that $E[\tilde{w}_b] > E[\tilde{w}_d]$:

$$\begin{aligned} E[\tilde{w}_d] &= b_d \frac{\theta}{1 + \theta} - c_d \frac{2\theta^2}{1 + 3\theta + 2\theta^2} \\ &= \bar{v} \frac{\theta}{1 + \theta} - \frac{\rho\tau_v^{-1}(N-1)}{N(N-2)} \frac{2\theta^2}{1 + 3\theta + 2\theta^2}, \\ E[\tilde{w}_b] &= b_b \frac{\theta}{1 + \theta} - c_b \frac{2\theta^2}{1 + 3\theta + 2\theta^2} \\ &= \left[\bar{v} - \frac{\rho\tau_v^{-1}\theta}{N(1 + \theta) - 1} \right] \frac{\theta}{1 + \theta} - \frac{\rho\tau_v^{-1}(N-1)}{2N[N(1 + \theta) - 1]} \frac{2\theta^2}{1 + 3\theta + 2\theta^2}. \end{aligned}$$

Thus,

$$E[\tilde{w}_b] - E[\tilde{w}_d] = \frac{\rho\tau_v^{-1}\theta^2}{(N-2)(1 + \theta)[N(1 + \theta) - 1]} > 0. \quad \square$$

Appendix D. Proof of Proposition 3.2

We showed in the proof to Proposition 3.1 that $E[\tilde{w}_b] > E[\tilde{w}_d]$. Clearly, a necessary condition for $U(\tilde{w}_d) > U(\tilde{w}_b)$ to occur (when $A > 0$) is that there are cases when $\text{Var}[\tilde{w}_b] > \text{Var}[\tilde{w}_d]$. For example, consider a N value that is large, we have:

$$\text{Var}[\tilde{w}_b] - \text{Var}[\tilde{w}_d] \approx \frac{\rho\tau_v^{-1}\bar{v}}{N} \frac{2\theta^3(1 + 6\theta + 9\theta^2 + 4\theta^3)}{(1 + 3\theta)(1 + \theta)^4(1 + 6\theta + 8\theta^2)} > 0. \quad \square$$

Appendix E. Proof of Proposition 3.4

To compare the hybrid market with the dealership market, we use Eqs. (20) and (21) to obtain the following:

$$E[U(\tilde{w}_b)] - E[U(\tilde{w}_d)] = \int_0^r [U(b'_bz - c'_bz^2) - U(b_dz - c_dz^2)] dz, \quad (\text{E.1})$$

i.e., the only difference between a hybrid market and a pure dealership market is entirely attributable to customer orders that are smaller than the cutoff level, r . Moreover, from Eq. (18), it is clear that as r approaches zero, b'_b approaches $b_d = \bar{v}$, whereas $c'_b < c_d$. This leads us to conclude that the integrand of the

above equation is strictly positive for sufficiently small values of r . For large values of r , the sign of the $E[U(\tilde{w}_h)] - E[U(\tilde{w}_d)]$ is ambiguous.

To compare the hybrid market with the limit-order book, we use Eqs. (20) and (22) to arrive at:

$$\begin{aligned} E[U(\tilde{w}_h)] - E[U(\tilde{w}_b)] &= \int_0^r [U(b'_bz - c_bz^2) - U(b_bz - c_bz^2)] dz \\ &\quad + \int_r^1 [U(b_dz - c_dz^2) - U(b_bz - c_bz^2)] dz. \end{aligned} \quad (\text{E.2})$$

The first term is always positive because $b'_b > b_b$. The second term, however, cannot be signed. Thus, the sign of the $E[U(\tilde{w}_h)] - E[U(\tilde{w}_b)]$ is ambiguous in general. $\square$

Appendix F. Proof of Proposition 4.1

Making use of the properties of CARA utility and normal random variables, we can write market maker i 's objective function as follows:

$$E_{\tilde{z}, \tilde{v}} \left[-\frac{1}{\rho} e^{-\rho(\tilde{v}x_i - TP_i)} \right] = E_{\tilde{z}} \left[-\frac{1}{\rho} e^{-\rho(\tilde{v}x_i - (\rho\tau_v^{-1}/2)x_i^2 - TP_i)} \right]. \quad (\text{F.1})$$

Defining $A(z)$ as the state variable and $p(z)$ as the control variable,³² we can analyze the problem using the following Lagrangian:

$$L(A, p, \lambda, z) = -\frac{g(z)}{\rho} e^{-\rho[\tilde{v}h(p) - (\rho\tau_v^{-1}/2)h^2(p) - A(z) + B(p)]} + \lambda p. \quad (\text{F.2})$$

The optimality condition is:

$$0 = \frac{\partial L}{\partial p} = g(z)[\tilde{v} - p - \rho\tau_v^{-1}h(p)] \frac{\partial h(p)}{\partial p} e^{-\rho[\tilde{v}h(p) - (\rho\tau_v^{-1}/2)h^2(p) - A(z) + B(p)]} + \lambda. \quad (\text{F.3})$$

The adjoint equation is:

$$-\frac{\partial \lambda}{\partial z} = \frac{\partial L}{\partial A} = -g(z) e^{-\rho[\tilde{v}h(p) - (\rho\tau_v^{-1}/2)h^2(p) - A(z) + B(p)]}. \quad (\text{F.4})$$

Defining

$$\lambda \equiv -A e^{-\rho[\tilde{v}h(p) - (\rho\tau_v^{-1}/2)h^2(p) - A(z) + B(p)]}, \quad (\text{F.5})$$

and Eq. (F.3) can be rewritten as:

$$\sum_{j \neq i}^N x'_j(p) = -\frac{A(z)/g(z)}{\tilde{v} - p - \rho\tau_v^{-1}x_i(p)}. \quad (\text{F.6})$$

³² Unless noted otherwise, the notation here will be the same as in Appendix A.

In addition, Λ must be solved from Eqs. (F.4) and (F.5):

$$\begin{aligned}\frac{\partial \Lambda(z)}{\partial z} &= \rho \Lambda(z) \left\{ [\bar{v} - p - \rho \tau_v^{-1} h(p)] \frac{\partial h(p)}{\partial p} \frac{\partial p(z)}{\partial z} - p(z) \right\} - g(z) \\ &= \rho \Lambda(z) \left\{ \left[\bar{v} - p(z) - \rho \tau_v^{-1} \frac{1}{N} \right] \frac{1}{N} - p(z) \right\} - g(z).\end{aligned}\quad (\text{F.7})$$

To find the sufficient condition, we use an Arrow type sufficiency result following Theorem 5 and Note 2 from Seierstad and Sydsaeter (1977). First

$$L^*(A, \lambda, z) = \max_p L(A, p, \lambda, z)$$

is well defined when the above optimization problem is quasi-concave in p . To show the quasi-concavity of p , i.e., $\partial L / \partial p$ is decreasing in p , we first note that

$$e^{-\rho[\bar{v}h(p) - (\rho\tau_v^{-1}/2)h^2(p) - A(z) + B(p)]} \quad (\text{E.8})$$

is decreasing in p provided $[\bar{v} - p - \rho\tau_v^{-1}h(p)]h'(p)$ is positive. Since this is true, it suffices to show that (see Eq. (F.3)) $[\bar{v} - p - \rho\tau_v^{-1}h(p)]h'(p)$ is decreasing in p . If we consider two price levels, $p(z)$ and $\hat{p}(z) > p(z)$, and require that:

$$\begin{aligned}&g(z)\{[\bar{v} - p - \rho\tau_v^{-1}h(p)]h'(p) - h(p)\} + \lambda(z) \\ &> g(z)\{[\bar{v} - \hat{p} - \rho\tau_v^{-1}h(\hat{p})]h'(\hat{p}) - h(\hat{p})\} + \lambda(z) \\ &= g(z)\{[\bar{v} - \hat{p} - \rho\tau_v^{-1}(z - Nx_j(\hat{p}))]h'(\hat{p}) - h(\hat{p})\} + \lambda(z) \\ &= g(z)\{[\bar{v} - \hat{p} - \rho\tau_v^{-1}(\hat{z} - Nx_j(\hat{p}))]h'(\hat{p}) \\ &\quad - h(\hat{p}) + \rho\tau_v^{-1}(\hat{z} - z)h'(\hat{p})\} + \lambda(z) \\ &= g(z)\{[\bar{v} - p(\hat{z}) - \rho\tau_v^{-1}h(p(\hat{z}))]h'(p(\hat{z})) - h(p(\hat{z})) \\ &\quad + \rho\tau_v^{-1}(\hat{z} - z)h'(p(\hat{z}))\} + \lambda(z),\end{aligned}$$

where, in the last step, we used the fact that $\hat{p}(z) = p(\hat{z})$ for some $\hat{z} < z$.

Using Eqs. (66) and (68) and defining $H(z) = \Lambda(z)/g(z)$, the above inequality can be rewritten as:

$$\begin{aligned}H(z) &> H(\hat{z}) + \rho\tau_v^{-1}(\hat{z} - z)h'(p(\hat{z})), \\ &= H(\hat{z}) - \frac{\rho\tau_v^{-1}(z - \hat{z})H(\hat{z})}{\bar{v} - \hat{p} - \rho\tau_v^{-1}x_i(\hat{p})},\end{aligned}\quad (\text{F.9})$$

which is similar to that in the mean-variance utility case. The extension to prices such that $p > p(0)$ and $p < p(1)$ is as before.

To complete the proof, we use Note 2 from Seierstad and Sydsaeter (1977) and consider the following problem

$$\max_A L^*(A, \lambda(z), z) + \lambda'(z)A. \quad (\text{F.10})$$

The first order condition for this problem is

$$\begin{aligned} \frac{\partial L^*}{\partial A} &= \frac{\partial L}{\partial A} + \lambda'(z) + \frac{\partial L}{\partial p} p'(A) \\ &= \frac{\partial L}{\partial A} + \lambda'(z) \\ &= 0 \end{aligned}$$

because the optimal $p(A)$ satisfies the first order condition of the control problem and hence $\partial L/\partial p$ is zero. The last equation above is the adjoint equation to the optimal control problem above and is satisfied by our candidate solution. The second order condition is

$$\frac{\partial^2 L}{\partial A^2} + \frac{\partial^2 L}{\partial A \partial p} p'(A) < 0. \quad (\text{F.11})$$

Since $\partial^2 L/\partial A^2$ and $\partial^2 L/\partial A \partial p$ are both negative, we need only check that $p'(A) > 0$. By examining the first order condition for the control variable p given the state variable A , this is easily verified. Hence Theorem 5 in Seierstad and Sydsaeter (1977) holds and equilibrium exists if the sufficient condition stated in the theorem holds.

To prove the final statement regarding the price comparison under CARA and under mean-variance utility we argue as follows. First we note that $\lambda'(z) < -\lambda'(z)$ and that $p'(z) < 0$, $\bar{v} - p - \rho\tau_v^{-1} > 0$. Also, $\lambda(0) = \lambda(0) = 0$. Thus $\lambda(z) < -\lambda(z) = 1 - G(z) \forall z$. Hence if for any z

$$p(z)|_{\text{CARA}} = p(z)|_{\text{MV}}, \quad (\text{F.12})$$

inspection of the first order condition implies that at such a point

$$\begin{aligned} (N-1)x'_j(p)|_{\text{CARA}} &= - \frac{\lambda(z)/g(z)}{\bar{v} - p(z)|_{\text{CARA}} - \rho\tau_v^{-1}(z/N)} \\ &> - \frac{(1 - G(z))/g(z)}{\bar{v} - p(z)|_{\text{MV}} - \rho\tau_v^{-1}(z/N)} \\ &= (N-1)x'_j(p)|_{\text{MV}}. \end{aligned}$$

Thus the slope under CARA is flatter (less negative) at such a point. In turn this implies at such a point that

$$p'(z)|_{\text{CARA}} < p'(z)|_{\text{MV}}, \quad (\text{F.13})$$

i.e., the bidding curve under CARA is steeper than the bidding curve under mean-variance utility.

Now remember that

$$p(z=1)|_{\text{CARA}} = p(z=1)|_{\text{MV}} = \bar{v} - \rho\tau_v^{-1}\frac{1}{N}. \quad (\text{F.14})$$

The rest of the proof is by contradiction. Suppose that for an interval of positive measure around $z=1$ one has:

$$p(z)|_{\text{CARA}} < p(z)|_{\text{MV}}. \quad (\text{F.15})$$

Then

$$\begin{aligned} (N-1)x'_j(p)|_{\text{CARA}} &= -\frac{\Lambda(z)/g(z)}{\bar{v} - p(z)|_{\text{CARA}} - \rho\tau_v^{-1}(z/N)} \\ &= -\frac{\Lambda(z)/g(z)}{\bar{v} - p(z)|_{\text{MV}} - \rho\tau_v^{-1}(z/N)} \\ &> -\frac{(1-G(z))/g(z)}{\bar{v} - p(z)|_{\text{MV}} - \rho\tau_v^{-1}(z/N)} \\ &= (N-1)x'_j(p)|_{\text{MV}}. \end{aligned}$$

Hence

$$p'(z)|_{\text{CARA}} < p'(z)|_{\text{MV}}, \quad (\text{F.16})$$

and, using the end point condition in Eq. (F.14), we get:

$$p(z)|_{\text{CARA}} > p(z)|_{\text{MV}}. \quad (\text{F.17})$$

which is a contradiction. Hence it must be the case that around $z=1$ we have

$$p(z)|_{\text{CARA}} > p(z)|_{\text{MV}}. \quad (\text{F.18})$$

But then it is impossible for the two price lines to cross at any z (doing so would violate the slope condition at such a point, Eq. (F.13)). Hence for all z , we must have

$$p(z)|_{\text{CARA}} > p(z)|_{\text{MV}}. \quad (\text{F.19})$$

This completes the proof of the theorem. $\square$

Appendix G. Derivation of results in Section 4.3

Suppose the market makers use the strategy, $x_j(p)$, $j=1, 2, \dots, N$. If the customer intends to sell a quantity $q < z$ with $z \in [0, 1]$, then his wealth after

trading is:

$$\hat{w} = \bar{v}(z - q) + qp + \sum_{j=1}^N \int_p^{\bar{p}} x_j(\psi) d\psi, \quad (\text{G.1})$$

where $\bar{p}$ is the intercept of market maker j 's demand schedule with the price axis.

Assuming that the customer's risk-aversion coefficient is the same as the market makers', i.e., $A = \rho/2$, then the maximization of the customer's expected utility which is the following,³³

$$\bar{v}(z - q) - \frac{\rho\tau_v^{-1}}{2}(z - q)^2 + pq + \sum_{j=1}^N \int_p^{\bar{p}} x_j(\psi) d\psi, \quad (\text{G.2})$$

leads to the first-order condition for the customer:

$$\begin{aligned} 0 &= -\bar{v} + \rho\tau_v^{-1}(z - q) + p + q \frac{\partial p}{\partial q} - \frac{\partial p}{\partial q} \sum_{j=1}^N x_j(p) \\ &= -\bar{v} + \rho\tau_v^{-1}(z - q) + p. \end{aligned} \quad (\text{G.3})$$

The market-clearing condition $q = \sum_{j=1}^N x_j(p)$ is used in the last step.

Thus far, we have proved that the trading strategy,

$$q(p, z) = z - \frac{\bar{v} - p}{\rho\tau_v^{-1}}, \quad (\text{G.4})$$

is the equilibrium strategy for the customer (i.e., optimal with respect to any trading strategy on the part of the dealers) so long as the market makers' quantity allocations are positive.

At the end of this appendix, we will show that there exists a unique cutoff customer order size, $\underline{z} \in (0, 1)$ such that the customer is a net seller if and only if $z \in [\underline{z}, 1]$. Should the customer decide to use the above strategy $q(p, z)$ when he observes a z value that is less than $\underline{z}$ (based on the incorrect assumption that the market makers get a positive amount), it can be shown that the market makers' quantity allocation would be negative in that case. This cannot happen in equilibrium and, therefore, the customer does not sell to the market makers when the order is sufficiently small (i.e., when $z < \underline{z}$).

The derivation of the market makers' equilibrium strategy is parallel to the proof in Appendix A. Because the customer retains the quantity $z - q(p, z)$, we can view the bidding process as an auction in which one of the bidders (we can label the customer as the 0th bidder) uses the strategy $x_0(z, p) = z - q(p, z)$ while the other bidders (the N dealers) use the strategy $x_j(p)$, $j = 1, \dots, N$ to compete for the total available quantity z . Therefore, except for allowing the summation index to go from 0 to N , the derivation in Appendix A can be

³³ Similar results obtain when the customer has a different risk-aversion coefficient.

duplicated to provide us the following characterization of market maker i 's equilibrium trading strategy:

$$\sum_{j=0, (j \neq i)}^N \frac{x_j(p)}{\partial p} = - \frac{[1 - G(z)]/g(z)}{\bar{v} - p - \rho\tau_v^{-1}x_i(p)}, \quad (\text{G.5})$$

where $g(z)$ and $G(z)$ to denote the pdf and cdf for the customer order size $\tilde{z}$ over $[\underline{s}, s]$. The same method of proof as in Appendix A can be used to provide sufficient conditions. It is easy to verify that these are satisfied in the linear strategy equilibrium considered below.

We now search for a linear strategy equilibrium of the form $\mu - \gamma p$. With this, the market-clearing condition is:

$$z = x_i + x_0 + (N - 1)(\mu - \gamma p). \quad (\text{G.6})$$

From Eq. (2), the pdf of $\tilde{z}$ over $[\underline{s}, 1]$ is given by:

$$g(z) = \frac{1}{\theta(1 - \underline{s})} \left(1 - \frac{z - \underline{s}}{1 - \underline{s}} \right)^{1/\theta - 1}, \quad (\text{G.7})$$

then the corresponding hazard ratio on $[\underline{s}, 1]$ is:

$$\begin{aligned} \frac{1 - G(z)}{g(z)} &= \frac{1 - [1 - (1 - (z - \underline{s})/(1 - \underline{s}))^{1/\theta}]}{\theta(1 - \underline{s})^{1/\theta - 1} (1 - (z - \underline{s})/(1 - \underline{s}))^{1/\theta - 1}} = \theta(1 - z) \\ &= \theta \left[1 - \frac{\bar{v} - p}{\rho\tau_v^{-1}} - x_i - (N - 1)(\mu - \gamma p) \right]. \end{aligned} \quad (\text{G.8})$$

Plugging Eqs. (G.4) and (G.8) into Eq. (G.5), we obtain:

$$\begin{aligned} \frac{1}{\rho\tau_v^{-1}} + (N - 1)\gamma &= \theta \frac{1 - (\bar{v} - p)/\rho\tau_v^{-1} - x_i - (N - 1)(\mu - \gamma p)}{\bar{v} - p - \rho\tau_v^{-1}x_i} \\ &= \theta \frac{1 - (\bar{v} - p)/\rho\tau_v^{-1} - N(\mu - \gamma p)}{\bar{v} - p - \rho\tau_v^{-1}x_i}, \end{aligned} \quad (\text{G.9})$$

which can also be written as:

$$x_k = \frac{1}{\rho\tau_v^{-1}} \left[\left(1 + \frac{\theta\sigma}{\rho\tau_v^{-1}} \right) (\bar{v} - p) + \theta\sigma N(\mu - \gamma p) - \theta\sigma \right], \quad (\text{G.10})$$

with

$$\sigma = \frac{1}{1/\rho\tau_v^{-1} + (N - 1)\gamma}. \quad (\text{G.11})$$

Thus:

$$\gamma = \frac{1}{\rho\tau_v^{-1}} \left[\left(1 + \frac{\theta\sigma}{\rho\tau_v^{-1}} \right) + \theta\sigma N\gamma \right], \quad (\text{G.12})$$

$$\mu = \frac{1}{\rho\tau_v^{-1}} \left[\left(1 + \frac{\theta\sigma}{\rho\tau_v^{-1}} \right) \bar{v} + \theta\sigma N\mu - \theta\sigma \right]. \quad (\text{G.13})$$

The solutions to the above equations are:

$$\gamma = \frac{N(1 + \theta) - 2 + \sqrt{N^2(1 + \theta)^2 - 4\theta}}{2(N - 1)\rho\tau_v^{-1}}, \quad (\text{G.14})$$

$$\begin{aligned} \mu &= \gamma\bar{v} - \frac{\theta}{(N - 1)\rho\tau_v^{-1}\gamma + 1 - N\theta} \\ &= \gamma\bar{v} - \frac{2\theta}{N(1 - \theta) + \sqrt{N^2(1 + \theta)^2 - 4\theta}} \\ &= \gamma \left[\bar{v} - \frac{2\rho\tau_v^{-1}\theta}{N(1 + \theta) + 2\theta + \sqrt{N^2(1 + \theta)^2 - 4\theta}} \right]. \end{aligned} \quad (\text{G.15})$$

We now verify the existence of a unique value of $\underline{z}$ such that $q(p, z) \geq 0$ when the market makers play the equilibrium strategies as described above. We can solve for the equilibrium price (from the market clearing condition) as follows:

$$\tilde{p} = \bar{v} - \frac{2N\theta/[N(1 - \theta) + \sqrt{N^2(1 + \theta)^2 - 4\theta}] + z}{1/\rho\tau_v^{-1} + N\gamma}. \quad (\text{G.16})$$

Thus, the customer's trading quantity is:

$$q(p, z) = z - \frac{2N\theta/[N(1 - \theta) + \sqrt{N^2(1 + \theta)^2 - 4\theta}] + z}{1 + \rho\tau_v^{-1}N\gamma}. \quad (\text{G.17})$$

It is straightforward to check that the following choice of threshold customer order size:

$$\underline{z} = \frac{2\theta}{N(1 + \theta) + 2\theta + \sqrt{N^2(1 + \theta)^2 - 4\theta}} \in (0, 1), \quad (\text{G.18})$$

is a unique number that satisfies the requirements that $q(p, z) \geq 0$ for $\tilde{z} \in [\underline{z}, 1]$. $\square$

References

- Anderson, R., Tychon, P., 1993. Competition among European financial markets: the case of cross-listed belgian equities. *Revue de la Banque/ Bank en Financiewezen* 7, 363–376.

- Back, K., Zender, J., 1993. Auctions of divisible goods: on the rationale for the treasury experiment. *Review of Financial Studies* 6, 733–764.
- Bernhardt, D., Hughson, E., 1997. Splitting orders. *Review of Financial Studies* 10, 69–102.
- Bernheim, D., Whinston, M., 1986. Menu auctions, resource allocation, and economic influence. *Quarterly Journal of Economics* 100, 1–31.
- Biais, B., 1993. Price formation and equilibrium liquidity in fragmented and centralized markets. *Journal of Finance* 48, 157–184.
- Biais, B., Foucault, T., Salanié, F., 1998. Floors, dealer markets and limit order markets. *Journal of Financial Markets* 3, 253–284.
- Biais, B., Martimort, D., Rochet, J.C., 2000. Competing mechanisms in a common value environment. *Econometrica* 68, 799–830.
- Booth, G., Iverson, P., Sarkar, S., Schmidt, H., Young, A., 1999. Market structure and bid–ask spreads: IBIS vs. NASDAQ. *European Journal of Finance* 5, 51–71.
- Brown, N., 1994. Information links and liquidity effects in parallel German and London equity markets, Working Paper.
- Christie, W., Schultz, P., 1994. Why do Nasdaq market makers avoid odd-eighth quotes? *Journal of Finance* 49, 1813–1840.
- Degryse, H., 1999. The total cost of trading Belgian shares: Brussels versus London. *Journal of Banking and Finance* 23, 1331–1355.
- De Jong, F., Nijman, T., Röell, A., 1995. A comparison of the cost of trading French stocks on the Paris Bourse and on SEAQ International. *European Economic Review* 39, 1277–1301.
- Glosten, L., 1989. Insider trading, liquidity, and the role of monopolist specialist. *Journal of Business* 49, 211–235.
- Glosten, L., 1994. Is the electronic open limit order book inevitable? *Journal of Finance* 49, 1127–1161.
- Hansch, O., Naik, N., Viswanathan, S., 1998. Do inventories matter in dealership Markets? Evidence from the London Stock Exchange. *Journal of Finance* 53, 1623–1656.
- Ho, T., Stoll, H., 1983. The dynamics of dealer markets under competition. *Journal of Finance* 38, 1053–1074.
- Huang, R., Stoll, H., 1996. Dealer versus auction markets: a paired comparison of execution costs on Nasdaq and the NYSE. *Journal of Financial Economics* 41, 313–357.
- Impenna, C., Maggio, P., Panetta, F., 1995. Innovazioni Strutturali nel mercato azionario: Gli Effetti della Contrattazione Continua. Banca d'Italia Working Paper.
- Kyle, A., 1989. Informed speculation with imperfect competition. *Review of Economic Studies* 56, 317–356.
- Maskin, E., Riley, J., 1989. Optimal multi-unit auctions. In: Hahn, F. (Ed.), *The Economics of Missing Markets, Information, and Games*. Clarendon Press, Oxford.
- Parlour, C., Seppi, D., 1997. Liquidity based competition for order flow. Carnegie-Mellon University Working Paper.
- Pirrong, C., 1999. The organization of financial markets: theory and evidence. *Journal of Financial Markets* 4, 329–358.
- Rock, K., 1996. The specialist's order book and price anomalies, unpublished working paper.
- Röell, A., 1992. Comparing the performance of stock exchange trading systems. In: Fingleton, F., Schoenmaker, D. (Eds.), *The Internationalization of Capital Markets and the Regulatory Response*. Graham and Trotman, London.
- Röell, A., 1998. Liquidity in limit order book markets and single price auctions with imperfect competition. Princeton University, Department of Economics Working Paper.
- Schmidt, H., Iverson, P., 1992. Automating German equity trading: bid–ask spreads on competing systems. *Journal of Financial Services Research*, 373–397.
- Seierstad, K., Sydsæter, K., 1977. Sufficient conditions in optimal control theory. *International Economic Review* 18, 367–391.

- Seppi, D., 1997. Liquidity provision with limit orders and a strategic specialist. *Review of Financial Studies* 10, 103–150.
- Viswanathan, S., Wang, J., 1996. Auctions with when-issued trading: a model of the U.S. treasury markets. Duke University Working Paper.
- Viswanathan, S., Wang, J., 2001. Inter-dealer trading in financial markets. Duke University Working Paper.
- Wang, J., Zender, J., Auctioning divisible goods. *Economic Theory*, forthcoming.
- Wilson, R., 1979. Auctions of shares. *Quarterly Journal of Economics* 93, 675–698.