Is the Electronic Open Limit Order Book Inevitable?
Author(s): Lawrence R. Glosten
Source: *The Journal of Finance*, Vol. 49, No. 4 (Sep., 1994), pp. 1127-1161
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/2329182
Accessed: 24/01/2011 20:02

THE JOURNAL OF FINANCE • VOL. XLIX, NO. 4 • SEPTEMBER 1994

Is the Electronic Open Limit Order Book Inevitable?

LAWRENCE R. GLOSTEN*

ABSTRACT

Under fairly general conditions, the article derives the equilibrium price schedule determined by the bids and offers in an open limit order book. The analysis shows: (1) the order book has a small-trade positive bid-ask spread, and limit orders profit from small trades; (2) the electronic exchange provides as much liquidity as possible in extreme situations; (3) the limit order book does not invite competition from third market dealers, while other trading institutions do; (4) If an entering exchange earns nonnegative trading profits, the consolidated price schedule matches the limit order book price schedule.

THIS ARTICLE PROVIDES AN analysis of an idealized electronic open limit order book. The focus of the article is the nature of equilibrium in such a market and how an open limit order book fares against competition from other methods of exchanging securities. The analysis suggests that an electronic open limit order book mimics competition among anonymous exchanges. As a result, there is no incentive to set up a competing anonymous dealer market. On the other hand, any other anonymous exchange will invite "third market" competition. These conclusions suggest that an electronic open limit order book of the sort considered here has a chance of being a center of significant trading volume. The analysis does not imply that an electronic limit order book will be, or should be the only trading institution. It does suggest some of the characteristics that an alternative institution should have in order to successfully compete with an electronic exchange. The results are obtained in a fairly general environment, and hence would appear to be robust.

The motivation for the article lies in recent developments in information processing technology, the interest in institutional innovation in the securities industry, and the uncertainty about future developments in trading

* Columbia University. Former versions of this article were immodestly titled "The Inevitability and Resilience of an Electronic Open Limit Order Book" and then too modestly titled "Equilibrium in an Electronic Open Limit Order Book." I have benefitted from the insights of Fischer Black, Puneet Handa, Pete Kyle, Bruce Lehmann, Matt Spiegel, Subra Subramanyam, and the comments of seminar participants at Baruch, Rutgers, New York University, the Atlanta Fed, University of Michigan, Northwestern University, University of Chicago, and Ohio State. Part of this research was done as a Visiting Economist at the New York Stock Exchange. The comments, opinions, and errors are those of the author only. In particular, the views expressed here do not necessarily reflect those of the directors, members or officers of the New York Stock Exchange, Inc.

institutions. Such systems as INSTINET, the "Wunsch Auction," and electronic trading on the regional exchanges represent different approaches to the use of information processing technology. The results in this article are indicative of the direction such developments might take. The analysis suggests that the open limit order book is a stable institution and, within the set of economic environments and trading structures considered, the only stable institution.

The model assumes away a number of frictions and costs that may well be important. The model deals with the architecture of the open limit order book only in general terms and does not address a host of potentially important technological issues—computing capacity, trade execution speed, display technology and clearing to name a few.[1] Certain other limitations will be discussed below in the concluding remarks.

There are a number of important antecedents to this work. Trading on private information is an important aspect of the analysis—without it, all of the propositions become trivial. As in Kyle (1985), investors may submit orders of any quantity, but, in contrast, orders arrive one at a time as in Glosten and Milgrom (1985). This combination of features recalls Easley and O'Hara (1987) and Glosten (1989). The design of the trading mechanism is, however, different from both of these models, and the environment is more general.

The model of the open limit order book and the specification of equilibrium are very similar to the limit order book analysis in Rock (1989). The most important difference is that the model here does not allow a specialist or market maker to disrupt trading against the book. A key feature of the Rock (1989) model is that a market maker can foist a second adverse selection problem onto those providing bids and offers—the book is only hit if the market maker decides to back away (because of order size) from a trade. A second difference is that the quantities traded in the Rock model are exogenous, whereas they are determined endogenously in this article. This allows an analysis of market breakdown and is very important for the analysis of competing exchanges.

The equilibrium of the model in this article is similar to the one of the model in Gale (1991). In that article, informed "hedgers" have the opportunity to trade more than once. In the equilibrium, large traders with extreme news trade twice, while small traders with less extreme news trade once. The two prices at which a large buyer buys are precisely the first and second lowest offers that would prevail in the open limit order book considered here.

The discussion in Black (1992) and its predecessors was a major inspiration for this analysis. In an earlier version (Black (1991)) an institution was developed that used taxes and subsidies to break the equivalence of the net price paid or received and revised expectations in response to a trade. This

[1] Harris (1990) provides an analysis of some of these and other issues. Also, see Domowitz (1991).

article shows that a similar structure of implicit taxes and subsidies can arise in the equilibrium considered here. In all versions, the "Black Market" requires an exchange official to set the terms of trade. In the market considered here, competing individuals determine the terms of trade.

The electronic open limit order book is modelled as a publicly visible screen providing bids and offers, each of which specify a price and a quantity. Transactions against the book pick off the limit orders at their limit prices. These market orders are presumed to be the result of rational optimization on the part of risk-averse and possibly informed investors, while bids and offers are assumed to reflect this. The source of bids and offers is a large population of (essentially) risk-neutral "patient traders." The large population and risk neutrality imply that equilibrium is characterized by a zero expected profit condition.

After setting up the economic environment, and analyzing the trades of investors who trade against the book of limit orders, the article presents an analysis of the bids and offers that will be provided. In an environment with discrete prices, the bids and offers submitted are seen to be related to, respectively, "lower tail" and "upper tail" conditional expectations. This is due to the "discriminatory" nature of the book—limit orders are picked off in succession. The possibility of information-motivated trade, as formulated here, implies that the schedule of offers is generally upward sloping—it costs more per share to purchase a large number of shares than to purchase a small number of shares. Furthermore, there is a positive small-trade bid-ask spread.

The open limit order book does as well as can be hoped at handling extreme adverse selection problems—if no liquidity is supplied by the open limit order book, then every other anonymous exchange would expect to lose money by staying open for trade. The reason for this is that the architecture of the open limit order book leads to an averaging of profits across trades—a feature shared with a monopolist specialist architecture.

The next propositions show that the open limit order book is uniquely immune to competing exchange "cream skimming" of orders when the only way to ascertain "cream" is with trade size—i.e. competing exchanges are anonymous. The key assumption here is that investors can costlessly split their orders among competing exchanges. The discriminatory design of the open limit order book implies that the book breaks up a trade into many smaller transactions (each at the lowest (highest) offers (bids)), and furthermore, the profits from such a breakup are competed away. Thus, a competing exchange cannot profitably allow investors to break up their order further. That is, the discriminatory limit order book mimics the competition among exchanges.

The subsequent sections defend the above assertions with a more rigorous analysis. Section I analyzes the equilibrium at a point in time by first examining the behavior of market order users (Section I.*A*), and then the behavior of limit order users (Section I.*B*). Section II explores some implications of the equilibrium, and Section III examines intermarket competition. Section IV discusses dynamic issues. Section V identifies and discusses limitations of the results and points to further analysis. Section VI concludes. All proofs are contained in the Appendix.

I. Equilibrium in the Electronic Market

It is assumed that all potential participants in the market have access to an electronic screen that provides an anonymous list of all limit orders, buy and sell, that have been entered. If an individual wishes to add a bid or offer to the market, this can be done costlessly. Furthermore, any bid or offer may be costlessly retracted at any time, except in the middle of the execution of a trade. Execution of a trade against the book occurs in a "discriminatory" fashion. That is, if a trade is large enough to execute against several limit orders at different prices, each limit order transacts at its limit price. For example, if there were two offers at 50 for 1,000 shares of each, and two offers at 51, each for 1,000 shares, a 4,000-share purchase would in effect lead to four transactions—two at 50 and two at 51. The marginal price for this 4,000-share trade would be 51, while the average price would be 50.5.[2]

Four assumptions are made to restrict the behavior of participants: 1) investors who trade against the book are rational and risk averse in that they choose their trade to maximize a quasi-concave function of their cash and share position; 2) there is the possibility of informed trade in that an investor's marginal valuation is affiliated with the future payoff of the security; 3) there are a large number of risk-neutral limit order submitters; 4) in the presence of more than one exchange, investors can costlessly and simultaneously split their orders among the exchanges.

The analysis takes place at a point in time. Though some expectations and probabilities are written as unconditional, they should be understood to be conditional on all past public information. Similarly, conditional probabilities and expectations should be understood to be conditional on the specific argument, as well as on all past information. The analysis thus looks at (i) the terms of trade provided conditional on all past public information; (ii) the trade made in response to these terms, conditional on all past information and possibly some private information; and (iii) subsequent revisions in expectations in response to this trade. After the trade, a new public information set is determined—the original public information, new public information, plus the trade that occurred. At that point, new terms of trade are determined in the same manner.

[2] One could also imagine a nondiscriminatory electronic limit order book. Analogous to a nondiscriminatory auction, a nondiscriminatory order book would transact all limit orders at the same price. There are reasons for considering the nondiscriminatory book, and these will be discussed below.

A. Investor Behavior

Bids and offers are submitted without knowing what the next arriving order will be. The next trader to come to market chooses the trade based on his or her privately known but generally unobservable characteristics—preferences, information, portfolio position, etc. The analysis uses the notation, ω, to indicate this vector of unobservable characteristics.

The terms of trade are determined by the list of bids and offers available. The schedule of bids and offers is denoted by the function $R'(q)$. For q positive (an investor purchase), $R'(q)$ is the ask price paid for the last share in a purchase of q shares. For q negative (an investor sale), $R'(q)$ is the bid price received for the last share in a sale of $-q$ shares. The "prime" notation is used to remind the reader that $R'(q)$ is a *marginal* price. For any q, $R(q)$ is defined to be the (Lebesgue) integral of $R'(\cdot)$ from zero to q. Thus (if all prices are positive), if q is positive, $R(q)$ is positive and represents the total amount paid for a purchase of q shares. If q is negative,$R(q)$ is negative and $-R(q)$ is the amount received for a sale of $-q$ shares.[3]

With this notation, the following assumption regarding investor behavior is offered.

ASSUMPTION 1: *An arriving investor with a vector of characteristics, ω, facing a schedule of bids and offers described by the function $R'(\cdot)$, chooses a quantity to trade, q, to maximize $W(-R(q), q; \omega)$. The function $W(c, q; \omega)$ is strictly quasi-concave in (c, q) and strictly increasing in c for all ω.*[4]

The first argument of W represents the change in the investor's cash position as a result of a trade, while the second argument represents the change in the investor's position in the security as a result of a trade. That W is strictly increasing in the first argument means that more cash is preferred to less. Quasi-concavity of W in (c, q) means that in the (c, q) plane, indifference curves are convex to the origin. As the following examples show, it is related to an assumption of risk aversion.

Formulation of examples, and the subsequent analysis of the equilibrium limit orders requires a specification of the probabilistic structure of the payoff from a position in the security in question. At time LD (a possibly random stopping time), the security will have a liquidation value of X^*_{LD}. Let F_t denote all the information, public and private, available at time t, and define X_t by $X_t = E[\exp(-r(LD - t))X^*_{LD}|F_t]$, where r is the appropriate continuously compounded discount rate. Finally, let H_t denote the public information available at time t, and define x_t by $x_t = E[\exp(-r(LD - t))X^*_{LD}|H_t] = E[X_t|H_t]$. If the private information is of an "unsystematic sort" (i.e., the

[3] It should be noted that $R'(q)$ may have discontinuities. Thus, while $R(q)$ must be continuous in q, it need not be differentiable, and hence while $R(q)$ is the integral of $R'(q)$, $R'(q)$ is not necessarily the derivative of $R(q)$.

[4] If W_i indicates the first partial derivative of W with respect to the ith argument and W_{ij} indicates the second partial derivative with respect to arguments i and j, then we require $W_1 > 0$, $W_2^2 W_{11} + W_1^2 W_{22} - 2W_1 W_2 W_{12} < 0$.

information does not change the discount rate), X_t is the "full information" value of the security while x_t is the value of the security given all public information.

Example: Myopic portfolio adjustment and consumption. Define $W(c, q; \omega)$ by:

$$W(c,q;\omega) = U_0(c^*(c,q,\omega)) + E[U(Y_T + (\nu+q)x_T - TC(\nu+q) + (\phi + c - c^*(c,q,\omega))(1+r_T); Y_T, x_T)|S].$$

In this case, the next arriving investor chooses q to maximize the expected (possibly state-dependent) utility of consumption now and at time T in the future. The security in question will have a value at time T of x_T, the investor has other sources of wealth represented by Y_T, has an initial position, ν, in the security in question, an initial cash position of ϕ, and chooses optimal consumption $c^*(c, q, \omega)$ now. Unwinding the position leads to transactions costs of $TC(\nu + q)$. The investor earns a risk-free return r_T over the T periods. Furthermore, the investor has a (possibly null) signal about the future random variables. The vector of unobservable characteristics consists of a specification of the utility function, the time horizon, the joint distribution of Y_T, S and x_T, the initial cash and security positions, the risk-free rate, and the nature of and realization of the signal S. Quasi-concavity of W is implied by concavity of U and convexity of the transaction cost function $TC(\cdot)$, while $W_1 > 0$ is implied by positive marginal utility of wealth. The formulation is myopic in the sense that the investor ignores future opportunities to trade. An informational motive for trade results from non-null S, while a "liquidity" motive for trade arises from suboptimal ν and ϕ given the random variables Y_T and x_T and/or a particular desire for or aversion to current consumption relative to future consumption.

Example: Dynamic portfolio adjustment. Define $W(c, q; \omega)$ by:

$$W(c,q;\omega) = E[U(Y_T + \phi + c - R_2(q_2) - \cdots - R_{T-1}(q_{T-1}) + (\nu + q + q_2 + \cdots + q_{T-1})x_T|S],$$

where q_i are the future optimal trades in the security, and R_i are the future terms of trade. In this case, the maximum depends upon the individual's expectations of the future terms of trade. Whether $W(\cdot, \cdot; \omega)$ is quasi-concave or not will depend upon how the investor believes future R_is will depend upon a current trade. For example, if the investor believes that future terms of trade will be unaffected by a current trade, then concavity of U will imply quasi-concavity of W. Such beliefs will also typically involve the investor planning on trading more than once. If this independence of a current trade and future bids and offers does not hold, then $W(\cdot, \cdot; \omega)$ may not be quasi-concave. For example, some expectations over future terms of trade and some utility functions may invite "destabilizing trade" (a sequence of small buys followed by a large sale, for example). In this case, quasi-concavity is unlikely to hold for all ω. This and other dynamic issues will be discussed further below.

Assumption 1 does rule out one specification that enjoys frequent academic consideration. That the marginal "utility" of cash is positive precludes the pure noise trader" specification in much of the "Rational Expectations Equilibrium" literature (for example, Hellwig (1980)), as well as the specification in Kyle (1985). While the general model admits a reasonably wide range of motives for trade, it still requires that investors care about the amount they pay for purchases or receive for sales.

The quasi-concavity assumption means that characterization of an investor's decision is conveniently derived. Essentially, the investor chooses a trade so that his or her "marginal valuation" equals the marginal price. The marginal valuation is given by:

$$M(q, R(q); \omega) = W_2(-R(q), q; \omega)/W_1(-R(q), q; \omega).$$

Quasi-concavity implies that in the neighborhood of any solution, the marginal valuation is decreasing in q. Since the institution requires that the marginal price function be nondecreasing, there can only be one solution to the marginal condition.

LEMMA 1: *Suppose that W is strictly quasi-concave, and that $R'(q)$ is any arbitrary nondecreasing marginal price function defined for q in the interval $[q_0, q_1]$ (q_0 may be negative infinity, and q_1 may be positive infinity). Define the marginal valuation of an investor with characteristics vector ω at a trade q and transfer $R(q)$, $M(q, R(q); \omega)$, by:*

$$M(q, R(q); \omega) = W_2(-R(q), q; \omega)/W_1(-R(q), q; \omega).$$

Then one of the following mutually independent and collectively exhaustive conditions holds:

(i) *$M(q, R(q); \omega) > R'(q)$ for all q in $[q_0, q_1)$;*
(ii) *$M(q, R(q); \omega) < R'(q)$ for all q in $(q_0, q_1]$;*
(iii) *There exists exactly one $q^*(\omega) \in [q_0, q_1]$ such that:*

$$q < q^*(\omega) \text{ implies } M(q, R(q); \omega) > R'(q)$$
$$q > q^*(\omega) \text{ implies } M(q, R(q); \omega) < R'(q).$$

The examples above the lemma illustrate some of the investor fundamentals that will imply quasiconcavity. The lemma illustrates the force of this assumption. The optimal trade of an investor can be characterized as the solution to a first-order condition. Strict quasi-concavity will make this solution unique. The characterization is provided in the following proposition.

PROPOSITION 1: *Suppose that W is strictly quasi-concave for all ω, and $R'(\cdot)$ is nondecreasing and defined for $q \in [q_0, q_1]$. Then an investor with a vector of* *characteristics ω will choose $D_R(\omega)$ as the trade where $D_R(\omega)$ is the unique solution to the following*:

(i) *if a solution to $M(q, R(q); \omega) = R'(q)$ exists, then $D_R(\omega)$ is this unique solution*;

(ii) *if the solution to the equation in (i) does not exist, but there is a point of discontinuity in R' at q^* and $M(q^*, R(q^*); \omega)$ lies between the limit from below q^* and the limit from above q^*, of $R'(\cdot)$, then $D_R(\omega)$ is q^*.*

(iii) *if neither (i) nor (ii) hold, then $D_R(\omega) = q_1$ if $M(q, R(q); \omega) > R'(q)$ for all q and $D_R(\omega) = q_0$ if*

$$M(q, R(q); \omega) < R'(q) \quad \textit{for all } q.$$

Before leaving the analysis of the individual investor, a corollary is provided that will be useful in the subsequent subsection. To the extent that investors have private information, limit order submitters may care about how individual investors value a share of the security. The following corollary shows the link between how investors value the security and the decisions that they make. The proof is immediate from Lemma 1 and Proposition 1.

COROLLARY 1: *If W is strictly quasi-concave, and $R'(\cdot)$ is any nondecreasing marginal price function that is left continuous for $q > 0$ and right continuous for $q < 0$, then the following two identities hold*:

(A) *for* $q > 0$, $\{\omega\colon D_R(\omega) \geq q\} = \{\omega\colon M(q, R(q); \omega) \geq R'(q)\}$;

(B) *for* $q < 0$, $\{\omega\colon D_R(\omega) \leq q\} = \{\omega\colon M(q, R(q); \omega) \leq R'(q)\}$;

Where $D_R(\omega)$ is defined in Proposition *1* above.

There may be marginal price functions decreasing in some interval that also satisfy the conclusions of the corollary. Any marginal price function that does satisfy the conclusions of the corollary shall be said to have the "single crossing" property. This will be important in the analysis of competing exchanges and market breakdown. What the property does is unambiguously link marginal valuations and trades.

B. Equilibrium Bids and Offers

The subsection above characterizes the behavior of investors taking the schedule of bids and offers as given. It is assumed that suppliers of liquidity —those who provide limit orders—recognize this behavior and take account of it in the provision of bids and offers. As stated in the introduction, this analysis focuses on the effects of asymmetric information. Rather than taking a particular parametric specification of information and division of information among potential investors, the assumption that defines the presence of private information encompasses a number of specific models.

The trading behavior of market order users is determined by their marginal valuation functions and the terms of trade offered. The anonymity of the electronic market implies that liquidity suppliers observe only an arriving investor's marginal valuation at the trade chosen. This suggests that, if there is private information that is of concern to liquidity suppliers, observing this point on the marginal valuation function must be, in general, informative. To avoid making assumptions about endogenous objects, it is assumed that any point on the marginal valuation function provides information about X, the current "full information value" of the security.[5] It will be assumed that all private information is "unsystematic," and hence a condition on conditional expected values is all that is needed.

ASSUMPTION 2: *For each q and R and m, define the "upper tail expectation" function, $V(m, q, R)$ to be the expectation of X conditional on the next arrival's marginal valuation at q and R being greater than or equal to m, and the "lower tail expectation" function, $v(m, q, R)$, to be the expectation of X conditional on the next arrival's marginal valuation at q and R being less than or equal to m*:

$$V(m, q, R) = E[X \mid M(q, R; \omega) \geq m];$$
$$v(m, q, R) = E[X \mid M(q, R; \omega) \leq m].$$

The functions $V(\cdot, \cdot, \cdot)$ and $v(\cdot, \cdot, \cdot)$ satisfy:

$$V(m, q, R) \geq E[X \mid M(q, R; \omega) = m] \geq v(m, q, R).$$

The economy exhibits strict adverse selection if the inequalities above are strict.[6]

A high marginal valuation (given R and q) could be due to the investor being short in the security; it could be due to a relative aversion to current consumption; or it could be due to the investor having another source of income negatively correlated with the security's return. Assumption 2 states that a possible explanation for a high marginal valuation is information indicating that X, the current full information value, is more likely to be large. It should be noted that the inequality must hold for each q and R, and hence it is not an assumption about endogenous objects.

The assumption is implied by the condition that the "point" conditional expectation, $E[X \mid M(q, R; \omega) = m]$, be increasing in m. The assumption is equivalent to the assumption that the functions $V(m, \cdot, \cdot)$ and $v(m, \cdot, \cdot)$ are both increasing in m. This, and another useful property of these functions, is proven in the following Lemma.

LEMMA 2: *Assuming strict concavity of the investors' objective functions, and given Assumption 2, the expectation of X conditional on the next arrival's marginal valuation at q and R being greater than or equal to m, $V(m, q, R) = E[X \mid M(q, R; \omega) \geq m]$ and the expectation of X conditional on the next*

[5] The random variable X_t was defined above as the discounted expected liquidation value conditional on all public and private information at time t. The analysis now focuses on a particular point in time and the subscript t is dropped.

[6] This assumption is in the spirit of the affiliation assumption in the auction literature (see for example Milgrom and Weber (1982)). In the case at hand, however, any quantity may be chosen, and hence the simple and elegant affiliation assumption of Milgrom and Weber is insufficient.

arrival's marginal valuation at q *and* R *being less than or equal to* m, $v(m, q, R) = E[X \mid M(q, R; \omega) \leq m]$ *are increasing in* m, *while the expectation,* $V(m, q, R + qm)$ *is increasing in* q *for* $q > 0$, *and the expectation,* $v(m, q, R + qm)$, *is increasing in* q *for* $q < 0$, *for all* R *and* m.

The first result follows immediately from the observation that the expectation conditional on the marginal valuation being greater than or equal to m is an average of expectations conditional on the marginal valuation being equal to m' for $m' \geq m$. In an environment with a single ask price m, $V(m, q, R + qm)$ is the expectation of X conditional on an investor choosing q or larger. By the strict concavity of the investors' objective functions, an investor who chooses q or larger must have a marginal valuation at q that is m or larger. Thus, the expectation conditional on an investor choosing q or larger exceeds the expectation conditional on an investor choosing q. The result follows.

The following examples illustrate Assumption 2, and are used throughout this article to illustrate the propositions.

Example: Consider the environment of Glosten (1989). The next arrival has an endowment w, which, from the point of view of limit order submitters, is normally distributed with mean zero. The full information value of the security, X, is normally distributed. The next arrival has seen a signal $S = X + \varepsilon$, with ε normally distributed with mean zero, independent of X. Finally, the next arrival maximizes the expected utility of future wealth, and the utility function is exponential with risk-aversion parameter r. Let σ be the standard deviation of X conditional on S. Standard calculations show that the marginal valuation is given by:

$$M(q, R; \omega) = E[X \mid S] - rw\sigma^2 - rq\sigma^2.$$

This example will be referred to below (call it the exponential-normal example), and it is convenient to choose some normalizations to minimize the number of parameters. If we interpret all conditional expectations and prices as deviations from the ex ante mean, we can choose the mean of X to be zero. Normalize the quantity units by setting $r\sigma^2 = 1$. Finally, let the variance of w be $\alpha < 1$ and set the variance of $E[X \mid S]$ equal to $1 - \alpha$. Roughly speaking, α is the proportion of the variance of trade explained by the liquidity motive. Then,

$$M(q, R; \omega) = \omega - q,$$

where $\omega = E[X \mid S] - r\sigma^2 w$, and, under the above assumptions, ω is a standard normal random variable. Furthermore, X and ω are correlated and $E[X \mid \omega] = (1 - \alpha)\omega$. Thus the following holds:

$$E[X | M(q, R, \omega) = m] = (1 - \alpha)(m + q).$$

If $\alpha < 1$, this is strictly increasing in m, and hence the assumption is satisfied.

Example: This example shows that Assumption 2 is not innocuous and can fail in a reasonable model of informed trade. The assumption can fail when extreme marginal valuations could only come from uninformed investors.

Suppose that there are informed agents and uninformed agents. Let U be a (zero, one) random variable that takes the value one if the next arrival is uninformed, and put $E[U] = \alpha$. Suppose that the uninformed have a marginal valuation given by $(\varepsilon - q)$. Informed have seen the realization of a signal, S, correlated with X, and they are risk neutral. Assume that U, ε, $E[X \mid S]$ are mutually independent, and $E[X] = 0$. Let $f(\cdot)$ denote the density of $E[X \mid S]$ and let $g(\cdot)$ be the density of ε. Then,

$$M(q, R; \omega) = (1 - U)E[X \mid S] + U(\varepsilon - q), \quad \text{and} \quad \omega = (U, S, \varepsilon).$$

Furthermore:

$$E[X|M(q, R; \omega) = m] = (1 - \alpha)f(m)m/[(1 - \alpha)f(m) + \alpha g(m + q)].$$

While increasing for m near zero, this conditional expectation need not be increasing for all m and q. For example, suppose that f and g are both uniform densities, but the support of f is strictly contained in the support of g. Then, for extreme m, and small (in absolute value) q, the conditional expectation above will be zero, and the assumption will not hold for all m and q. With f and g uniform, this will be referred to as the uniform example.

Note that the above two examples entail marginal valuations that are independent of the amount paid or received for a trade of q. This was, of course, due to the constant absolute risk aversion and the absence of wealth effects in the marginal valuation. Examples using other utility functions that exhibit wealth effects can be constructed, although they tend to be difficult to manipulate.

To derive the equilibrium among competing suppliers of liquidity (limit order submitters), the following assumption is made.

ASSUMPTION 3: *Let N be the number of potential limit order submitters. Private information is unsystematic in that each limit order submitter maximizes expected trading profit given only publicly available information. That is, a liquidity supplier provides bids and offers to maximize $E[P - XQ]$ where P represents the liquidity supplier's (signed) proceeds from the next arrival, and Q is the (signed) quantity provided by the liquidity supplier to the next arrival. Each liquidity supplier can submit any number of bids and offers. A limit order can be for any positive quantity. Competing limit orders at one price are executed in a pro rata fashion. Equilibria will considered for the limit as N goes to infinity.*

We can think of these liquidity suppliers as "patient" or "value" traders in that their only interest in trading is expected profit. It might be reasonable to think of this population as consisting of managers of reasonably large portfolios, both institutional and individual. Since the portfolios are large, even participation in a sizeable trade does not make a substantial difference in the diversification of the portfolios. Such an interpretation calls into question the consideration of a large population, but one should think of this analysis as a base case. More will be said on this issue below.

While most of the analysis in this article considers the case in which the set of allowable prices is the continuum, understanding the equilibrium is facilitated by first considering the more realistic case in which prices are restricted to a discrete set. The continuous price equilibrium is then the limiting equilibrium as the discreteness in prices goes to zero. The details of the derivation of the discrete price equilibrium are provided in the Appendix; the following provides an outline of the logic.

Let the set of allowable prices be $P = \{\dots p_{-1}, p_0, p_1, \dots\}$ where this set is arranged in increasing order. Let p_0 be the allowable price closest to the ex ante mean of X That is, $p_{-1} < E[X] < p_1$. It seems reasonable, and will be proven below, that no liquidity supplier offers quantities at p_{-1} or below or bids for quantities at p_1 or higher. Given this set up, the strategy for each liquidity supplier consists of a specification of $\{q_i^A, q_i^B\} \geq 0$ where q_i^A is the quantity offered at price i and q_i^B is the quantity bid at price i. Quantities of zero are to be interpreted as no bid or offer provided. The analysis seeks the Nash equilibrium of the game in which liquidity suppliers expect investors to behave as derived in the subsection above. Each liquidity supplier observes the bids and offers of all other liquidity suppliers and chooses his or her profit-maximizing response.

The Appendix shows that with an infinite number of limit order providers, the Nash equilibrium is characterized by the following zero-profit condition for prices at which positive quantities are offered:

$$\int_{AQ_{i-1}}^{AQ_i} (p_i - V(p_i, d, R_{i-1} + p_i(d - AQ_{i-1})) \times P\{M(d, R_{i-1} + p_i(d - AQ_{i-1})) \geq p_i\} = 0 \tag{1}$$

where AQ_{i-1} is the quantity offered at prices lower than p_i, and R_{i-1} is the total cost of these shares. That is, a positive quantity is offered at p_i as long as p_i is, on average, at least the "upper tail expectation" of X conditional on a market order trading at the price p_i. On the other hand, by Lemma 2, if p_i is less than the upper tail expectation conditional on the arrival of an order large enough to pick off the first share offered at p_i, $p_i < V(p_i, AQ_{i-1}, R_{i-1})$, then $p_i < V(p_i, AQ_{i-1} + q, R_{i-1} + qp_i)$ for all positive q. If this holds, no shares will be offered at p_i. Proposition 2 summarizes the equilibrium derived from these two observations.

PROPOSITION 2: *Given the maintained assumptions, the following describes the equilibrium*:

(i) *If* $p < V(p,0,0)$ *for all* $p \in P$, *then no offers are provided. If* $p > v(p,0,0)$ *for all* $p \in P$, *then no bids are provided.*
(ii) *If there exists a* $p \in P$ *satisfying* $p > V(p,0,0)$, *then the lowest ask*, A_1 *is the smallest such* p. *If there exists a* $p \in P$ *satisfying* $p < v(p,0,0)$, *then the highest bid*, B_1 *is the largest such* p.

(iii) *If the expression for the ask side first-order condition, equation* (1), *with* $p_i = A_1$, $R_{i-1} = 0$ *is positive for all* q, *then an infinite quantity will be offered at* A_1. *Otherwise, the quantity offered at* A_1 *will be the solution to the zero-profit condition. If the expression for the bid side first-order condition with* $p_i = B_1$ *is positive for all* q, *then an infinite quantity will be bid at* B_1. *Otherwise the quantity offered at* B_1 *will be the solution to the first-order condition.*

(iv) *If positive quantities are offered at k different ask prices, and letting* AQ_k^* *equal the aggregate quantity offered at the k ask prices and letting* R_k *equal the amount paid for the quantity* AQ_k^* *then*:

(a) *If* $\{p \in P\colon p > V(p, AQ_k^*, R_k)\}$ *is empty, then there are no higher offers.*
(b) *Otherwise,* A_{k+1} *is* $\min\{p \in P\colon p > V(p, AQ_k^*, R_k)\}$
(c) *If the integral in* (1) *with* $p_i = A_{k+1}$ *is nonnegative for all* q, *then an infinite quantity is offered at* A_{k+1}.
(d) *Otherwise,* Q_{k+1}^* *is the solution to the first-order condition.*

If positive quantities are bid at k different bid prices, and letting BQ_k^* equal the aggregate quantity bid at the k bid prices and setting R_k equal to the amount received for the quantity BQ_k^* then:

(a) If $\{p \in P\colon p < v(p, -BQ_k^*, -R_k)\}$ is empty, then there are no lower bids.
(b) Otherwise, B_{k+1} is $\max\{p \in P;\ p < v(p, -BQ_k^*, -R_k)\}$.
(c) If the first-order condition with $p_i = B_{k+1}$ is nonnegative for all q, then an infinite quantity is bid at B_{k+1}.
(d) Otherwise, Q_{k+1}^* is the solution to the bid side first-order condition.

Some fairly general characteristics of the equilibrium fall out of the above derivation, and these are provided in Proposition 3. Consideration of these general characteristics gives some insight into the driving forces of the equilibrium.

PROPOSITION 3: *Assume that* $V(m, q, R)$ *is strictly increasing in* m, *while* $E[X \mid M(q, pq; \omega) = p]$ *is continuous in* q.

(i) *If the market is open, then for* ε *small but positive,*

$$A_1 > V(A_1, 0, 0) > v(B_1, 0, 0) > B_1; \quad \textit{and}$$
$$A_1 > E[X|M(\varepsilon, \varepsilon A_1; \omega) = A_1];$$
$$B_1 < E[X|M(-\varepsilon, -\varepsilon B_1; \omega) = B_1].$$

(ii) *If there are offers at k different ask prices, and bids at k different bid prices, then for* ε *positive but small*:

$$E[X|D = AQ_{k-1} + \varepsilon] < E[X|D \geq AQ_{k-1}^*] < A_k < E[X|D \geq AQ_k^*];$$
$$E[X|D = -BQ_{k-1} - \varepsilon] > E[X|D \leq -BQ_{k-1}^*] > B_k > E[X|D \leq BQ_k^*];$$

Part (i) of Proposition 3 shows that if the economy exhibits strict adverse selection, then the limit order book will have a positive bid-ask spread no matter what the set of allowable prices is; the set of prices can be made arbitrarily fine, and the small-trade bid-ask spread will persist. The reason for this is the possible trading on private information. An individual that provides an offer at the smallest ask price, will transact on every trade. Not only will he or she get a portion of small trades, but on all large trades, the total quantity offered will be taken. This means that in order to place an offer at the smallest ask, the individual has to be concerned with the informational implications of all investor purchases. Similarly, an individual placing a limit order at the largest bid needs to be concerned with the informational implications of all investor sales.

The first part of the proposition also shows that limit order submitters profit from small investor purchases and sales. The second part of the proposition stresses the importance of the "upper tail" expectations for the determination of offers, and the "lower tail" expectations for the determination of bids. The proposition also shows that if the realized trade is just greater than AQ_{k-1}, then an offer at A_k will be profitable.

Part (i) of Proposition 3 has a further implication. If the equilibrium does not provide an infinite quantity at any ask price, then every offer has a zero expected profit. But if each limit order breaks even, the book in aggregate expects to break even. That is, in expectation, the average price received by the book, $R(q)/q$ equals the revised expectation. Since small trade are profitable, some larger trades must be unprofitable. That is, for small trades the average price paid by a buying investor exceeds the revised expectation, while for some larger trades the revised expectation is greater than the average price paid by a buying investor.

Part (ii) of Proposition 3 points out an interesting feature of the market. Suppose an order for $AQ_{k-1} + \varepsilon$ arrives. This will clear out all the offers at A_1 through A_{k-1}, and part of the orders at A_k. The revised expectation in response to this realized trade lies strictly between B_1 and the now lowest ask price at A_k. Thus, there are no offers lying exposed below the revised expectation, and no bids lying exposed above the revised expectation. It is not necessarily the case that offers need to be canceled after this trade. Even though the model assumes constant vigilance on the part of limit order submitters, constant monitoring need not be necessary to avoid unfavorable trades.[7]

Examples: Before proceeding to a further analysis of the electronic open limit order book, it is perhaps informative to examine some examples of the above general analysis. First consider the normal-exponential example introduced above. Recall that $E[X \mid M(q, R; \omega) = m] = (1 - \alpha)(m + q)$. Thus, if f

[7] Limit orders are still exposed to movements due to public information arrivals. Perhaps as Black (1992) suggests limit orders could be "marked to market" by moving all limit orders up or down in response to a, respectively, positive or negative public announcement.

is the standard normal density and F is the standard normal distribution function, $V(m, q, R)$ is given by:

$$V(m,q,R) = (1-\alpha)f(m+q)/(1-F(m+q)).$$

As long as α is positive, there exists a solution to $p = V(p, q, R)$ for all q. Thus, the order book will, in principle, provide terms of trade for arbitrarily large orders. In fact, if the set of prices is coarse enough, and α is large enough, an infinite quantity will be offered at A_1.

It can be seen that the lowest offer is nonincreasing in α. That is, the small-trade spread tends to increase in the severity of the adverse selection problem.[8]

Figures 1 and 2 illustrate, respectively, the derivation and description of the equilibrium when $\alpha = 0.8$. There are three distinct offer prices—0.25, 0.375 and 0.5. Finite quantities are offered at the first two prices, while, in principle, an infinite quantity is offered at 0.5.[9]

The second example provides a somewhat different equilibrium. Recall the uniform example discussed above:

$$M(q,R;\omega) = (1-U)E[X \mid S] + U(\varepsilon - q),$$

where U, $E[X \mid S]$, and ε are mutually independent, $E[U] = \alpha$, and suppose that $E[X \mid S]$ and ε are both uniformly distributed on $[-L, L]$. In this case, for $L > m > 0$, $q \geq 0$:

$$V(m,q,R) = (1-\alpha)(L^2 - m^2)/ \\ \left[2(1-\alpha)(L-m) + 2\alpha(L-m-q)I_{\{q \leq L-m\}}\right],$$

where I_E is the indicator function of the set E. In particular, $V(m, 0, 0) = (1 - \alpha)(L + m)/2$. As long as the set of prices is not too coarse and/or α is large enough, some quantity will be offered. All that is required is that there be an allowable price in the interval $((1-\alpha)L/(1+\alpha), L)$. However, arbitrarily large trades are not possible in this environment. At any ask an infinite quantity is not offered, and if q exceeds $L(1-(1-\alpha^2)^{0.5})/\alpha$, the function $V(m, q, R)$ lies above m for all m. Thus, after the book has provided a quantity up to the above limit or higher, no subsequent offers will arrive. The exact quantity provided depends upon the allowable price set and the other parameters. That the quantity offered is finite is true regardless of the allowable price set.

In this example, α measures the importance of liquidity trade, just as in the previous example. As in the previous example, the lowest ask price is decreasing in α, while the maximum quantity offered is increasing in this parameter.

[8] Some care is needed in interpreting this result. Recall that a particular normalization was used in this example to minimize the number of parameters. Thus, a change in α represents a simultaneous change in the variance of the arrival's endowment, the variance of the value of the security and the precision of the information.

[9] The offering of an infinite quantity at some price seems to be a feature of the normal-exponential model with discrete prices.

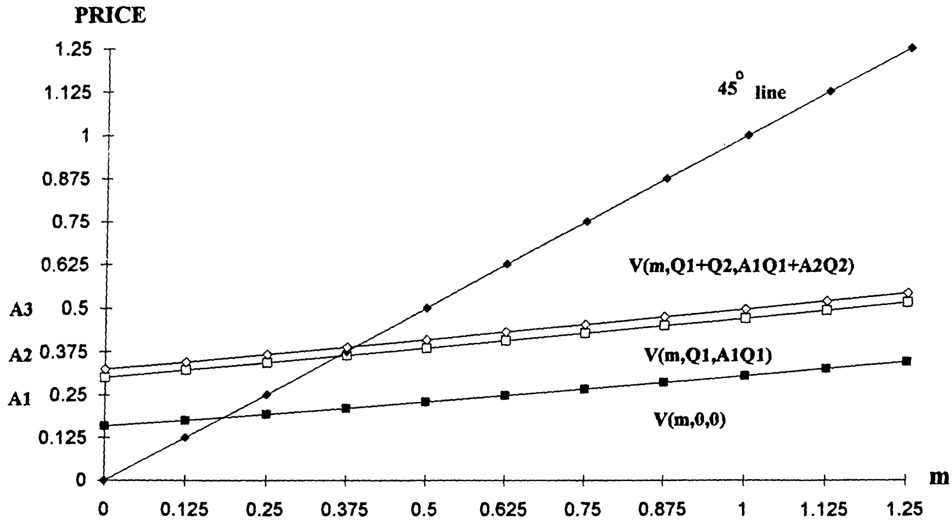

Figure 1. Determination of the equilibrium offers for the exponential-normal model. The function $V(m, q, R(q)$ is the expectation of the terminal payoff, X, conditional on an arriving investor who pays $R(q)$ for q shares having a marginal valuation greater than or equal to m. The marginal valuation, $M(q, R(q); \omega)$ is the amount that an investor paying $R(q)$ for q shares would be willing to pay for an additional share. It is given by $M(q, R(q); \omega) = \omega - q$, where ω is a standard normal random variable, and $E[X \mid \omega] = (1 - \alpha)\omega$. Thus, $V(m, q, R(q)$ is given by:

$$V(m, q, R(q)) = E[X \mid M(q, R(q); \omega) \geq m] = (1 - \alpha)f(m + q)/(1 - F(m + q)),$$

where $f(\cdot)$ and $F(\cdot)$ are, respectively, the standard normal density and distribution function. The adverse selection parameter, α, is set at 0.8 (i.e., 20% of trade is motivated by private information) and the set of allowable prices is 1/8's. Approximately one unit is offered at 0.25, a small amount is offered at 0.375 and an arbitrarily large amount is offered at 0.5. The details of the calculations are provided at the end of the Appendix.

Finally, the above example can be modified to show that there are situations in which the market will not open; i.e., no bids or offers will be provided. Suppose that ε is uniformly distributed on $[-L_u, L_u]$, while $E[X \mid S]$ is uniformly distributed on $[-L, L]$, and $L > L_u$. Then it can be verified that $V(p, 0, 0)$ exceeds p for all p if $\alpha < 2(L - L_u)/(2L - L_u)$. That is, if the adverse selection problem is severe enough (α is small), and the liquidity motive for trade is relatively limited (L_u is small), the market will close down.

For the remainder of the analysis, it will be convenient to drop the assumption that only a discrete set of prices is allowed. While admittedly unrealistic, the mathematics is simplified tremendously. It should be noted that relatively few of the characteristics derived above in the general analysis and the specific examples relied on the particular set of allowable prices. The passage to continuous prices will be accomplished by taking limits of the discrete analysis above as the set of prices becomes finer. Thus, one may think of the continuous price case as a mathematically convenient approximation to the more realistic step function marginal price schedule.

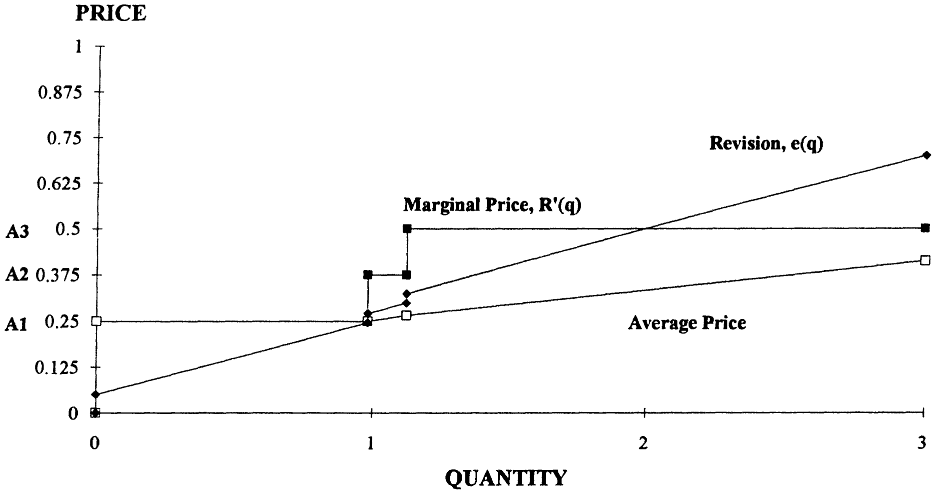

Figure 2. Illustration of the exponential-normal model equilibrium with discrete prices. The marginal price, $R'(q)$ is calculated in Figure 1: 0.98 units are offered at 0.25, 0.14 units are offered at 0.375, and an arbitrarily large amount is offered at 0.5. The average price is $R(q)/q$, and the revision in expectations, $e(q)$, is given by $e(q) = (1-\alpha)(q + R'(q))$. Note that $e(q)$ lies below the average price for $q < 0.98$ and hence small buys are profitable for the limit order submitters. Since the investor's marginal valuation, $M(q, R(q); \omega) = \omega - q$, and ω is a standard normal random variable, roughly 73%, $(F(0.98 + 0.25) - F(0.25))/(1 - F(0.25))$, of market order purchases are less than 0.98. For $q > 2$, the revision in expectations exceeds the marginal offer of 0.5 (which exceeds the average price). Roughly 1.5% of the market order purchases will exceed 2 units.

The Appendix provides the details of the limiting argument. In the discrete case, offers are approximately "upper tail" expectations and bids are approximately "lower tail" expectations. In the continuous price case, marginal offers equal "upper tail" expectations, while marginal bids equal "lower tail" expectations. The remaining conditions in Proposition 4 insure that the equilibrium picks out the lowest offers and highest bids satisfying the expectation condition.

PROPOSITION 4: *For $Q > 0$, the marginal price function $R'(Q)$ must satisfy*:

$$R'(Q) = V(R'(Q), Q, R(Q)) = E[X \mid M(Q, R(Q); \omega) \geq R'(Q)]$$
$$V_1(R'(Q), Q, R(Q)) \leq 1;\ R'_+(0) = \inf\{p\colon p > V(p, 0, 0)\},$$

where $R'_+(0)$ is the limit of $R'(q)$ as q goes to zero from above.

For $Q < 0$, the marginal price function must satisfy:

$$R'(Q) = v(R'(Q), Q, R(Q)) = E[X \mid M(Q, R(Q); \omega) \leq R'(Q)]$$
$$v_1(R'(Q), Q, R(Q)) \leq 1;\ R'_-(0) = \sup\{p\colon p < v(p, 0, 0)\},$$

where $R'_-(0)$ is the limit of $R'(q)$ as q goes to zero from below.

A finite solution to this system will exist for some interval of quantities, if $m > V(m,0,0)$ for some interval of m's and $m < v(m,0,0)$ for some interval of m's. Furthermore, $R'_{+}(0) > R'_{-}(0)$, and for $\epsilon > 0$ but small,

$$E[X \mid D = -\varepsilon] > R(-\varepsilon)/(-\varepsilon) \text{ and } E[X \mid D = \varepsilon] < R(\varepsilon)/\varepsilon.$$

Examples: In the exponential-normal example, we have

$$R'(q) = (1-\alpha)f(R'(q)+q)/(1-F(R'(q)+q)) \quad \text{for } q > 0$$
$$R'(q) = -(1-\alpha)f(R'(q)+q)/F(R'(q)+q) \quad \text{for } q < 0.$$

The equilibrium can be illustrated in a neater form by deriving the equilibrium trade by an individual of type z. Denote by $q(z)$ the solution to $z - q(z) = R'(q(z))$. Then, for $q(z) > 0$:

$$z - q(z) = (1-\alpha)f(z)/(1-F(z)).$$

The solution will be positive as long as z exceeds z^*, the solution to $z^* - (1-\alpha)f(z^*)/(1-F(z^*)) = 0$. For $z < -z^*$, the solution is given by:

$$q(z) = z + (1-\alpha)f(z)/F(z) < 0.$$

Note that z^* is the limit of $R'(q)$ as $q > 0$ goes to zero. Note also that $q(z)$ is increasing in α. That is, as the severity of the adverse selection declines, the marginal price function declines and, in equilibrium, investors make larger trades. The marginal price function, $R'(q)$ and the revision in expectations, $e(q)$, can be found numerically by graphing (respectively) $(q(z), z - q(z))$ and $(q(z), (1-\alpha)z)$ for various z's. This is done in Figure 3 for the case $\alpha = 0.8$.

For the uniform distribution example, $R'(q)$ is the solution to a quadratic equation. Depending upon q, the quadratic equation has two roots, one root, or no roots. If two roots are available, the partial derivative condition $V_1 \leq 1$ requires taking the smaller root. The lack of a root indicates that a marginal price is not offered for that quantity. Using the expression for V developed above in the previous discussion of the example (with $L_u = L$), $R'(q)$ for $q > 0$ is:

$$R'(q) = \left\{L - \alpha q - (\alpha^2 q^2 - 2\alpha q L + \alpha^2 L^2)^{0.5}\right\}/(1+\alpha)$$

for $q < L(1-(1-\alpha^2)^{0.5})/\alpha$. Note that the total quantity offered is increasing in α, while the marginal price schedule is decreasing in α. As noted above, if $L_u < L$, and α is small enough, there will be no offers less than L and the market will close down.

II. Further Characteristics of The Electronic Market

One characteristic of a trading mechanism that may be important is its ability to consistently provide some liquidity. The ability of the monopolist specialist system to provide liquidity is the focus of Glosten (1989). The key

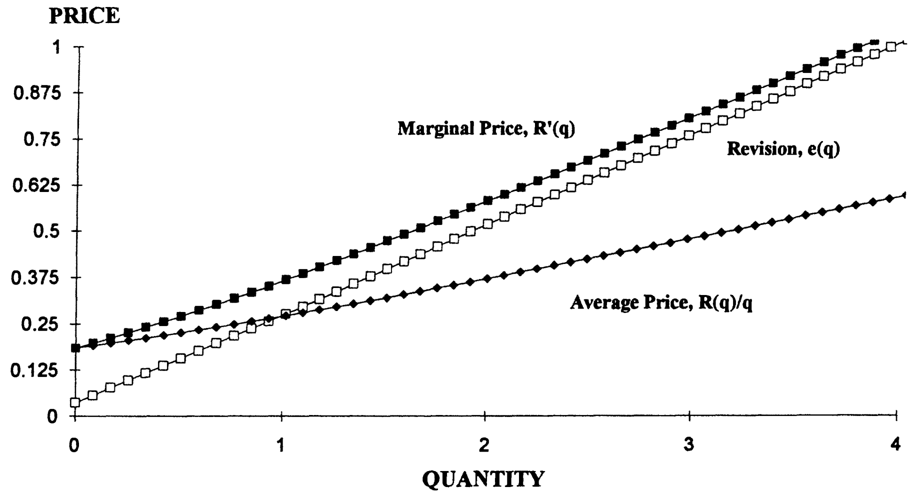

Figure 3. Illustration of the continuous price exponential-normal model equilibrium. $R'(q)$ is the marginal price schedule, $e(q)$ is the revision in expectations due to a trade of q, $e(q) = e[X \mid M(q, R(q); \omega) = R'(q)]$, and $R(q)/q$ is the average price schedule for $q > 0$. The picture for $q < 0$ (bids) is symmetric. The calculations are for the exponential-normal model with $\alpha = 0.8$: the investor's marginal valuation is given by $M(q, R(q); \omega) = \omega - q$, ω is a standard normal random variable and $E[X \mid \omega] = \{1 - a)\omega$,. $R'(q)$ is the solution to $R'(q) = (1 - a)f(q + R'(q))/(1 - F(q + R'(q))$ where $f(\cdot)$ and $F(\cdot)$ are, respectively, the standard normal density and distribution function. Note that the lowest offer is strictly positive, as is the smallest revision in expectations, and for large q, the revision in expectations exceeds the average price. The functions $R'(q)$, $R(q)/q$, and $e(q)$ are convex, but approximately linear for large q.

property that allows a specialist to keep the market open when the competitive mechanism closes down is the ability of the specialist to average profits across trades. Notice, that this is a feature that the electronic market considered here shares with the specialist system. A reasonable question is whether this electronic market does it as well. The answer to this question is a restricted yes. If the electronic market provides no liquidity (formally there is no finite solution for $R'(q)$, or the only solution precludes trade with probability one), then any other market mechanism that has a "nice" marginal price function will expect to lose money. Thus, a large set of markets will be open in an environment only if the electronic exchange would be open in that environment.

PROPOSITION 5: *Suppose that there is no finite fixed point, m, $m = V(m, 0, 0)$ so that the electronic market will not open, and assume an economy in which marginal valuations are independent of cash positions so that $V(m, q, R)$ is independent of R. Then any other price schedule that has the single crossing property* (*see the discussion following Corollary* 1) *will expect to lose money.*

For the electronic market to open, all that is required is that liquidity suppliers be willing to make a small trade. Any other exchange, if open, would have to make this small trade, plus trades that are worse from an informational perspective. Thus, if the liquidity suppliers are unwilling to provide quotes in the electronic market, any other market would be unlikely to provide terms of trade.

One can measure liquidity in a variety of ways. Based on the size of the small-trade spread, one might be tempted to say the electronic market is not liquid. Indeed, it is possible to specify an economic environment in which a nondiscriminating (or single price) electronic market has no small-trade spread. This is the example of competitive pricing in Glosten (1989). However, such a market might close down too quickly. The above proposition states that if the measure of liquidity is resilience in the face of severe adverse selection problems, then the electronic market as conceived here is as good as one can do.

If the electronic market is open for some quantity, then a monopolist specialist would keep the market open for some quantity as well. Thus, in the normal-exponential example, both the monopolist specialist and the electronic open limit order book would be open for all quantities and in all environments (as long as α exceeds 0). In the modified uniform example (presented following Proposition 3), both the electronic open limit order book and a myopic monopolist specialist will close if the adverse selection is too severe.[10] The proposition raises the possibility that an electronic market may be able to reap the benefits of competition, while at the same time preserving the monopolist specialist liquidity in the face of severe adverse selection problems. The normal-exponential example that has been considered above indicates that, in at least one environment, this statement is true. Proposition 6 provides the details.

PROPOSITION 6: *Consider the normal-exponential example. No trader is worse off, and many are strictly better off with the open limit order book when compared to a monopolist specialist.*

III. Competition Among Exchanges

This section of the article considers competition among exchanges, and asks how susceptible the electronic exchange and other conceivable exchanges are to entry of competitors. For this analysis, the article considers a wide open regulatory environment in which anyone can offer to make a market in the security. Furthermore, setting up such a "market" is costless. On the investor side, market orders can be costlessly split up among "exchanges." It turns out that this assumption is a very powerful one and is a driving force behind the results. This is put formally as Assumption 4.

ASSUMPTION 4: *In the presence of more than one exchange, an investor can costlessly and simultaneously send separate orders to each exchange. A com-*

[10] A monopolist specialist may not close if what is learned from trade reduces the subsequent adverse selection problem. See the discussion in the conclusion of Glosten and Milgrom (1985). This issue is also addressed in Leach and Madhavan (1993).

peting exchange can be costlessly established and supplies a marginal price schedule that satisfies the single crossing property (*see the discussion following Corollary* 1).[11]

The first question to be asked is whether, given the existence of the electronic exchange, any potential entrant would be willing to enter. The standard Nash assumption is made—the entrant takes the marginal price function of the electronic market as given. It might appear that, since small trades are profitable for the electronic market, there will be an incentive to offer a price schedule to capture these small trades and skim the cream. This will not work because if small orders find it profitable to go to the competing exchange, then all investors will find it profitable to send some part of their order to the competing exchange. Even if the quantity accepted by the competing market is limited, it would still get a portion of all trades. The structure of the proof is as follows: since investors optimally split their orders, the marginal price received will be the marginal price in the electronic exchange. This marginal price is the upper tail expectation if there were only the electronic market. However, this artificial upper tail expectation is less than the actual upper tail expectation if the quantity traded in the competing market is positive, since upper tail expectations are increasing in quantity (in a world with no wealth effects). Thus, the competing market will consistently receive marginal prices that are less than the upper tail conditional expectations. However, expected profit is a weighted average of the marginal price, less the upper tail conditional expectation.

PROPOSITION 7: *Assume an economy in which marginal valuations are unaffected by cash positions so that $V(m, q, R)$ is independent of R. Suppose $R'_e(q)$ satisfying $R'_e(q) = V(R'_e(q), q, R_e(q))$ is the marginal price schedule in the electronic exchange. Assuming this price schedule fixed, an entrant with a marginal price schedule satisfying the single crossing property will expect to make nonpositive trading profits.*

The proposition asserts that, in a sense, the electronic market is competition proof. One Nash equilibrium is that there will be no entrance. The proposition is almost, but not quite, trivial. After all, an entrant supplying a competing nondecreasing schedule could as easily provide this schedule by participating in the limit order book. The assertion of equilibrium in the limit order book implies that there are no profit opportunities, and that any such effort would lead to negative profits. The slight addition is the allowance of marginal price schedules with a downward sloping portion, as long as the single crossing property is satisfied. What the proposition provides is the first hint that the competition in the discriminatory limit order book mimics the competition among exchanges. This point will, it is hoped, become clearer with subsequent results.

Reference to the proof above suggests that should an entrant come in, unless the limit orders change, limit order submitters will lose money as well.

[11] That the exchange must post a price schedule rules out "quote matching" type competition. See Glosten (1991).

Thus one equilibrium is no entrance. In fact, there will be other equilibria. For example, two competing open limit order books, each offering half the liquidity provided by a single limit order book, will be an equilibrium. The result will be terms of trade identical to those provided by a single order book. The next proposition shows that this is more generally true: if the entrant makes nonnegative profits, the composite price schedule provided by the two markets replicates the price schedule that would be determined if there were only the electronic exchange. The proof uses the same approach as above. If there are two exchanges, the marginal price received in the competing exchange will be driven by the marginal price in the open limit order book. But this is determined to be an upper tail conditional expectation taking into account the existence of the other exchange. Thus, in every case, as long as the competing exchange does not undercut for small trades, the marginal price equals the upper tail expectation. But this is precisely the equilibrium when there is only one order book. The non-negative profit assumption rules out undercutting at small trades.

PROPOSITION 8: *Suppose that there is an equilibrium in which a competing market enters and supplies a marginal price schedule $R_c'(q)$, satisfying the single crossing property. Then there is an equilibrium in which the total revenue function defined by,*

$$R_T(q) = R_e(q_e) + R_c(q_c) \quad \textit{and} \quad q_c + q_e = q$$

is equal to $R(q)$ the schedule determined when there is only the electronic market.

The above two propositions state that if there is a great deal of competition in the provision of limit orders, any additional competition is either unprofitable or redundant. The question that remains to be answered, however, is whether this result is due merely to the great deal of competition that has been assumed, or does the actual architecture of the discriminatory limit order book play a role? The next proposition shows that the architecture is important. It is the particular zero-profit condition determined by the architecture of the discriminating limit order book that discourages further competition. Specifically, any other exchange that expects non-negative profits, but does not replicate the electronic exchange, will invite entrants.

PROPOSITION 9: *Consider an exchange with marginal price functin $R'(q)$, and suppose that for some interval of q's it does not equal the electronic exchange marginal price schedule. Suppose further that this schedule has nonnegative expected trading profits, and satisfies the single crossing property. Then, holding this schedule constant, there exists a competing schedule that will earn positive profits.*

The idea of the proof is that, if an entrant offers a small quantity, every investor with marginal valuation greater than or equal to the price offered will be interested in trading with the entrant. Thus, the cost of supplying the offer is the conditional upper tail expectation. By hypothesis, the price is greater than the upper tail expectation, and the entrant expects to make money. The proof of this proposition shows that while the electronic exchange is not open to cream skimming, any other exchange is.

The proposition implies that the particular design of the electronic market is important; it is not just the competition among a large number of liquidity suppliers that leads to the resilience of the electronic exchange. For example, an alternative design of an electronic market would be a "nondiscriminating" exchange. Liquidity suppliers submit limit bids and offers for quantities of the security. If a market order to purchase q units arrives, then the first limit orders totaling q all transact at the price of the highest offer to transact. Equilibrium among the large number of liquidity suppliers dictates that the price for an order of size q, $P(q)$, satisfy $P(q) = E[X \mid Q = q]$. Since $R(q)$ is given by $P(q)q$, we have $R(q)/q = E[X \mid M(q, R(q); \omega) = R'(q)]$.

In the event that there is no private information, both designs will yield the same result—all bids and offers will stack up at $E[X]$. If there is private information, however, the two designs will lead to different marginal price functions. Recall that with private information, the original specification of the electronic market had $R'_+(0) > E[X]$. Taking limits of the above expression for the alternative design "nondiscriminating" exchange, $R'(0) = E[X \mid M(0,0) = R'(0)]$. In some environments (for example, the normal-exponential example) the solution to this is $R'(0) = E[X]$. Thus, the alternative design will have $R'(q) < E[X \mid Q \geq q]$ for q small. Since the exchange earns zero profits on average, for larger q the opposite inequality must hold. The above proposition demonstrates that such an exchange will invite competition.

It should be added that the analysis in Glosten (1989) shows that the nondiscriminatory exchange will break down if the adverse selection problem is too severe. Thus, the analysis has suggested two reasons for preferring a discriminating design: it is less likely to break down, and it does not invite competitive reaction. The comparison is not unambiguous, however, since the nondiscriminating form will tend to offer lower spreads for small quantities.

IV. Dynamic Issues

The analysis of the order book concerns the development of the book at a point in time. However, as Black (1992) has argued, characterizations of equilibrium may be flawed if dynamic issues are ignored. This section provides no general answers to the Black critique, but some examples are suggestive.

The simplest destabilizing (or bluffing) strategy to consider is the following: buy using a market order and then reverse the trade using limit orders. If one could be assured of there being a buy order following the initial purchase, then this would clearly be a profitable strategy. The initial buy would push up the market's expectations and the average price received would exceed the average price paid. The problem, of course, is that one cannot be assured that the next trade will be an investor buy. Furthermore, one would expect informed investors' expectations to be less influenced by the bluff than the remainder of the market. Consequently, the probability of an informed sell as assessed by the bluffer will be larger than the probability assessment of the uninformed market. That is, the bluffer will find it relatively unlikely that the next trade will be a purchase and, consequently, the expected average price received from using limit orders may be less than the average price paid using the market order.

The above logic can be illustrated more rigorously using an example. Suppose the environment is as follows: there are essentially risk-neutral informed traders who know the future payoff of the security, and this payoff is either zero or one. The proportion of such traders is $1 - \alpha$. There are extremely risk-averse uninformed traders, half of whom are long one unit of the security and half of whom are short one unit. The extreme risk aversion implies that the "shorts" are willing to pay almost any price to buy one unit, and the "longs" are willing to receive almost any price to sell one unit. This environment is like the example in Glosten and Milgrom (1985). The equilibrium will involve market orders for one unit only. One unit will be offered at the expected value of the payoff conditional on a market buy order, while the expected value of the payoff conditional on a market sell will be bid for one unit.

Define N to be the time (in number of transactions) of the first market buy. Denote by A_n the ask price for the nth transaction. To evaluate the profitability of the bluff described above we need to calculate the expected value of $-A_0 + A_N$ conditional on knowing that the time-zero transaction is an uninformed buy. First note that since A_N is a revised expectation (it is the ask at the time of a market buy), the expected value of A_N conditional on an initial buy is A_0. The expectation conditional on an initial buy is an average of the expectation conditional on an uninformed buy (the expectation when A_0 was set) and the expectation conditional on an informed buy (one). The former of these two is smaller and hence the expected profit from the strategy is negative.

Consider, now, the exponential-normal example. A feature of this example is that the revision in expectations function, $e(q)$, is positive and not infinitesimal for infinitesimal q (see Figure 3). This might suggest that the equilibrium would elicit the following bluff: make a very large number of very small buys and then reverse using a single market sell order. Absent any other trades or public announcements, the large number of buys would push the market's expectations up substantially. However, the bluffer should expect that both information, and the arrival of other traders will tend to reverse the effects of the bluff. This is because the bluffer knows that the market's expectations have been artificially pushed up, and, hence, public announcements will, on average, provide correcting information. Furthermore, other informed traders should be less sensitive to the bluff than liquidity suppliers, and will, on average, provide correcting trades. Whether or not this completely erases the expected profitability of the bluff is an open question.

The next set of questions concerns the strict dichotomy between those who supply limit orders and those who use market orders. An informed trader may have two reasons for using a market order: (1) public announcements will tend to "depreciate" the value of the information and hence "patience" is costly; (2) competing informed traders, using market orders will tend to be on the same side of the market as the informed limit order user—they will tend to move the price against the informed limit order user, and he or she will assess a relatively smaller probability of execution. In the example above, one can show that if α is small enough (there are enough informed market order users), or the depreciation rate of private information is large enough, then informed traders will prefer to use market orders.

In a similar vein, if market order users use market orders gradually, they might just as well use limit orders. Note that in the exponential-normal example, a trader who chooses a trade to equate marginal price and marginal valuation will not wish to trade again immediately if the new schedule is merely translated by the change in expectations. Even though expectations will not move all the way to the marginal price, the posttrade marginal valuation at a zero new trade will lie within the new bid-ask spread (see Figure 3). This is because the distance between the revision in expectations function, $e(q)$, and the marginal price schedule, $R'(q)$,is greatest at q equal to zero. Thus, it does not appear that market order traders will wish to trade gradually.

It is probably true that some "liquidity" traders would use limit orders. This is particularly true if access to the book were very inexpensive. If the model were to allow this, it is possible to make rough predictions about the results. Consider the discrete price analysis. There is now no longer any reason to expect a zero-profit condition to hold at every price where there is a positive quantity. A liquidity trader may be willing to experience negative expected trading profits in return for a more optimally balanced portfolio and consumption. However, if there are positive profits at some price, one would expect the patient traders to step in to remove those profits. This would suggest that the resulting marginal price function would offer larger aggregate quantities at each price than the schedule considered here. However, the arguments of Cohen, Maier, Schwartz, and Whitcomb (1981) would suggest that the positive small-trade spread would not disappear. If there is value to immediacy, certain execution with small transaction costs will dominate the uncertain execution and losses to informed traders resulting from a limit order strategy.

V. Extensions and Speculations

The assumption of a large number of "patient traders" providing limit orders is unlikely to be met in reality. After all, providing limit orders is, in fact, not costless since it requires some monitoring to insure that orders are not left exposed after, for example, a public information release. As the discussion of the discrete price case suggests, the quantity competition that results in this sort of environment does not lead to the "Bertrand" conclusion that $N = 2$ is large. Of course, if there are a small number of liquidity suppliers, then there is an incentive for others to provide quotes. It is probably cheapest, however, for such liquidity suppliers to merely join the book by providing limit orders, and compete directly with the "patient traders," rather than establish a new exchange.

The analysis here describes the equilibrium in a "full" electronic limit order book. Note, however, that the profitability of low offers is unaffected by the existence or lack of higher offers. Thus, the lowest equilibrium offer is independent of the terms of higher offers.[12]

As the uniform example illustrates, it is not difficult to come up with reasonable examples that do not conform to the "affiliation" Assumption 2. If this assumption fails to hold, it may mean that the resulting pattern of bids and offers is roughly upward sloping but involves many "flat" spots—prices at which a large quantity is bid or offered.

The restriction of the analysis to anonymous exchanges is important. It is possible, and perhaps likely, that exchange floors provide the sort of information that allows either (1) some further determination of who does and does not have information, or (2) the possibility of disciplining via future penalties those who make information-based trades. Indeed, Admati and Pfleiderer (1991) argue that (1) can occur via "sunshine trading." Benveniste and Wilhelm (1992) argue that (2) is an important role of the specialist and floor traders. Specialists insist that these other sources of information are important for the smooth running of the New York Stock Exchange. Perhaps this floor information is important for some trades, unimportant for others. An important area of research is first, to determine the importance of this other information and second, to determine if the securities industry can simultaneously enjoy the benefits of competition and liquidity that an open limit order book appears to provide with the information benefits that a floor may provide.[13]

VI. Conclusion

After setting up a reasonably general model of investor behavior, the article develops some characteristics of the equilibrium in an electronic market when there are a large number of limit order submitters. It is shown that the equilibrium involves an "upper (lower) tail" conditional expectation

[12] Note that this is not a feature of a nondiscriminating limit order book, and hence limit order submitters in a nondiscriminating book face the risk that the profitability of their orders may be harmed by changes in other orders. Thus, the discriminating order book may attract more orders than a nondiscriminating order book in the realistic case in which the book may not have time to fill up.

[13] Junius W. Peake of Peake/Ryerson has suggested in private conversation that "floor information" could be represented in an electronic market.

in the determination of offers (bids). While exhibiting a small-trade spread, the open limit order book provides as much liquidity as can be expected in extreme adverse selection environments. The article suggests that if there is a large population of potential liquidity suppliers, and if the actual costs of running an exchange are small, then among exchanges that operate continuously and anonymously, and supply nice marginal price schedules, the electronic exchange is the only one that does not tend to engender additional competing exchanges.

Simultaneous trading of equities on the London Stock Exchange (a dealer market) and the Paris Bourse (an electronic open limit order book) would seem to refute the immunity characteristics derived in this analysis. However, the structure of the London Stock Exchange provides something outside the analysis in this article. Specifically, trading on the London Stock Exchange need not be anonymous. More generally then, the results regarding competing exchanges might usefully be interpreted in the following way: with an electronic open limit order book a competing exchange may well survive, but to survive it must provide something outside of the analysis in this article.

Appendix

Proof of Lemma 1: Suppose that (i) and (ii) do not hold, and suppose there exists $q(\omega)$ such that:

$$M(q(\omega), R(q(\omega)); \omega) = R'(q(\omega)).$$

The derivative of $M(q, R(q); \omega)$ evaluated at $q(\omega)$ is:

$$(W_1^2 W_{22} + W_2^2 W_{11} - 2W_1 W_2 W_{12})/W_1^3 < 0, \quad \text{by strict quasi-concavity.}$$

Thus, since $R'(\cdot)$ is nondecreasing, if $M(q, R(q); \omega)$ and $R'(q)$ ever cross, M crosses from above (i.e., if $q < q(\omega)$, $M(q, R(q); \omega) > R'(q)$). If there is no solution, $q(\omega)$, then either condition (i) or (ii) is satisfied, or there is a discontinuity in $R'(q)$ at q_0 and M passes through this discontinuity. Since $R'(q)$ is nondecreasing, any discontinuity must involve a jump up. If M goes through this discontinuity it must do so from above. If whenever M crosses R' it does so from above, then the two functions can cross at most once and conclusion (iii) holds. Q.E.D.

Proof of Proposition 1: The derivative of $W(-R(q), q; \omega)$ with respect to q is:

$$\begin{aligned} -W_1(-R(q), q; \omega)R'(q) &+ W_2(-R(q), q; \omega) \\ &= W_1(-R(q), q; \omega)[M(q, R(q); \omega) - R'(q)]. \end{aligned}$$

After observing that W_1 is strictly positive by assumption, and applying the uniqueness results of Lemma 1, the result is immediate. Q.E.D.

Proof of Lemma 2: First note that if Y is a random variable with density f and distribution function F:

$$E[X \mid Y \geq y](1 - F(y)) = \int_y^{\infty} E[X \mid Y = t] f(t)\, dt$$

$$E[X|Y \leq y]F(y) = \int_{-\infty}^{y} E[X|Y = t] f(t)\, dt$$

Taking the derivatives of the above with respect to y shows that:

$$(d/dy)E[X \mid Y \geq y] = f(y)\{E[X \mid Y \geq y] - E[X \mid Y = y]\}/(1 - F(y)),$$
$$(d/dy)E[X \mid Y \leq y] = f(y)\{E[X \mid Y = y] - E[X \mid Y \leq y]\}/F(y).$$

Given Assumption 2 with $Y = M(q, R; \omega)$ shows that $V(m, q, R)$ and $v(m, q, R)$ are increasing in m. For the second part of the proposition, define $Q_{Rm}(\omega)$ as the optimal trade of an investor with characteristic vector ω but with cash position reduced by R facing a fixed price m for any quantity. Such a "schedule" is nondecreasing, and hence by Corollary 1 above:

$$E[X|Q_{Rm}(\omega) \geq q] = E[X|M(q, R + qm; \omega) \geq m] = V(m, q, R + qm)$$
$$E[X|Q_{Rm}(\omega) \leq q] = E[X|M(q, R + qm; \omega) \leq m] = v(m, q, R + qm).$$

Also, $E[X \mid Q_{Rm}(\omega) = q] = E[X \mid M(q, R + qm; \omega) = m]$. Thus, we have by Assumption 2:

$$E[X|Q_{Rm}(\omega) \geq q] \geq E[X|Q_{Rm}(\omega) = q] \geq E[X|Q_{Rm}(\omega) \leq q].$$

By the demonstration above, both $E[X \mid Q_{Rm}(\omega) \geq q]$ and $E[X \mid Q_{Rm}(\omega) \leq q]$ are increasing in q. That is, both $V(m, q, R + qm)$ and $v(m, q, R + qm)$ are increasing in q. Q.E.D.

Proof of Proposition 2: The analysis will deal with the derivation of the equilibrium on the offer side. The analysis for the bid side can be easily derived from this analysis. Let q_i be the quantity offered at the ith price by a typical liquidity supplier. Let Q_i be the total quantity offered by all N liquidity suppliers at the ith price, and let AQ_i be the total quantity offered by all N liquidity suppliers at the ith price and lower. Finally, define R_i by $R_i = p_0 Q_0 + \cdots + p_i Q_i$,the amount paid for a purchase of AQ_i. Since the set of allowable prices is discrete, the marginal price function will be a step function. Thus, even if cross-sectionally the marginal valuation functions are continuously distributed, the probability that D, the quantity traded at the next arrival, is equal to AQ_i may be positive. In particular:

$$P\{D = AQ_I\} = P\{p_i \leq M(AQ_i, R_i; \omega) < p_{i+1}\}.$$

Denote the density of D, for $AQ_{i-1} < d < AQ_i$ by $f_i(d)$. Note that the above probabilities and densities are functions of the actual bids and offers provided. If a trade arrives strictly between AQ_{i-1} and AQ_i, then the excess over AQ_{i-1} needs to be allocated among those supplying offers. It is assumed that the allocation is pro rata according to the size of the offer provided. With this specification, the expected profit to the liquidity supplier who offers $\{q_i\}$ while others offer $\{Q_i - q_i\}$ is:

$$\sum_{i=0}^{\infty} q_i(p_i - E[X \mid D \geq AQ_i)P\{D \geq AQ_i\} + \sum_{i=0}^{\infty}\left[\frac{q_i}{Q_i}\right]\int_{AQ_{i-1}}^{AQ_i}(d - AQ_{i-1})(p_i - E[X \mid D = d])f_i(d) \tag{A1}$$

If a liquidity supplier offers q_i, then all of this quantity will be transacted at price p_i if a trade comes in for AQ_i or greater. If this happens, the revised value of the share is $E[X \mid D \geq AQ_i]$, and this happens with probability $P\{D \geq AQ_i\}$. If a trade comes in for an amount strictly between AQ_{i-1} and AQ_i, say d, then $d - AQ_{i-1}$ will be allocated in a pro rata fashion. The revised expectation will be $E[X \mid D = d]$. Integrating over all such d's weighted by the density provides the expected profit in this event. Sum over all possible prices to obtain the expected profit from the choice of q's. To obtain the first-order condition that Q_i must satisfy, take the derivative of the above expression with respect to q_i. This yields:

$$\begin{aligned}(p_i - E[X \mid D \geq AQ_i])P\{D \geq AQ_i\} &+ \frac{Q_i - q_i}{Q_i^2} \\ &\times \int_{AQ_{i-1}}^{AQ_i}(d - AQ_{i-1})(p_i - E[X \mid D = d])f_i(d) \\ &+ \sum_{j=1}^{\infty}\left\{q_j\frac{d}{dq_i}\left(p_j - E\left[X \mid D \geq AQ_j\right]\right)P\{D \geq AQ_j\}\right. \\ &\left. + \frac{q_j}{Q_j}\frac{d}{dq_i}\int_{AQ_{j-1}}^{AQ_j}(d - AQ_{j-1})\left(p_j - E[X \mid D = d]\right)f_i(d)\right\} = 0\end{aligned} \tag{A2}$$

This condition must hold for all liquidity suppliers, so sum this derivative over all liquidity suppliers and divide by N.

If $Q_i > 0$ but finite:

$$\begin{aligned}(p_i - E[X \mid D \geq AQ_i])P\{D \geq AW_i\} \\ + \frac{N-1}{N}\frac{1}{Q_i}\int_{AQ_{i-1}}^{AQ_i} \times (p_i - E[X \mid D = d])(d - AQ_{i-1})f_i(d) + \frac{K}{N} = 0\end{aligned} \tag{A3}$$

The term K/N indicates a number of individual terms reflecting the effect of adding a unit of quantity more at p_i on the probability and profitability of trades larger than AQ_{i-1}. As N gets large, this term vanishes. After integrating the second term in equation (3) by parts, substituting $V(p_i, d, R_{i-1}$

$+ p_i(d - AQ_{i-1}))$ for $E[X \mid D \geq d]$ and ignoring terms of order $1/N$, it is found that if $Q_i > 0$, but finite:

$$\int_{AQ_{i-1}}^{AQ_i} (p_i - V(p_i, d, R_{i-1} + p_i(d - AQ_{i-1})) \times P\{M(d, R_{i-1} + p_i(d - AQ_{i-1}) \geq p_i\} = 0 \qquad \text{(A4)}$$

By Lemma 2, if $p_k < V(p_k, 0, 0)$, then $p_k < V(p_k, q, qp_k)$ for all positive q, and the first-order condition can never be satisfied at p_k. The second-order condition at a price p_i with a positive quantity is found by taking the derivative of the initial first-order condition, equation (A2), summing across all liquidity suppliers, dividing by N and ignoring terms of order $1/N$. This yields: $P\{D_j \geq AQ_j\}(p_j - V(p_j, AQ_j, R_j))/Q_j < 0$. The results of Lemma 2 imply that a point that satisfies the first-order condition also satisfies the second-order condition.

Proof of Proposition 3: The first inequality in (i) follows immediately from the definition of A_1. The second inequality follows from Assumption 2, and the third inequality follows from the analogous definition of B_1. The second set of inequalities follow from Assumption 2 and continuity. The same arguments apply for part (ii). Q.E.D.

Proof of Proposition 4: The analysis will deal with the ask side; the analysis for the bid side is completely analogous. The limit as q goes to zero from above of $R'(q)$ is:

$$R'_+(0) = \inf\{p: p > V(p, 0, 0)\} \quad \text{if the set is nonempty.} \qquad \text{(A5)}$$

If the set is empty, there are no offers provided. Now suppose that offers totaling Q are available. The following limiting argument will indicate the conditions that $R'(Q)$ and $R(Q)$ must satisfy. Suppose that $R'(Q) + \varepsilon$ is the next allowable price, and further that a positive quantity will be offered at $R'(Q) + \varepsilon$. Following the development above, this implies that:

$$R'(Q) + \varepsilon > V(R'(Q) + \varepsilon, Q, R(Q)). \qquad \text{(A6)}$$

Let the quantity offered at $R'((Q) + \varepsilon$ be εq. Then the first-order condition must be equal to 0:

$$\frac{\displaystyle\int_{Q}^{Q+\varepsilon q} (R'(Q) + \varepsilon - V(R'(Q) + \varepsilon, t, R(Q) + (t - Q)(R'(Q) + \varepsilon))) \times G(Q, t, R'(Q) + \varepsilon)\, dt}{\varepsilon q}$$

Where $G(Q, t, R'(Q) + \varepsilon) = P\{M(t, R(Q) + (t - Q)(R'(Q) + \varepsilon); \omega) > R'(Q) + \varepsilon\}$. Taking the limit as ε goes to zero yields:

$$R'(Q) = V(R'(Q), Q, R(Q)).$$

It is also required that $1 > [V(R'(Q) + \varepsilon, Q, R(Q)) - R'(Q)]/\varepsilon$ (from equation (A6). Taking limits yields the additional condition that $(\partial/\partial p)V(p, Q, R(Q))|_{p=R'(Q)} \leq 1$. There may still be more than one solution. The solution is pinned down by condition (A5).

Proof of Proposition 5: For any $R(\cdot)$ the expected profit is:

$$\int_{-\infty}^{\infty} dP\{Q_R \leq q\}(R(q) - qE[X \mid Q_R = q])$$

Integrate by parts to get:

$$\int_{0}^{\infty} P\{Q_R \geq q\}(R'(q) - E[X \mid Q_R \geq q])\, dq$$
$$+ \int_{-\infty}^{0} P\{Q_R \leq q\}(E[X \mid Q_R \leq q] - R'(q))\, dq$$

This follows since:

$$(d/dq)P\{Q_R \geq q\}E[X \mid Q_R \geq q] = -E[X \mid Q_R = q](d/dq)P\{Q_R \leq q\}, \quad \text{and}$$
$$(d/dq)P\{Q_R \leq q\}E[X \mid Q_R \leq q] = E[X \mid Q_R = q](d/dq)P\{Q_R \leq q\}$$

(see the proof of Lemma 2). Under the hypothesis of the proposition,

$$m < V(m,0,0) \leq V(m,q,R(q)) \text{ for all } m.$$

The second inequality follows from Lemma 2 and the absence of wealth effects. Then in particular, for the $R'(q)$ considered here:

$$R'(q) < V(R'(q), q, R(q)) = E[X|M(q, R(q); \omega) \geq R'(q)] = E[X \mid Q_R \geq q],$$

where the last equality follows from the single crossing property. Thus, this R leads to negative expected profits. Q.E.D.

Proof of Proposition 6: Under the normalization chosen above, the certainty equivalent of a trader of type ω, making optimal trade $q(\omega)$ is given by: $CE(\omega) = \omega q(\omega) - 0.5q(\omega)^2 - R(q(\omega))$. The derivative of this is given by:

$$CE'(\omega) = q(\omega) + q'(\omega)(\omega - q(\omega) - R'(q(\omega))) = q(\omega)$$

since $q(\omega)$ satisfies the first-order condition for optimality. Since the certainty equivalent is zero when the optimal quantity traded is zero, the certainty equivalent evaluated at ω is the integral from 0 to ω of $q(t)$. A monopolist will set a marginal price schedule so that the quantity traded by a investor of type ω is given by:

$$\begin{aligned} q_m(\omega) &= \alpha\omega - (1 - F(\omega))/f(\omega),\ \omega > \omega_m \\ &= \alpha\omega + F(\omega)/f(\omega),\ \omega < -\omega_m \\ &= 0 \quad \text{otherwise;} \end{aligned}$$

where $\alpha\omega_m - (1 - F(\omega_m)0/f(\omega_m) = 0$. The full details of this derivation are in Glosten (1989); a sketch is provided below. In contrast, the electronic market determines $q_e(\omega)$ as:

$$\begin{aligned} q_e(\omega) = {} & \omega - (1-\alpha)f(\omega)/(1-F(\omega)),\ \omega > \omega^* \\ & \omega + (1-\alpha)f(\omega)/F(\omega),\ \omega < -\omega^* \\ & 0, \quad \text{otherwise.} \end{aligned}$$

It can be shown that $0 < f(t)[(f(t)/(1-F(t))) - t]/(1-F(t)) < 1$. Hence, $\omega^* < \omega_m$ and for $\omega > \omega^*$, $q_e(\omega) > q_m(\omega)$ and for $\omega < -\omega^*$, $q_e(\omega) < q_m(\omega)$. Thus, for ω outside of $[-\omega^*, \omega^*]$ the certainty equivalent is strictly larger with the electronic market.

Derivation of Monopolist Solution in Proposition 6

If the monopolist chooses a marginal price schedule $R'(\cdot)$, then the choice of a trader, $Q(z)$, is determined by $R'(Q(z)) = z - Q(z)$. The monopolist's problem on the offer side is to maximize:

$$\int_{t_m}^{\infty} [R(Q(z)) - (1-\alpha)zQ(z)]f(z)\,dz$$

Integrate the first term by parts, noting that $R(Q(t_m)) = 0$, and $R'(Q(z)) = z - Q(z)$ to get:

$$\int_{t_m}^{\infty} (1 - F(z))Q'(z)(z - Q(z)) - (1-\alpha)zQ(z)f(z)\,dz$$

Integrate the first term by parts again to get:

$$\int_{t_m}^{\infty} f(z)\left[\alpha z Q(z) - \frac{1-F(z)}{f(z)}Q(z) - 0.5Q(z)^2\right]dz$$

The maximizing $Q(z)$ is as claimed.

Proof of Proposition 7: Call $R_c'(q)$ the marginal price schedule in the competing market, and Q_c, a random variable, the next trade at the competing exchange. After integrating by parts, the expected profit to the entrant is:

$$\int_0^{\infty} P\{Q_c \geq q\}(R_c'(q) - E[X \mid Q_c \geq q])\,dq$$

$$+\int_{-\infty}^{0} P\{Q_c \leq q\}(E[X \mid Q_c \leq q] - R_c'(q))\,dq$$

Where Q_c is the quantity chosen in the entering market. Consider only the offer side. If $R_c'(0) > R_e'(0)$ then if $Q_c > 0$, $Q_e > 0$, where Q_e is the quantity chosen in the electronic market. Furthermore, $R_c'(Q_c) = R_e'(Q_e)$ and hence $Q_e = R_e'^{-1}(R_c'(Q_c)$. To simplify the notation, define $q_e = R_e'^{-1}(R_c'(q))$, $q_T = q_e + q$ and $R_T = R_c(q) + R_e(q_e)$. That is, q_e is the trade made in the electronic market when q is traded in the competitive market, q_T is the total trade, while R_T is the total amount paid for a purchase of q_T shares. By the single crossing property, the events $\{Q_c \geq q\} = \{Q_e \geq q_e\}$, and furthermore:

$$\begin{aligned} E[X \mid Q_c \geq q] &= E[X \mid Q_e \geq q_e] \\ &= E[X | M(q_T, R_T; \omega) \geq R'_c(q)] \\ &= V(R'_c(q), q_T, R_T) = V(R'_e(q_e), q_T, R_T) \\ &\geq V(R'_e(q_e), q_e, R_e(q_e)) = R'_e(q_e) = R'_c(q). \end{aligned}$$

The last inequality follows from the fact that $q_T > q_e$ and the use of Lemma 2 in the case of no wealth effects. For any q such that $R'_c(q) \geq R'_e(0)$, the term in the integral is nonpositive. Suppose that $R'_c(q) < R'_e(0)$. Then, for some q:

$$\begin{aligned} R'_c(q) - E[X \mid Q \geq q] &= R'_c(q) - E[X | M(q, R_c(q); \omega) \geq R'_c(q)] \\ &= R'_c(q) - V(R'_c(q), q, R_c(q)). \end{aligned}$$

Since $R'_c(q) < R'_e(0)$ and $V(R'_c(q), q, R_c(q)) \geq V(R'_c(q), 0, 0))$ this term is not positive since $R'_e(0)$ is the smallest m with $m \geq V(m, 0, 0)$. Q.E.D.

Proof of Proposition 8: If both q_c and q_e are positive for some q, then they are determined by: $q_c + q_e = q$ and $R'_c(q_c) = R'_e(q_e)$. Thus, $R'_T(q)$ equals $R'_e(q_e)$, and:

$$\begin{aligned} R'_T(q) = R'_e(q_e) &= E[X | M(q, R_T(q); \omega) \geq R'_e(q_e)] \\ &= V(R'_e(q_e), q_T, R_T(q)) = V(R'_T(q), q, R_T(q)). \end{aligned}$$

That is, $R_T(q)$is a solution to $R'_T(q) = V(R'_T(q), q, R_T(q))$. One such solution is the electronic open limit order book solution, $R(q)$. The entrant cannot set $R'_c(0) < R'_e(0)$ and expect to make nonnegative profits, for if he or she did, some marginal prices would be below upper tail expectations while other marginal prices would equal upper tail expectations. Q.E.D.

Proof of Proposition 9: Suppose, without loss of generality, that the schedule diverges from the electronic exchange schedule on the offer side. If

$$\int_0^\infty P\{Q \geq q\}(R'(q) - E[X \mid Q \geq q]) \geq 0$$

but $R'(q)$ is not the electronic exchange marginal price schedule, then there exists q^* with $R'(q^*) > E[X \mid Q \geq q^*]$. Consider the following strategy of an entrant. Set $P = R'(q^*)$ and announce that up to Q units will be sold at price P. The expected profit from this strategy is:

$$QP\{Q_c = Q\}(P - E[X \mid Q_c = Q] + \int_0^Q dP\{Q_c < q\}q(P - E[X \mid Q_c = q])$$

where Q_c is the random quantity picked in the competing market. From the investors maximization problem $\{\omega \mid Q_c = Q\} = \{\omega \mid M(Q + q^*, R(q^*) + PQ; \omega) \geq P\}$, and hence $E[X \mid Q_c = Q] = V(P, Q + q^*, R(q^*) + PQ)$. Divide the expression for profits by Q and let Q go to zero. The first term vanishes, the second becomes $(P - V(P, q^*, R(q^*))P\{M(q^*, R(q^*); \omega) \geq P\} = (R'(q^*) - V(R'(q^*), q^*, R(q^*))P\{M(q^*, R(q^*); \omega) \geq R'(q^*)\} > 0$. Thus, for some $Q > 0$ expected trading profits will be positive. Q.E.D.

Determination of the Equilibrium in Figure 1

The function $V(m, q, R(q)) = (1 - \alpha)f(m + q)/(1 - F(m + q))$ where $f(\cdot)$ and $F(\cdot)$ are, respectively, the standard normal density and distribution. A1 is the solution to $A1 = \min\{p\colon p > (1 - \alpha)f(p)/(1 - F(p))\}$ and is found to be 0.25. The first-order condition to determine Q_i is:

$$\int_{AQ_{i-1}+p_i}^{AQ_{i-1}+Q_i+p_i} p_i(1 - F(t)) - (1 - \alpha)f(t)\,dt = 0$$

The integral can be evaluated as:

$$\begin{aligned}\{(1 - F(p_i + AQ_{i-1} + Q_i))(1 - \alpha + p_i(p_i + AQ_{i-1} + Q_i\\ -f(p_i + AQ_{i-1} + Q_i)/(1 - F(p_i + AQ_{i-1} + Q_i)))\}\\ -\{(1 - F(p_i + AQ_{i-1}))(1 - \alpha + p_i(p_i + AQ_{i-1} - f(p_i + AQ_{i-1})/\\ (1 - F(p_i + AQ_{i-1})))\} = 0.\end{aligned}$$

Letting $\alpha = 0.8$, $AQ_{i-1} = 0$, and $p_i = 0.25$, the above can be solved numerically to get Q1 = 0.98. The next highest offer, A2 is the solution to $A2 = \min\{p\colon p > (1 - \alpha)f(p + 0.98)/(1 - F(p + 0.98))\}$ and is found to be 0.375. Letting $AQ_{i-1} = 0.98$ and $p_i = 0.375$ and solving numerically, leads to Q2 = 0.14. The next highest offer is the solution to $A3 = \min\{p\colon p > (1 - \alpha)f(p + 1.12)/(1 - F(p + 1.12))\}$ and is found to be 0.5. Letting $AQ_{i-1} = 1.12$ and $p_i = 0.5$, it can be verified that the left side of the first-order condition is positive for all Q_i, that is $-(1 - F(1.62))(0.2 + 0.5(1.62 - f(1.62)/(1 - F(1.62)) > 0$, and hence the equilibrium involves an infinite amount offered at 0.5.

REFERENCES

Admati, A., and P. Pfleiderer, 1991, Sunshine trading and financial market equilibrium, *Review of Financial Studies* 4, 443–481.

Benveniste, Lawrence M., Alan J. Marcus, and William J. Wilhelm, 1992, What's special about the specialist? *Journal of Financial Economics* 32, 61–86.

Black, Fischer, 1991, Trading in equilibrium with bluffing, credits and debits, Working paper, Goldman Sachs, New York, NY.

———, 1992, Equilibrium exchanges, working paper, Golman Sachs, New York, NY.

Cohen, K., S. Maier, R. Schwartz, and D. Whitcomb, 1981. Transactions costs, order placement strategy and the existence of the bid-ask spread, *Journal of Political Economy* 89, 287–305.

Domowitz, Ian (1991). Automating continuous double auctions in practice: automated trade execution in financial markets, Working paper, Northwestern University.

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

Gale, Ian, 1991. Sequential trade and information revelation, Working paper, University of Wisconsin.

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

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

———, 1991, Asymmetric information, the third market and investor welfare, Working paper, Columbia University.

Harris, Lawrence, 1990, Liquidity, trading rules and electronic trading systems, Monograph Series in Finance and Economics, New York University Salomon Center, New York.

Hellwig, Martin, 1980, On the aggregation of information in competitive markets, *Journal of Economic Theory* 22, 477–498.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.

Leach, J. C., and A. N. Madhavan, 1993, Price experimentation and security market structure, *Review of Financial Studies* 6, 375–404.

Milgrom, Paul R., and Robert Weber, 1982. A theory of auctions and competitive bidding, *Econometrica* 50, 1089–1122.

Rock, Kevin, 1989. The specialist's order book, Working paper, Harvard University.