Journal of Financial Economics 32 (1992) 61-86. North-Holland

What's special about the specialist?

Lawrence M. Benveniste, Alan J. Marcus, and William J. Wilhelm
Boston College, Chestnut Hill, MA 02167, USA

Received July 1991, final version received March 1992

Exchange members claim that the professional relationships that evolve on exchange floors yield benefits not easily duplicated by an anonymous exchange mechanism. We show that longstanding relationships between brokers and specialists can mitigate the effects of asymmetric information. Moreover, a specialist who actively attempts to differentiate between informed and uninformed traders can achieve equilibria that Pareto-dominate an equilibrium in which the two types of trades are pooled. Our model also elucidates the role of block trading houses in mitigating information problems in the block market.

1. Introduction

The chairman of the New York Stock Exchange, William Donaldson, recently asserted:

> When you have a human being in the middle of a trade working for you and a crowd of other buyers and sellers you can get the benefits of better bids and offers. (*Business Week*, November 5, 1990, p. 121)

This claim is representative of the widely-held belief among exchange members that the professional relationships that evolve on exchange floors yield benefits not easily duplicated by an anonymous exchange mechanism. Although exchange members certainly have an interest in promoting this view of the market, our analysis demonstrates that these claims may be more than self-interested

Correspondence to: William J. Wilhelm, Jr., Wallace E. Carroll School of Management, Boston College, Chestnut Hill, MA 02167, USA.

*We would like to thank Murray Teitelbaum, James Shapiro, and the Specialist's Association of the New York Stock Exchange for access to the exchange and for providing insight into the institutional details bearing on our analysis. We have also received helpful comments from Fischer Black, Robyn McLaughlin, Erik Sirri, Robert Taggart, and the referee. The third author gratefully acknowledges the support of a Boston College Faculty Fellowship.

attempts to maintain the status quo. We show, in fact, that the specialist system can be viewed as a market mechanism that improves both the welfare of exchange members and the terms of trade for public customers by reducing the incentives to exploit informational asymmetries. The main results of our analysis are:

(1) Equilibrium in a market where informed and uninformed traders are anonymous and therefore pooled can be Pareto-dominated in the sense that both classes of traders could be charged lower bid–ask spreads without reducing specialist profits. Further, there is *always* a level of informational asymmetry that results in such an inefficient equilibrium.

(2) Equilibrium in a market where the specialist actively differentiates between informed and uninformed traders and has the power to sanction traders exploiting private information improves on the terms of trade faced by uninformed traders in the pooling equilibrium.

(3) If the pooling equilibrium is Pareto-inefficient, this active specialist will improve the terms of trade for informed as well as uninformed traders.

(4) Although even this separating equilibrium can be inefficient, the active specialist can, with sufficient sanctioning power, achieve an efficient equilibrium.

(5) The power to sanction those who exploit private information can also keep the market open and achieve an efficient equilibrium under conditions of severe adverse selection in which the pooling market would otherwise close.

The potential for trade motivated by private information has long been recognized as contributing to the spread between dealer bid and ask prices [see Bagehot (1971)]. Glosten and Milgrom (1985), among others, view the spread as a passive mechanism by which the specialist passes the costs of trading with informed traders through to uninformed traders. The ability to exercise discretion over the size [see Easley and O'Hara (1987) and Glosten (1989)] and timing [see Admati and Pfleiderer (1988, 1989)] of orders as well as the potential for credibly announcing the nature of their trades [see Admati and Pfleiderer (1990)] can give uninformed traders some relief from these costs. Unfortunately, even with these mechanisms at their disposal, uninformed traders are likely to face costs associated with having their orders pooled with those of informed traders.

Common to these analyses is the implicit assumption that the specialist trades directly with public customers of the exchange. In fact, the majority of the trading volume on the New York Stock Exchange (NYSE) is funneled through floor brokers acting as agents for public customers. Floor brokers are easily identified and trade repeatedly with the specialist. We contend that this previously ignored layer of intermediation on exchange floors offers an opportunity for the specialist to reduce the costs of asymmetric information.

The second element of our argument is that asymmetric information imposes a burden on the exchange as well as on uninformed traders if the volume of uninformed (liquidity) trading is sensitive to the bid–ask spread. Brokers representing information trades may reap private benefits while sharing the costs of a lower volume of liquidity trading with the remainder of the exchange membership. In our model, the specialist, acting for the exchange membership, enforces sanctions that focus the burden of this externality on the offending broker. The exchange membership willingly subjects itself to this discipline because it recognizes that the long-term benefits of doing so outweigh the short-term private benefits of exploiting private information. Although each broker would find it privately beneficial to conceal the informational motivation for a trade, each is willing to be bound to a system that punishes such behavior. The specialist in our model acts as an enforcer of the agreement that protects the common good.

Although it may be relatively difficult for specialists to identify in advance floor brokers exploiting information on behalf of their principals, the stability and relatively small size of the trading community limit brokers' ability to systematically exploit private information by increasing the probability that those doing so will be identified after the fact. As Benveniste and Spindt (1990) emphasize in the context of the primary market for corporate equity, a continuing business relationship, such as that maintained by brokers and specialists, allows for the imposition of sanctions on those found to have exploited private information.[1] Thus the specialist is able to actively promote cooperation, mainly in the form of information sharing, among members of the trading community by sanctioning those who choose not to cooperate.

Since the key to the broker/specialist relationship is the specialist's ability to identify and trade repeatedly with the broker, we formalize the preceding argument by comparing two market mechanisms that differ only in the degree to which they protect the participants' anonymity. In the first, traders are assumed to retain absolute anonymity. In this environment (which characterizes virtually the entire market microstructure literature), privately informed traders have every incentive to maximize the gains from private information. In the second, we assume that all traders are represented by brokers and that brokers who exploit private information may be identified as having done so. Therefore, they must weight the risk of *ex post* sanctions by the specialist against any private gains associated with their actions.

[1]Although the market maker's role in promoting cooperation in the exchange of information is informal, it is widely recognized within the trading community. Cox and Rubinstein (1985, pp. 80–81) note: 'From perhaps bitter past experience, market makers learn to identify likely information traders and protect themselves by giving more conservative quotes in response.' We describe several additional sanctions and incentives that the market maker can use to induce revelation of private information.

An anonymous trading environment necessarily implies pooling of liquidity traders and information traders. As does earlier work, we find that in a pooling equilibrium, the specialist's inability to distinguish between members of the two classes of traders leads to a positive bid–ask spread (reflecting a transfer from uninformed to informed traders) that increases with the expected value of such information.

Identification and sanctioning of information traders provides the leverage necessary for the specialist to improve on the terms of trade that would result from pooling informed and uninformed traders. If the specialist can impose sanctions on a broker discovered to have exploited private information, he or she can always charge brokers representing uninformed traders a smaller bid–ask spread than they would face when pooled with informed traders. The threat of sanctions induces informed brokers to reveal their private information and allows the specialist to reduce losses by imposing an explicit or implicit charge on the informed traders. Since the specialist passes losses on to uninformed liquidity traders through the bid–ask spread, lower losses permit a reduction in the spread faced by uninformed traders.

Less obvious and more important is that under certain circumstances the specialist's leverage over informed brokers can actually lead to improvement in their terms of trade as well. The intuition behind this result rests on an assumption similar to that used by Admati and Pfleiderer (1988), that uninformed trading demand is discretionary. If the reduction in the bid–ask spread charged to uninformed brokers swells the volume of liquidity trading (and thus specialist revenues) sufficiently, the specialist will be able to charge informed brokers a lower spread than would be possible in the pooling equilibrium. Thus the specialist's ability to improve the terms of trade for informed traders is contingent on the elasticity of liquidity volume with respect to transaction costs as well as on the value of private information.

Our analysis is in the spirit of recent work by Glosten (1989) and Gammill (1990) that investigates the role of institutional design in reducing the burden of asymmetric information. Unlike Glosten's results, ours apply to both specialist and competitive market-making mechanisms and suggest an alternative explanation for claims that the specialist system is a superior form of market organization. Gammill demonstrates that the ability to knowingly accept losses on individual trades allows the specialist to heighten competition among informed traders by compensating (that is, offering a more attractive price to), or facing losses from, only the first to reveal private information. We extend Gammill's analysis by showing that in addition to a 'carrot' in the form of compensation for private information, the specialist's ability to sanction brokers provides a 'stick' that can be used to further mitigate the specialist's adverse selection problem. In addition, the dependence of our results on repeated trade among a small contingent of identifiable trading partners provides a more satisfying explanation for the existence of floor exchange mechanisms.

The paper is organized in four sections. In section 2 we provide the institutional background necessary to establish two points: (1) that specialists trade repeatedly with a small and stable pool of brokers and (2) that they have the power to enforce cooperation in the form of information sharing on the exchange floor. The model developed in sections 3 and 4 establishes formally that these two features of exchange floors can weaken the adverse effects of asymmetric information. We conclude with a discussion of the implications of the model for policy issues related to the design and regulation of securities exchanges.

2. Institutional background

The trading floors of the major securities exchanges are widely considered to be among the most competitive of markets. A visitor would, therefore, probably find the collegial atmosphere among market participants surprising. Certainly an important element of this atmosphere is simply the relatively small size of the trading communities. As of May 31, 1989, the NYSE membership, for example, included 1,366 members with voting rights and distributive rights to the exchange's assets and an additional 63 members who maintained either physical or electronic access to the trading floor through payment of an annual membership fee [Shapiro (1989)]. The actual trading community is dominated by specialists (432) and floor brokers (846).[2] Floor brokers act primarily as agents for public customers, delivering orders for a stock to the trading crowd, whereas the specialist acts as the 'broker's broker', providing both dealership and brokerage sevices. In return for these services the exchange grants the specialist an exclusive right to make a market for the stock.

Approximately three-fourths of all orders are submitted through the exchange's computerized order-routing system (SuperDot), and thus bypass the intermediary services of the floor broker.[3] These orders, however, account for only 10%–20% of trading volume. Thus it is primarily large, and more likely information-driven, trades for which the broker's order-execution skills are in demand.

The longstanding professional relationships that evolve on the trading floor are a natural consequence of the repeated transactions between exchange members. These relationships and associated reputation effects can induce

[2] In addition to specialists and floor brokers, the trading community included seventeen registered competitive market makers and eight registered floor traders. Shapiro (1989) notes that not all members are active in the trading process and that the size and composition of the trading community are somewhat variable. Although statistics are not maintained, our discussions with exchange officials suggest that the trading community is quite stable.

[3] As will become apparent later, it is noteworthy that the specialist can identify both the firm submitting orders through SuperDot and the nature of such orders (program trade, principal vs. customer orders, etc.).

cooperation even among self-interested parties. For example, it is not uncommon for members to accommodate one another by participating in transactions that are not in their immediate best interests. Quite frequently, for example, a broker will arrive with an order and find that the price quoted is the result of a limit order. Finding this price satisfactory, the broker may seek execution of the order only to find that his or her entire order cannot be filled against the limit order. Rather than forcing the broker to accept a less attractive price on the remainder of the order, the specialist will offer to complete it at the quoted price. To the extent that the specialist's own quotes represent his or her willingness to add to or reduce inventory, such behavior represents a concession to the broker for which the specialist receives no immediate compensation.

Perhaps more surprising is that the cooperation works both ways: brokers commonly share information about forthcoming order flows with specialists. Through a client's order placement strategy, a broker may gain access to private information about the intrinsic value of the traded stock.[4] Less obvious but perhaps more important is Grossman's (1990) observation that brokers serve as repositories for information about 'unexpressed' demand: with existing technology the entire range of an investor's state-contingent demands cannot be easily or costlessly conveyed to the market. Through explicit declaration or a long-standing business relationship, however, a broker may have knowledge of the conditions under which a client will be in the market and the nature of such contingent demands.

Whatever its source or nature, private information presents a strategic opportunity for the broker to obtain better execution for a principal's order at the expense of the specialist and other traders. Establishing a reputation for quality execution through strategic use of private information represents a private benefit to the broker and thus may produce a strong incentive to exploit private information. The information sharing commonly observed is surprising, therefore, and certainly at odds with the behavior posited in the market microstructure literature.

But such behavior follows logically from an institutional structure designed to alleviate problems associated with private information. Individual NYSE members must recognize that trading on private information produces an externality in the form of wider equilibrium bid–ask spreads, and thus reduced trading volume. We contend that the design of the NYSE's market mechanism may be viewed as a mitigating institutional adaptation. The specialist serves the common good by enforcing an informal agreement among brokers to share information in order to reduce information-induced bid–ask spreads that can impede

[4]The threat of trading on private information is sufficient in the options markets to cause the Chicago Board Options Exchange (CBOE) to compel brokers, at the request of market makers, to reveal the identity of the client for which a trade is being executed [see Cox and Rubinstein (1985, pp. 80–81)].

liquidity trading volume. Although it may be relatively difficult for specialists to identify in advance brokers who are exploiting information, the size and stability of the typical trading crowd on an exchange floor limit brokers' ability to exploit private information systematically by increasing the probability that those doing so will be identified after the fact. Moreover, the continuing business relationships between brokers and specialists make it possible for specialists to sanction brokers who exploit private information.

Specialists have a range of subtle but effective means of sanctioning brokers. In addition to choosing or declining to fill the remainder of an order executed against the limit order book, as described earlier, specialists have other ways of exercising discretion over the prices at which they trade with individual brokers. For example, the quotation cited in footnote 1 suggests specialists may simply provide a less attractive price schedule to brokers with reputations for trading on private information. Hasbrouck's (1988) finding that approximately 15% of the NYSE transactions in his sample were orders executed at the midpoint of the bid–ask spread suggests another dimension of pricing discretion. Although this may be a result of the simultaneous arrival of buy and sell orders, Hasbrouck argues that more often such trades arise as a result of the broker's negotiating a better price than the current quotes. In fact, rule 123.41 of the NYSE constitution and rules requires that the broker exercise due diligence to execute an order at the best price possible. Negotiation for a price within the quoted spread, however, is at the discretion of the specialist.

The specialist also provides brokers with services that if withheld would represent a substantial opportunity cost to the broker. For example, specialists occasionally 'work' large orders for brokers [see Sirri (1989)]. In this role the specialist acts much the same as an upstairs block positioner [see Burdett and O'Hara (1987)] disclosing the quantity of the order and the identity of the buyer or seller to the entire trading crowd in an effort to obtain an offsetting order. Finally, by virtue of his or her position at the center of trade, the specialist becomes a clearinghouse for market information that if provided to the broker may lead to better execution of a client's order. For example, if the specialist is aware that a broker has an order pending, he may be willing to page the broker upon finding that an offsetting order is imminent.

In summary, the floor broker's role as intermediary between public customers and the specialist allows specialists to identify and trade repeatedly with a small set of partners. Specialists thus have considerable power to sanction brokers identified as having exploited private information.[5] In the following section we demonstrate how these features of a floor exchange mechanism permit improvement in the terms of trade in the presence of asymmetric information.

[5]Although it is impossible to measure directly the extent to which specialists retaliate against 'cheaters', our discussions with several specialists reveal that detected cheaters are in fact subsequently given less consideration by the specialist.

3. The model

The market comprises a risk-neutral specialist and brokers who represent market orders driven by either liquidity needs or private information.[6] The specialist may or may not be competitive but is bound to minimal trading profits by either competition or exchange monitoring.

Trading occurs at discrete times and any new information bearing on the security's true value is publicly announced following each round of trade. Transaction prices refer to bid or ask prices quoted by the specialist for the current round of trade.

We assume initially that all traders have the same information, and thus assign the asset an initial intrinsic value of p^*. The single source of uncertainty in the market is the possibility of a random shock to p^* before a round of trade that becomes public information only after the trading round is completed. The random shock, which occurs with probability π, will lead to a revision of the asset's intrinsic value to either $p^* + \alpha$ or $p^* - \alpha$, with each revision having equal probability. Thus, the asset's intrinsic value is assumed to follow a random walk with an end-of-round distribution as follows:

Intrinsic value	Probability
$p^* + \alpha$	$\pi/2$
p^*	$1 - \pi$
$p^* - \alpha$	$\pi/2$

Traders place orders simultaneously in each trading period, meaning that all trades are batched. Since all orders are assumed to be executed through the specialist, order imbalances lead to additions to or reductions in the specialist's inventory of the asset. We assume that all agents are risk-neutral and that the interest rate is zero, so there is no cost to carrying inventory.

The volume of liquidity trades in each trading round is assumed to depend on the absolute difference between transaction (bid and ask) prices and the security's intrinsic value. Because liquidity traders' trading is motivated by cash needs, we assume they are solving an unmodeled problem of consumption and savings in which transaction cost is compared with the benefit of current consumption. This results in liquidity orders that are elastic in the cost of transacting.

[6]The assumption that all trades are the result of market orders is motivated by the fact that once an order is placed in the limit order book, the specialist is unable to exercise any discretion in its execution. Limit orders receive first priority for execution against incoming orders and are executed at the designated limit price. Thus our model does not give the specialist a role in dealing with information asymmetries associated with such orders.

Formally, aggregate shares demanded or supplied by liquidity traders depend on the ask price $p_d = p^* + s_d$ or bid price $p_s = p^* - s_s$. The assumed schedules are

$$q_d(s_d) = q^* - \delta(p_d - p^*) = q^* - \delta s_d\,,$$

$$q_s(s_s) = q^* - \delta(p^* - p_s) = q^* - \delta s_s\,.$$

We could add a zero-mean stochastic term to each of these schedules, but because of the risk-neutrality assumption, such an addition would not affect our results. The sensitivity of volume to the spread, δ, is the same in each schedule. This specification is made entirely for notational convenience, as it assures that the market for liquidity trades will clear for identical values $s_d = s_s = s$. Therefore, the bid–ask spread is $2s$. Given this symmetry, however, our focus will be on deriving equilibrium values for s. With this in mind we henceforth refer to s as the spread.

We assume the existence of a stable proportion of informed traders who learn of a shock to p^* when it occurs. Thus, with probability π informed traders will participate in a round of trade. We also assume that informed traders are constrained by an aggregate position limit (either long or short) equal to q_i shares. This assumption is necessary to rule out infinite demand for securities, which otherwise would result from our assumption of risk neutrality. Long positions are taken when good information is forthcoming and informed traders anticipate selling prices one period forward that exceed current buying prices. Short positions are taken when bad information is forthcoming. Informed traders unwind their positions as soon as their private information is made public and is impounded in security prices.

4. Equilibrium and the specialist

The specialist provides the mechanism for trading by establishing transaction prices for each round and taking the opposite side of all trades. We assume (1) that the specialist is not informed beyond publicly available information, (2) that the specialist is constrained to earn zero profits, and (3) that all trades are made with the specialist. Thus, equilibrium is characterized by the specialist's setting prices using only publicly available information and earning zero expected profit. Without the ability to distinguish information traders from liquidity traders, the specialist solves a purely statistical problem requiring determination of the spread that sets expected profits to zero. A specialist solving this problem will be referred to as passive. As we have noted, however, repeated trading between a specialist and floor brokers gives the specialist leverage he or she can use to elicit information about a trader's motives. This power allows the specialist to set different terms of trade for different individuals.

A specialist following this strategy will be referred to as active. The following two subsections compare market outcomes under the two types of specialists.

We use the following notation. Anticipating the possibility of distinct terms of trade offered to the two classes of traders, we let s_l denote the spread offered to liquidity traders and s_i the spread to informed traders. Thus, liquidity traders face bid and ask prices of $p_{lb} = p^* - s_l$ and $p_{la} = p^* + s_l$. Similarly traders identified as informed face bid and ask prices $p_{ib} = p^* - s_i$ and $p_{ia} = p^* + s_i$. The posting of two spreads is not intended to be taken literally. Rather, we envision a more informal arrangement in which the specialist exercises discretion in establishing alternative terms for traders who reveal private knowledge of the security's value. For example, the specialist's posted price is good only for a stipulated quantity. Negotiations over prices for larger quantities, or the willingness to grant a price concession (see section 2), will depend on the specialist's perception of the nature of the trade.

The volume of liquidity trades that cross as a function of the spread, s_l, is

$$V(s_l) = q^* - \delta s_l \,.$$

Noting that each liquidity trade crossed results in a net revenue of $2s_l$ for the specialist, the specialist's total revenue from liquidity trades is

$$R(s_l) = 2s_l(q^* - \delta s_l)\,. \tag{1}$$

The specialist anticipates informed orders for q_i shares when good information is forthcoming and informed sales of q_i shares when bad information is forthcoming. The cost per share to the specialist is the mispricing by α net of the spread charged on the trade, s_i. Total costs following a price shock, therefore, are $q_i(\alpha - s_i)$. Weighting the cost by the probability that information is forthcoming, the expected cost borne by the specialist in each period as a function of the informed traders' spread is

$$C(s_i) = \pi q_i(\alpha - s_i)\,. \tag{2}$$

We turn next to characterizing the equilibria.

4.1. *Equilibrium with a passive specialist*

The objective of this section is to establish a benchmark equilibrium that is consistent in its predictions with models that currently characterize the literature [for example, Glosten and Milgrom (1985)]. The benchmark equilibrium is characterized by a passive specialist, who by definition does not distinguish between the two classes of traders. The implication of such behavior is that the specialist sets an identical spread $s = s_i = s_l$, and thus identical bid and ask prices, for the two classes of investors.

Equilibrium with a passive specialist is defined by a pooling spread, s_p, for which the specialist's expected profits from trading are zero. Assembling the components of profit given above, we can characterize the candidate spreads by the equation

$$R(s_p) = C(s_p) = 0\,. \tag{3}$$

Eq. (3) can be expanded by substituting for $R(s_p)$ and $C(s_p)$ to obtain the quadratic equation

$$2s_p(q^* - \delta s_p) - \pi q_i(\alpha - s_p) = 0\,. \tag{4}$$

Given our cost and revenue functions, there are usually two zero-profit spreads.[7] Competition dictates that the smaller of the two will emerge in equilibrium. By taking the smaller of the solutions to (4) we can identify the pooling equilibrium spread as

$$s_p = \frac{q^*}{2\delta} + \frac{\pi q_i - [(\pi q_i + 2q^*)^2 - 8\alpha\pi q_i\delta]^{1/2}}{4\delta}\,. \tag{5}$$

The pooling equilibrium is presented graphically in fig. 1.

Inspection of (5) reveals that the pooling spread is a product of the threat to the specialist of trading with informed traders. If α, q_i, or π is zero, indicating no threat from information trading, the last term in square brackets is zero, and so is the bid–ask spread. If, on the other hand, the value of private information is very high and the resulting adverse-selection problem is severe, the specialist may not be able to charge a large enough spread to satisfy the zero-profit condition and keep the market open.[8] For values of α that allow market equilibrium, one can show that s_p is increasing in α and decreasing in q^*. The positive relationship between s_p and α reflects the specialist's need to set a wider spread when faced with larger expected losses to informed traders. Increases in the volume of liquidity trading (reflected in q^*), all else being equal, increase profits for any spread and allow the specialist to quote a narrower spread in equilibrium.

To complete our characterization of the equilibrium spread, we demonstrate that the equilibrium spread may be efficient or inefficient, where inefficiency is defined as follows.

[7]It may be impossible to satisfy the zero-profit condition when the adverse-selection problem is severe. Glosten (1989) points out that while a competitive specialist will close the market under such circumstances, the ability of a monopolist specialist to average profits across trades and time can improve the terms of trade by keeping the market open.

[8]Specifically, if $\alpha > (\pi q_i + 2q^*)^2/8q_i\pi\delta$, the market will close.

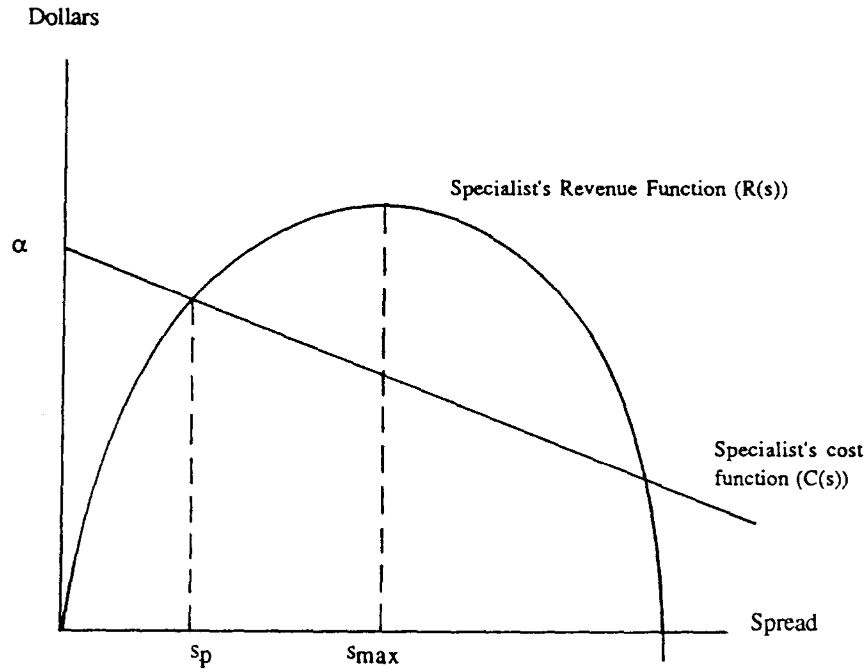

Fig. 1. Graphic representation of equilibrium with a passive specialist. $C(s)$ and $R(s)$ represent the specialist's cost and revenue functions. s_{max} is the revenue-maximizing spread. Intersections of $C(s)$ and $R(s)$ represent feasible solutions to the specialist's problem, but competition requires that s_p be the pooling equilibrium spread.

Definition. A pooling spread s_p that satisfies (4) is inefficient if and only if there are spreads s_l and s_i such that

(i) $s_l < s_p$,

(ii) $s_i < s_p$,

(iii) $C(s_i) \leqslant R(s_l)$.

The significance of an inefficient spread is that there are alternative spread *pairs* (s_l,s_i) that simultaneously improve the positions of both classes of traders, that is, they provide more profit for informed traders and a lower spread for liquidity traders, while not imposing positive expected cost on the specialist. The following lemma characterizes inefficient spreads.

Lemma 1. A spread s_p that yields zero expected profit to the specialist is inefficient if and only if $R'(s_p) < 0$. Furthermore, $R'(s_p) < 0$ if and only if $s_p > q^*/2\delta$, where $q^*/2\delta$ is the spread that maximizes revenue from liquidity traders, denoted s_{max}.

Proof. Candidate spreads must satisfy the zero-profit condition

$$2s_l(q^* - \delta s_l) - \pi q_i(\alpha - s_i) = 0\,.$$

Equilibrium occurs at a point on the zero-profit locus defined by this equation. By definition, the equilibrium is inefficient if it occurs at a point for which $\partial s_i/\partial s_l > 0$. Direct calculation using the implicit function rule shows that this occurs if and only if $s_l > q^*/2\delta$. Since $R'(s_p) = 0$ at $s_p = q^*/2\delta$ and $R''(s_p) < 0$, $q^*/2\delta$ is the revenue-maximizing spread and spreads larger than $s_{\max}$ are inefficient. In this range, $R'(s_p) < 0$.

Thus the pooling spread set by the passive specialist is inefficient when both intersections of $R(s_p)$ and $C(s_p)$ in fig. 1 occur to the right of $s_{\max}$. The conditions under which the passive specialist will set an efficient spread are given in Theorem 1:

Theorem 1. The pooling equilibrium spread is inefficient if and only if

$$\alpha > s_{\max}[1 + (q^*/\pi q_i)]\,. \tag{6}$$

Proof. From Lemma 1, any spread larger than $s_{\max}$ is inefficient. Since $s_{\max} = q^*/2\delta$, the condition for s_p to exceed $s_{\max}$ is that the second term on the right-hand side of (5) is positive, which is satisfied when $(\pi q_i + 2q^*)^2 - 8\alpha\delta\pi q_i < (\pi q_i)^2$. This condition can be rearranged to yield (6).

Theorem 1 demonstrates that the pooling equilibrium is more likely to be inefficient when the value of private information is high in relation to the revenue generated by liquidity trading (as measured by q^*). To understand this result, note that when expected losses to informed traders are small, the specialist easily offsets them through the spread charged to liquidity traders. As such costs rise, however, the potential for earning offsetting revenue is capped at the liquidity-revenue-maximizing spread $s_{\max}$. If the specialist realizes net losses even at this spread, the only way to achieve zero profits is to increase the spread further, reducing revenue from liquidity traders, but reducing losses to informed traders by even more, at least within a neighborhood of $s_{\max}$. This results in an inefficient equilibrium with $s_p > s_{\max}$. When expected losses to informed traders are extremely large, the market can break down altogether, because the spread that drives liquidity trading to zero is not large enough to fully offset the specialist's expected losses. In this case, $C(s)$ lies strictly above $R(s)$ in fig. 1, implying the absence of a zero-profit pooling equilibrium.

Although the preceding discussion suggests that the market equilibrium is most likely to be inefficient under conditions in which an equilibrium may not exist, that is, when α is large, Theorem 2 establishes that there are values of α for which both the market will remain open and the resulting equilibrium will be inefficient.

Theorem 2. If there is a nonzero probability of trade driven by private information, there always exist values of α for which a pooling equilibrium exists and is inefficient.

Proof. The proof requires only that the upper bound on α for the market to remain open (see footnote 7) exceed the right-hand side of (6), or that $(\pi q_i + 2q^*)^2 > 4q^*(q^* + \pi q_i)$. This is satisfied as long as $\pi q_i > 0$.

In summary, the pooling equilibrium achieved by the passive specialist shows characteristics similar to those of other asymmetric information models in the literature. The unique feature of our analysis is the ability to make statements about the efficiency of the equilibrium and the conclusion that there are always levels of information asymmetry for which the market will remain open and the resulting equilibrium will be inefficient. As we demonstrate in the following section, a floor exchange mechanism in which the specialist has powers beyond those granted the passive specialist can generally improve on inefficient equilibria.

4.2. Equilibrium with an active specialist

An active specialist seeks to separate informed from liquidity-motivated trades. Traders are now explicitly represented by brokers, who are many fewer and therefore come to the market more often. This extra layer of intermediation permits the specialist to discriminate between informed and liquidity trades to the extent that brokers are themselves able to do so.

The active specialist may be distinguished from the passive specialist in three dimensions. First, the active specialist may observe after the fact whatever information was available to the broker at the time a trade was executed. For example, a broker might observe an increase in demand volume before the specialist and conclude that a trade is motivated by information. The specialist may eventually learn that the broker knew this at the time of the trade.

Second, the active specialist may sanction brokers identified as having represented an informed trade as liquidity-motivated. For example, the specialist can withhold favorable terms for some future period if the broker is caught misrepresenting.

Third, we will show that the active specialist can increase the value of a typical broker's franchise by offering terms of trade that depend on the broker's representation of trading motive, be it liquidity- or information-based. In essence, the active specialist can charge different spreads and exploit the elasticities in the liquidity trading schedules to elicit optimal trading volume in the presence of information asymmetries.

Trading takes place in two stages. In the first stage, informed and liquidity traders submit market orders to brokers based on the terms of trade they expect to face, and brokers observe a (noisy) signal of the motivation for these orders. In the second stage, brokers represent orders to specialists as either information- or liquidity-motivated and are assigned a spread of s_i or s_l corresponding to their characterization of the order.

As noted, brokers do not receive perfect signals about their clients' trading motives. To capture the uncertainty, we call λ the probability that the broker's inference about the motivation for the trade is correct. For simplicity, we assume that the probability of a correct inference is the same whether the trade is in fact liquidity- or information-motivated. Given this symmetry, $\lambda = 0.5$ implies that signals are wholly uninformative [Merton (1981)]. Therefore, we assume that $0.5 \leqslant \lambda \leqslant 1$. We further assume that with probability γ the specialist will observe the noisy signal the broker receives at the time of the trade. What matters is that the specialist potentially may infer when a broker thought he was executing an informed trade.[9] It is *not* relevant to our analysis whether the broker was correct; it matters only that there is some probability that the specialist can eventually infer whether the broker intended to be truthful.

Brokers are assumed to derive profit from two sources. The first is the commission charged to execute a trade, and thus this component of profits is proportional to trading volume. Given the elasticities of supply and demand for liquidity trades, broker profits are a decreasing function of the spread liquidity traders expect to face. Further, since the incentive to gather information depends on the potential profits from such activity, the volume of information trades also should be (at least in the long run) a decreasing function of the spread informed traders expect to face. We assume that brokers also capture some fraction of the profits resulting from a well-executed trade. In our model, good execution means that a broker obtains a lower spread for an informed client by disguising the motivation for the trade.

The problem for the specialist, acting as an agent of the exchange, is to quote the spread pair (s_l, s_i) that maximizes the broker's expected profits subject to the zero-profit constraint and to incentive-compatibility constraints that induce brokers to reveal their signals truthfully. We could define an explicit function for broker profits and solve for the optimal (s_l, s_i) pair. Instead, we take the more general approach of identifying the range of pairs that is consistent with potential equilibria. This approach allows us to establish the boundaries within which it is possible for the active specialist to achieve an equilibrium that dominates that achieved by the passive specialist.

[9] In reality, the specialist must decide which brokers have cheated using statistical inference from a history of trades. Inference is necessarily imperfect, and some brokers might escape detection. We capture this aspect of the relationship in our one-period model by assuming the probability that the broker is detected is positive, but less than 1.0.

The specialist's revenue depends on trading volume. Recognizing the potential for misidentification by their brokers, risk-neutral traders base their trading decisions on the probability-weighted average of the spreads (s_l, s_i) quoted by the specialist. The expected spreads faced by liquidity and informed traders are therefore

$$s_l^e = \lambda s_l + (1 - \lambda)s_i \,, \tag{7a}$$

$$s_i^e = \lambda s_i + (1 - \lambda)s_l \,, \tag{7b}$$

We define the function by which expected spreads (s_l^e, s_i^e) map back into actual spreads (s_l, s_i) for a given value of λ to be[10]

$$(s_l, s_i) = S(s_l^e, s_i^e; \lambda) \,. \tag{8}$$

Clearly, with uninformative signals ($\lambda = 0.5$), both types of traders would receive identical spreads, and the pooling equilibrium would result.

Since both traders and the specialist are assumed to be risk-neutral, the zero-profit condition requires the specialist to set (s_l, s_i) to satisfy

$$2s_l^e(q^* - \delta s_l^e) - \pi q_i(\alpha - s_i^e) = 0 \,. \tag{9}$$

The first term on the left side of (9) uses the expected spread s_l^e because liquidity trading volume is determined by the average spread paid for liquidity trades. Similarly, the second term uses s_i^e because that is the average spread specialists receive from informed traders.

The spread pair quoted by the specialist, (s_l, s_i), must also satisfy individual rationality conditions for both traders and brokers. Individual rationality for an informed trader requires that

$$(\mathrm{IR}_i): \quad s_i^e \leqslant \alpha \,.$$

(IR_i) states that informed buyers or sellers will not trade if the expected transaction cost exceeds the value of their information. Individual rationality for liquidity traders is expressed implicitly in the liquidity supply-and-demand schedules, $V(s_l^e)$.

[10]Solving (7a) and (7b) for s_i and s_l for $\lambda > 0.5$ yields

$$s_i = \frac{\lambda s_i^e - (1 - \lambda)s_l^e}{\lambda^2 - (1 - \lambda)^2} \,, \qquad s_l = \frac{\lambda s_l^e - (1 - \lambda)s_i^e}{\lambda^2 - (1 - \lambda)^2} \,.$$

Using l'Hopital's rule for $\lambda = 0.5$, we obtain $s_i = s_l = (s_i^e + s_l^e)/2$.

In contrast to traders, brokers make decisions based on actual spreads. Traders must account for the possibility that the broker and therefore the specialist will misidentify their motives. Brokers, in contrast, know the spreads their clients will receive once they infer and report trading motives. Moreover, their compensation for quality execution is a function of the spreads actually obtained for informed clients.

For it to be incentive-compatible for a broker to reveal truthfully an inference that a trade is liquidity-motivated, the liquidity spread cannot exceed the informed spread, that is,

$$(\mathrm{IC}_l): \quad s_i \geqslant s_l \, .$$

If this inequality were reversed, brokers would have an incentive to increase volume by representing liquidity trades as informed.

Similarly, if a broker who infers that an order is information-motivated is to reveal that inference, the benefit derived from misrepresenting the order as liquidity-motivated must be outweighed by the expected cost of doing so. If the broker represents a trade as liquidity-motivated, the specialist will execute it at the liquidity spread s_l. Likewise, if the broker represents a trade as information-motivated, it will be executed at the informed spread s_i. When given an order identified as motivated by private information, the specialist requires the premium $(s_i - s_l)$ in addition to the cost s_l that would have been borne had the trade been represented as liquidity-motivated. Thus the benefit from misrepresenting a trade perceived to be information-motivated as a liquidity trade is $(s_i - s_l)$.

To counter the gains from misrepresentation, the specialist threatens to punish brokers identified as having misrepresented their signals, which the specialist is able to do with probability γ. We assume that brokers so identified face two costs. First, over some interval the specialist will treat all of the broker's orders as information-motivated. If the broker is to retain the business of liquidity traders during this interval, competition will demand that he bear the cost $(s_i - s_l)$ on trades that he infers are liquidity-driven. Assuming that N such orders arrive, the cost to the broker is $N(s_i - s_l)$. Consistent with the discussion in section 2, we also assume that the specialist will be less willing to share information with or make concessions to the broker in the future, which imposes a further cost C.

Incentive compatibility for a broker who believes a trade is information-driven therefore requires that the benefit from misrepresenting the signal not exceed the expected cost of detection, or

$$(\mathrm{IC}_i): \quad s_i - s_l \leqslant \gamma[N(s_i - s_l) + C] \, .$$

Fig. 2 illustrates the specialist's problem. The vertical axis represents either the quoted informed spread, s_i, or the expected informed spread, s_i^e. Similarly,

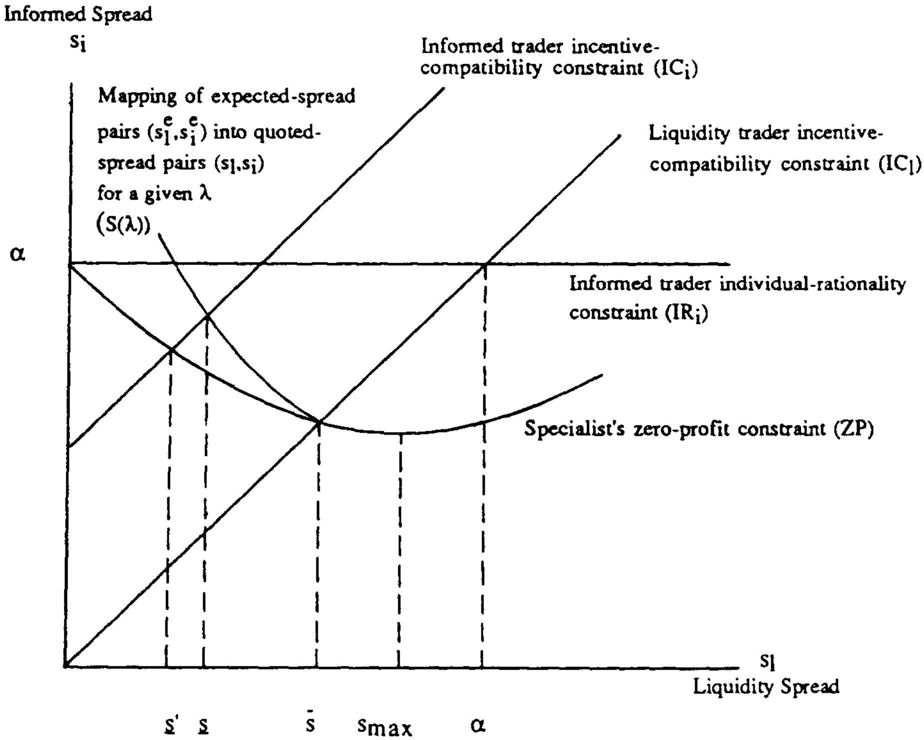

Fig. 2. Graphic representation of equilibrium with an active specialist. The specialist's zero-profit constraint (ZP) and the individual rationality constraint faced by informed traders (IR_i) determine the feasible expected-spread pairs (s_l^e, s_i^e). Individual rationality for informed traders requires that (s_l^e, s_i^e) lie below (IR_i). (ZP) represents the locus of expected-spread pairs (s_l^e, s_i^e) satisfying the specialist's zero-profit constraint. $S(\lambda)$ represents the mapping of expected-spread pairs (s_l^e, s_i^e) into quoted-spread pairs (s_l, s_i) for a given λ. Incentive compatibility for liquidity traders requires that the equilibrium quoted-spread pair (s_l, s_i) lie to the left of (IC_l). Incentive compatibility for informed traders requires that (s_l, s_i) lie below (IC_i). Admissable solutions to the specialist's maximization problem for a given λ lie on $S(\lambda)$, and thus are bounded by $\underline{s}$, which represents the quoted liquidity-trade spread for which (IC_i) is binding, and $\bar{s}$, which is equal to the pooling equilibrium spread s_p. For $\lambda = 1$, the lower bound on the quoted liquidity-trader spread is $\underline{s}'$.

the horizontal axis represents the terms of trade, quoted and expected, for liquidity traders. Solving the zero-profit condition (9) for s_i^e yields

$$\text{(ZP):}\quad s_i^e = \alpha - [2s_l^e(q^* - \delta s_l^e)/\pi q_i]\,. \tag{10}$$

Whereas the broker's incentive-compatibility conditions are determined by actual spreads, the specialist's zero-profit condition is determined by expected spreads. Again, this is because brokers who report signals know whether they will receive a spread of s_i or s_l. In contrast, traders may be misclassified, and therefore base their trading decisions on expected spreads. Trading volume and the specialist's revenue depend on s_i^e and s_l^e.

Because the specialist's expected profits are quadratic in s_l^e and linear in s_i^e, the locus of (s_l^e, s_i^e) pairs satisfying the zero-profit constraint (ZP) is a parabola bounded by (IR_i). The minimum point along the zero-profit locus occurs at $s_{\max}$, corresponding to the value of s_l^e that maximizes expected revenue from liquidity trades. As before, $s_{\max} = q^*/2\delta$. Combinations of s_l^e and s_i^e to the right of $s_{\max}$ are strictly inefficient. In this region, increasing the expected spread faced by liquidity traders reduces the specialist's revenue because the elasticity of liquidity trading demand with respect to the expected spread in this range exceeds one. Both liquidity and informed traders can be made better off in this region by lowering both s_i^e and s_l^e along the zero-profit locus.

The locus of points labeled $S(\lambda)$ represents the mapping using the function $S(.,.;\lambda)$ defined in (8) of the zero-profit locus of expected spreads into the corresponding locus of quoted-spread pairs (s_l,s_i) for a given value of λ. At the intersection of (ZP) with (IC_l), $s_i^e = s_l^e$, and thus $(s_l^e,s_i^e) = (s_l,s_i)$ for all λ. That is, when the broker observes a noiseless signals of the motivation for trades, quoted spreads and expected spreads are identical. More generally, for $\lambda < 1$, the locus of zero-profit actual-spread pairs will lie above and to the left of the locus of zero-profit expected-spread pairs in accordance with (7).

The boundaries defined by (IC_l) and (IC_i) are linear in s_l. (IC_l) constrains the solution to the specialist's maximization problem to (s_l^e,s_i^e) pairs to the left of (IC_l). (IC_i) constrains the solution to (s_i,s_l) pairs to the right of the boundary it defines.[11] Thus, the specialist is constrained to quoting spread pairs along the segment of $S(\lambda)$ bounded by (IC_l) and (IC_i) in fig. 2 given the value of λ as the informativeness of the broker's signal. This corresponds to expected liquidity spreads between $\underline{s}$ and $\bar{s}$.

This geometric intuition is formalized in the following lemma:

Lemma 2. Let $\bar{s} = s_p$ and let $\underline{s}$ be the smallest value of the expected liquidity spread consistent with both (IC_i) and the zero-profit constraint.[12] *If a zero-profit pooling equilibrium exists, the equilibrium value of s_l^e is in the set bounded by*

$$\underline{s} \leqslant s_l^e \leqslant \bar{s}\,.$$

Proof. Since incentive compatibility for a broker representing a liquidity trader requires $s_l \leqslant s_i$, the upper bound is established by solving for s_l when (IC_l) is

[11] The intercept for (IC_i) occurs at $s_i = (C\gamma)/(1-\gamma N)$. Thus for large values of N (IC_i) can lie below (IC_l), in which case (s_l, s_i) pairs to the left of its boundary are incentive-compatible. Under such circumstances the specialist's leverage over brokers is so great that the broker will always truthfully reveal the content of the signal he observes.

[12] Solving (10) for s_i, substituting this result into (IC_i), and solving for the smallest root when (IC_i) is binding yields

$$\begin{aligned}\underline{s} = s_{\max} + \{\pi q_i(1-\lambda) - [(\lambda\pi q_i - \pi q_i - 2\lambda q^*)^2 \\ - 8\delta\lambda\pi q_i\{\lambda\alpha - (2\lambda-1)[s_l + (C\gamma/(1-\gamma N))]\}]^{1/2}\}/4\lambda\delta\,. \qquad (11)\end{aligned}$$

binding, in which case the active specialist's problem reduces to that of the passive specialist. Therefore $s_l^e = \bar{s} = s_p$. For nonnegative s_l^e, the lower bound is established by solving for s_l when (IC_i) is binding and noting that $(\underline{s}, s_i^e) = S^{-1}(s_l, s_i; \lambda)$. Because the zero-profit locus is convex, all values of s_l^e between $\underline{s}$ and $\bar{s}$ are feasible.

Lemma 2 has an obvious welfare implication for liquidity traders, which we characterize in Theorem 3.

Theorem 3. The expected spread faced by liquidity traders under an active specialist is never greater than the spread set by a passive specialist, or

$$s_p \geqslant s_l^e .$$

Proof. The result follows directly from Lemma 2.

The intuition in this result is that the ability to discriminate between brokers representing liquidity and informed traders and to sanction those who attempt to disguise the motives for their trading activity gives the specialist the leverage necessary to weaken the adverse-selection problem and thereby improve the terms of trade for liquidity traders. Further, we show in Lemma 3 that as the broker's ability to infer the motivation for a trade increases, the specialist's ability to improve the terms for liquidity traders is enhanced.

Lemma 3. $\partial \underline{s}/\partial \lambda \leqslant 0$.

Proof. For given (s_l^e, s_i^e), $\partial s_l/\partial \lambda \leqslant 0$ and $\partial s_i/\partial \lambda \geqslant 0$. Therefore, the (s_l^e, s_i^e) pair for which (IC_i) is binding for a given value of λ will satisfy (IC_i) with strict inequality for higher values of λ. Therefore, as λ increases, the lower bound $\underline{s}$ can be reduced.

For the extreme case that $\lambda = 1$, $S(\lambda)$ and (ZP) are identical, so that the lower bound on s_l^e falls to $\underline{s}'$ in fig. 2.

Although it would seem likely that the active specialist's ability to extract informed traders' private information would worsen their terms of trade, we show in Theorem 4 that there are conditions under which the active specialist can actually improve the terms for these traders.

Theorem 4. If a pooling equilibrium exists and is inefficient and if $\gamma > 0$, *there is an admissible pair of spreads* (s_l, s_i) *such that*

(*i*) $s_l^e < s_p$,

(*ii*) $s_i^e < s_p$.

Proof. Satisfying (i) and (ii) simultaneously requires $\partial s_l^e/\partial s_i^e > 0$ along the zero-profit locus. Differentiating (9) with respect to s_i^e yields $-(2q^* - 4\delta s_i^e)/\pi q_i$, which is greater than zero when $s_i^e > q^*/2\delta$. Since $q^*/2\delta = s_{\max}$, (i) and (ii) are satisfied when s_p is inefficient.

Theorem 4 follows from the definition of an inefficient spread: one at which the ability to discriminate between liquidity and informed traders would improve the terms of trade for both. The spread quoted by a passive specialist is more likely to be inefficient when the information asymmetry is most severe, so the specialist's leverage over brokers is most likely to lead to better terms for informed traders precisely when their informational advantage is greatest. Severe informational asymmetry leads to inefficient pooling spreads; the ability to charge lower spreads to liquidity traders in this environment actually generates increased revenue that can be used to reduce the information spread.

To complete our characterization of the active-specialist equilibrium, we investigate the efficiency of the expected-spread pair (s_l^e, s_i^e) corresponding to the active specialist's quoted-spread pair (s_l, s_i) to determine whether, as for some pooling equilibria, a Pareto improvement is possible. From Lemma 2 it follows that if the pooling equilibrium spread is efficient, the expected-spread pair under the active specialist will also be efficient. Theorem 5 shows, however, that even if the pooling equilibrium is inefficient, the active-specialist equilibrium still may be efficient.

Theorem 5. Market parameters that result in an inefficient equilibrium in a pooling environment will result in an efficient equilibrium in a separating environment if the specialist's ability to sanction brokers identified as having exploited private information satisfies

$$\frac{C\gamma}{I - \gamma N} \geqslant s_{\max}\left[\frac{\lambda(\pi q_i + q^*) - \pi q_i}{\lambda(\pi q_i \alpha - 2\pi q_i) + \pi q_i}\right] - s_l \,. \tag{12}$$

Proof. Since broker profits increase when both s_l and s_i are reduced and because inefficiency implies that both s_l and s_i can be reduced while still satisfying the specialist's zero-profit constraint, the active specialist produces an efficient outcome whenever $s_{\max}$ is in the feasible region. This requires $\underline{s} < s_{\max}$. Setting $\underline{s} = s_{\max}$ and simplifying yields the stated condition.

Thus the active specialist, in contrast to the passive specialist, always produces Pareto-efficient terms of trade if given enough leverage over brokers repesenting information trades. Observe that s_l is a decreasing function of λ. Further, Lemma 3 implies that the active specialist needs less leverage to achieve a Pareto-efficient outcome as the quality of the broker's signal improves (i.e., as λ increases).

Finally, when a pooling equilibrium exists, an active specialist also will keep the market open. Further, given sufficient leverage over brokers, an active specialist will keep the market open under conditions for which a pooling equilibrium is nonexistent. From (11), the condition that must be satisfied for an equilibrium to exist with an active specialist is

$$(\lambda\pi q_i - \pi q_i - 2\lambda q^*)^2 - 8\delta\lambda\pi q_i\{\lambda\alpha - (2\lambda - 1)[s_l + (C\gamma/(1 - \gamma N))]\} \geqslant 0\,. \tag{13}$$

Since $(2\lambda - 1) \geqslant 0$, for a given set of parameters for which the passive specialist closes the market, the active specialist will always keep the market open if he has enough leverage $[C\gamma/(1 - \gamma N)]$ over brokers.

Comparing this result with Glosten's (1989) analysis of a monopolist specialist's ability to keep a market open in the presence of severe adverse selection highlights the difference in our models. Glosten's monopolist specialist facilitates the revelation of private information by keeping the market open and bearing the burden of adverse selection under conditions in which a competitive specialist would simply refuse to trade. The specialist's burden is transferred to uninformed traders through the collection of monopoly rents. In the terminology of our model, this represents a passive response to asymmetric information. In our model, given the power to sanction brokers exploiting private information, the active specialist can induce revelation of information that if kept private would otherwise lead to the market's being closed. By distributing the burden of adverse selection across both informed and uninformed traders the active specialist can improve the terms of trade for each class.

Before leaving the formal model, we should address an issue of time consistency. We have asserted that the incentive-compatibility condition for informed trades, which allows traders to retain some profits from their private information, will be honored by the specialist. But once brokers for informed traders reveal their status, the specialist could avoid losses on such trades by reneging on the posted price. This raises the following question: is the incentive-compatibility condition for informed traders incentive-compatible for specialists? In a repeated-trading setting such as the one applicable here, we believe it is.

First consider a market with competing specialists. If an individual specialist attempted to appropriate all profits from informed traders, then there would be no incentive to reveal information, and that specialist could sponsor only a pooling market. But we know that the pooling equilibrium is always dominated for liquidity traders by a separating equilibrium. Therefore, no liquidity trader would participate in that specialist's market, and, left only with information traders, the specialist would be driven from the market by a classic lemons problem. So, while there might be short-run incentives to renege, long-run considerations would force the specialist to honor commitments to allow information traders to profit on their trades.

In the monopolist specialist setting of the NYSE, the specialist is prevented from reneging on the commitment to allow informed traders to retain some profit (that is, is bound to honor the quoted spread s_i) by various exchange rules and procedures. Specialists are exchange members, and benefit from the increased volume that results from honoring the conditions that allow for the separating equilibrium. Moreover, specialists are evaluated annually by brokers, and these evaluations help to determine priority in the distribution of new listings. Therefore, the specialist's willingness to honor the commitment to allow informed traders to profit affects his standing among both other specialists and brokers, and ultimately affects future profit opportunities.[13]

Finally, although we believe that the most important element of private information might be that pertaining to (unexpressed) demand envisioned by Grossman (1990) and Benveniste and Spindt (1990), our model has some implications for the incentives to gather costly fundamental information on firms' future cash flows. Because an active specialist will charge a lower spread to informed (as well as liquidity) traders than would emerge in an inefficient pooling equilibrium, the active specialist may actually improve incentives for information gathering. Therefore, if information trading is sufficient to make the pooling market inefficient in the sense defined above, active separation of informed and liquidity traders can increase the *informational* efficiency of the market as well. If, on the other hand, the pooling equilibrium would be efficient, the separating equilibrium results in a larger spread to informed traders and reduces incentives for information gathering.

5. Discussion

The main implication of our model is that the benefits of a floor exchange mechanism will be greatest when the potential for privately informed trading is greatest and when liquidity traders are most sensitive to transaction costs. This suggests that floor exchange mechanisms will yield their greatest benefits in markets for assets such as stocks and stock options where the potential for private information is greatest. Unfortunately, this does not explain why the NASD has chosen to maintain a computerized dealer market for the most heavily traded over-the-counter stocks. Glosten and Gammill suggest that in fact these stocks may not be subject to the degree of information asymmetry associated with NYSE or American Stock Exchange (AMEX) stocks. Alternatively, it may be that liquidity demand is relatively inelastic in the over-the-counter market, in which case the ability of the active specialist to achieve the

[13] This leverage is strengthened to the extent that specialist firms incur significant fixed costs in the course of doing business. The fact that the average specialist firm makes markets for approximately 30 stocks supports the hypothesis of economies of scale.

welfare gains discussed earlier will be limited. If this is true, we would expect to observe a relatively large adverse-selection component in the bid–ask spreads quoted for NASDAQ National Market System stocks; an implication borne out in the empirical evidence presented by Stoll (1989) suggesting that 43% of the quoted spread is attributable to asymmetric information.

At the opposite extreme are markets characterized by significant (inelastic) hedging demand and less information-driven activity such as currency and index derivatives or government bonds. Consistent with the predictions of our model, the major participants in the government bond market, for example, are aggressively experimenting with a computerized exchange mechanism in which the layer of intermediation provided by brokers is eliminated (*Wall Street Journal*, August 28, 1991, p. A1).

We therefore conclude that a floor exchange mechanism is best suited to the 'middle ground'. An active specialist will be unable to achieve significant welfare benefits in a market in which liquidity demand is relatively inelastic. On the other hand, high liquidity demand, low information markets might best be served by a market structure that presents relatively low entry barriers to potential (uninformed) contributors to market liquidity. In either case, a computerized exchange mechanism may prove a viable alternative to a floor exchange. Floor exchange systems appear to have the greatest comparative advantage in markets where liquidity and information traders are most likely to confront one another and liquidity traders are most sensitive to transaction costs.

Despite our conclusion that the floor exchange mechanisms maintained by the major U.S. stock exchanges can improve the terms of trade for public customers, exchange officials are increasingly concerned about recent indications of vulnerability to off-board trading activity. Within the context of our model this is perhaps less surprising than it might at first appear. A significant proportion of the off-board activity involves trading of stock portfolios by large institutions. As the analysis in section 4 suggests, liquidity traders always prefer to be distinguished from information traders. Trading an entire portfolio of stocks gives a trader a means of credibly announcing the liquidity motivation for the trade.[14] Thus, off-board trading of stock portfolios simply represents an attempt by liquidity traders to isolate themselves from the effects of asymmetric information.

The problem that arises is that not all liquidity traders have access to these alternative exchange mechanisms. When portfolio traders leave the exchange floors for alternative markets, the burden of asymmetric information is borne by the smaller pool of liquidity traders, perhaps dominated by individual investors,

[14]Admati and Pfleiderer (1990) investigate the welfare effects of 'sunshine trading' mechanisms. Within the context of their model, we can be viewed as asking how the specialist can improve the lot of the uninformed and informed traders remaining after the uniformed traders who can credibly announce their motivations have selected out of the market.

who have no alternative to the exchange floor. Were the shift to off-board trading of portfolios to become extreme, the adverse selection faced by the remaining pool of liquidity traders could be enough to derive them from the market, leading to the collapse of the central marketplace.[15] Thus, attempts by institutional liquidity traders to separate themselves from informed traders have important implications for both individual liquidity traders and, to the extent that centralized trading contributes to price discovery [Schwartz (1988, p. 499)], the informational efficiency of the market.

Finally, although our model is cast in terms of a specialist system, it should be clear that it applies generally to environments with repeated trading and opportunities for *ex post* sanctions. For example, Burdett and O'Hara (1987) note that a block house acts as both a broker and a market maker. A common belief is that in its brokerage role, the block trader is frequently called upon to certify that a block sale is not information-motivated. A block house that develops a reputation for cheating on behalf of its clients will lose credibility and consequently find it difficult to form buying syndicates for future blocks. This in turn weakens the ability of the block trader to negotiate for a block seller, and thus the ability to compete in the block market. Hence, the reputation effects among the small community of identifiable block houses serve precisely the same role in that market as the sanctioning power of the active specialist in our model. In both cases, *ex post* sanctions reduce the incentive to aid clients in exploiting differential information and ultimately reduce the adverse-selection component of the bid–ask spread.

As Seppi (1990) and Keim and Madhavan (1991) note, the lack of anonymity in the block trading market can be used by the block house to mitigate asymmetric information problems. The basis for this conjecture appears to be the importance of repeat trade between block sellers and block houses. Seppi (1990) notes that block sellers are frequently called on to commit themselves not to trade the stock for some period during which the block house is presumably trying to liquidate its inventory of the stock. Failure to honor this commitment (referred to as 'bagging the street') is generally punished by refusal to position future blocks for the seller. Presumably, the block house could use the continuing business relationship with the block seller as leverage to induce revelation of the motivation for a block sale, much as the active specialist is able to induce brokers to reveal private information on the exchange floor.

[15]The exchange's introduction of basket trading and the SuperDot system offer alternatives to off-board trading of stock portfolios. Although activity in the market basket has been weak since its inception, SuperDot has been an active avenue for program trading of portfolios. Since the level of immediacy provided by SuperDot is essential to index arbitrage and other dynamic trading strategies, it seems unlikely that off-board mechanisms that substitute reduced execution costs for immediacy will draw this source of liquidity demand from the exchange.

References

Admati, Anat R. and Paul Pfleiderer, 1988, A theory of intraday patterns: Volume and price variability, Review of Financial studies 1, 3–40.

Admati, Anat R. and Paul Pfleiderer, 1989, Divide and conquer: A theory of intraday and day-of-the-week mean effects, Review of Financial Studies 2, 189–224.

Admati, Anat R. and Paul Pfleiderer, 1990, Sunshine trading and financial market equilibrium, Working paper (Stanford University, Stanford, CA).

Bagehot, Walter, 1971, The only game in town, Financial Analysts Journal 22, 12–14.

Benveniste, Lawrence M. and Paul A. Spindt, How investment bankers determine the offer price and allocation of new issues, Journal of Financial Economics 24, 343–362.

Burdett, Kenneth and Maureen O'Hara, 1987, Building blocks: An introduction to block trading, Journal of Banking and Finance 11, 193–212.

Cox, John C. and Mark Rubinstein, 1985, Options markets (Prentice-Hall, Englewood Cliffs, NJ).

Easley, David and Maureen O'Hara, 1987, Price, trade size, and information in securities markets, Journal of Financial Economics 19, 69–90.

Gammill, James F., 1989, The organization of financial markets: Competitive versus cooperative market mechanisms, Working paper (Harvard Business School, Boston, MA).

Glosten, Lawrence R., 1989, Insider trading, liquidity, and the role of the monopolist specialist, Journal of Business 62, 211–235.

Glosten, Lawrence R. and Paul R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 14, 71–100.

Grossman, Sanford J., 1990, The informational role of upstairs and downstairs trading, Working paper (University of Pennsylvania, Wharton School, Philadelphia, PA).

Hasbrouck, Joel, 1988, Trades, quotes, inventories, and information, Journal of Financial Economics 22, 229–252.

Keim, Donald B. and Ananth Madhavan, The upstairs market for large-block transactions: Analysis and measurement of price effects, Working paper (University of Pennsylvania, Wharton School, Philadelphia, PA).

Merton, Robert C., 1981, On market timing and investment performance, Part I: An equilibrium theory of value for market forecasts, Journal of Business 54, 323–361.

Seppi, Duane, 1990, Equilibrium block trading and asymmetric information, Journal of Finance 45, 73–94.

Shapiro, James E., 1989, The NYSE market system: Background and issues, Discussion paper (New York Stock Exchange, New York, NY).

Schwartz, Robert A., 1988, Equity markets: Structure, trading, and performance (Harper and Row, New York, NY).

Sirri, Erik R., Bid/ask spread, price, and volume in a specialist market, Working paper (University of California, Los Angeles, CA).

Stoll, Hans R., 1989, Inferring the components of the bid–ask spread: Theory and empirical tests, Journal of Finance 44, 115–134.