Volume Title: The Industrial Organization and Regulation of the Securities Industry

Volume Author/Editor: Andrew W. Lo, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-48847-0

Volume URL: http://www.nber.org/books/lo__96-1

Conference Date: January 19-22, 1994

Publication Date: January 1996

Chapter Title: Brokerage, Market Fragmentation, and Securities Market Regulation

Chapter Author: Kathleen Hagerty, Robert L. McDonald

Chapter URL: http://www.nber.org/chapters/c8101

Chapter pages in book: (p. 35 - 62)

2 Brokerage, Market Fragmentation, and Securities Market Regulation

Kathleen Hagerty and Robert L. McDonald

2.1 Introduction

A striking fact about the organization of modern financial markets—and one of the great interest to market regulators and exchanges—is the prevalence of market fragmentation, that is, multiple mechanisms or locations for trading a security. A share of common stock, for example, may be traded on one of many organized exchanges, through dealers away from an exchange, in another country, or indirectly through a derivative financial contract, which in turn may be traded on an exchange or through a dealer.

To the extent that securities markets provide a central trading location serving to minimize the search cost of finding a counterparty, fragmentation is a puzzle. On the other hand, market participants often have private information, either about the "true value" of the traded security, or about their trading motives.[1] In markets with asymmetric information, informed traders earn a profit at the expense of the uninformed traders.[2] Therefore there is clearly an incentive to create mechanisms that mitigate (for at least some subset of participants) costs created by the existence of private information. One obvious way for uninformed traders to minimize these costs is to trade in a nonanonymous

Kathleen Hagerty is associate professor of finance at the Kellogg Graduate School of Management at Northwestern University. Robert L. McDonald is the Erwin P. Nemmers Distinguished Professor of Finance at the Kellogg Graduate School of Management at Northwestern University and a research associate of the National Bureau of Economic Research.

The authors are grateful for comments from Larry Harris, Geoffrey Miller, seminar participants at the University of Houston and Massachusetts Institute of Technology, and participants at the NBER conference.

1. As Fischer Black (1990) has observed, having no information and knowing you have no information can be valuable private information.

2. Kyle (1985), Glosten and Milgrom (1985), and Admati and Pfleiderer (1988) are examples of models like this.

market. Nonanonymous arrangements often have the appearance of a fragmented market.

Although forms of fragmentation have changed over time with technological and new product developments, fragmentation has been a perennial issue of interest to regulators and market participants. There have been a number of academic papers on topics related to fragmentation (see, for example, Chowdhry and Nanda [1991]; Madhavan [1993]; Pagano [1989]; and Röell [1990]). Harris (1992) in particular discusses fragmentation in detail and argues that it is an outgrowth of different traders' having different trading needs (for example, immediacy versus price improvement).

This paper discusses fragmentation in the context of two observations: first, brokers expend resources inducing investors to trade; second, most securities markets are replete with mechanisms permitting firms to capture order flow, including both implicit and explicit payments for order flow. We argue that these observations together suggest a reason for persistent fragmentation: when order flow arises from brokerage activity, the information characteristics of that order flow are known to the broker, or at least better known to the broker than to an average market maker. There is then an incentive for the broker to serve as a counterparty to the trade rather than just as a brokerage conduit. In order to avoid issues related to dual trading and front running, we assume that the brokers do not trade for their own account on information extracted from the customer order flow nor do they hold any uncrossed trades in their portfolio. This differs from the focus of papers such as Fishman and Longstaff (1992), Roell (1990), and Grossman (1992).

We assume that there exist both perfectly informed and uninformed traders, who have a choice of two ways to trade: with a broker who knows the trader's informedness, or anonymously with a market maker. We model market makers as in Easley and O'Hara (1987)—they are risk-neutral competitive agents who set a bid and ask price and accept all one-share orders at those prices. Brokers, however, serve a quite different function. They accept orders from traders, charging a bid-ask spread that is possibly type-dependent, cross buys against sells, and then export to a market maker any remaining order flow, paying the market bid-ask spread. Unlike market makers, brokers in our model face no price risk. And since brokers do not retain shares, they risklessly earn the bid-ask spread on any orders they net against one another.

In this setting, we compare monopolist and competitive brokers. In all cases, this kind of brokerage activity increases the bid-ask spread charged by market makers. The two kinds of broker treat their customers quite differently, however. The monopolist broker simply charges customers the market bid-ask spread. Competitive brokers, on the other hand, charge different spreads to the two types of customers, and in general *both* kinds of customers will be charged a lower spread than that set by the market maker. The ability of the broker to net orders, coupled with competition, forces the broker to charge each type of customer a price reflecting the contribution of that kind of order to the expected net order export. Perhaps surprisingly, perfectly informed investors will in general be charged a price less than the full bid-ask spread.

We find that fragmentation may be a reflection of increased price competition, and that the fragmented and competitive system provides better prices for customers than the less-fragmented monopolistic broker case. Order flow intermediaries in this setting foster competition at the brokerage level but raise the bid-ask spread in the central market.

As an example, consider a corporation that receives advice from trading firms and then undertakes a hedging transaction in the over-the-counter (OTC) derivatives market. By virtue of understanding the customer's trading motives, the brokers understand the information content of the order. When the customer obtains quotes on the deal, brokers will bid in a way that reflects their expected cost of hedging (or "exporting") the trade. Obviously this expected cost will be greater if the broker expects other orders in the same direction.

The perspective on fragmentation in this paper can be contrasted with the view that fragmentation reflects skimming the best customers from existing markets. First, brokers in this paper optimally accept all customers, informed and uninformed. Second, while we do not model the market participation decision of potential traders, our model is consistent with a world in which brokers can increase order flow by expending resources. The mechanisms that result in fragmented markets are also the mechanisms that give brokers the incentive to generate business. This view suggests that there is no "silver bullet" trading system that, if implemented, would attract all order flow. Rather, it suggests that central markets and brokerage markets serve different needs for different investors. This argument is very similar to aspects of Merton (1992) and Harris (1992).

Obviously, there are dimensions to fragmentation other than those we emphasize. In particular, agency problems—which we ignore—may be severe in practice. Our goal in this paper is not to exhaustively explore legal, regulatory, and practical issues associated with fragmentation, but rather to provide a framework for thinking about the link between fragmentation and market liquidity.

Section 2.2 introduces the model of the security market, and sections 2.3 and 2.4 look at the equilibrium bid-ask spreads and distributions of traders across brokerage firms and the central market in monopoly and perfectly competitive environments, respectively. Section 2.5 discusses policy implications. Proofs of the lemmas and propositions are in the appendix.

2.2 The Model

In this section, we describe the trading environment for a single risky asset. The assumptions and analysis are most similar to Easley and O'Hara (1987). The value of the risky asset is given by $\theta = \theta_H$ with probability .5 or θ_L with probability .5.

2.2.1 Traders

All traders are constrained to buy one unit or sell one unit of the asset. There are two types of traders, traders who are perfectly informed, of whom there are N_1, and traders who are uninformed, of whom there are N_0. When a trader is perfectly informed, he buys if $\theta = \theta_H$ and sells if $\theta = \theta_L$ with probability 1. An uninformed trader buys one unit with probability .5 and sells one unit with probability .5. A fraction ρ of uninformed traders are discretionary; the rest are nondiscretionary. The discretionary uninformed traders, of whom there are ρN_0, can choose where to trade. The nondiscretionary traders, of whom there are $(1 - \rho)N_0$, must trade in the central market. Fully informed traders can trade wherever they choose.

2.2.2 Market Makers

The central market can be thought of as the floor of an exchange or as a dealer market. In this market, trades are submitted simultaneously and anonymously to a risk-neutral market maker. Prior to observing order flow, the market maker posts a bid, b, and an ask, a, at which he will satisfy all orders.

The bid and ask prices are set so that the market maker earns zero expected profits. Let B_H equal the expected number of buys in the market if $\theta = \theta_H$ and let B_L be the expected number of buys in the market if $\theta = \theta_L$. In order for the market maker to earn zero expected profits, the ask must satisfy

$$0 = .5B_H(a - \theta_H) + .5B_L(a - \theta_L).$$

Solving for a, we get

$$a = \frac{B_H\theta_H + B_L\theta_L}{B_H + B_L}. \tag{1}$$

The bid, which is derived in an analogous way, is equal to

$$b = \frac{S_H\theta_H + S_L\theta_L}{S_H + S_L}, \tag{2}$$

where S_L and S_H are the expected number of sell orders if $\theta = \theta_L$ and $\theta = \theta_H$, respectively. The number of buys and sells in the market is affected by both the traders who trade directly in the central market and the net exports of the brokerage firms. Given symmetry between good and bad information states and the symmetry between uninformed traders' propensity to buy or sell, it follows that $B_H = S_L$ and $B_L = S_H$. Therefore, the bid-ask spread can be written as

$$a - b = \frac{(B_H - B_L)(\theta_H - \theta_L)}{B_H + B_L}. \tag{3}$$

2.2.3 Brokers

A brokerage firm is a profit-maximizing trading firm. By assumption, brokers know a customer's type. In general, we can imagine that brokers learn a customer's type as part of marketing brokerage services. For example, a customer could be a firm with hedgeable exposure (e.g., currencies, interest rates, commodity prices), but where management lacks knowledge of financial hedging products, and is uncertain how to measure its own exposure. Because of correlations among input and output prices, it is often not obvious what derivatives position would constitute a hedge. While helping the firm determine the appropriate hedge, the broker learns the motives for the trade and the market views of the customer.

The broker is assumed not to invest for its own account. Brokers are able to immediately and frictionlessly export net order flow to the central market; thus brokers never bear inventory price risk. Obviously, in practice, broker-dealers will hold some inventory since, although they may provide immediate execution to customers, they will not be able to export large net order flow immediately without a price penalty. This in turn will make them sensitive to the information of customers: other things equal, an order from a more informed customer is costlier to accept.

As in the central market, the customers of a brokerage firm either buy one unit or sell one unit. After the buys and the sells are netted, the firm exports the unmatched buys or sells to the market. Note that, whereas individual traders trade only one share, the broker is able to anonymously trade many shares at once. This means that the brokerage firms earn a spread on any matched orders and must pay a spread on any unmatched orders.

The requirement that net order flow be exported to the central market can be motivated by the assumption that brokers face capital requirements and have limited capital. In this case, they would be forced to export net order flow or, similarly, run a hedged book. Many broker-dealers do in fact operate this way. For example, if a broker-dealer serves as counterparty in long-lived transactions involving OTC derivatives, customers will be concerned about broker-dealer credit ratings and long-term viability; hedging the book thus enhances the ability of the firm to engage in these transactions. The requirement that brokers not hold shares means that they are unable to trade on information they may glean from customer order flow.

Let n_0, $n_0 \leq \rho N_0$, and n_1, $n_1 \leq N_1$, denote the number of uninformed and informed customers, respectively, of a brokerage firm. Let S_0 and S_1 denote the number of sell orders by uninformed and informed customers, respectively, and let B_0 and B_1 denote the number of buy orders by uninformed and informed customers, respectively, where $B_0 + S_0 = n_0$ and $B_1 + S_1 = n_1$. Let $B = B_0 + B_1$ and $S = S_0 + S_1$. The ask and bid prices, a_i and b_i, are charged by a brokerage firm to customers of type i, $i = 0, 1$. The profit of a brokerage firm is given by

$$\tilde{\pi} = a_0B_0 - b_0S_0 + a_1B_1 - b_1S_1 - \begin{cases} -a(B - S) & \text{if } B - S > 0, \\ +b(S - B) & \text{if } S - B > 0, \end{cases}$$

and expected profits are therefore

$$\pi(n_0, n_1) = E(\tilde{\pi}) = \tfrac{1}{2}(a_0 - b_0)n_0 + \tfrac{1}{2}(a_1 - b_1)n_1 - aE(B - S \mid B - S > 0) \\ \text{Prob}(B - S > 0) + bE(S - B \mid S - B > 0)\text{Prob}(S - B > 0)$$

Since $E(B - S \mid B - S > 0)\text{Prob}(B - S > 0) = E(S - B \mid S - B > 0)\text{Prob}(S - B > 0)$, we can rewrite the expected broker profit as

$$\pi(n_0, n_1) = \tfrac{1}{2}(a - b)[\lambda_0 n_0 + \lambda_1 n_1 - \phi(n_0, n_1)], \tag{4}$$

where

$$\lambda_i = (a_i - b_i)/(a - b),$$

and

$$\phi(n_0, n_1) = 2E(B - S \mid B - S > 0)\text{Prob}(B - S > 0). \tag{5}$$

λ_i is the fraction of the market bid-ask spread that the broker charges to a customer, and $.5\phi$ is the expected number of trades exported to the market by the brokerage firm. Note that the expected broker profit is proportional to the market bid-ask spread, and is decreasing in the number of orders exported to the central market.

In the following sections, we first look at the mathematics of the net exports and then at the effect of the broker market on the distribution of traders across the central market and brokerage firms and the effect on the spreads that traders must pay. Two market structures are considered, a monopoly brokerage firm and perfectly competitive brokerage firms.

2.2.4 Understanding Net Order Exports

Since the focus of the paper is on order netting at the broker level, it is crucial to understand the behavior of $\phi(n_0, n_1)$, the expected order export. We are interested in two properties of ϕ: (1) how does the expected order export change when new customers are added, and (2) how does the expected order export change when one type of customer is replaced by the other. Let W denote the total number of customers, that is, $W = n_0 + n_1$; when there is no possibility of confusion, we will write $\phi(n_0)$, suppressing the argument $n_1 \equiv W - n_0$. For future reference, we now state some facts about ϕ.

LEMMA 1. (1) The expected net order flow is nondecreasing if one customer of either type is added, that is,

$$\phi(n_0, W - n_0) \geq \phi(n_0, W - n_0 - 1),$$

$$\phi(n_0, W - n_0) \geq \phi(n_0 - 1, W - n_0).$$

(2) If an uninformed customer is added, the expected net order flow per customer is nonincreasing:

$$\frac{\phi(n_0, W - n_0)}{W} \leq \frac{\phi(n_0 - 1, W - n_0)}{W - 1}.$$

(3) If an informed customer is added, the expected net order flow per informed customer is nonincreasing:

$$\frac{\phi(n_0, W - n_0)}{W - n_0} \leq \frac{\phi(n_0, W - n_0 - 1)}{W - n_0 - 1}.$$

For a given mix of customers, adding customers increases the expected net order export but at a decreasing rate, in a sense lemma 1 makes precise. It may be surprising that adding perfectly informed investors does not always increase net order flow by one. To see this, consider the case where $n_0 > 0$ and $n_1 = 0$. At this point, adding a single informed investor is exactly like adding an uninformed investor because the informed investor's trade is uncorrelated with the net trade of the uninformed investors. Adding a second informed investor contributes to net order flow by more than an uninformed investor since the informed trader's trade is partially correlated with the trades of the n_0 uninformed investors and the one informed investor. As we add informed investors, the marginal contribution to order flow increases. At the point where $n_1 \geq n_0$, additional informed investors increase net order flow one for one. This provides intuition for the second result:

LEMMA 2. $\phi(n_0, W - n_0)$ is decreasing and convex in n_0.

The intuition for both properties is straightforward. Increasing n_0 by 1 substitutes an uninformed customers for an informed customer. Conditional on net order flow being positive, the magnitude of the net order flow is smaller, the smaller the informed (and hence positively correlated) component of order flow.

Convexity implies that substituting informed for uninformed order flow increases the conditional expectation at an increasing rate. Substituting a single informed order for one uninformed order adds an order that is uncorrelated. Substituting an informed order for an uninformed order when there is a mix of informed and uninformed orders replaces an uncorrelated order with a partially correlated order and thus increases the conditional expectation.

2.3 Monopoly Brokerage Market

2.3.1 The Profit of the Monopoly Firm

Consider a market with a single brokerage firm. Traders are willing to trade with the brokerage firm as long as the firm matches the market spread. (This implies that the traders do not take into account the effect that their choice has on the equilibrium spread.) From examination of the profit function, (4), it is obvious that the monopoly broker sets $\lambda_0 = \lambda_1 = 1$; that is, he charges customers the full market spread. No customers would be attracted at a higher price, and the broker would needlessly give up profits at a lower spread. The brokerage firm chooses the number of informed and uninformed customers it wishes to have. From (4), therefore, the broker's maximization problem is to choose n_0 and n_1 so as to maximize

$$\pi(n_0, n_1) = \tfrac{1}{2}[a(n_0, n_1) - b(n_0, n_1)][n_0 + n_1 - \phi(n_0, n_1)]. \tag{6}$$

The choice of n_0 and n_1 affects the brokerage firm's profits in two ways. First, it affects how well matched are the buys and the sells. Other things equal, the greater the matching of trades, the more profitable the brokerage firm. Second, the choice of n_0 and n_1 affects the ratio of informed to uninformed in the central market, which in turn affects the size of the market bid-ask spread. The monopolistic broker takes this into account in selecting customers.

2.3.2 Determination of the Market Bid-Ask Spread

The market bid-ask spread for a given customer mix, (n_0, n_1), is given in the following proposition.

PROPOSITION 1. For a given customer mix (n_0, n_1), the market bid-ask spread is given by

$$a(n_0, n_1) - b(n_0, n_1) = \frac{N_1(\theta_H - \theta_L)}{(N_0 + N_1) - n_0 - n_1 + \phi(n_0, n_1)}.$$

If $n_0 \leq n_1$, the spread simplifies to

$$a(n_0, n_1) - b(n_0, n_1) = \frac{N_1(\theta_H - \theta_L)}{(N_0 + N_1) - n_0}.$$

COROLLARY 1. The market spread is increasing in n_0 and nondecreasing in n_1.

The number of buys and sells in the market is affected by the net exports of the brokerage firm. As the number of uninformed customers increases, more uninformed trades are absorbed and crossed internally by the broker, which means they are never seen by the market maker. With effectively fewer uninformed traders, the market maker will raise the bid-ask spread. When the broker accepts a perfectly informed trade, the bid-ask spread either increases (if $n_0 > n_1$) or remains unchanged. If $n_0 > n_1$, an informed trader switching from the market maker to the broker does not change the expected difference in the number of "correct" and "incorrect" trades (e.g., $B_H - B_L$), but there is a decrease in the total number of trades the market maker sees (because the broker absorbs some trades). When $n_0 \leq n_1$, an additional informed trade switching from the market to the broker does not affect the expected number of trades in any state.

Substituting the bid-ask spread into (6) yields the broker's expected profit for a given customer mix (n_0, n_1):

$$\pi(n_0, n_1) = .5 \frac{N_1(\theta_H - \theta_L)(n_0 + n_1 - \phi(n_0, n_1))}{(N_0 + N_1) - n_0 - n_1 + \phi(n_0, n_1)}$$

2.3.3 Equilibrium in the Market with a Monopoly Brokerage Firm

Brokers in general want more orders, since both the bid-ask spread and the expected number of crossed orders are increasing in $n_0 + n_1 - \phi(n_0, n_1)$, and $n_0 + n_1 - \phi(n_0, n_1)$ is nondecreasing in scale. Thus, brokers accept as many orders as possible, up to the point where $n_0 = n_1$. This is summarized in the following proposition:

PROPOSITION 2. (1) The broker sets $n_0 = \rho N_0$. (2) The broker accepts informed trades up to the point where $n_1 = n_0 = \rho N_0$ and is indifferent about accepting further trades. (3) The expected profit is equal to

$$\pi(n_0, n_1) = .5 \frac{N_1(\theta_H - \theta_L)(\rho N_0)}{(N_0 + N_1) - \rho N_0}.$$

The equilibrium spread is maximized since the discretionary uninformed traders have been absorbed by the broker and the number of informed traders is unaffected by the presence of the broker. All traders, whether they trade with the broker or in the central market, are worse off relative to a setting where there is only a central market due to the higher spread. It is also interesting to note that, unlike many microstructure models, profits are nondecreasing in the number of informed traders. This is because unmatched informed trades are not held in inventory and hence do not impose any cost on the broker.

Note that the broker earns positive profits from uninformed customers. Thus, it is in the broker's interest to induce uninformed traders to trade. While outside the model, suppose that the brokerage firm could attract uninformed traders who were not participating in the market. These traders could be thought of as people who can be induced to trade if the brokerage firm makes some kind of marketing effort. In this case, the firm would not just be skimming from the central market but increasing the participation in the market as a whole. The addition of new customers lowers the spread, but the effect of netting is such that many new customers would be required for the spread to be as low as it would have been had there been no brokerage firm. With endogenous order flow (which we do not model), there is an ambiguous effect on the market spread of broker crossing of customer orders.

2.4 Competitive Brokers

In the competitive case, we assume that the composition of orders between uninformed and informed investors is selected by the broker; however, brokers are exogenously constrained to not accept more than W orders each. Generally speaking, the larger the scale of the broker, the more netting is achieved on a per-customer basis, and a larger broker could therefore charge customers lower prices. The restriction on firm size is necessary to ensure a competitive outcome, and it could reflect diseconomies of scale to brokerage activities. However, we show later that, in the presence of what we term "order flow intermediaries," the limit on firm scale is moot. We also assume that there is free entry into brokerage and to simplify the solution (i.e., to avoid an integer problem), we assume there are Mn_0^* uninformed customers in the pool available to brokers, and $M(W - n_0^*)$ informed customers.

2.4.1 Analysis of Equilibrium

Unlike in the monopoly case, where all customers pay the same price, in the competitive case we model the broker as posting a type-specific price for customers, and brokers are permitted to choose the fraction of informed and uninformed customers. Customers either accept or reject the posted price. Clearly, given the free-entry condition, in equilibrium all brokers must post the same type-specific price. The complication associated with type-specific prices is that customers with different information contribute differently to netting of orders. Further, the contribution to netting depends on both the size of the firm and the customer composition. Since the broker does not actually hold the order, however, the information content of a particular order does not affect the broker *except* as it affects netting. This differs from standard microstructure models such as Kyle (1985), where risk arising from holding the order flow is the reason prices depend on the information content of the order flow.

Unlike in the monopolistic case, we assume that the number of customers and brokerage firms is sufficiently large that, in setting prices, brokers do not take into account their effect on the market bid-ask spread. The problem for the brokerage firm is to set prices for customers that take account of the effect on order flow export. As noted above, λ_i in general depends on both the scale and customer composition of the firm. For example, suppose the firm has only uninformed customers. Clearly, the more such customers, the lower the cost per customer of order flow exports, and the lower the average competitive charge to a customer.

From (4), we can see that if brokers earn zero profits, we have

$$(7) \qquad \phi(n_0, W - n_0) = n_0\lambda_0(n_0, W - n_0) + (W - n_0)\lambda_p(n_0, W - n_0).$$

We now have the apparatus necessary to define an equilibrium. Recall that, by assumption, competitive brokerage firms can have only W customers. Brokers announce type-specific prices for orders, and customers give their order to the broker with the lowest price. Therefore, given the free-entry condition, all brokers with a positive order flow from a given type charge the same price for a given type. We use the Nash equilibrium concept, in that brokers take as given the prices of other brokers.

DEFINITION 1. An equilibrium is a type-specific price schedule $\{\lambda_0, \lambda_1\}$ such that (1) brokers earn nonnegative expected profits, taking as given prices of other brokers; (2) newly entering brokers earn nonpositive expected profits; (3) customers are no worse off dealing with brokers than transacting directly at the market bid-ask spread.

The next proposition is a direct consequence of the convexity of ϕ.

PROPOSITION 3. All brokers select the same mix of informed and uninformed customers.[3]

Note that in no case can the equilibrium price for a given type exceed the zero-profit price for a broker that accepted customers only of that type. Otherwise, it would be profitable for a broker to specialize in customers of that type. Thus, we have

PROPOSITION 4. Type-specific prices can never be greater than the prices charged by a firm specializing in customers of a given type: $\lambda_0(n_0) \leq \lambda_0(0)$; $\lambda_1(n_0) \leq \lambda_1(W)$.

Convexity of ϕ implies that, given equilibrium prices, all brokers seek to have the same mix of informed and uninformed customers. Given a potential equilibrium, however, we must verify that at the posited prices it does not pay for the broker to accept a different mix of customers. In particular, if n_0 is the equilibrium customer mix, it must be the case that profits are nonpositive at the same prices but with a different customer mix, that is,

$$(8) \quad (n_0 + j)\lambda_0(n_0) + (W - (n_0 + j))\lambda_1(n_0) \leq \phi(n_0 + j) \qquad -n_0 \leq j \leq W - n_0.$$

Because ϕ is convex, an equilibrium price schedule exists, and we present an example price schedule below. The proof of existence involves showing the following:

LEMMA 3.1. At an equilibrium price schedule, if it is not optimal to switch the customer mix for $j = \pm 1$, then it is not optimal to switch for any larger j. Thus, in verifying that (8) is satisfied for a given n_0, it suffices to check the cases $j = \pm 1$.

LEMMA 3.2. Convexity of ϕ is a necessary condition for the existence of an equilibrium price schedule.

LEMMA 3.3. Equilibrium $\lambda_0(n_0, n_1)$ and $\lambda_1(n_0, n_1)$ schedules are given by

$$\lambda_0(W, 0) = \phi(W, 0)/W; \lambda_1(0, W) = \phi(0, W)/W;$$

$$(9) \qquad \lambda_1(n_0, W - n_0) = \frac{\phi(n_0, W - n_0)}{W} + \frac{n_0}{W}[\phi(n_0 - 1, W - n_0) - \phi(n_0, W - n_0 - 1)];$$

3. Note that, because of the assumption that there are MW brokerage customers, the only way there can be two different customer mixes is if n_0^1 and n_0^2 differ by more than one.

$$(10)\quad \lambda_0(n_0, W - n_0) = \frac{\phi(n_0, W - n_0)}{W} + \frac{W - n_0}{W}[\phi(n_0, W - n_0 - 1) - \phi(n_0 - 1, W - n_0)].$$

Since the convexity of ϕ is used to construct a price schedule, and since convexity is necessary for an equilibrium to exist, we have

PROPOSITION 5. There exists an equilibrium price schedule if and only if ϕ is convex.

From lemma 2, ϕ is convex. Since we have assumed that the numbers of traders accessible to brokers is a multiple of W, we have

COROLLARY 2. It is an equilibrium for there to be M brokerage firms, each with n_0^* uninformed and $W - n_0^*$ informed customers.

There are several interesting features of the equilibrium. First, if a mix of customers exists, then it is optimal for all firms to have a mix of customers. It would not be optimal, for example, to specialize in uninformed customers if other firms had a mix of informed and uninformed customers. The intuition for this follows from the properties of an equilibrium price schedule: for a fixed number of total customers, as brokers choose to trade with fewer customers of a given type, the price charged customers of that type declines. For example, the marginal value of an uninformed customer is greater, the smaller the number of uninformed customers (and hence the greater the number of informed customers). Similarly, adding an informed customer to a group of uninformed customers is not costly because the informed customer is uncorrelated with existing order flow. However, adding an informed customer to a large number of informed customers is more expensive.

The value of having uninformed customers is greater for a broker who also has informed customers, so firms that have both types would be able to undercut the price of firms that had only uninformed customers. Effectively, customers of different types cross-subsidize one another.

If all customers who can deal with brokers do so, and if brokers charge equilibrium fees as outlined above, then customers are at least as well off dealing with brokers as with market makers, and brokers have no incentive to attract a different mix of customers, given prices charged by other brokers.

It is worth emphasizing that, since customers end up receiving a better bid-ask spread from the broker than from the market maker, equilibrium requires that all customers who can use a competitive broker, do use a competitive broker. Otherwise, customers of the same type would pay different prices for the same order.

2.4.2 Example

Figures 2.1 and 2.2 present a numerical example, with the competitive brokerage firm having a scale of ten customers (i.e., $W = 10$). This example dem-

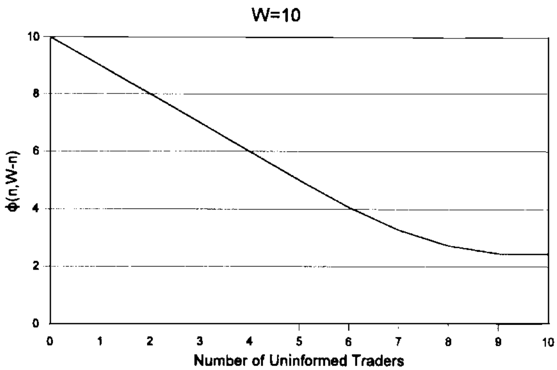

Fig. 2.1 Change in expected net order export, $\phi(n, W - n)$, as fraction of uninformed customers increases
Note: Assumes ten customers.

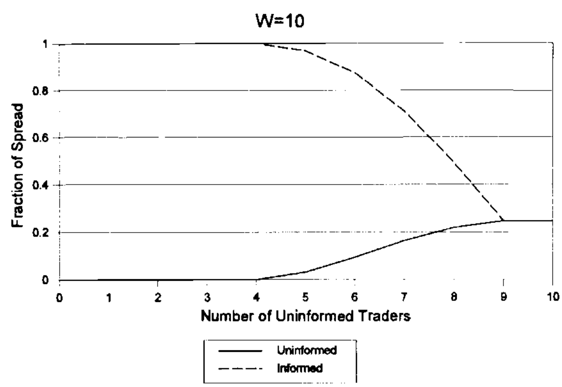

Fig. 2.2 Fraction of market spread charged to each type of customer, $\lambda_0(n, W - n)$ and $\lambda_1(n, W - n)$, as fraction of uninformed customers increases
Note: Assumes ten customers.

onstrates how customers of different types effectively cross-subsidize each other in the competitive equilibrium. Figure 2.1 displays $\phi(n_0, n_1)$ for all different possible mixes of customer types. Figure 2.2 displays the competitive spreads charged to each type of customer, computed using the schedule derived in the appendix in the proof of lemma 3. The cross-subsidization is evident in the first informed customer being charged the same spread as uninformed customers. As the customer mix moves from uninformed to informed, the spread charged the informed rises and that charged the uninformed falls. Once there are more informed than uninformed customers, the uninformed are charged a zero spread since *ex ante* the contribute nothing to the firm's net order exports.

2.4.3 Effect on the Market Bid-Ask Spread

As in the monopolistic case, $\phi(n_0, W - n_0)$ becomes linear when $n_0 < W/2$, with the result that the broker is indifferent about accepting additional uninformed traders. Thus, as in the monopolistic case, there is a corner solution at $n_0 = W/2$. Although the maximization problems faced by the monopolist and the competitive brokers are different, the implications for the market bid-ask spread are similar. Since both kinds of brokers net orders and export the residual, exported order flow represents more information than the original order flow.

The difference between the monopolist and the competitive broker stems from the fact that competitive order flow exports are greater by

$$\sum_{i=1}^{M} \phi_i(n_0^*, W - n_0^*) - \phi(Mn_0^*, M(W - n_0^*)).$$

Thus, the increase in the spread is smaller in the competitive case. Notice, however, that this implies that, for a given spread, brokerage customers do not pay the lowest possible fees because there is less than full netting at the broker level. This suggests that it would be profitable for a broker to enter as a "broker's broker," crossing broker net order flow.

2.4.4 Competitive Order Flow Intermediaries

Because the expected order export per customer declines with the number of customers, there is a natural economy to scale for brokers. If there are competitive order flow intermediaries, however, scale economies can be achieved through the purchase and sale of order flow. For example, suppose there are M retail brokers who on average export $\phi(n_0, W - n_0)$ shares each. Unconditionally, this export has a 50 percent chance of being either a buy or a sell. Assuming that the order flow broker knows the customer characteristics of the retail broker, expected profits are

$$M[.5a\phi(n_0, W - n_0) - .5b\phi(n_0, W - n_0)] - .5(a - b)\phi(Mn_0, M(W - n_0)).$$

With entry, the order flow broker's expected profits must be zero. This means that in order to attract business the order flow broker charges the retail broker a discount from the market spread. On a per-firm basis, this discount is

$$.5(a - b)[\phi(n_0, W - n_0) - \phi(Mn_0, M(W - n_0))/M].$$

Since $\phi(Mn_0, M(W - n_0))/M$ is decreasing in M, the order flow broker provides a bigger discount the greater the number of customers.

Notice that with this discount added to the broker profit function (4), the broker sets a spread for customer orders that is equivalent to having MW customers. As long as there are no transaction costs or frictions, there can be many layers of order flow brokers, each buying order flow, aggregating it, and giving to the preceding broker a discount reflecting the benefits of aggregation.

In comparing the monopolistic equilibrium and competitive equilibrium where there are order flow intermediaries, we find that both produce the same effect on the spread, but customers receive better prices in the competitive case because brokers bid for orders and pass along any cost advantages. The interesting point is that the competitive case is more "fragmented," in the sense that there are many more separate mechanisms by which shares are exchanged. In this setting, however, fragmentation is associated with benefits for customers because it reflects increased competition for orders.

2.5 Discussion and Policy Implications

In this section, we discuss some policy implications of the model, as well as its robustness.

2.5.1 OTC versus Exchange-Oriented Markets

We have focused on the implications of order netting for market spreads. A natural question is, to what kind of market does this model apply? One interpretation of the model is that brokers are also serving as principals in a trade, and customer prices and orders are determined individually in negotiations with the broker without any requirement that they be exposed to the central market. This description is suggestive of the OTC market for custom financial products, in which broker-dealers originate the trade, serve as counterparty, and generally hedge the transaction in some related central market.

The model also has applicability for thinking about equity markets, however. A common description of trade in listed equities in the United States characterizes stock trading as involving the submission of an order to a central exchange, where it is bid for by a variety of market participants. While this does appear to describe how a significant fraction of orders are handled, the National Association of Securities Dealers (1991) report on order flow inducement practices makes clear that there are many ways for firms to acquire order flow.

- *Explicit payment for orders.*
- *Agreements for exchange of order flow.* For example, on Nasdaq, broker A may direct orders for stock X to broker B, who in turn directs orders for stock Y to broker A. Such order-preferencing arrangements are often explicit.
- *Vertical integration of brokers and specialists.* For a given stock, the Intermarket Trading System reports posted bids and asks from a variety of locations. However, any market maker can take an order for his own account by matching the best quoted price. Thus, a broker for securities firm XYZ can send orders to market maker A, who is owned by XYZ. The market maker can then accept the order on behalf of the brokerage firm by matching the best current price.

All of these practices effectively provide ways for brokers to selectively choose to be counterparty to an order.

In addition, Angel (1994), using data from the New York Stock Exchange (NYSE), shows that different kinds of order flow (e.g., retail versus proprietary trades) systematically receive different amounts of price improvement over the spread, as would be suggested by our model. We conclude that the model is at least partially descriptive of both exchange-centered and OTC markets.

2.5.2 Policy Implications

The interesting policy implication is that increased fragmentation reflects increased competition for orders, generating better prices for customers and no worse prices in the central market. Of course, this results from a comparison of central market prices between two settings with broker-dealers. Since we do not model the decisions of customers to trade in the first place, welfare comparisons with just a central market (which typically will yield a lower spread) are not possible. It seems safe to speculate, however, that in general there will always exist parameters where such comparisons are ambiguous, and the purpose of analysis such as this is to point out sources of costs and benefits.

Although we do not model the generation of order flow, it is obvious that once a broker is permitted to also serve as market maker, the incentives to generate order flow are increased: in addition to generating commissions, the broker can earn some portion of the spread as well. This raises the issue of endogenous order flow generation, or "order flow discovery." In order to capture order flow, broker marketing efforts may include customer education about markets and about personal or corporate financial issues. The result may be increased trading and a welfare improvement for customers. The offsetting effect, of course, is that liquidity is typically reduced for traders who use the central market.

This view of fragmentation also has implications for the empirical literature on trade execution quality. It is well-documented (e.g., Lee [1993]; Petersen and Fialkowski [1992]) that execution quality for a given stock differs among stock exchanges in the United States with the NYSE typically providing execution at least as good as regional exchanges. One common interpretation of poorer performance on regional exchanges is that it reflects an agency problem resulting from vertical integration of brokers and market makers; the incentive of the broker to seek the best price is compromised. However, it could also reflect a payment to the broker in exchange for marketing services that would not have been performed in the first place had the broker not been able to route the trade to a particular market maker.[4]

There are several implications for policy. First, given the existence of asymmetric information and the broker's superior knowledge about the quality of order flow, it is inevitable that brokers will try to capture order flow. Preventing

4. This discussion ignores other dimensions of trade execution, such as immediacy and depth.

capture of order flow would not be unambiguously beneficial. Even mechanisms designed to ensure the best execution price, such as a consolidated limit order book (CLOB), may not guarantee best execution in a broad sense (by reducing the ability to capture order flow, the CLOB reduces the incentive by brokers to provide marketing services). In addition, the creation of derivative securities and offshore trading provide ways to bypass a CLOB.

From a regulatory perspective, it is desirable to increase competition among brokers, and a key to doing this is making sure that customers have enough information to make an informed choice among different brokerage and trade execution practices. Since marketing services most benefit precisely those customers least able to make an informed choice, this is likely to be difficult.

Appendix

Using assumptions from the text, the explicit expression for ϕ (the expected number of net buy orders conditional on buy orders exceeding sell orders) is

$$\text{(A1)} \quad \begin{aligned} \phi(n_0, n_1) = 2\Bigg[& \sum_{i=ceil[W/2-n_1]}^{n_0} \left(i + n_1 - \frac{W}{2}\right)\binom{n_0}{i}(.5)^{n_0} \\ & + \sum_{i=ceil[W/2]}^{n_0} \left(i - \frac{W}{2}\right)\binom{n_0}{i}(.5)^{n_0}\Bigg], \qquad \text{if } n_0 > n_1; \\ \phi(n_0, n_1) = n_1, & \qquad \text{if } n_0 \leq n_1, \end{aligned}$$

where $ceil(x)$ is the smallest integer greater than or equal to x and $W = n_0 + n_1$. Since informed orders can be either buys or sells, there are different terms in the first expression accounting for these two cases.

The following recursive relationships may be verified by direct calculation:

$$\text{(A2)} \quad \phi(n_0, n_1) = \phi(n_0 - 1, n_1) + \tfrac{1}{2}\eta(n_0 - 1, n_1);$$

$$\text{(A3)} \quad \phi(n_0, n_1 + 1) = \phi(n_0, n_1) + \sum_{i=ceil(W/2)-n_1}^{floor(W/2)} \binom{n_0}{i}.5^{n_0};$$

where

$$\begin{aligned} \eta(n_0, n_1) &= .5^{n_0}\left[\binom{n_0}{\frac{W}{2} - n_1} + \binom{n_0}{\frac{W}{2}}\right], & W \text{ even}, \\ &= 0, & W \text{ odd}, \end{aligned}$$

is the probability of exactly zero buy orders with x uninformed and y informed customers. Note that when W is even, removing one uninformed customer leaves ϕ unchanged.

Proof of Lemma 1. (1) From equations (A2) and (A3), ϕ is increasing in both n_0 and n_1. (2) From (A2), $\phi/(n_0 + n_1)$ is obviously decreasing in n_0 when $n_0 + n_1$ is even. For odd values, $\phi/(n_0 + n_1) = \phi/W$ is nonincreasing in n_0 if and only if $\frac{1}{2}W\eta(n_0, n_1) < \phi(n_0, n_1)$. For given W, ϕ is decreasing and η is increasing in n_0, so it is sufficient to consider the case $n_1 = 0$. It is straightforward to show that $n_0\eta(n_0, 0) > (n_0 + 2)\eta(n_0 + 2, 0)$, which implies that $n_0\eta\,(n_0, 0)$ is decreasing in n_0. Considering $n_0 = 3$ as a base case, $\phi(3, 0) = 1.5$ and $.5 \times 3\eta(3, 0) = .75$. Thus, the condition holds for all greater n_0. (3) The claim is that $(n_1 - 1)\phi(n_0, n_1) \leq n_1\phi(n_0, n_1 - 1)$. Using (A3), this can be rewritten as

$$\phi(n_0,\ W - n_0) \geq (W - n_0) \sum_{W/2-n_1}^{W/2-1} \binom{n_0}{i}.$$

Using equation (A1), this inequality holds if

$$\sum_{W/2}^{n_0}\left(i + \frac{W}{2} - n_0\right)\binom{n_0}{i} + \sum_{n_0-W/2}^{W/2-1}\left(i - \frac{n_0}{2}\right)\binom{n_0}{i} > 0.$$

Using the fact that

$$\binom{n_0}{x} = \binom{n_0}{n_0 - x}$$

and comparing terms for a given combinatorial factor shows that the inequality is satisfied.

Proof of Lemma 2. To demonstrate convexity of ϕ, we need to show that

$$\phi(n_0 - 1,\ W - (n_0 - 1)) + \phi(n_0 + 1,\ W - (n_0 + 1)) - 2\phi(n_0,\ W - n_0) > 0.$$

Consider the case where W is odd (the case where W is even is similar and easier). Using the recursion relationships (A2) and (A3) we can reduce this inequality to

$$.5^{n_0-1}\sum_{i=W/2-n_1}^{W/2-1}\binom{n_0 - 1}{i} + .5^{n_0}\sum_{i=W/2-(n_1-1)}^{W/2-1}\binom{n_0}{i} > \eta(n_0 - 1,\ n_1) - \eta(n_0,\ n_1 - 1).$$

Using the fact that

$$\binom{N}{j} = \binom{N-1}{j} + \binom{N-1}{j-1};\quad \binom{N}{N} = \binom{N-1}{N-1};\quad \binom{N}{0} = \binom{N-1}{0},$$

the inequality can be shown to hold.

Proof of Proposition 1.

$B_H \quad = (N_1 - n_1) + .5(N_0 - n_0) + E[\text{\# of exported broker buys} \mid \theta_H)$

$B_L \quad = .5(N_0 - n_0) + E[\text{\# of exported broker buys} \mid \theta_L)$

$$
\begin{aligned}
B_H + B_L &= (N_1 - n_1) + (N_0 - n_0) + E[\#\text{ of exported broker buys} \mid \theta_H] \\
&\quad + E[\#\text{ of exported broker buys} \mid \theta_L] \\
&= (N_1 - n_1) + (N_0 - n_0) + E[B - S \mid B - S > 0, \theta_H] \\
&\quad Pr(B - S > 0 \mid \theta_H)\, E[B - S \mid B - S > 0, \theta_H] \\
&\quad Pr(B - S > 0 \mid \theta_H) + E[B - S \mid B - S > 0, \theta_L] \\
&\quad Pr(B - S > 0 \mid \theta_L) \\
&= (N_1 - n_1) + (N_0 - n_0) + 2E[B - S \mid B - S > 0] \\
&\quad Pr(B - S > 0) \\
&= (N_1 - n_1) + (N_0 - n_0) + \phi(n_0, n_1) \\
B_H - B_L &= (N_1 - n_1) + E[\#\text{ of exported broker buys} \mid \theta_H] \\
&\quad - E[\#\text{ of exported broker buys} \mid \theta_L]
\end{aligned}
$$

Expanding the expression on the right-hand side and using the fact that

$$
Pr\left(B_0 > \frac{n_0 + n_1}{2}\right) = Pr\left(B_0 < \frac{n_0 - n_1}{2}\right),
$$

we get

$$
\begin{aligned}
B_H - B_L = N_1 &+ n_0\left(Pr\left(B_0 > \frac{n_0 + n_1}{2}\right) - Pr\left(B_0 \geq \frac{n_0 - n_1}{2}\right)\right) \\
&+ 2E\left[B_0 \mid B_0 \geq \frac{n_0 - n_1}{2}\right] Pr\left(B_0 \geq \frac{n_0 - n_1}{2}\right) \\
&- 2E\left[B_0 \mid B_0 > \frac{n_0 + n_1}{2}\right] Pr\left(B_0 > \frac{n_0 + n_1}{2}\right).
\end{aligned}
$$

If $n_1 \geq n_0$, then $B_H - B_L = N_1$, since the $Pr(B_0 \geq (n_0 - n_1)/2) = 1$ and $Pr(B_0 > (n_0 + n_1)/2) = 0$. If $n_1 < n_0$, then when we expand the expression for $B_H - B_L$ we get

$$
B_H - B_L = N_1 - \sum_{i=(n_0-n_1)/2}^{(n_0+n_1)/2} (n_0 - 2i) \binom{n_0}{i} .5^{n_0} = N_1.
$$

The spread given in the text is found by substituting $B_H - B_L$ and $B_H + B_L$ into (3) and using the fact that $\phi(n_0, n_1) = n_1$ when $n_1 \geq n_0$.

Proof of Corollary 1. (1) The market bid-ask spread is increasing in n_0 if

$$
n_0 - \phi(n_0, n_1) > n_0 - 1 - \phi(n_0 - 1, n_1),
$$

which is equivalent to $\phi(n_0, n_1) < 1 + \phi(n_0 - 1, n_1)$. This follows from (A2).

(2) The market bid-ask spread is nondecreasing in n_1 if

$$n_1 - \phi(n_0, n_1) \geq n_1 - 1 - \phi(n_0, n_1 - 1),$$

which is equivalent to $\phi(n_0, n_1) \leq 1 + \phi(n_0, n_1 - 1)$. This follows from (A3).

PROOF OF PROPOSITION 3. Suppose that in equilibrium brokers selected two different customer mixes, n_0^1 and n_0^2. For this to be an equilibrium, it must be that $\lambda_i(n_0^1) = \lambda_i(n_0^2) = \lambda_i$. If profits are to be zero, the λ_i must satisfy

$$\phi(n_0^1) = n_0^1\lambda_0 + (W - n_0^1)\lambda_1,$$

and

$$\phi(n_0^2) = n_0^2\lambda_0 + (W - n_0^2)\lambda_1.$$

Now consider a broker who enters and selects quantity $\hat{n}_0$ such that $n_0^1 < \hat{n}_0 < n_0^2$. Let $\delta = (\hat{n}_0 - n_0^1)/(n_0^2 - n_0^1)$. By convexity,

$$\phi(\hat{n}_0) < \delta\phi(n_0^2) + (1 - \delta)\phi(n_0^1) = \hat{n}_0\lambda_0 + (W - \hat{n}_0)\lambda_1,$$

hence there are positive profits to entry.

PROOF OF LEMMA 3.1. Suppose it is optimal to deviate by j customers but not by one. Then we have

$$(m + j)\lambda_0(n_0) + (w - (m + j))\lambda_1(n_0) > \phi(n_0 + j),$$

and

$$(m + 1)\lambda_0(n_0) + (w - (m + 1))\lambda_1(n_0) < \phi(n_0 + 1).$$

These inequalities together imply

$$\phi(n_0 + j) + (j - 1)\phi(n_0) - j\phi(n_0 + 1) < 0,$$

which violates convexity.

PROOF OF LEMMA 3.2. Suppose that brokers lose money either by switching to mix $m + j$ or $m - i$.

$$(n_0 + j)\lambda_0(n_0) + (W - (n_0 + j))\lambda_1(n_0) < \phi(n_0 + j)$$

$$(n_0 - i)\lambda_0(n_0) + (W - (n_0 - i))\lambda_1(n_0) < \phi(n_0 - i)$$

Using the fact that in equilibrium

$$n_0\lambda_0(n_0) + (W - n_0)\lambda_1(n_0) = \phi(n_0), \tag{A4}$$

we have

$$\phi(n_0) + j[\lambda_0(n_0) - \lambda_1(n_0)] < \phi(n_0 + j)$$

and

$$\phi(n_0) - i[\lambda_0(n_0) - \lambda_1(n_0)] < \phi(n_0 - i),$$

which implies

$$i\phi(n_0 + j) + j\phi(n_0 - i) > (i + j)\phi(n_0).$$

PROOF OF LEMMA 3.3. Straightforward computation verifies that this schedule satisfies (8), so that, given W, it is not optimal to select a different n_0. It is also necessary to verify that, given these prices, the broker will not accept fewer than W customers. Suppose the equilibrium is to accept W customers, n_0 of them uninformed. The broker fails to make money by accepting one less uninformed customer if

$$\lambda_0(n_0, W - n_0)(n_0 - 1) + \lambda_1(n_0, W - n_0)(W - n_0) \leq \phi(n_0 - 1, W - n_0).$$

Using (A2) and (A4), we can rewrite this as

$$\text{(A5)} \qquad 0 \leq \lambda_0(n_0, W - n_0) - \tfrac{1}{2}\eta(n_0 - 1, W - n_0).$$

Similarly, it is unprofitable for the broker to accept one fewer informed customer if

$$\lambda_0(n_0, W - n_0)n_0 + \lambda_1(n_0, W - n_0)(W - n_0 - 1) \leq \phi(n_0, W - n_0 - 1).$$

Again using (A2), (A4), and (9), we can write this as

$$0 \leq \lambda_0(n_0, W - n_0) - \tfrac{1}{2}\eta(n_0 - 1, W - n_0).$$

We conclude that given the schedule (9), it is optimal to accept one less uninformed customer if and only if it is optimal to accept one less informed customer. We wish to verify that it is not optimal to accept one less customer. Using (A2) and (9), (A5) can be rewritten as

$$\text{(A6)} \qquad n_0\phi(n_0 - 1, W - n_0) + (W - n_0)\phi(n_0, W - n_0 - 1) - (W - 1)\,\phi(n_0, W - n_0) \geq 0.$$

For simplicity, we will consider just the case of W even (when W is odd, the above expression can be shown to hold with equality). Using (A2), (A6) can be rewritten

$$(W - n_0)\phi(n_0, W - n_0 - 1) \geq (W - n_0 - 1)\phi(n_0, W - n_0),$$

which follows from lemma 1.

References

Admati, Anat, and Paul Pfleiderer. 1988. A Theory of Intraday Trading Patterns. *Review of Financial Studies* 1:3–40.

Angel, James. 1994. Who Gets Price Improvement on the New York Stock Exchange? Working paper, Georgetown University.

Black, Fischer. 1990. Bluffing. Working paper.

Chowdhry, B., and V. Nanda. 1991. Multi-Market Trading and Market Liquidity. *Review of Financial Studies* 4:483–511.
Easley, David, and Maureen O'Hara. 1987. Price, Trade Size, and Information in Securities Markets. *Journal of Financial Economics* 19:69–90.
Fishman, Michael, and Francis Longstaff. 1992. Dual Trading in Futures Markets. *Journal of Finance* 47:643–71.
Glosten, Lawrence, and Paul Milgrom. 1985. Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents. *Journal of Financial Economics* 14:71–100.
Grossman, Sanford. 1992. The Informational Role of Upstairs and Downstairs Trading. *Journal of Business* 65 (4):509–28.
Harris, L. 1992. Consolidation, Fragmentation, Segmentation and Regulation. Working paper, University of Southern California.
Kyle, Albert S. 1985. Continuous Auctions and Insider Trading. *Econometrica* 53: 1315–35.
Lee, Charles. 1993. Market Integration and Price Execution for NYSE-Listed Securities. *Journal of Finance* 48:1009–38.
Madhavan, Ananth. 1993. Trade Disclosure and the Integration of International Security Markets. Working paper, University of Pennsylvania.
Merton, Robert C. 1992. Operation and Regulation in Financial Intermediation: A Functional Perspective. Working paper, Harvard Business School.
National Association of Securities Dealers. 1991. Inducements for Order Flow: A Report to the Board of Governors, National Association of Securities Dealers, Inc. Washington, DC: National Association of Securities Dealers.
Pagano, Marco. 1989. Endogenous Market Thinness and Stock Price Variability. *Review of Economic Studies* 56:269–88.
Petersen, Mitchell A., and David Fialkowski. 1992. Price Improvement: Stocks on Sale. Working paper, Center for Research in Security Prices, University of Chicago.
Röell, Ailsa. 1990. Dual-Capacity Trading and the Quality of the Market. *Journal of Financial Intermediation* 1:105–24.

Comment Lawrence E. Harris

This paper examines a formal model of brokerage crossing markets in which residual order flows are cleared in a central competitive dealer market. One purpose of the paper, to judge from its title, introduction, and conclusion, is to obtain some results about fragmented markets. Although the paper also addresses other issues, I will confine my remarks to this one.

The authors conclude that fragmentation may be a reflection of increased price competition. This conclusion is obtained in the following sense: brokerage crossing markets that are fragmented and competitive provide better prices for (uninformed) customers than do monopolistic brokerage crossing markets. The result follows from assumed inelasticities of customer demands to trade

Lawrence E. Harris is professor of finance and business economics at the University of Southern California.

and from zero-profit conditions applied to competitive brokers and to competitive dealers. These assumptions ensure that the pie of wealth that can be distributed among market participants is of fixed size. The monopolistic broker market structure provides inferior prices because the brokerage reduces the size of the pie by taking out monopoly profits.

Although I found the result interesting, it does not completely address the problem that I think about when I hear the adjective *fragmented* placed before the noun *markets:* I would like to know whether any broker should be allowed to cross orders internally. This question seems to be at the heart of many of the current controversies about fragmented markets.

This paper does not attempt to answer this question. To do so would require a welfare analysis that would have to consider how the interests of informed traders, uninformed traders, and securities industry intermediaries should be weighed relative to each other. Issues relating to the external benefits of price discovery and liquidity would also need to be considered. In addition, various agency problems and other issues too numerous to mention here would affect the analysis. These issues are all beyond the scope of this paper.

The paper does, however, provide a very important result about price discrimination among diverse traders. The authors formally prove that competitive brokers who can discriminate between informed and uninformed traders will, and must, charge different commissions to the two types of traders. Any crossing broker who tries to charge an intermediate price would get only informed traders. Pursuing this pricing strategy would be unprofitable.

This conclusion is very important because such discrimination can be effected only in a fragmented market. It cannot be provided in an anonymous central market to which all orders are routed.

The fragmentation that we see in the U.S. equity markets reflects this price discrimination. Many dealers pay brokers for order flows from retail traders who are widely believed to be uninformed. Competition among brokers will pass these payments through to the customer in the form of reduced commissions. The uninformed order flow will thereby pay lower transaction costs than they would pay in a completely anonymous market.

Having raised the issue of payment for order flow, I would like to finish by discussing a related public policy issue concerning best execution. Should brokers be required to search for best price for small orders when they can obtain the best wholesale price plus some payment for order flow? If the payments for order flow reflects the appropriate discount from the anonymous market price, the answer should be no.

Now, consider what happens if we require the broker to search anyway. Will the broker do it? Only if the benefits of an improved price can be measured by the customer. If the price improvement is not recognized by the customer, no competitive broker will search beyond the anonymous market price. To do so would incur search costs that would not produce recognizable benefits to the firm.

I believe that most retail clients cannot effectively audit their agent's search for best execution. They do not trade frequently enough, they do not trade with enough different brokers, and they cannot easily collect the information necessary to determine whether the trade prices are consistently good or bad. Retail clients know their commission costs, but they do not know how good their price is relative to what is appropriate for their order type.

The existing system thus helps solve the agency problem. By requiring brokers to search further for best price, we only exacerbate the agency problem because we force brokers to compete to provide immeasurable services. Agency problems are solved by improving measurement.

We might imagine that we could help small uninformed traders by requiring that their orders be consolidated to a single market that provides prices appropriate to them. The results in this paper suggest that market cannot exist without the brokers' active participation. They must discriminate among the orders. Otherwise, informed traders will try to use the uninformed traders' market. A broker must have an incentive to participate. Payment for order flow provides these incentives.

Comment

Geoffrey P. Miller

Kathleen Hagerty and Robert McDonald provide a theoretical exploration of one aspect of the phenomenon of market fragmentation and offer suggestions about the implications for public policy.

As a law professor with a strong interest in the regulation of financial markets and institutions, I found the paper to be stimulating and thoughtful, but ultimately not very informative about the proper direction of public policy. The conditions set up in the authors' model are so general and abstract that they don't allow for much purchase on real-world institutions. Beyond this, even under the constraints of the model, the authors are not able to draw unambiguous policy conclusions. This may not be a particularly telling critique of a piece of pure theory, but it is important to make the observation because those who are constructing social policy need all the help they can get from theory in an area as rapidly evolving and complex as this.

Perhaps I could summarize my basic critique by recommending that the authors revise the title of their paper—instead of "Brokerage, Market Frag-

Geoffrey P. Miller is the Kirkland and Ellis Professor, director of the Program in Law and Economics, and chairman of the Center for the Study of Central Banks at the University of Chicago Law School; a consultant for the Federal Reserve Bank of Chicago; and editor of the *Journal of Legal Studies*.

mentation, and Securities Market Regulation," they should strike all the words after "Market Fragmentation" and insert an "and" after "Brokerage"—the paper, in other words, might be more satisfying if it were simply about "Brokerage and Market Fragmentation," without venturing into the field of legal regulation.

The authors' general view of market fragmentation is that it might not be such a bad thing. Brokers receiving order flows from informed and uninformed customers can distinguish between them, and find that it is a profitable strategy to act as counterparties rather than as mere brokers when dealing with uninformed customers. Thus, trading gets diverted away from the centralized exchange. Because of competition among brokers, the benefits of this strategy are shared with customers. Both brokers and customers benefit from this kind of market fragmentation.

This model has some plausibility, but it leaves important questions unanswered.

The authors provide an account of market fragmentation that does not depend on the off-market transaction free riding on the price discovery function of a centralized market. But their account is not inconsistent with the free-riding theory, suggesting, at most, that there might be forces *other* than free riding that drive off-exchange trading.

The authors acknowledge the potential real-world importance of agency costs in this setting, but they abstract this factor out of their model. Yet the conflict of interest that arises when a broker has the opportunity to act as a counterparty is obvious. The model the authors propose places brokers in the role of protecting themselves against the costs of dealing with informed traders by directing a portion of the informed trades to the central exchange while keeping much of the uninformed trade for themselves. Yet it seems at least equally likely that it will be securities firms that are the informed traders relative to their customers, and that customers face the losses associated with dealing with their brokers as counterparties.

The authors recognize that market fragmentation in their model has potentially deleterious consequences for centralized markets. This includes a reduction in liquidity of the primary market and associated increases in bid-ask spreads. Presumably, under their model, you would also observe an increase in the proportion of informed to uninformed traders in the centralized market as a result of market fragmentation. Thus, uninformed traders dealing on the centralized market face a higher probability of trading with an informed counterparty. It might be worth exploring whether the Hagerty-McDonald model predicts a kind of snowball effect or variant of Gresham's law, as informed traders drive out uninformed traders on the centralized market. In any event, public policy analysis should look at the costs of market fragmentation on the centralized market and compare, if possible, these costs with the benefits realized in collateral markets.

The authors' model does not replicate the richness of observed markets.

They present a picture of trading that occurs either on a centralized market or with brokers, but they do not model the complex array of off-market trading that occurs in practice. Of course, this is a function of the constraints of formal modeling, which is difficult enough even in stylized settings, much less in settings that reflect the muddiness of real-world practice.

Perhaps more troubling, the model appears to have greatest application in settings where the prices and orders are determined in individual negotiations with a broker. This is more descriptive of over-the-counter markets for custom products such as derivatives than for organized equity trading. But such OTC markets have never been characterized by extensive trading on centralized exchanges, so with respect to these products we are not dealing with market fragmentation—these markets are fragmented to begin with. The authors are correct that their model has some application for equity trading, but they could usefully develop this point further.

This paper is premised on the assumption that brokers can distinguish the uninformed and informed traders in the order flow. This assumption might stand further support, given the possibility that informed traders might present themselves as uninformed traders in order to obtain the benefits of the favorable treatment that broker-dealers are capable of giving to uninformed traders. There may be ways that broker-dealers can make this distinction; for example, the broker's knowledge of the customer, or the size of the trade. But the authors might further elaborate this point.

In general, the paper asks an interesting question and makes an important contribution to the literature, but does not offer substantial insights into the identification of optimal social policy or the formulation of desirable legal regulation.

Authors' Reply

Both Lawrence Harris and Geoffrey Miller express the wish that we had engaged in a broader analysis of trade-crossing and market fragmentation. Before addressing their specific comments (many of which we are sympathetic with), we would like to place our paper in perspective. A crude but largely accurate characterization of the microstructure literature is that it assumes that all orders transact at one price in one location, although perhaps with competitive market makers (exceptions to this characterization were noted in our paper). Recent years have seen the rapid growth of off-exchange trading systems, including some that explicitly purchase order flow, cross offsetting orders, and then export to the central market (in effect hedging) any residual order flow.

We see our paper as a first attempt at understanding the economic effects of this kind of trade crossing. As we indicate in the paper, we think the stylized model is actually applicable to a wide variety of real market practices. Our model suggests that, even in a world where all traders trade one share, the bid-ask spread is at best a crude indicator of prices paid and received by investors. The model predicts that heterogeneous treatment of customers should be widespread. Somewhat surprisingly, competitive order-crossing brokers will seek to deal with both informed and uninformed traders.

Harris is concerned by the omission in our paper of a welfare analysis. In particular, how do we trade-off the interests of informed investors, uninformed investors, market intermediaries, and other market participants broadly defined? In some sense, the answer to this question represents the Holy Grail of security market policymaking. We did not attempt a welfare analysis in the paper, and our guess is that this kind of question is, in the end, unanswerable. Knowledgeable price discrimination by brokers will help uninformed traders and hurt those traders who pay the full market spread. We are pessimistic that there will ever be an unambiguous ranking of systems that hurt one group and help another. We think of efficiency questions as being more interesting, which leads to the next point.

Both Harris and Miller point out a closely related limitation in our analysis, one we agree is important. Markets in our model do not discover prices. There is no way for us to analyze the effect of brokers free riding off a central market price. If the siphoning of traders away from the central market reduces the informativeness of prices, financial markets will do a poorer job of providing signals useful in resource allocation. It is an interesting question whether complete reporting of trades makes prices sufficiently informative, even if the trades occur at disparate locations. There are reasons for thinking reporting alone is not sufficient, which is why this question is so important.

In our setting, the information content in orders is ultimately preserved via the export of net orders. It seems reasonable to speculate that trade crossing would reduce central market liquidity, however, since the market maker sees lower trade volume. This might be offset by the availability of greater liquidity from the crossing brokers, who know their customers and their trade motives well and hence will make deeper markets for them.

While we also agree with Miller that we have little to contribute in the way of active suggestions for market regulation, our paper does sound a cautionary note for regulators. The market in which all participants trade in one place at one price is not necessarily the market preferred by all traders, and there is no compelling reason for thinking it best in any sense. We view our contribution as describing a benefit associated with fragmentation which, while not a policy prescription, should have value for thinking about policy.

This Page Intentionally Left Blank