Insider Trading, Liquidity, and the Role of the Monopolist Specialist

Lawrence R. Glosten

Lawrence R. Glosten
Northwestern University

Insider Trading, Liquidity, and the Role of the Monopolist Specialist*

The purpose of this article is to investigate a rationale for the specialist system on the New York Stock Exchange (NYSE). A key characteristic of the specialist system is that there is one individual in each stock, the specialist, that has sole access to information about the trading process, and this information provides some monopoly power. It is shown that this monopoly power may, in some environments, reduce some of the social costs associated with trading on private information.

Promotional literature published by the New York Stock Exchange points to the "specialist system" as a superior form of market organization, citing the "increased liquidity" and the maintenance of a "fair and orderly market" attributable to the specialist. On the face of it, the claim seems preposterous. One's intuition might suggest that any model of rational market making

Trading on private information creates inefficiencies because there is less than optimal risk sharing. This occurs because the response of market makers to the existence of traders with private information is to reduce the liquidity of the market. The institution of the monopolist specialist may ease this inefficiency somewhat by increasing the liquidity of the market. While competing market makers will expect a zero profit on every trade, the monopolist will average his profits across trades. This implies a more liquid market when there is extensive trading on private information.

would show that a monopolist will optimally provide a less liquid market than competing market makers. In fact, Ho and Stoll (1981) show that competing market makers will lead to a more liquid market in the sense that the average bid/ask spread will be smaller with competing market makers.

A possible response is that the NYSE monitors the performance of specialists, and therefore, in fear of retribution, specialists provide a more liquid market than they would if unregulated. However, if this is the only argument, the obvious rebuttal is, If a regulated monopolist can supply what would be supplied by competitors, why should regulatory resources be expended when the seemingly more efficient alternative would be to replace the specialist system with competing market makers?

The possible advantage of the specialist system was suggested in Glosten and Milgrom (1985), and the intuition is provided by the observation that a monopolist specialist can average profits over time. For example, the analysis of Glosten and Milgrom suggests that there may be instances in which competing market makers will not quote, bid, or ask prices. If the adverse selection problem is too extreme, then no matter what bid and ask are set, each market maker will expect to lose money on a trade. The consequence is that the market shuts down until enough public information arrives to relieve the adverse selection problem. Now imagine a situation in which no public information arrives until the informational event occurs. In this case, after the informed traders receive their information the market will remain closed until the information is revealed.

A monopolist specialist may also close the market in such a situation, but he need not. By keeping the market open, the specialist learns some of the information of the informed. This may reduce the adverse selection problem, making subsequent trades more profitable. With a protected position, the specialist is able to realize the greater profits that offset the losses earned early on in the trading. The result is that both liquidity traders and informed traders are made better off relative to the competing market maker system.

This article extends this intuition by considering an environment in which trades are not restricted to unit amounts as in Glosten and Milgrom (1985). In this case, the market maker (monopolist or competitor) quotes a price schedule that specifies the price per share as a function of the quantity to be traded. While competing market makers are forced to set a price schedule that leads to a conditional expected profit of zero (conditioned on the quantity traded), the monopolist specialist maximizes expected profits over all possible quantities. This extra degree of freedom, the ability to average profits across trades at a point in time, reinforces the above observation that the specialist can average his profits across time. The result is that in some environments the monopolist will provide a more liquid market. This reinforcement may

be important. It is reasonable to assume that there are times in which the specialist enjoys a monopoly position and times in which this monopoly power is curtailed by other traders. Thus, the ability of the specialist to average profits over time may be restricted.

The model considered bears some resemblance to other models of trade with asymmetric information. Like Glosten and Milgrom (1985), an environment is considered where traders arrive one at a time, motivated by liquidity and/or information considerations. As noted above, however, trades may be of different quantities, and hence there is a relation between this model and the one of Easely and O'Hara (1987).[1] This feature also recalls the model of Kyle (1985). In contrast to Kyle's model, liquidity demand is not perfectly inelastic and only in the analysis of competing market makers is the zero-profit assumption made. Furthermore, all traders are assumed to know the price they will pay when they submit their demands. The problem of the optimal dynamic strategy of an informed trader is not considered as in Kyle. Furthermore, the model ignores the "manipulation" strategies considered by Ketterrer (1987).

Like Gould and Verrecchia (1985), part of this article considers the behavior of a monopolist specialist facing informed agents. There are several important differences in the analysis. In the Gould and Verrecchia model, the specialist has some private information and faces agents with private information. Furthermore, the specialist is constrained to quoting a single price. Despite this lack of a "spread" in the Gould and Verrecchia model, it is still possible to talk about the liquidity of the market. Since the risk-sharing motive of the trader is assumed to be common knowledge and since the specialist is risk neutral, the amount of risk sharing measures the liquidity of the market.

Though Gould and Verrecchia (1985) do not work out the zero-profit solution for comparison with the monopolist solution, it can be shown (at least in the uninformed specialist case) that the monopolist specialist will provide less risk sharing than a specialist who sets the single price that will yield zero-expected profits. Essentially, given that only a single price is quoted, both a zero-profit specialist and a monopolist specialist are able to average profits across trades. Thus, in the Gould and Verrecchia model there is no advantage associated with a monopolist. However, it is unlikely that competing market makers would charge a single price—rather they would establish a price schedule, in which case the Gould and Verrecchia model is not easily extended to the case of competing specialists.

In the context of insurance markets, Rothschild and Stiglitz (1978) raise the possibility that—but state that it is impossible to say whether

---

1. The model of competitive market makers is identical to the independently derived model in Madhavan (1987). He uses the model to examine continuous markets and call markets.

—asymmetric information and the noncompetitive nature of insurance markets are related. However, subsequent work by Stiglitz (1977) would seem to provide some evidence that, in certain environments, an equilibrium among competing insurers will not exist while a monopolist will provide insurance in such environments. This result is in the same spirit as the liquidity arguments presented here.

The intuition that a monopolist specialist could average profits across time is central to the analysis of Gammill (1986). He proposes a specific market structure in which there is a "large-trade" trading mechanism and a "small-trade" trading mechanism. Small trades transact at exogenously specified bid and ask prices, whereas large trades are accommodated by a monopolist specialist. In certain environments, the informed choose to trade via the large-trade mechanism and reveal their information. The specialist loses on these trades, but makes up his losses in the subsequent small-trade market.

While this model requires a monopolist specialist to achieve the temporal averaging of profits, it is not clear that a monopolist would choose the trading mechanism posited. Furthermore, the exogeneously specified "small-trade" market seems somewhat ad hoc. This makes difficult a comparison of the proposed trading mechanism with a competitive market maker system, and hence Gammill's article (1986) does not (indeed was not designed to) directly address the impact of a monopolist specialist on the performance of the market.

This article is organized as follows. In the next section, modeling issues are discussed, and the model itself is displayed. It is shown that there will be instances in which competing market makers will close down the market. The intuition for this is as follows: start with the hypothesis that small trades are not generated by extreme information. In such a case, the spread for small trades will be near zero. The break-even spread for large trades, however, will be large if there is a high probability of extreme information. If this spread is too large, then those with extreme information will in fact wish to make small trades, violating the initial hypothesis. Extending this argument, every market maker's hypothesis about the behavior of traders is refuted and the market closes down.

The equilibrium of the competitive model is used to show that investors would vote, ex ante, for a law restricting insider trading. Intuitively, the fact that the zero-profit price schedule must be upward sloping implies that asymmetric information leads to a reduction in the liquidity of the market. Ex ante, an investor does not know whether he will profit from information or pay a premium to escape a risky position. Hence he sees the consequence of allowing insider trading as being a reduction in the liquidity of the market.

The subsequent subsection examines the behavior of a monopolist specialist. It is shown that a monopolist optimally earns a positive

profit on frequent small trades, but loses on the infrequent large trades. Essentially, the monopolist keeps the price of large-investor buys relatively low to prevent pooling of extreme and nonextreme information at low quantities. Consequently, he subsidizes the large trades so that the small, profitable trades remain.

There are two conclusions of this section. First, the monopolist will always keep the market open since he can achieve this cross-subsidization of trades. Consequently, in situations in which asymmetric information is perceived to be a significant problem, the monopolist specialist is able to maintain a more liquid market. Second, even in some environments in which competitive market makers would keep the market open, market participants would vote for a monopolist specialist. This is proved by showing that the benefit of a monopolist specialist is related to a parameter that measures the severity of the asymmetry of information. The benefit starts out positive (when the market closes down) and ends up negative (when there is no asymmetry of information). The article concludes with some speculation on the meaning of the results, extensions of the analysis, and a summary of the major results.

I. Model

As stated in the introduction, the purpose of this article is to examine a rationale for the specialist's existence. Of course, the specialist is not a total monopolist—he faces competition from limit order submitters, other floor traders, and specialists in other securities (see Hagerty 1986). In fact, some of the advantage that the specialist has is informational. He, and he alone, knows what is in the "specialist's book." Furthermore, he is in a position to know more about market activity since most trades will be crossed with him. Indeed, many formal and informal discussions of the position of the specialist assert that his monopoly position derives from this information (see Leffler and Farwell 1963; Ho and Stoll 1981; and Stoll 1984).

This observation can be interpreted several ways. Some have taken it as a motivation for analyses of a monopolist with private information (see Gould and Verrecchia 1985; and Altug 1984). This article takes the point of view that the information available from the specialist's book, and the fact that he sees the result of all trades, is information not about the future profitability of the firm or the future realizations of its stock price, but rather about uncertainty in the trading process. Thus, this information gives the specialist a monopoly position, but it is not information that traders off the floor of the exchange care about (except to the extent that it might affect their order placement strategy, i.e., the choice between limit order and market order). Furthermore, the observation that specialists do not typically engage in security analysis

strengthens the position that a more appropriate model of the specialist assumes that he has monopoly power but no private information about the cash flows of the security.

A model embodying private information and rational expectations should provide for "liquidity motivated" trade—trade unrelated to information—and the inability of the market makers to distinguish informed and uninformed traders. Without the former, there will be no trade. Without the latter, market makers will refuse to trade with the informed and always trade with the uninformed, in which case there will be no informed trade. The indistinguishability of informed and uninformed is modeled here by assuming that every trader has both an informational as well as liquidity motivation for trading. To the extent that a trader's endowment is, ex ante, optimal and his information is extreme, he is trading for information reasons. To the extent that his information is not extreme and his portfolio is out of balance, he is trading for liquidity reasons.

Market makers (monopolist or competitor) are presumed to be risk neutral. Indeed, the analysis of Gould and Verrecchia (1985) would suggest that market makers would be relatively less risk averse than the typical investor. However, this implies that inventory costs induced by risk aversion are ignored in this model. This may bias the case in favor of the monopolist system since presumably the ability to handle a given "market" inventory is increased by increasing the number of market makers. This bias should be recognized, but removing it would not alter the basic results. It will merely reduce the set of environments in which the monopolist organization is optimal.

Finally, the dynamic order-placement strategy of the traders is ignored. It is easiest to interpret the model in the following way. Each agent becomes informed and comes to market immediately with an endowment of shares known only to him. He plans to trade only once and chooses the quantity to maximize his expected utility given the posted price schedule. Subsequent traders have endowments that any market maker considers to be independent of the past.

The analysis of a market maker's behavior will consider a response to a single proposed trade without regard to the dynamic strategies the market maker could adopt. Given the assumed behavior of traders, there is no loss in generality in this for the case of competing market makers. Free entry of market makers and a risk-neutrality assumption imply that the equilibrium in the competing market-maker setting is independent of expectations of future trades. A monopolist specialist, on the other hand, will adopt a dynamic strategy and consideration of the single trade case ignores this. The model should be interpreted as a description of what occurs at a particular point in time. Then, except for revisions in the relevant distributions, the situation is repeated.

Though the analysis will be carried out with a specific example of

preferences and information, it is easier to set up the nature of the model somewhat more generally. The paper concentrates on the trading in one security, and for tractability it is assumed that this security is the only risky security in a typical investor's portfolio. An investor arriving at a point in time is presumed to have preferences over distributions of future random wealth represented by the utility function of wealth, $U(\cdot)$, which is increasing and concave. There are two securities, one risky and one risk free. The trader has an endowment of W_0 dollars in cash and W shares of the risky asset. The risky asset will have a payoff of $X(w)$ in state w while the risk-free asset has payoff R. In addition to the endowment, the investor has received information, S, about the payoff, X. When the investor comes to market, he faces a pricing schedule, $P(\cdot)$, with the interpretation that a (signed) trade of Q will lead to a transfer of $QP(Q)$ from the trader to the specialist. Thus, if Q is positive, the investor pays a price of $P(Q)$ per share, and if Q is negative he receives $P(Q)$ per share. Given his position in the other securities and the price function $P(\cdot)$, the investor chooses Q to solve

$$\max E[U(X(W + Q) + W_0 R - P(Q)QR)|S,W_0,W]. \tag{1}$$

Solution to this maximization leads to a trade:

$$Q = \hat{Q}_P(S,W_0,W), \tag{2}$$

where the subscripted P indicates that the chosen quantity depends upon the pricing function P.

A typical market maker is assumed to be risk neutral, facing no inventory costs or constraints and ignorant of the endowment vector and the realization of the information. He does know the cross-sectional distribution of these variables. This specification leads to the solution among competing market makers, P_c, which, as long as it exists, must satisfy

$$E[X|Q] = P_c(Q). \tag{3}$$

That is, any trade of Q shares must lead to a zero conditional expected profit. If P_c were set so that for some trades a profit were expected, then a competing market maker would have an incentive to beat the price for those trades.[2]

There may be many solutions to (3) since the conditional expected value function depends upon the price schedule that is established. Intuitively, if P_c is one schedule satisfying (3) and P' is another schedule that is above P_c for $Q > 0$, then an investor facing P' will trade a smaller quantity than if P_c were the schedule. Hence a given Q leads to

---

2. It is assumed that either (1) traders cannot split their orders among market makers or (2) they can split their orders, but such splitting is observed by all market makers so that N orders of size Q/N placed by one trader are informationally equivalent to one order of size Q.

a greater revision in expectations when P' is the schedule. This greater revision in expectations may just match the higher price schedule. It seems natural to assume that competition among competing market makers will determine the smallest (largest) price schedule for $Q > (<)$ 0 satisfying (3).[3]

If there is only one specialist, then it is assumed that the pricing function, P_m, must satisfy (again, as long as it exists)

$$P_m(\cdot) = \arg \max E[P(\hat{Q}_P(S,W_0,W))\hat{Q}_P(S,W_0,W) - X\hat{Q}_P(S,W_0,W)]. \tag{4}$$

The specification of the objective function in (4) deserves further comment. It is conceivable (indeed in the example considered it will be shown to be the case) that for some Q the expected profit, conditional on that Q, will be negative. In specifying the objective function as the expected profit across possible quantities, the tacit assumption is that the specialist quotes a price schedule and then is required to honor those quotes for the trade, no matter what Q is requested. Of course, in reality there is no posting of a price schedule—the only prices posted are a bid and ask for small trades. If a trader wishes to trade a larger quantity, then quotes must be obtained for larger quantities. In any event, once the specialist supplies quotes for various quantities, he is in fact bound to honor those quotes, whereas the trader is free to choose the quantity he wishes to trade based on those quotes. The fact that the specialist "moves first," in the sense that he must honor his quotes, is the motivation for choosing the objective function in (4).

Price functions satisfying (2) and (3) or (4) need not exist. In fact, the subsequent analysis will show that there are situations in which the competitive pricing function does not exist. This happens when there is very little uncertainty surrounding the endowments and relatively great uncertainty surrounding the information about X. In this instance, the adverse selection problem that market makers face is so severe that there is market breakdown—all market makers refuse to make a market.

This may or may not happen in the case of the monopoly specialist, and in the example considered here it does not. Since the specialist maximizes expected profits, he may expect to lose for some Q and profit on other trades and still expect a positive profit. His monopoly position allows him to average his profits cross-sectionally, that is, across possible trades.

The specification of the competitive and monopoly environments in (1)–(4) are so general as to preclude all but the weakest analysis summarized in the above two paragraphs. To obtain more results, more

3. Of course, schedules satisfying (3) may not be ordered, in which case there is no basis for choosing among the equilibria. This will not be the case with the example considered.

structure must be imposed, and the time-honored example from the rational expectations literature is chosen. Assume that X is normally distributed, S is a noisy observation of X, and the utility function of a typical trader is exponential. Furthermore, W is, from a market maker's point of view, normally distributed with zero mean.[4]

Formally, the model is specified as follows:

$$U(y) = -\exp(-\rho y), \quad \rho > 0, \tag{5a}$$

$$X \sim N(m, 1/\pi_x), \tag{5b}$$

$$W \sim N(0, 1/\pi_w), \text{ independent of } X, \tag{5c}$$

$$S = X + \varepsilon, \varepsilon \sim N(0, 1/\pi_s), \text{ independent of } X, \tag{5d}$$

$$R = 1. \tag{5e}$$

Thus, an arriving investor has exponential utility with risk-aversion parameter ρ. The payoff on the security is normally distributed with mean m and precision (one over the variance) of π_x. From a market maker's point of view, W is normally distributed with mean zero and precision π_w. The information received by the trader consists of a noisy observation of X, the precision of the signal conditional on X being π_s. Finally, the return on cash is normalized to zero.

A. Competitive Market-Maker Equilibrium

In order to derive the equilibrium with competing market makers, we have either to find a price schedule $P(Q)$ satisfying $E[X|Q] = P(Q)$, when Q is the optimal Q chosen by an arriving trader, or show that no such schedule exists. We begin by assuming that a pricing schedule exists, and that it is differentiable and defined for all Q. The terminal wealth of a trader facing price schedule $P(\cdot)$ and choosing quantity Q with initial endowment W is thus

$$-P(Q)Q + (W + Q)X. \tag{6}$$

The conditional normality of X given S and exponential utility imply that expected utility maximization is equivalent to maximization of the certainty equivalent

$$CE(W,S,Q,P(Q)) = -P(Q)Q + (W + Q)E[X|S]$$
$$- .5\rho(W + Q)^2 \text{var}(X|S). \tag{7}$$

4. This specification of the liquidity motivation for trade is also used in Diamond (1985). Also notice that the specification of a zero mean implies that the risk represented by the random variable X is in zero net supply. Thus, the effects of the risk-neutral market makers as insurers is minimized. That is, the market makers facilitate risk sharing between agents and they do not expect to end up with a positive (or negative) inventory of shares. Furthermore, the risk neutrality of the market makers does not affect the average price of the security.

After noting that $E[X|S]$ and $\text{var}(X|S)$ are given by

$$E[X|S] = (\pi_x m + \pi_x S)/(\pi_x + \pi_s); \ \text{var}(X|S) = 1/(\pi_x + \pi_s), \quad (8)$$

and assuming that the price function is twice differentiable with first and second derivatives given by $P'(\cdot)$ and $P''(\cdot)$, respectively, the first- and second-order necessary conditions for the investor's optimal Q are given by

$$\{[P'(\hat{Q})\hat{Q} + P(\hat{Q}) - m](\pi_x + \pi_s)/\pi_s\} + \rho \hat{Q}/\pi_s$$
$$= S - m - (\rho W/\pi_s) = Z - m, \tag{9}$$

$$- P''(\hat{Q})\hat{Q} - 2P'(\hat{Q}) - [\rho/(\pi_x + \pi_s)] < 0. \tag{10}$$

Notice that if an arriving trader announces Q, then the market maker can calculate the value of the left side of (9), and hence he knows $Z = S - (\rho W/\pi_s) = X + \varepsilon - (\rho W/\pi_s)$, a noisy observation of X. Let π_z denote the precision of this observation conditional on X (i.e., one over the variance of $\varepsilon - [\rho W/\pi_s]$). Then, a market maker can calculate $E[X|Z]$:

$$E[X|Z] = m + \pi_z Z/(\pi_x + \pi_z). \tag{11}$$

Substituting for Z from (9) in (11) provides an expression for $E[X|Q]$. Equating this to $P(Q)$ leads to a differential equation that $P(\cdot)$ must satisfy. The family of solutions is presented in the following lemma, which is proven in the Appendix.

LEMMA 1.   Any differentiable pricing function that satisfies (9) and (3) is given by

$$P(Q) = m + (\rho \pi_w \pi_s/N)[Q/(2\alpha - 1)] + K[\text{sign}(Q)]|Q|^\gamma, \quad (12)$$

where

$$N = \rho^2 \pi_x + \pi_w \pi_s(\pi_s + \pi_x),$$

$$\alpha = \rho^2 \pi_x/N \neq .5, \ \gamma = \alpha/(1 - \alpha),$$

K unrestricted, or

$$P(Q) = m + k_0 Q - k_1 Q \log(|Q|),$$

$$\alpha = .5,$$

k_0 unrestricted,

$$k_1 = \rho/(\pi_x + \pi_s).$$

The price schedules in lemma 1 were constructed assuming that (9) characterized the decision of an investor. In fact, not all the price schedules satisfy the second-order condition in (10). It can be verified that in order for the second-order condition to be satisfied, α must exceed .5, and K, the unspecified constant, must be nonnegative. The

assumed competition among market makers will thus determine the linear price schedule as the only equilibrium. The results concerning the equilibrium among competing market makers are presented in the following proposition.

PROPOSITION 1. If

$$\alpha = \rho^2 \pi_x / [\rho^2 \pi_x + \pi_w \pi_s (\pi_x + \pi_s)] > .5,$$

then there is a unique competitive equilibrium pricing shcedule given by

$$P_c(Q) = m + \{\rho \pi_w \pi_s / [(2\alpha - 1)N]\}Q, \tag{13}$$

$$N = \rho^2 \pi_x + \pi_w \pi_s (\pi_s + \pi_x).$$

The Q chosen by an arriving investor is derived from (9). The equilibrium Q, $Q_c(Z)$ is given by

$$Q_c(Z) = (Z - m)\pi_s(2\alpha - 1)/\rho. \tag{14}$$

If $\alpha \leq .5$, then there is no equilibrium pricing schedule and the market closes down. The proof for proposition 1 is in the Appendix.

Whether or not the market is open is determined by α, which is a function of the parameters $\rho, \pi_x, \pi_s, \pi_w$. Each of these determines how extreme a position the investor wants to take in response to a given signal. The more extreme a position the investor wishes to take, the more the market makers have to protect themselves by making the price schedule steeper. At some point (when $\alpha = .5$), market makers are unable to protect themselves and the market closes down.

Some discussion of the equilibrium concept is in order here. The price schedules that market makers can pick have been limited to those satisfying the second-order condition in (10). This insures that the equilibrium, when it exists, is a separating equilibrium. It has been pointed out to me by Ailsa Roell that, if the schedules are not so restricted, the linear price schedule that has been derived above is not an equilibrium. Specifically, if the linear schedule above were offered, a market maker could offer a competing schedule that would cause the investors with Z's in some interval to all choose one quantity. If the competing schedule is chosen so that this one quantity is large enough, the alternative schedule will yield nonnegative expected profits. Based on the results in Rothschild and Stiglitz (1978), it is probably true that an unrestricted Nash equilibrium does not exist for a model with a continuum of types of investors.

A continuum of types of investors is used here so that the inference problem of market makers can be solved easily. I have verified that there is a discrete model with the property that a Nash equilibrium exists as long as the informational asymmetry is not too great, but does not exist when there is too much private information. The restriction

on the allowable price schedules in the continuum model is imposed so
that the results of the continuum model mirror the results of the much
more complex discrete model.

The nonexistence of equilibrium when the informational motive for
trade outweighs the liquidity motive for trade is due to the Nash equi-
librium assumption. The results of Spence (1978) suggest that (in a
discrete model), even when the information asymmetries are large, a
"reactive equilibrium" may exist. Given that market makers must
commit to a schedule (i.e., are not able to back away from unprofitable
trades) I do not find the reactive equilibrium concept attractive for
modeling competing market makers. For example, suppose that a Nash
equilibrium does not exist, but a reactive one does. This reactive equi-
librium involves expected losses on some trades and profits on others.
Suppose that a market maker posts the reactive equilibrium price
schedule. Then the market maker is committed to this schedule and
hence cannot back away from the unprofitable trades if another market
maker posts a price schedule which takes the profitable trades. Conse-
quently, it seems reasonable that such a competing price schedule
would appear and hence the reactive equilibrium price schedule does
not seem reasonable in this environment.

Investors with relatively extreme endowments will be trading for
portfolio rebalancing reasons, while investors with endowments near
zero will be trading for information reasons. Intuitively, one would
expect that, taking an ex post point of view, the market makers, on
average, gain from investors with relatively extreme endowments and
lose, on average, from investors with endowments near zero. This is
verified in the following lemma, the result of which will also be useful
in the subsequent proposition.

LEMMA 2.   A trader's equilibrium, expected profit (equivalently, the
negative of the market makers' profit) conditional on the endowment,
but ex ante to the signal, is given by

$$E(\hat{Q}_c(z)\{X - P_c[\hat{Q}_c(z)]\}|W) = A(1 - \pi_w W^2), \qquad (15)$$

where A is a nonnegative constant. See Appendix for proof.

The fact that the price schedule is upward sloping suggests that the
existence of informed traders imposes a cost on the market. The reduc-
tion in "liquidity" means that there will be inefficient risk sharing. In
fact, if the market closes down there will be no risk sharing. An inter-
esting question at this point is whether or not, ex ante, investors would
vote for a law limiting the extent of informed trading. If informed
trading were outlawed, then competitive market makers would post the
efficient (in terms of risk sharing) price schedule that is independent of
quantity traded.

Before receiving a signal, and without knowing what his endowment
will be when he wishes to trade, an investor does not know whether he

will be able to profit from his information (have an endowment near zero) or end up limiting his trade because of the positive slope of the price schedule. Since the market makers will, on average, break even, one would think that, ex ante, traders would rather have a prohibition on trading based on information. This is shown in the following proposition. The intuition for this result would appear to be somewhat general, and the proof is divorced as much as possible from the specific example considered here.

PROPOSITION 2. Assume that $U'''(\cdot) < 0$, that $e(W) = E[\hat{Q}_c\{X - P_c(\hat{Q}_c)\}|W]$ is concave and symmetric about $W = 0$. Further assume that $f_w(\cdot)$, the density of W, is symmetric. Then, the ex ante utility with asymmetric information (expectation is taken over W,S,X),

$$E[U(-P_c(\hat{Q}_c(Z))\hat{Q}_c(Z) + \{W + \hat{Q}_c(Z)\}X)],$$

is strictly less than the ex ante maximized utility without private information, $E[U(Wm)]$. The Appendix contains the proof for proposition 2.

The proposition is true, not because of resources wasted in the acquisition of information (information is costless in this model) but because of the imperfect risk sharing due to the existence of private information. The result is thus analogous to the results of Hirschleiffer (1971) and Marshall (1974) that private information can frustrate optimal risk-sharing. If information were costly, the result would continue to hold and the interpretation would be similar to that in Diamond (1985). The advantage of reconsidering this result in this framework is that the source of the inefficiency is obvious and (at least in principle) observable. Private information decreases the liquidity of the market, and it is this lack of liquidity that provides the result.

B. Monopolist Specialist Equilibrium

Given that any law restricting information acquisition can only be imperfectly enforced, the next question to ask is whether or not the costs of informed trading are a function of institutional design. In this section it is shown that, in certain circumstances, a specialist with an institutionally protected position can increase the liquidity of the market. Given the past literature on the topic, the environments in which this should hold must be those in which asymmetric information is thought to be a significant problem. Otherwise, the usual inefficiency attributed to a monopolist will hold. A monopolist specialist in an environment with symmetric information will restrict trade, leading to inefficient risk sharing.

It is convenient to recast the investor's choice and the monopolist's maximization problem in the following way. Notice from (9) that, given any price schedule, the investor's optimal trade will be a function of Z and observables, say $\hat{Q}_P(Z)$. The specialist's problem is to find a sched-

ule $P(\cdot)$ that will maximize his expected profits. For a particular schedule, $P(\cdot)$, define the functions $Q(\cdot)$ and $R(\cdot)$ by

$$Q(z) = \hat{Q}_P(\sigma_z z + m),$$

$$R(z) = P(\hat{Q}_P(\sigma_z z + m))\hat{Q}_P(\sigma_z z + m), \ P(\hat{Q}_P(\sigma_z z + m))\hat{Q}_P(\sigma_z z + m) \tag{16}$$

where σ_z^2 is the variance of Z,

$$\sigma_z^2 = [\rho^2 \pi_x + \pi_w \pi_s (\pi_x + \pi_s)]/\pi_w \pi_x \pi_s^2,$$

and z is Z normalized to have mean zero and unit variance. Then the specialist's optimal choice of $P(\cdot)$ amounts to a choice of functions $Q(\cdot)$ and $R(\cdot)$, which maximize his expected profit subject to an incentive compatibility constraint. Formally, his problem is

$$\max E[R(z) - XQ(z)], \tag{17}$$

subject to

$$CE[W,S,Q(z),R(z)] \geq CE[W,S,Q(z'),R(z')],$$

for all z', $z = [S - (\rho W/\pi_s) - m]/\sigma_z$, and $CE(\cdot)$ is given in (7).

The strategy for solving this problem is to first assume that the constraint can be summarized by the local condition for maximization. This leads to functions $R(\cdot)$ and $Q(\cdot)$ that are optimal given the first-order condition. However, this solution is not feasible because there is a discontinuity that implies that the local representation of the constraint in (17) does not, in fact, characterize the constraint. A modification of the solution to the first problem is optimal and feasible, however.

Assuming that $R(\cdot)$ and $Q(\cdot)$ are differentiable with derivatives $R'(\cdot)$ and $Q'(\cdot)$, and defining $D(z)$ by $D(z) = R(z) - mQ(z)$, the local representation of the constraint in (17) is

$$\begin{aligned} D'(z) &= Q'(z)[\pi_s \sigma_z z - \rho Q(z)]/(\pi_s + \pi_x), \\ &= Q'(z)V[z,Q(z)]. \end{aligned} \tag{18}$$

Note that $V(z,Q(z))$ is the marginal valuation of an investor of type z, optimally trading $Q(z)$. It is independent of $D(z)$ due to the absence of wealth effects in exponential utility.

Under the distributional assumptions of the model, z is normally distributed with zero mean and unit variance. Furthermore, X and z are correlated and

$$E[X|z] = m + \pi_z \sigma_z z/(\pi_x + \pi_z) = m + e(z); \tag{19}$$

where π_z is given by $\pi_z = \pi_w \pi_s^2/(\rho^2 + \pi_w \pi_s)$. Let $f(\cdot)$ indicate the standard normal density and let $F(\cdot)$ be the associated distribution

function. Then the maximization problem can be written in terms of the state variables $Q(z)$ and $D(z)$ and the control variable $u(z) = Q'(z)$:[5]

$$\max \int f(z)[D(z) - Q(z)e(z)], \tag{20}$$

subject to

$$Q'(z) = u(z); \quad D'(z) = u(z)V[z,Q(z)].$$

The proof of proposition 3, in the Appendix, establishes that the solution to this problem is

$$Q(z) = (\pi_s\sigma_z/\rho)(\alpha z - \{[1 - F(z)]/f(z)\}) \quad \text{for } z > 0$$
$$= (\pi_s\sigma_z/\rho)[\alpha z + F(z)/f(z)] \quad \text{for } z < 0, \tag{21}$$
$$\alpha = \rho^2\pi_x/[\rho^2\pi_x + \pi_w\pi_s(\pi_x + \pi_s)].$$

While $Q(\cdot)$ given in (21) has a positive derivative almost everywhere, at $z = 0$ there is a discontinuity:

$$Q_+(0) = -\pi_s/[2\rho f(0)] < \pi_s[2\rho f(0)] = Q_-(0).$$

It can be verified that, in this case, (18) does not completely characterize the constraint in (17). In fact, if $z > 0$ is the true characterization of an arriving investor and $Q(z) < 0$, then the investor will prefer to act as if $-z$ were his characteristic. However, $Q^*(\cdot)$ given by

$$Q^*(z) = Q(z) \quad \text{for } |z| > z^*$$
$$= 0 \quad \text{otherwise,} \tag{22}$$

z^* satisfying $z^* > 0$ and $Q(z^*) = 0$, is optimal for the specialist and does satisfy the constraint in (17). The excess revenue function $D(z)$ can be found by integrating (18) to get

$$D(z) = Q(z)\{\pi_s\sigma_z z/(\pi_x + \pi_s) - \rho Q(z)/[2(\pi_x + \pi_s)]$$
$$- \pi_s\sigma_z/(\pi_x + \pi_s)\int Q(t)dt/Q(z)\}, \tag{23}$$

where the integral is from z^* to $|z|$.

PROPOSITION 3. When the arriving investor has signal S and endowment W, then the equilibrium trade will be $Q^*(z)$ at price $R(z)/Q(z)$, where z is given by $z = [S - m - (\rho W/\pi_s)]/\sigma_z$, and $Q^*(z)$ is given by (21) and (22), and $R(z) = D(z) + mQ^*(z)$, where $D(z)$ is given by (23). See Appendix for proof.

The first thing to note from proposition 3 is that the monopolist will keep the market open as long as z^* is finite; that is, α is positive. Given the results of the analysis of competitive market makers, this might suggest that there are trades for which the specialist will expect to lose

---

5. The maximization in (21) is correct as long as $Q(z)$ is monotone and strictly monotonic on any interval where $Q(z)$ is nonzero. This will be shown to be the case.

money. In fact, the specialist will always lose money on extreme trades. This can be seen by considering the behavior of $R(z)$ for large z.[6] For large z, $[1 - F(z)]/f(z)$ is near zero and hence, for large z, $Q^*(z)$ is approximately $\pi_s\sigma_z\alpha z/\rho$. Furthermore,

$$\int Q^*(t)/Q^*(z) = \frac{\{.5z - .5z^{*2}/z - (1/\alpha z)\int[1 - F(t)]/f(t)\}}{\{1 - [1 - F(z)]/[\alpha zf(z)]\}}$$

is approximately $z/2$ for large z. Thus, for large z, $R(z)/Q(z)$ is approximately

$$m + [\pi_s\sigma_z(1 - \alpha)z]/2(\pi_x + \pi_s)$$

$$= m + .5(\sigma_z\pi_w\pi_s^2z)/[\rho^2\pi_x + \pi_w\pi_s(\pi_x + \pi_s)].$$

In contrast, $E[X|z]$ is given by

$$E[X|z] = m + (\pi_w\pi_s^2\sigma_zz)/(\rho^2\pi_x + \pi_w\pi_s\pi_x + \pi_w\pi_s^2),$$

and hence for large z, $R(z)/Q(z) < E[X|z]$, and the specialist loses money on large trades.

Extreme z's are relatively unlikely, and hence the specialist still expects to profit on average. In fact, the specialist's expected profit is given by

$$\{[\rho^2\pi_x + \pi_w\pi_s(\pi_x + \pi_s)]/\pi_w\pi_x(\pi_x + \pi_s)\} \int_{z^*}^{\infty} f(t)\{\alpha t - [1 - F(t)]/f(t)\}^2.$$

That the specialist chooses to make expected losses on extreme trades, while perhaps surprising, can be understood by examination of the proof of proposition 3. The expected benefit of marginally increasing the quantity traded by a trader of type z is the marginal revenue at that z times the density at z less the expected reduction in the marginal valuations of traders with types larger than z. The expected marginal cost is the conditional expected value of X times the density at z. At large z's, the probability of an investor arriving with a higher type is small, and hence the marginal benefit at z is approximately the marginal revenue at z. Since the conditional expected value is increasing in z, for large z's the marginal revenue must be increasing. But this implies that for large z's the marginal revenue is above the price, and, hence, the price is below the conditional expected value; that is, the specialist optimally expects to make negative profits on large trades.

Another way to see the intuition is to consider an alternative strategy for the specialist. Suppose the specialist were to eliminate all the non-profitable trades by specifying an arbitrarily high price for quantities above some level. This would cause all the high z traders to pool at lower quantities, reducing the profitability of the low-quantity trades.

6. Large positive z will be considered, z negative and large in absolute value will lead to identical results since it is easy to check that $Q^*(z) = -Q^*(-z)$ and $D(-z) = -D(z)$.

Since the specialist is operating in the elastic part of the demand curve, the reduction in the expected revenue accompanying this pooling is greater than the reduction in the expected cost, and the expected profit is lower.

Since $Q^*(z)$ is incentive compatible, it must be that $Q^*(z)$ is nondecreasing. It was shown in the proof of proposition 3 that for $z > z^*$, $Q^{*\prime}(z) > 0$ and $Q^{*\prime\prime}(z) < 0$. Since $Q^{*\prime}(z) > 0$ for $z > z^*$, there is a function $Q^{-1}(\cdot)$ such that if $Q^*(z) = q$, $Q^{-1}(q) = z$. Thus, we can define the price function $P_m(q)$ analogous to the competitive price schedule by

$$P_m(q) = R[Q^{-1}(q)]/q.$$

There is a discontinuity in $P_m(\cdot)$ at $q = 0$. The limit as q decreases to zero is $m + z^*\pi_s\sigma_z/(\pi_x + \pi_s)$, whereas the limit from below is $m - z^*\pi_s\sigma_z/(\pi_x + \pi_s)$. The difference, $2z^*\pi_s\sigma_z/(\pi_x + \pi_s)$, can be termed the "zero-quantity spread." This spread is of interest because it is the calculable quantity in this model that is closest to the observed quoted spread. This spread will depend upon π_s, a measure of how much private information there is, but the spread is positive no matter what π_s is, including $\pi_s = 0$. Interestingly, the zero-quantity spread is not, in general, a monotone function of π_s. This suggests that examining quoted spreads for evidence of insider trading may be misguided. What is required is an examination of how much large trades move the price. Using the results above for the behavior of $R(z)$ and $Q^*(z)$ for z large, it is easy to see that $P_m(q)$ for q large is approximately $m + q(\pi_w\pi_s/\rho\pi_x)$, which is clearly increasing in π_s.

II. Monopolist/Competitor Comparison

The results of the analysis of the competing market-makers model and the monopolist-specialist model lead immediately to the central welfare result: in some environments, the monopolist-specialist system will be weakly preferred by all traders and strictly preferred by some traders. The set of environments for which this is true is the set of environments for which $\alpha = \rho^2\pi_x/[\rho^2\pi_x + \pi_s\pi_w(\pi_x + \pi_s)] \leq .5$. If $\alpha \leq .5$, then the competing market-maker market will not open, whereas the specialist will not suspend trading. Since the zero trade is feasible in either market, traders who trade a nonzero amount with the specialist will strictly prefer that there be a specialist market. This proves the following proposition.

PROPOSITION 4. If $\alpha \leq .5$, then all potential traders weakly prefer the monopolist-specialist system to competing market makers, and some potential traders strictly prefer the monopolist specialist.

As the analysis of competing market makers suggested, the parameter α essentially measures the severity of the adverse selection problem. When α is one, there is no private information, whereas when α is

less than or equal to .5, the adverse selection problem is so severe that the competitive market closes down. Of course, the monopolist specialist will not dominate in all environments. In particular, if there is no private information, then everyone would prefer competing market makers. This can be seen by noting that when there is no private information, competitive market makers will set a zero spread, whereas the monopolist will set a positive spread. Thus, in the competitive case, there will be complete risk sharing, whereas it will be limited in the monopolist-specialist case.

The next proposition shows that after a reparameterization of the model, there is an $\hat{\alpha} > .5$ such that, ex ante, the monopolist specialist is preferred when $\alpha < \hat{\alpha}$, whereas competing market makers are preferred if $\alpha > \hat{\alpha}$. This verifies the intuition that the superiority of the monopolist position derives solely from informational problems. If those problems are not great then the monopolist-specialist system is inferior to a system with competing market makers.

PROPOSITION 5. Let V_m denote the ex ante expected utility of a trader when there is a monopolist specialist, and let V_c denote the ex ante expected utility when there are competing market makers. There is a reparameterization of the model with the parameters π_w, π_x, π_s, and ρ replaced by independent parameters γ_i, $i = 1 - 3$ and $\alpha = \rho^2 \pi_x / N$, $N = \rho^2 \pi_x + \pi_w \pi_s (\pi_x + \pi_s)$. Consider changes in α keeping the γ's constant. Then there is an $\alpha > .5$ such that $V_m > V_c$ when $\alpha < \hat{\alpha}$ and $V_m > V_c$ when $\alpha > \hat{\alpha}$. A sketch of the proof for this proposition is provided in the Appendix.

III. Concluding Remarks

The focus of this article has been the inefficiencies created by allowing trading on private information and the institutional response to such inefficiencies. Trading on private information creates inefficiencies because it leads to less than optimal risk sharing. This occurs because the response of market makers to the existence of traders with private information is to reduce the liquidity of the market. This reduction in the liquidity of the market reduces the amount of trade and hence the amount of risk sharing.

The institution of the monopolist specialist may ease this inefficiency somewhat by increasing the liquidity of the market. Since the monopolist can average his profits, he need not make a profit on every trade. In fact, it is shown in this example that optimally he will not. This implies a more liquid market as long as there is extensive trading on private information. If there is not, then the result is that the monopolist specialist does worse from a welfare perspective than do competing market makers.

If this analysis is to be believed as an explanation for the existence of the specialist system, there should be some evidence that trading on

private information is a significant problem. Glosten and Harris (1988) present evidence that part of the spread can be attributed to private information. One can also interpret trading suspensions as a reaction to the existence of private information. Finally, the Securities and Exchange Commission (SEC) has been successful in prosecuting a number of insider trading cases.

If this analysis is correct, one of the sources of institutional design differences must be the extent of asymmetric information. In the case of options and futures markets, one can make a reasonable argument that asymmetric information is not a large problem. Trading in financial futures is probably not substantially motivated by speculation on private information. While there may be private information about future commodity spot prices, the individuals likely to have that information are the producers of the commodities, and it is probably true that their insurance motive is large enough to swamp any information-motivated trade.

For options markets, the argument is not so obvious, particularly because it is often asserted that traders with private information make extensive use of the options market. However, given that there is a specialist in the stock market, there may be no gain to having a monopolist specialist in the option market. Market makers on the options exchanges can typically hedge using the underlying stock. Thus, if a trader has information that a stock is undervalued, he may buy the undervalued call from a market maker. If the option market maker can buy the (undervalued) stock before word of the option transaction reaches the floor of the stock exchange, then the risk that the trade was information motivated can be transferred from the option market maker to the stock market maker.[7]

Of course, the specialist can also use options to hedge in the other direction. The point of this discussion is to argue that given a monopolist in one market, there is no gain and a probable loss to having a monopolist in the other market. On the other hand, if both markets had competitive market makers then, according to the above model there would be times in which both markets would shut down. In such an instance there may be a gain to having a monopolist in one of the two markets.

The larger task is to explain the coexistence of the over-the-counter (OTC) market, with competing dealers, and the NYSE, with the specialist system. At first glance it might be thought that asymmetric information would be more of a problem for the smaller firms traded on the

---

7. Of course, following the analysis of this model, the price may move against the option market maker making the hedging trade, and rational option market maker will take this into account in his quotes. However, both a monopolist and competing market maker can do this equally well, so there is no benefit to having a monopolist market maker in the options market.

OTC market. However, it may be that insider trading is not the most serious source of informational asymmetries. Given that market makers do not typically spend time analyzing the security they deal with and given the monitoring of insider trading by the SEC, superior analytical ability may be the larger day-to-day source of private information. In this case, it is not obvious that the smaller stocks would be more prone to information-motivated trade. This is particularly true given that the dispersion of ownership is probably less with the OTC stocks, and hence the gains to doing security analysis may be smaller.

No doubt, one reason that the specialist system began was historical accident, and this article does not claim to supply the only reason for its existence. Certainly, current owners of seats on the exchange might object to a change in the system and hence any change could be costly. What the analysis in this article suggests is that changing is not obviously socially optimal even if the changes were not costly.

There are a number of areas of research suggested by the results and the limitations of this analysis. It has been argued (see Manne 1966) that there are benefits to allowing informed trading, and the result of proposition 1 completely ignores these. What the proposition says is that these benefits, if any, are bought at a cost—the liquidity of the market suffers. This suggests that a general equilibrium analysis that considers both the liquidity issues as well as the benefits of insider trading might do much to answer the question of what restraints should be placed on informed trading. A hypothesis is that trading on inside information that will be revealed in a short period of time hurts the liquidity of the market without a corresponding increase in the efficiency of firms' productive decisions. On the other hand, information due to superior analysis may increase the productive efficiency of firms by making stock prices reflect more of the information available to the firm. The increase in efficiency may be worth the concomitant decrease in the liquidity of the market.

This article takes a fairly simple view of the role of the specialist as a dealer and ignores the brokerage function. In fact, limit order submitters may be a nontrivial source of competition for the specialist. The extent to which limit order submitters and floor traders compete with the specialist and what effect that has on the results here would be of interest. Along these same lines, it was argued in the initial discussion of the model that knowledge of the specialist's book conferred knowledge about trading uncertainty, not knowledge of future prices of the stock. This depends on the assumption that informed traders do not submit limit orders but rather use market orders. The order-placement strategy of informed traders would be another interesting line of research.

The model in this article does not consider the dynamic order-placement strategies of investors. The intuition for the results of this article suggests that ignoring this feature would not have a substantial effect

on the general result, but it certainly does have an effect on the specific results. The simultaneous consideration of the optimal order strategy as well as the optimal dynamic strategy of the specialist would be interesting.

With a more detailed analysis of the strategies open to market participants, it would be possible to consider the question of optimal market design. While the analysis presented here suggests a feature that an optimal institution would exhibit (averaging of profits across trades) it is premature to use this model to address the issue of optimality.

To sum up, this article has two major results. First, informed trading imposes a welfare cost in that it reduces the liquidity of the market. Second, the existence of the specialist system may be in part explained by this cost. An unregulated specialist may provide a more liquid market than competing market makers, and hence some of the welfare loss due to informed trading may be negated. In analyzing these questions, a model was developed that may hold some methodological interest. In particular it was shown that examination of the small-trade spread may not supply much information about the perceived presence of informed traders.

Appendix

Proof of Lemma 1

Substituting for Z from (9) into (3) yields

$$P(Q) = \frac{\pi_x m}{\pi_x + \pi_z} + \frac{\pi_z}{\pi_x + \pi_z} \left\{ [P'(Q)Q + P(Q)] \frac{(\pi_x + \pi_s)}{\pi_s} - \frac{(\pi_x m - \rho Q)}{\pi_s} \right\}.$$

Define $g(Q)$ by

$$g(Q) = P(Q) - m - Q\rho\pi_w\pi_s/D,$$

where

$$D = \rho^2\pi_x - \pi_w\pi_s(\pi_x + \pi_s) \quad \text{if } D \neq 0.$$

After substituting for

$$\pi_z = \pi_w\pi_s^2/(\rho^2 + \pi_w\pi_s),$$

we have

$$\rho^2\pi_x/Q = \pi_w\pi_s(\pi_x + \pi_s)g'(Q)/g(Q).$$

For $Q > 0$,

$$g(Q) = KQ^\gamma, \ \gamma = \rho^2\pi_x/[\pi_w\pi_s(\pi_x + \pi_s)],$$

with K unrestricted. A similar analysis holds for $Q < 0$. If $D = 0$, then the expression above can be written

$$(\pi_x + \pi_s)[g'(Q)Q - g(Q)]/Q^2 + \rho/Q = 0; \quad g(Q) = P(Q) - m.$$

This has the solution given in the statement of the lemma. Q.E.D.

Proof of Proposition 1

First, the uniqueness of the linear price function is shown when $D > 0$. Assume that a pricing schedule is given as in lemma 1 with $K > 0$. Now consider a continuous, differentiable price schedule that is equal to this price function for $Q \leq 1$, and linear for $Q > 1$. To ensure differentiability make the slope of the linear portion equal to the slope of the original schedule at $Q = 1$. Evaluation of the expected profit will show that this second price schedule will yield positive profits for trades larger than 1. On the other hand, K cannot be taken to be negative, for then the second-order condition will be violated. Thus, the only undominated pricing schedule when $D > 0$ is linear.

Suppose $D < 0$. Then the derivatives of the solution given in lemma 1 are

$$P'(Q) = (\rho\pi_w\pi_s/D) + K\gamma|Q|^{\gamma - 1},$$

$$P''(Q)Q = K\gamma(\gamma - 1)|Q|^{\gamma - 1}.$$

Note that, if $D < 0$, $\gamma < 1$. The second-order condition is

$$|Q|^{\gamma - 1} > \{\rho[\rho^2\pi_x + \pi_w\pi_s(\pi_x + \pi_s)]\}/[K\gamma(\gamma + 1)(-D)]. \tag{A1}$$

Since $D < 0$, and $\gamma < 1$, the second-order condition is satisfied for $-Q' < Q < Q'$ where Q' satisfies (A1) with equality. Consider the expected profit to the market makers on trades of Q'. A trader will trade Q' if Z is larger than some Z'. But then, $E[X|Q'] > E[X|Z'] - P(Q')$, and the market makers' profit is negative on trades of Q contrary to the hypothesis. Thus, if $D < 0$, there is no schedule where the market maker breaks even. The analysis of $D = 0$ is similar. Q.E.D.

Proof of Lemma 2

If the market closes down ($D \leq 0$), then $\hat{Q}_c = 0$ and $A = 0$ in the statement of the lemma. If $D > 0$, then the result is obtained by a straightforward analysis of the expressions in proposition 1, assuming that S and W are independent and X and W are independent. Q.E.D.

Proof of Proposition 2

First note that if there is no private information (i.e., $\pi_s = 0$), competing market makers will set $P(Q) = m$ for all Q. In that case, the optimal trade is $-W$ and the ex ante expected utility is $E[U(Wm)]$. Assume this expectation exists. When there is private information, the optimal trade is $\hat{Q}$ and the price schedule is $P_c(\cdot)$. The realized wealth of a typical trader is $Wm + W(X - m) + \hat{Q}[X - P_c(\hat{Q})]$, and the ex ante utility is $E[U(Wm + W[X - m] + \hat{Q}[X - P_c(\hat{Q})])]$. This expectation may not exist. If it does not, the proof is complete. Suppose it does. A sufficient condition for the ex ante utility without private information to exceed the ex ante utility with private information is

$$E(U'(Wm)\{W(X - m) + \hat{Q}[X - P_c(\hat{Q})]\}) \leq 0.$$

Given the independence of X and W, this can be rewritten $E[U'(Wm)e(W)] \leq 0$. Note that

$$E[e(W)] = E[\hat{Q}\{X - P_c(\hat{Q})\}] = E[\hat{Q}\{E[X|\hat{Q}] - P_c(\hat{Q})\}] = 0.$$

Since $e(W)$ is presumed to be symmetric and concave and since $E[e(W)] = 0$, there is w' such that $e(w) < 0$ for $|w| > w'$ and $e(w) > 0$ for $|w| < w'$. Letting $f_w(\cdot)$

be the density of W and assuming $f_w(\cdot)$ is symmetric implies

$$E[U'(Wm)e(W)] = \int_{-\infty}^{-w'} e(w)f_w(w)[U'(wm) + U'(-wm)] +$$

$$\int_{-w'}^{0} e(w)f(w)[U'(wm) + U'(-wm)].$$

Since $U'''(\cdot) \geq 0$, $U''(x)$ is increasing in x, and hence $U'(wm) + U'(-wm)$ is decreasing in w. Thus, for $w \leq -w'$, $U'(wm) + U'(-wm) \geq U'(w'm) + U'(-w'm)$ and for $-w' < w \leq 0$, $U'(wm) + U'(-wm) \leq U'(w'm) + U'(-w'm)$, and hence

$$E[U'(Wm)e(W)] \leq [U'(w'm) + U'(-w'm)]\int_{-\infty}^{0} e(w)f_w(w) = 0.$$

Q.E.D.

Proof of Proposition 3

The Hamiltonian associated with (20) is—$Q(z)$ and $D(z)$ are state variables, $Q'(z) = u(z)$ is the control—

$$f(z)[D(z) - e(z)Q(z)] + \lambda_1(z)u(z) + \lambda_2(z)u(z)V[z,Q(z)].$$

The conditions the maximum must satisfy are

    i. $\lambda_1(z) + \lambda_2(z)V[z,Q(z)] = 0$;

    ii. $-\lambda_1'(z) = -e(z)f(z) + \lambda_2(z)u(z)V_2[z,Q(z)]$;

    iii. $-\lambda_2'(z) = f(z)$.

The transversality conditions imply $\lambda_2(z) = 1 - F(z)$ for z positive and $-F(z)$ for z negative. Differentiating condition (i) after substituting for $\lambda_2(z)$ and using (ii),

$$f(z)V[z,Q(z)] - [1 - F(z)]V_1[z,Q(z)] = e(z)f(z).$$

Substituting for $V[z,Q(z)]$ yields $Q(z) = (\pi_s\sigma_z/\rho) \{\alpha z - [1 - F(z)]/f(z)\}$ for z positive. The procedure for z negative is similar. This $Q(z)$ is not monotonic. However, if we add the constraint $z[Q(z) - Q(0)] \geq 0$, the resulting $Q(z)$ is monotonic. The symmetry of the problem implies that $Q(0) = 0$. Q.E.D.

Sketch of the Proof of Proposition 5

Since all functions of z are symmetric around 0, the discussion will proceed with $z > 0$. The quantities traded in the competitive and monopoly cases are of the form

$$Q(z) = (\pi_s\sigma_x/\rho)q(z),$$

where the competitive and monopolist $q's$, q_c and q_m, respectively, are given by $q_c(z) = (2\alpha - 1)z$, $q_m(z) = \alpha z - [1 - F(z)]/f(z)$ for $z > z^*$, 0 otherwise. The certainty equivalent is given by:

$$CE_i = wm + \pi_s\sigma_z wz/(\pi_x + \pi_s) +$$

$$.5\rho w^2/(\pi_x + \pi_s) + \pi_s^2\sigma_z^2/[\rho(\pi_x + \pi_s)]\int_0^z q_i(t)dt,$$

where $i = c,m$. Denote the coefficient on the integral by γ_1/ρ and the integral by $G_i(z,\alpha)$. Let CE_c and CE_m denote, respectively, the equilibrium certainty equivalents for the competitive and monopoly cases. Our intention is to evaluate.

$$E[\exp(-\rho CE_m) - \exp(-\rho CE_c)],$$

the gain in utility from using the competitive system. Note that the expectation is taken over both z and w. We know that for $\alpha = .5$ this gain is negative, and for $\alpha = 1$ it is positive. By continuity there is an $\alpha > .5$ such that the difference in expected utility is 0.

After some tedious calculations it can be verified that

$$E[\exp(-\rho CE_i)|z] = K\exp(-\gamma_0 z - \gamma_2 z^2 - \gamma_1 G_i(z,\alpha)) = KH(z,\gamma,\alpha),$$

where K depends on all the parameters, and the γ's can be chosen independent of α. The difference in expected utilities can thus be written

$$KE[H(z,\gamma,\alpha)\exp\{\gamma_1[G_c(z,\hat{\alpha}) - G_c(z,\alpha)]\} (\exp\{\gamma_1[G_c(z,\alpha) - G_m(z,\alpha)]\} - 1)].$$

Note that $G_c(z,\alpha) - G_m(z,\alpha) = \Delta(z,\alpha)$ is increasing in α. Furthermore, there is a $t(\alpha)$ such that for $z < t(\alpha)$, $\Delta(z,\alpha)$ is positive while it is negative for $z > t(\alpha)$. Also note that $G_c(z,\hat{\alpha}) - G_c(z,\alpha) = (\hat{\alpha} - \alpha)z^2$. For $\alpha > \hat{\alpha}$, $\Delta(z,\alpha) > \Delta(z,\hat{\alpha})$. By integrating separately over the regions in which $\Delta(z,\hat{\alpha})$ is positive $[z < t(\hat{\alpha})]$ and negative $[z > t(\hat{\alpha})]$, one can verify that the difference in utilities is at least as big as

$$KE\{H(z,\gamma,\alpha)\exp[\gamma_1(\hat{\alpha} - \alpha)t(\hat{\alpha})^2](\exp[\gamma_1\Delta(z,\hat{\alpha})] - 1)\} = 0.$$

A similar analysis holds for $\alpha < \hat{\alpha}$. Q.E.D.

References

Altug, Sumru. 1985. The effect of insider trading by a dominant trader in a simple securities market model. Working paper. Minneapolis: University of Minnesota, Center for Economic Research.

Diamond, Douglas. 1985. Optimal release of information by firms. *Journal of Finance* 40 (September): 1071–94.

Easely, David, and O'Hara, Maureen. 1987. Price, trade size and information in securities markets. *Journal of Financial Economics* 19 (September): 69–90.

Gammill, James F., Jr. 1986. Financial market design when traders have private information. New York: Columbia University, Graduate School of Business, Center for the Study of Futures Markets.

Glosten, Lawrence R., and Harris, Lawrence E. 1988. Estimating the components of the bid/ask spread. *Journal of Financial Economics* 21 (May): 123–42.

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

Gould, John P., and Verrecchia, Robert E. 1985. The specialist as economic agent. *Journal of Political Economy* 93 (February): 66–83.

Hagerty, Kathleen. 1986. Equilibrium bid-ask spreads in markets with multiple assets. Working paper. Evanston, Ill.: Northwestern University, J. L. Kellogg Graduate School of Management.

Hirschleifer, Jack. 1971. The private and social value of information and the reward to inventive activity. *American Economic Review* 61 (September): 561–74.

Ho, Thomas, and Stoll, Hans R. 1981. Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics* 9 (March): 47–73.

Ketterrer, Juan. 1987. Market games with inside information. Working paper. Minneapolis, Minn.: University of Minnesota, Economics Department.

Kyle, Albert. 1985. Continuous auctions and insider trading. *Econometrica* 53 (November): 1315–35.

Leffler, George L., and Farwell, Loring C. 1963. *The Stock Market.* 3d ed. New York: Ronald Press.

Madhavan, Ananth N. 1987. Trading mechanisms in securities markets. Working paper. Ithaca, N.Y.: Cornell University.

Manne, H. 1966. *Insider Trading and the Stock Market.* New York: Free Press.

Marshall, John M. 1974. Private incentives and public information. *American Economic Review* 64 (June): 373–90.

Rothschild, M., and Stiglitz, J. 1978. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. In P. Diamond and M. Rothschild (eds.), *Uncertainty in Economics.* New York: Academic Press.

Spence, Michael. 1978. Product differentiation and performance in insurance markets. *Journal of Public Economics* 10 (December): 427–47.

Stiglitz, J. 1977. Monopoly, non-linear pricing and imperfect information: The insurance market. *Review of Economic Studies* 44 (October) 407–30.

Stoll, Hans R. 1984. Alternative views of market-making. Working paper. Nashville, Tenn.: Vanderbilt University, Owen Graduate School of Management.