Security Market Microstructure

Lecture 4: The Inventory Models and the Earlier Thoughts about Informational Trading

Instructor: Chyi-Mei Chen
Room 1102, Management Building 2
(Tel) 3366-1086
(Email) cchen@ccms.ntu.edu.tw

1. Most earlier models of dealership markets are inventory-based models. We shall review these inventory models first, and then mention briefly the information-based models.

2. Demsetz (1968) seems to be the first work emphasizing order imbalance. In that paper, the author emphasizes that trading has a time dimension and at each point in time, there may be an imbalance in buy and sell orders. The presence of dealers allows submitted orders to be executed immediately. The bid-ask spread is therefore a price that public investors must pay in order to obtain *immediacy* in order execution.

3. Unlike Demsetz (1968) where the focus is on the trading desires of individual traders, Garman (1976) switches the spot light to market clearing mechanisms. Hence Garman's research is considered by many to be the first formal analysis of market microstructure. Two mechanisms were considered in Garman's paper, a double auction and a monopolistic dealer exchange. With the latter mechanism, a monopolistic dealer must first set an ask price and a bid price, receive and execute orders from public traders, and the bid and ask prices are chosen to maximize the dealer's profit at each point in time subject to the constraint that bankruptcy or failure must not take place (which would occur if the dealer ran out of the traded security or cash). In Garman's model, the bid-ask spread exists in order that specialist will not be ruined with probability one. That is, market viability dictates the existence of a bid-ask spread.

   Garman's model exhibits the following features:
   -monopolistic risk neutral specialist;

-liquidity traders only;
-market orders following Poisson processes;
-specialist can only set prices once and for all;
-specialist is facing the 'ruin' problem.

More precisely, let $I_c(t)$ and $I_s(t)$ be the dealer's cash balance and stock inventory at time t, and $N_a(t)$ and $N_b(t)$ be respectively the numbers of shares sold to and bought from public traders during the time interval $[0, t]$. Let $I_c(0)$ and $I_s(0)$ be given. Then we have

$$I_c(t) = I_c(0) + p_a N_a(t) - p_b N_b(t),$$

and

$$I_s(t) = I_s(0) + N_b(t) - N_a(t).$$

Let $\lambda_a(p_a)$ and $\lambda_b(p_b)$ be the parameters of respectively the buy order and sell order Poisson processes (which are taken as primitives instead of being derived from individual optimization problems), with $\lambda'_a < 0 < \lambda'_b$. The dealer's cash inflow and outflow processes are hence also Poisson, with parameters $p_a\lambda_a$ and $p_b\lambda_b$ respectively.

Let $Q_k(t)$ be the probability that the dealer has k units of cash at time t, and $R_k(t)$ be the probability that the dealer has k units of stock at time t. For sufficiently small time interval Δt, we have from the assumed Poisson processes[1]

$$Q_k(t) = Q_{k-1}(t - \Delta t)[\lambda_a(p_a)p_a\Delta t][1 - \lambda_b(p_b)p_b\Delta t]$$

$$+Q_{k+1}(t - \Delta t)[\lambda_b(p_b)p_b\Delta t][1 - \lambda_a(p_a)p_a\Delta t]$$

$$+Q_k(t - \Delta t)[1 - \lambda_a(p_a)p_a\Delta t][1 - \lambda_b(p_b)p_b\Delta t]$$

implying that

$$\frac{\partial Q_k(t)}{\partial t} = Q_{k-1}(t)[\lambda_a(p_a)p_a] + Q_{k+1}(t)[\lambda_b(p_b)p_b] - Q_k(t)[\lambda_a(p_a)p_a + \lambda_b(p_b)p_b],$$

[1]This expression was copied from Maureen O'Hara's 1995 book. Note, however, that unless $p_a = p_b = 1$, the change in cash balance would not be plus or minus one when the dealer absorbs a buy or sell order at time t.

and hence (with $\overline{p}$ denoting the mean of p_a and p_b)

$$\lim_{t\to+\infty} Q_0(t) \begin{cases} \sim [\frac{\lambda_b(p_b)p_b}{\lambda_a(p_a)p_a}]^{\frac{I_c(0)}{\overline{p}}}, & \text{if } \lambda_a(p_a)p_a > \lambda_b(p_b)p_b; \\ 1, & \text{otherwise.} \end{cases}$$

Similarly, one can show that

$$\lim_{t\to+\infty} R_0(t) \begin{cases} \sim [\frac{\lambda_a(p_a)}{\lambda_b(p_b)}]^{I_s(0)}, & \text{if } \lambda_a(p_a) < \lambda_b(p_b); \\ 1, & \text{otherwise.} \end{cases}$$

Hence Garman concludes that to avoid a sure failure, it must be that

$$p_a\lambda_a > p_b\lambda_b$$

and

$$\lambda_a < \lambda_b.$$

This implies a spread.

The intuition is rather clear: in order for the dealer to not run out of the stock aymptotically, the dealer should expect more sell orders than buy orders in a given period of time, and hence we need $\lambda_b > \lambda_a$; but if $p_a \leq p_b$, then the dealer should expect to run out of cash asympotically, so that we must also require that $p_a > p_b$. In other words, to guarantee that the dealer can make the market indefinitely, the dealer that starts with a fixed amount of cash and a fixed number of shares of the stock at time 0 must set the ask price higher than the bid price.

4. Amihud and Mendelson (1980) extend Garman's study to allow price adjustments. Prices change according to inventory positions. The main features of Amihud and Mendelson's model include:
-monopolistic risk neutral specialist;
-liquidity traders only;
-market orders following Poisson processes;
-specialist can set prices continuously where the inventory is the state process;
-specialist does not face the 'ruin' problem; that is, the inventory has exogenous upper and lower bounds.

Unlike in Garman (1976), spread in Amihud and Mendelson simply reflects monopoly power of the specialist. One can show that the spread tends to zero, as competition gets in and increases. Inventory plays the role of a 'buffer.' Amihud and Mendelson obtain three main results: (i) bid and ask decrease with a dealer's inventory; (ii) there exists an optimal inventory level, and departing from this level induces a dealer to adjust bid and ask prices in order that inventory can move back to that level; (iii) spread is always positive.

5. Stoll (1978) takes a different view about the bid-ask spread: market makers are those selling insurance to liquidity traders, and the spread is risk premium. Stoll's model includes the following features:
-market makers are competitive;
-market makers are risk averse;
-two-period model, similar to the CAPM setting;
-assets have 'true prices' observable to market makers;
-spread is nothing but compensation for market makers taking positions that make their portfolio deviate from their ideal positions;
-all traders are liquidity traders;
-no ruin problems for a dealer.

Stoll's observation is that, in order for a dealer to be willing to perform his or her function, engaging in market making should not lower his or her satisfaction. Hence, the bid-ask spread can be derived so that a dealer feels indifferent about to or not to engage in market making. Then, Stoll obtained a formula for the spread, and demonstrated that the dealer's initial wealth, risk attitude, inventory position may all affect the spread.

Consider M risk-averse market makers operating at date 0 for N risky securities. Borrowing and lending at zero interest rate is allowed to public traders also at date 0. Security i pays a dividend $\tilde{v}_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$ per share at date 1. A market maker is endowed with a non-random wealth W_0 at date 0, and he seeks to maximize $E[-e^{-A\tilde{W}}]$, where $\tilde{W}$ is his date-1 terminal wealth, and $A > 0$ his measure of absolute risk aversion. Let $p_i(x_i)$ be the date-0 price for security i when the market maker is called upon to absorb a market order x_i in security i. In the absence of the market order, each market maker's date-0 portfolio given

$(p_1(0), p_2(0), \cdots, p_N(0))$ is denoted by $(q_1^*, q_2^*, \cdots, q_N^*)$.

Now, suppose that exactly one market maker is called upon to absorb an order of $x_i = q_i^* - q_i$, so that his position in security i becomes

$$q_i^* - (q_i^* - q_i) = q_i,$$

then in order to keep him indifferent about taking or not taking this order, the price of security i must adjust so that

$$E[-e^{-A[\sum_{j=1}^{N}(\tilde{v}_j - p_j(0))q_j^*]}] = E[-e^{-A[\sum_{j\neq i}(\tilde{v}_j - p_j(0))q_j^* + (q_i^* - x_i)(\tilde{v}_i - p_i(x_i))]}],$$

and hence for all $j = 1, 2, \cdots, N$,[2]

$$p_j(x_j) = p_j(0) + \frac{A}{2} x_j \sigma_j^2.$$

It is clear that for all $b > 0 > s$, the ask price $p_j(b)$ exceeds $p_j(0)$, which exceeds the bid price $p_j(s)$. Given the order size $|x|$, the bid-ask spread

$$p_j(|x|) - p_j(-|x|) = A|x|\sigma_j^2,$$

which increases with $|x|$, A, and σ_j^2 for obvious reasons.

6. In the preceding section, we have assumed that each market maker starts with a non-random initial wealth W_0 at date 0. Now, assume that each market maker starts with a position I_i in security i. Upon absorbing an order x_i, a market maker's position in security i becomes $I_i - x_i$. In this case, using the results obtained in the preceding section, we have

$$p_j(I_j) = p_j(0) - \frac{A}{2} I_j \sigma_j^2,$$

and

$$p_j(x_j) = p_j(I_j) + \frac{A}{2} x_j \sigma_j^2.$$

[2]It can be easily verified that the following equation does not hold in general. Stoll (1978) used several mutually inconsistent assumptions before reaching such a result. The CAPM equation is assumed to hold always (even though dealer i is not allowed to unload x_i to other dealers), and x_i is so small that a linear approximation for $-e^{-A\tilde{W}}$ is valid. To see that the following formula makes little sense, note that x_i will affect the the variance of $\tilde{W}$ by changing its co-variance components generated by the positions taken in the risky assets.

Thus the market maker's quotes decrease with his initial inventory, other things equal. This is very intuitive: given a larger I_j, the risk-averse market maker becomes more eager to reduce the inventory of the stock, and thus wish to reduce more the ask price (which helps attract more buy orders) and the bid price (which helps attract fewer sell orders).

7. In Stoll (1978), there is one single post-trade period, and hence the dealer does not face the uncertainty over 'how long' he must hold the inefficient portfolio. Ho and Stoll (1981) extend Stoll (1978) into a continuous time framework. Three results are obtained with this innovation: (i) Spread depends on the time horizon of the dealer–the risk reduces when the remaining time is less. In the limiting case where time remaining is essentially zero, the dealer behaves as if he were risk neutral; (ii) spread depends also on the size of trade, risk attitude of the dealer, and the riskiness of the transacted security; (iii) spread placement is related to inventory but not to the size of the spread.[3]

8. Stoll (1978) assumes that there is one single market maker that will absorb the newly arriving market order after the securities markets have reached a competitive equilibrium.[4] Now we consider the case where the M competitive dealers can share the task of aborbing the arriving market order. We shall consider a call market that adopts a uniform pricing rule.[5]

   If the dealers or market makers behave as price-takers, then such a model is essentially the same as an REE model. Let us deal with this case in this section, and take up the case of strategic dealers in the next section. We assume from now on that in addition to the riskless asset paying zero interest, there is a single risky security traded at date 0, which pays a one-time dividend $\tilde{v} \sim \mathrm{N}(\mu, \sigma^2)$.

[3]The next three sections are copied from Chapter 5 of De Jong and Rindi (2009).

[4]Note that the single market maker absorbing the market order has an incentive to trade with the rest $M-1$ dealers in order to improve risk bearing. By symmetry, trade will occur so that the M dealers will hold the same portfolio in the new equilibrium. Hence Stoll's approach does not ensure a new equilibrium of the securities markets.

[5]In a limit-order book market or a continuous dealership market a discriminatory pricing rule is typical. We shall discuss the latter in a subsequent note.

Suppose that there are M dealers, with I_j being dealer j's endowed position in the traded risky security at date 0. Again dealers have the preferences described in Stoll (1978), with A being their common measure of absolute risk aversion. Let $\tilde{x}$ be the market order submitted by public traders at date 0, and q_j the net trade by dealer j at date 0.

Given the date-0 security price p, dealer j seeks to

$$\max_q E[-e^{-A[I_j\tilde{v}+(\tilde{v}-p)q_j]}]$$

so that dealer j's limit oder or demand function is

$$q_j = \frac{\mu - p}{A\sigma^2} - I_j,$$

where the second term is the inventory effect, or more generally, the hedging demand.

The market-clearing condition requires that

$$\sum_{j=1}^{M} q_j + \tilde{x} = 0,$$

and hence the equilibrium price functional is

$$\tilde{p} = \mu - A\sigma^2[\frac{\tilde{x} - \sum_{j=1}^{M} I_j}{M}].$$

Thus the equilibrium net trade for dealer j is

$$q_j = \frac{\sum_{j=1}^{M} I_j}{M} - I_j - \frac{\tilde{x}}{M},$$

or equivalently, dealer j's equilibrium security holding is

$$I_j + q_j = \frac{[\sum_{j=1}^{M} I_j] - \tilde{x}}{M}.$$

In equilibrium, given the order size $|x|$, the ask price and bid price are respectively

$$p^A = \psi + \frac{A\sigma^2}{M}|x|,$$

$$p^B = \psi - \frac{A\sigma^2}{M}|x|,$$

where the mid-quote

$$\psi = \mu - A\sigma^2 \frac{\sum_{j=1}^M I_j}{M}$$

decreases with the dealers' average initial inventory. The bid-ask spread increases with A, σ^2 and $|x|$, and it decreases with M, because a larger M alleviates risk bearing for each single dealer.

We have assumed thus far that $\tilde{x}$ is an exogenous random variable. Now, let us assume that it is submitted by a rational investor X endowed with initial position $\tilde{L}$ in the risky security, with a CARA utility function for date-1 wealth, and with a signal $\tilde{s} = \tilde{v} + \tilde{e}$, where $(\tilde{v}, \tilde{e}, \tilde{L})$ are totally independent and jointly Gaussian with means $\mu_v > 0$, 0 and 0, and with variances σ_v^2, σ_e^2, and σ_L^2.

Now, let x be investor X's date-0 net trade in the risky security and

$$\tilde{W}_x = (x + \tilde{L})\tilde{v} - xp$$

be investor X's date-1 terminal wealth. It follows that[6]

[6]The following formulae, which can be found on page 13 of Brunnermeier's book, are used to derive the conditional moments below:

$$\mathrm{E}[\tilde{x}|\tilde{s}_j = s_j; j = 1, 2, \cdots, n] = \mu_x + \frac{1}{\tau_x + \sum_{j=1}^n \tau_j} \sum_{j=1}^n \tau_j (s_j - \mu_x),$$

$$\mathrm{var}[\tilde{x}|\tilde{s}_j = s_j; j = 1, 2, \cdots, n] = \frac{1}{\tau_x + \sum_{j=1}^n \tau_j},$$

where

$$\tilde{s}_j = \tilde{x} + \tilde{e}_j, \ \forall j = 1, 2, \cdots, n,$$

and $(\tilde{x}, \tilde{e}_1, \tilde{e}_2, \cdots, \tilde{e}_n)$ are totally independent and jointly normally distributed, with

$$\mathrm{E}[\tilde{x}] = \mu_x, \ \mathrm{var}[\tilde{x}] = \frac{1}{\tau_x},$$

$$\mathrm{E}[\tilde{e}_j] = 0, \ \mathrm{var}[\tilde{e}_j] = \frac{1}{\tau_j}, \ \forall j = 1, 2, \cdots, n.$$

$$E[\tilde{W}_x|s,L] = (x+L)[\mu_v + \frac{\tau_e}{\tau_v+\tau_e}(s-\mu_v)] - xp,$$

$$\text{var}[\tilde{W}_x|s,L] = \frac{(x+L)^2}{\tau_v+\tau_e},$$

where τ_v and τ_e are the precisions of $\tilde{v}$ and $\tilde{e}$ respectively.

Let ρ be investor X's measure of absolute risk aversion. Then investor X seeks to

$$\max_x E[\tilde{W}_x|s] - \frac{\rho}{2}\text{var}[\tilde{W}_x|s].$$

The first-order condition gives rise to investor X's date-0 demand for the risky security, which is

$$x = \frac{1}{\rho}[(\tau_v+\tau_e)(\mu_v-p) + \tau_e(s-\mu_v)] - L.$$

Note that in equilibrium the uninformed dealers know p, and they expect correctly that the functional form of p is

$$p = \Gamma + \lambda x,$$

for some positive constants Γ and λ, and hence dealers can recover x, or the sufficient statistic

$$\Omega \equiv \tilde{s} - \frac{\rho}{\tau_e}\tilde{L} = \tilde{v} + \tilde{u},$$

where the noise term

$$\tilde{u} = \tilde{e} - \frac{\rho}{\tau_e}\tilde{L}.$$

Thus in equilibrium dealer j's demand must solve the following maximization probelm:

$$\max_{q_j} E[W_j|\Omega,p] - \frac{A}{2}\text{var}[W_j|\Omega,p],$$

which implies that

$$q_j = \frac{1}{A}[(\tau_v+\tau_u)(\mu_v-p) + \tau_u(\Omega-\mu_v)] - I_j, \quad \forall j = 1,2,\cdots,M.$$

It follows from the market-clearing condition

$$0 = \tilde{x} + \sum_{j=1}^{M} q_j$$

that

$$p = \mu_v + \frac{\frac{M}{A}\tau_u(\Omega - \mu) + \frac{\tau_e}{\rho}(s - \mu_v)}{\frac{M}{A}(\tau_v + \tau_u) + \frac{\tau_e}{\rho}},$$

or equivalently,

$$p = \mu_v + \frac{A + \frac{\rho M \tau_u}{\tau_e}}{M\tau_v(1 - \frac{\tau_u}{\tau_e})}x - \frac{A}{M\tau_v(1 - \frac{\tau_u}{\tau_e})}\sum_{j=1}^{M} I_j.$$

It can be verified that given the size $|x|$, the equilibrium spread increases with $\frac{\tau_u}{\tau_e}$ (which is less than one), because when the latter ratio gets higher the order $\tilde{x}$ conveys more precise information about $\tilde{v}$ to the dealers, and hence the price becomes more sensitive to changes in $\tilde{x}$.

9. We have assumed in the preceding section that dealers are price-takers. In this section, let us assume that dealers are strategic. The rest formulations are the same as in the preceding section, except that the distribution of the market order $\tilde{x}$ is again exogenously given. Assume that $\tilde{v} \sim \mathrm{N}(\mu, \sigma^2)$. We shall find a symmetric equilibrium in which the price is linear in the dealers' inventories.

   In equilibrium, dealer j conjectures that his rivals all adopt the following pricing rule:

$$q_h = \xi - \beta p - \gamma I_h, \ \forall h \neq j,$$

   where ξ, β and γ are positive constants. From the market-clearing condition,

$$\tilde{x} + q_j + \sum_{h \neq j} q_h = 0,$$

dealer j can thus infer his inverse "residual supply function"

$$\tilde{p} = \frac{1}{\beta(M-1)}[(M-1)\xi - \sum_{h \neq j} \gamma I_h + q_j + \tilde{x}],$$

so that, by replacing the inverse residual supply into his objective and computing the first-order condition, we can obtain dealer j's best response[7]

$$q_j = \frac{1}{\frac{1}{\beta(M-1)} + A\sigma^2}[\mu - p] - \frac{A\sigma^2}{\frac{1}{\beta(M-1)} + A\sigma^2} I_j.$$

---

[7]Here, being able to submit a demand function, dealer j can make his net trade q_j depend on the realized price p, and note that p fully reveals to dealer j the statistic $\gamma \sum_{h \neq j} I_h - x$, where x is the realization of $\tilde{x}$ and I_h is the realization of dealer h's pre-trade inventory $\tilde{I}_h$. By choosing net trade q_j, dealer j's terminal wealth given the realization $\gamma \sum_{h \neq j} I_h - x$ is

$$\tilde{W}_j = (I_j + q_j)\tilde{v} + \frac{q_j}{\beta(M-1)}[\gamma \sum_{h \neq j} I_h - x - q_j - (M-1)\xi],$$

so that, conditional on the realized $\gamma \sum_{h \neq j} I_h - x$,

$$\mathrm{E}[\tilde{W}_j | \gamma \sum_{h \neq j} I_h - x] - \frac{A}{2}\mathrm{var}[\tilde{W}_j | \gamma \sum_{h \neq j} I_h - x] = (I_j + q_j)\mu + \frac{q_j}{\beta(M-1)}[\gamma \sum_{h \neq j} I_h - x - q_j - (M-1)\xi] - \frac{A}{2}(I_j + q_j)^2\sigma^2.$$

The q_j that maximizes the preceding dealer j's objective function must satisfy the following necessary and sufficient first-order condition:

$$\mu + \frac{1}{\beta(M-1)}[\gamma \sum_{h \neq j} I_h - x - 2q_j - (M-1)\xi] - A(I_j + q_j)\sigma^2 = 0$$

$$\Rightarrow q_j = \frac{\mu - AI_j\sigma^2 + \frac{\gamma \sum_{h \neq j} I_h - x - (M-1)\xi}{\beta(M-1)}}{\frac{2}{\beta(M-1)} + A\sigma^2} \Leftrightarrow q_j = \frac{1}{\frac{1}{\beta(M-1)} + A\sigma^2}[\mu - p] - \frac{A\sigma^2}{\frac{1}{\beta(M-1)} + A\sigma^2} I_j.$$

Indeed, using

$$q_j = \frac{\mu - AI_j\sigma^2 + \frac{\gamma \sum_{h \neq j} I_h - x - (M-1)\xi}{\beta(M-1)}}{\frac{2}{\beta(M-1)} + A\sigma^2}, \ \forall j = 1, 2, \cdots, M,$$

and the market-clearing condition

$$-\tilde{x} = \sum_{j=1}^{M} q_j = \frac{1}{\frac{2}{\beta(M-1)} + A\sigma^2}[M\mu + (\frac{\gamma}{\beta} - A\sigma^2)\sum_{j=1}^{M} I_j - \frac{M\xi}{\beta} - \frac{M}{\beta(M-1)}\tilde{x}],$$

which must hold for all realized $\sum_{j=1}^{M} I_j$ and all realized $\tilde{x}$, we have

$$\frac{M}{\beta(M-1)} = \frac{2}{\beta(M-1)} + A\sigma^2 \Rightarrow \beta = \frac{M-2}{(M-1)A\sigma^2},$$

$$\gamma = \beta A\sigma^2,$$

and

$$\xi = \beta\mu.$$

Now, in a symmetric equilibrium, we must have

$$\xi = \frac{\mu}{\frac{1}{\beta(M-1)} + A\sigma^2},$$

$$\beta = \frac{1}{\frac{1}{\beta(M-1)} + A\sigma^2},$$

and

$$\gamma = \frac{A\sigma^2}{\frac{1}{\beta(M-1)} + A\sigma^2}.$$

Solving the above system of equations, we obtain

$$\beta = \frac{M-2}{(M-1)A\sigma^2},$$

$$\gamma = \beta A\sigma^2,$$

and

$$\xi = \beta\mu.$$

In equilibrium, we have thus

$$q_j = \frac{M-2}{M-1}[\frac{\mu - p}{A\sigma^2} - I_j], \quad \forall j,$$

and

$$p = \psi + \frac{A\sigma^2(M-1)}{M-2}\frac{\tilde{x}}{M}.$$

The equilibrium bid-ask spread corresponding to order size $|\tilde{x}|$ is therefore

$$2\frac{A\sigma^2(M-1)}{M-2}\frac{|\tilde{x}|}{M}.$$

As can be seen, the bid-ask spread is wider than in the case where dealers are price-takers. Here strategic dealers take into account the impact of their own trades on the market-clearing price, and hence in equilibrium they offer less liquidity to the market.

10. The above models are referred to as the *inventory models*, where dealers' pre-trade inventories play a crucial role in determining the bid-ask spreads. Before we end this note, we briefly remark on earlier contributions to the theory of informed trading.

    There are only a few authors studying the impact of information asymmetry on asset trading prior to 1985. The origin of these information-based models is Bagehot (1971; Financial Analysts Journal), which makes two important points: (i) there is a difference between market gains and trading gains; and (ii) for market making to be viable, that liquidity traders must lose money to other traders who are not hit by liquidity shocks. Consequently, as Bagehot argued, information alone may lead to the presence of a bid-ask spread.

11. The first information-based model is due to Copeland and Galai (1983; JF), where a risk-neutral monopolistic specialist must choose the ask price A and bid price B to

$$\max -\pi_I[\int_A^{+\infty}(v-A)f(v)dv + \int_0^B (B-v)f(v)dv]$$
$$+(1-\pi_I)[\pi_{BL}(A-E(v)) + \pi_{SL}(E(v)-B)],$$

    where v is the true value of the stock with density $f(\cdot)$, and v will be revealed right after the trade, π_I is the probability that the next trader is informed (who knows v), $1-\pi_I$ is the probability that the next trader is a liquidity trader, who might buy, sell, or do nothing with respectively probabilities π_{BL}, π_{SL}, and $1-\pi_{BL}-\pi_{SL}$. These probabilities can be affected by the bid and ask prices chosen by the market maker. An overly widened bid-ask spread can discourage the liquidity trader from coming forward, and hence the market maker's optimization problem has a solution.

    The most important point made here is that information can be the sole reason that a bid-ask spread exists, although in equilibrium the size of the bid-ask spread also reflects the monopoly power of the market maker.

12. The seminal Golsten and Milgrom model (1985; JFE), which we shall review in a subsequent note, differs from Copeland and Galai (1983) in

two dimensions—Glosten and Milgrom consider Bertrand competitive market makers in a dynamic trading model (where trade itself may convey information).

With dynamic trading, a monopolistic market maker may optimally incur losses by attracting more trades in earlier periods and then make more profits in later periods after adverse selection problems become alleviated, whereas Bertrand-competitive market makers must still break even period by period, which results in the transaction price being an unbiased estimator for its future intrinsic value.

In Golsten and Milgrom (1985), market makers are risk-neutral and Bertrand-competitive. They must announce bid and ask prices before public traders arrive. The latter arrive sequentially (one at a time). Traders and market makers differ in time preferences so that trade can occur. Each trader has utility function

$$\rho x V + c,$$

where ρ is a discount factor, x is the holding of the risky asset whose intrinsic value is V (representing the future consumption) and c is current consumption. The median of ρ over the public traders is assumed to be 1, which is also the market makers' common parametric value. A trader is also endowed with private information. The paper produces the following main results. First, spread exists to reflect the adverse selection problem, and ask and bid prices straddle the common knowledge level of price. Second, transaction prices form a martingale. This implies serially uncorrelated returns. Third, a bound on the size of spread is obtained, which has testable implications. Fourth, markets are 'eventually' strong form efficient and the inventory process is 'eventually' driftless. Finally, the spread is related to the precision of private information, the likelihood of informed trading, and the intensity of liquidity trading.

References

1. Amihud, Y., and H. Mendelson, 1980, Dealership market: market making with inventory, *Journal of Financial Economics*, 8, 31-53.

2. Bagehot, W., 1971, The only game in town, *Financial Analysts Journal*, 27, pp. 12-14 and page 24.

3. Copeland, T., and D. Galai, 1983, Information effects and the bid-ask spread, *Journal of Finance*, 38, 1457-1469.

4. De Jong, F., and B. Rindi, 2009, *The Microstructure of Financial Markets*, Cambridge University Press.

5. Demsetz, H., 1968, The cost of transacting, *Quarterly Journal of Economics*, 82, 33-53.

6. Garman, M., 1976, Market Microstructure, *Journal of Financial Economics*, 3, 257-275.

7. Glosten, L., and P. Milgrom, 1985, Bid, ask, and transaction prices in a specialist market with heterogeneous informed traders, *Journal of Financial Economics*, 13, 71-100.

8. Ho, T., and H. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, *Journal of Financial Economics*, 9, 47-73.

9. Stoll, H., 1978, The supply of dealer services in securities markets, *Journal of Finance*, 33, 1133-1151.