

The Microeconomics of Market Making

Maureen O'Hara and George S. Oldfield*

Abstract

This paper examines the influence of risk aversion on the pricing policies of a market maker for securities. It is shown that a market maker's bid-ask spread can be decomposed into a portion for the known limit orders, a risk-neutral adjustment for expected market orders, and a risk adjustment for market order and inventory value uncertainty. It is demonstrated that a risk-averse market maker may set a smaller spread than a risk-neutral specialist. Finally, this paper demonstrates the pervasive role of inventory in affecting both the placement and size of the spread.

I. Introduction

Most security transactions involve the services of a specialized financial intermediary known as a market maker. A market maker is a dealer who conducts a two-sided auction for securities by standing ready to trade on either side of the market for his or her own account when an order arrives.¹ A buyer may purchase securities at the market maker's ask price and a seller may sell securities at the bid price. A market maker also may maintain a list of orders at various prices. If an excess of buy or sell orders arrives, a market maker accommodates traders through inventory adjustments and adjusts bid and ask prices to balance the order flow. Traders compensate a market maker for these services by paying the bid-ask spread.

In this paper, the pricing policy of a risk-averse market maker is analyzed. Because the market maker faces uncertainty, both with respect to the future order flow and to the future value of the stock, his or her choice of bid and ask prices constitutes a complicated decision problem. This problem becomes particularly complex if, as the authors believe, the market maker's risk preferences must also

* Both authors, S. C. Johnson Graduate School of Management, Cornell University, Ithaca, NY 14853. The authors thank David Easley and Vithala Rao for their valuable assistance. They also acknowledge the extremely helpful comments of anonymous *JFQA* referees. An earlier version of this paper was presented at the Econometric Society Meetings, Cornell University, June 1982. Financial support was provided by the Johnson Graduate School of Management Financial Research Institute and the National Science Foundation, Grant No. IST-8408770.

¹ A market maker is a simple type of financial intermediary. An intermediary typically maintains a long position in one type of security and a short position in a different type. It buys primary issues and sells its own liabilities. Market makers buy and sell the same security. The two services are closely related because a market maker may issue a liability to finance an inventory and an intermediary may buy and sell essentially identical securities.

be considered. This price setting problem is analyzed in an explicit multiperiod framework. Unlike previous analyses, this one requires neither a specific stochastic order flow process nor a known “intrinsic” value of the stock. This formulation permits us to identify exactly how transaction and inventory value uncertainty each influence market prices.

This work contributes three major insights into the market making process. First, it is shown that the market maker’s bid-ask spread can be decomposed into three parts: a portion for the known-limit orders (orders submitted for execution at a specific price), a risk-neutral adjustment for the expected market orders (orders submitted for immediate execution at the current quote), and a risk adjustment for market order and inventory value uncertainty. Second, it is demonstrated that, as a result of market order and inventory value uncertainty, a risk-averse market maker may set a smaller spread than a risk-neutral specialist. This has the important implication that a risk-averse market maker may not always be driven out of the market by a risk-neutral specialist (or, indeed, by a computerized clearing process). Finally, the influence of the market maker’s inventory position on pricing is investigated. Contrary to the finding of Ho and Stoll [10], Zabel [14], and Bradfield [4], this paper shows that the market maker’s inventory affects *both* the placement and size of the spread.

One important aspect of this work is the inclusion of numerous institutional characteristics in the depiction of the market clearing process. By drawing on previous work (see [11], [7], [9], [12], [1], [13], [6]) this paper considers realistic factors such as limit orders, nontrading periods, and auction characteristics in the analysis. This provides a more detailed setting for this theory and a more accurate understanding of the market making process. Perhaps more important, this analysis of the market clearing process permits more accurate modeling of the manner in which prices are formed. It is demonstrated not only how prices are set in a repetitive auction market but also how the auctioneer is compensated for price setting activities. As a result, this work has important theoretical and empirical implications for the process of price formation in securities markets. It is emphasized, though, that the results derive from the specific analytic framework defined by this paper.

In the next section, the rules and assumptions are specified that define one possible market making arrangement. Special attention is given to the microeconomic aspects of auction procedures, inventory value problems, and the intertemporal pricing policies the market maker must devise. Section III presents general solutions for ask and bid prices within a dynamic programming framework. It is also assumed that the market maker has constant absolute risk aversion and explicit solutions for bid and ask prices are found. The last section is a summary.

II. Markets and Market Making

Trading may be arranged on organized exchanges in many different ways. One important way in which trading institutions differ is that trading may be either continuous or intermittent [12]. With intermittent trading, the market maker accumulates orders and then executes market orders and qualified limit orders simultaneously at a specified trading time. This type of market is termed a

call auction market. A New York Stock Exchange (NYSE) specialist conducts such a call market when the exchange opens. A continuous auction is an alternative exchange arrangement. In a continuous auction, orders are booked or executed when they arrive. Limit orders can accumulate (if such orders are allowed) only because they do not get crossed or executed against an offsetting order during trading. A NYSE specialist conducts a continuous auction during a trading day.

In addition to the auction methods described above, trading may take place through negotiation. In this process, the essential terms of a trade are worked out in advance, and the buyer and seller typically know the price and quantity before the trade is executed. Block trades (or trades of a large volume of a particular security) are commonly arranged in this way.² Individual trading, or "retail" trading, on the NYSE rarely involves negotiation because transactions costs (particularly search costs) are lower if the trade goes through the specialist's auction process.

Because trading arrangements differ among exchanges, there is no one correct depiction of the market clearing process. This paper analyzes the behavior of a single market maker under one trading arrangement.³ The daily market is modeled as a cross between a call market and a continuous auction. A large number of trading periods are specified during a day, where each trading period is a type of short-lived call market. In each call market, the market maker sets prices at the beginning of the trading period and executes all trades at the period's end. This differs from a standard call market in which both price-setting and execution occur at the end of the period. The advantage to this paper's formulation is that it allows incorporation of several important features of continuous trading. For example, traders can "hit the quote" by submitting market orders to trade at the specialist's current price. In addition, by specifying a large number of call markets in a trading day, the dealer is allowed to adjust prices rapidly in response to the order flow. However, while this trading framework approximates continuous trading, this approximation is not exact. For example, as is characteristic of call markets, limit orders are allowed to trade at better prices than they were written. Because of the order flow treatment, traders' optimal order strategies may differ between call and continuous markets. This point is analyzed in [6].

It is assumed that the market maker sets bid and ask prices to maximize the expected utility of trading profits over an infinite horizon of trading days, $j = 1, 2, \dots$. The dealer's problem is

$$(1) \quad \max E \left[\sum_{j=0}^{\infty} \alpha^j U \left(\sum_{i=1}^n (\tilde{\pi}_{ji}) \right) \right],$$

where U is a Von-Neumann Morganstern utility function that is increasing, concave, bounded, and twice differentiable ($U'' < 0$), α is a discount rate $0 \leq \alpha < 1$,

² Although NYSE rules require that all trades take place on the floor of the exchange, the specialist's role in a block trade may be minimal. For an analysis of block trading, see [5].

³ Although it is assumed there is only one market maker mediating trades in this particular market, this does not necessarily make him or her a monopolist. Limit order traders can compete with the specialist and there is some competition across specialists if the securities they transact are close substitutes.

and $\tilde{\pi}_{jt}$ is the trading profit in period t of day j . The market maker may receive utility from other exogenous factors, but it is assumed that these factors have a zero covariance with his trading activity and, hence, they are not considered.

Each day has n trading periods and the market maker maximizes utility over an infinite number of trading days. This paper's model is thus similar to the models developed by Bradfield [3] and by Zabel [14] in that it views the market maker's dynamic program as an infinite series of finite n period intervals. In this paper's model, each daily n period interval is linked by an overnight market. Each period within a day is a brief time interval so only a few (or zero) trades are executed in a period.

At each period's start, the market maker knows the limit orders (both buys and sells, quantities and prices) submitted for that trading interval. The market maker also has expectations over the market orders he or she will receive at each possible price quote. With this information, the market maker sets bid and ask prices. The market maker then executes all appropriate limit orders and any market orders that arrive during the period at the quoted prices. The period ends, inventory is computed, and the market maker gets limit orders for the next trading interval.

A market maker's order flow comprises both limit orders and market orders. This paper's auction protocol specifies competitive rules so bidders are induced to bid their reservation prices (see [11]). The market maker books an incremental limit order at the reservation price but executes all qualified limit orders at the quote. The linear cumulative order functions are defined as integrals of the incremental orders

$$(2) \quad \begin{aligned} \alpha^L - \gamma^L a_t &= \int_{a_t}^{\bar{a}} q_a(a_t) da_t, \quad \text{limit buy function;} \\ \beta^L + \phi^L b_t &= \int_{\underline{b}}^{b_t} q_b(b_t) db_t, \quad \text{limit sell function.} \end{aligned}$$

In these expressions, α^L , β^L , γ^L , and ϕ^L are parameters of the cumulative order flow functions. a_t and b_t are the ask and bid prices, with \bar{a} the highest possible reservation buy price and \underline{b} the lowest reservation sell price. The integrands $q_a(a_t)$ and $q_b(b_t)$ represent the incremental quantity bid or tendered at each possible ask or bid price. Solving the integral equations (2) gives incremental orders $q_a(a_t) = \gamma^L$ and $q_b(b_t) = \phi^L$.⁴ These cumulative linear order functions characterize the specialist's book at the beginning of period t during a trading day.⁵

⁴ In words, each buyer bids for the same incremental number of shares (γ^L) and each seller tenders the same incremental number of shares (ϕ^L). Thus, although each incremental order is the same size, the cumulative number of shares ordered is downward sloping (buys) and upward sloping (sells). Since the auction is competitive, any limit buyer who bids for $q_a(\hat{a}_t)$, $\hat{a}_t \geq a_t$, gets his or her quantity at the ask price a_t . A similar interpretation applies to limit sell orders. Note that we set $q_a(\bar{a}) = 0$, $q_b(\underline{b}) = 0$.

⁵ The analysis here is appropriate when limit orders stay in the market maker's book for one period only. They disappear at period's end so unexecuted orders (sell orders above b_t and buy orders below a_t) do not persist into the next period. This analysis can also incorporate orders that remain for

A period's market order flow is composed of both price-dependent orders and liquidity-induced orders. This order flow is represented as

$$(3) \quad \begin{aligned} \tilde{A}_t^m &= \alpha^m - a_t \gamma^m + \tilde{w}_t \\ \tilde{B}_t^m &= \beta^m + b_t \phi^m + \tilde{\varepsilon}_t. \end{aligned}$$

The symbols α^m , β^m , γ^m , and ϕ^m are parameters; \tilde{w}_t and $\tilde{\varepsilon}_t$ are random variables. The random variables incorporate both deviations from the market maker's expected price-dependent market orders and the nonprice-dependent market orders that may arise from liquidity trading.⁶ It is assumed that both \tilde{w}_t and $\tilde{\varepsilon}_t$ are independently and identically distributed over time, with $E[\tilde{w}_t] = 0$ and $E[\tilde{\varepsilon}_t] > 0$. A period's total buy and sell order flows are \tilde{A}_t and \tilde{B}_t ,

$$(4) \quad \begin{aligned} \tilde{A}_t &= \alpha - a_t \gamma + \tilde{w}_t \text{ if } \alpha^L - a_t \gamma^L \geq 0, \\ &\quad \alpha^m - a_t \gamma^m + \tilde{w}_t \text{ otherwise;} \\ \tilde{B}_t &= \beta + b_t \phi + \tilde{\varepsilon}_t \text{ if } \beta^L + b_t \phi^L \geq 0, \\ &\quad \beta^m + b_t \phi^m + \tilde{\varepsilon}_t \text{ otherwise;} \end{aligned}$$

where $\alpha = \alpha^L + \alpha^m$; $\gamma = \gamma^L + \gamma^m$; $\beta = \beta^L + \beta^m$; and $\phi = \phi^L + \phi^m$. The conditions $\alpha^L - a_t \gamma^L \geq 0$, $\beta^L + b_t \phi^L \geq 0$ mean that the market maker does not accept limit orders for negative quantities. A specialist must select a_t and b_t such that he or she never sells at bid nor buys at ask prices.

This paper's specification of the market maker's order flow permits characterization of how the specialist uses the supply and demand information provided by the limit orders to set prices. In addition, by incorporating both limit orders and market orders, this analysis demonstrates how the market maker's pricing

several periods in the limit order book, but the analytics become tedious. The analysis continues in the text because it captures the essence of the problem. In general though, quoted prices depend upon the state of the book because this gives aggregate supply and demand parameters. If limit orders persist for several periods, a decay process must be specified so that old orders naturally disappear and the order book does not expand without bound. In fact, in the NYSE, limit orders may be specified for days, weeks, or months. They do not have infinite lives. While old orders remain, the book contains discontinuities wherever previous period quotes eliminated some orders. The market maker knows the book and thus can do the maximization in a piece-wise fashion. The parameters for each segment are the sum of parameters for periods in which the orders occurred, adjusted for decay. Finally, there must be a rule that all orders or none get executed at a quoted price to evaluate discontinuity points. A NYSE specialist's rule is more complex than the all or none requirement. The specialist must execute a whole order (if all or none) but is not required to execute all orders at a price. The priority rule is price, quantity, arrival time. Some orders can be skipped if the order quantity does not suit the specialist's strategy for filling orders on the other side of the market. In our set up, the market maker cannot skip orders at a given price. The priority is simply price.

⁶ If transactions costs were low, it is likely that most market orders would be liquidity-induced. That is, if transactions costs occur only when an order executes, then individuals could simply place limit orders corresponding to their supply and demand schedules. In this case, market orders would be used to guarantee immediate execution, and limit orders alone would be price dependent. If transactions costs were significant, however, traders might wait to see the quote before deciding to submit a market order. Our model incorporates both trading scenarios, but is perhaps most appropriate for cases in which transactions costs are low. As Cohen, Maier, Schwartz, and Whitcomb [6] demonstrate, transactions costs introduce almost a strategic element into traders' choice of order form. Because this paper's focus is on the specialist's pricing problem, such order strategy problems are not considered.

strategy depends on both the known and the unknown order flow. This analysis is simplified, however, in that it does not consider the interaction between the dealer's current price quotes and future limit orders. This restricts its ability to characterize fully the role of the specialist's book or to analyze the order strategies of traders (see [14], [4], and [6] for more in-depth analysis).

The following example illustrates this paper's market making protocol. Suppose the market maker's book of limit orders at a period's start is given in columns 1 and 2 below. These columns give the quantities and prices of limit orders (incremental and total) that are placed at the beginning of the period. Thus, a buy order for 200 shares is placed at 50 1/4 but 400 shares (200 at 50 1/4 and 200 at 50 3/8) is the total limit demand at 50 1/4. If buy and sell limit orders are placed at the same price, they are simply crossed. Thus, they do not affect the analysis and do not appear in the example. Columns 3 and 4 give the specialist's expected market buy and sell orders.

Prices	(1)		(2)		(3)	(4)
	Limit Buy Orders		Limit Sell Orders		Expected Market Buy Orders	Expected Market Sell Orders
	Incremental	Total	Incremental	Total		
50	200	800			300	200
50 1/8	200	600			200	200
50 1/4	200	400			100	200
50 3/8	200	200			100	200
50 1/2			100	100	100	300
50 5/8			100	200	100	400
50 3/4			100	300	100	500

Suppose the market maker sets the quotes at 50 1/4 ask, 50 1/8 bid. If no market orders arrive, the two limit buy orders, 200 shares at 50 1/4 and 200 shares at 50 3/8, are executed at 50 1/4. This reflects the dutch auction aspect of the market. Now suppose that at the dealer's quotes market orders arrive for 200 shares buy and 300 shares sell. Total transactions in the period are thus 600 shares sold to traders at 50 1/4 and 300 shares bought from traders at 50 1/8. The trading profit is \$15,112.50 and inventory declines 300 shares. As is characteristic of call markets, all the limit buy orders executed go for 50 1/4 although 200 shares are in the limit order book at 50 3/8. Similarly, any limit sell orders below 50 1/8 would be executed at 50 1/8. The bidders capture the rent.

Because of the problem's long-run nature, the market maker must consider the effect of actions taken today on future expected utility. In particular, as the above example illustrates, the market maker may acquire a positive or negative inventory position as a result of trading activity. Since this position is taken into the next trading day, inventory acquired in the current trading day affects the market maker's future operations. Inventory, therefore, represents the state variable of this dynamic system. Using standard dynamic programming arguments (see [2]) the market maker's infinite horizon problem given by (1) can be expressed as

(5)
$$\max E \left(U \left(\sum_{t=1}^n (\tilde{\pi}_t) \right) + V(\tilde{I}_n) \right),$$

where, since we start at day $j = 1$, the index j is dropped from $(\tilde{\pi}_j)$. The market maker's inventory position at the end of the day is I_n and V is the market maker's derived value for inventory. V is assumed concave, increasing, twice differentiable ($V'' < 0$) with α assumed in V . The value function incorporates the effect of current actions on future expected utility, given that future actions are chosen optimally. The V function thus represents the market maker's utility of the inventory position taken into the next trading day.⁷ The market maker's overall utility is a function of both the cash flow during the day ($\sum_{t=1}^n \pi_t$) and this inventory position taken into the future.

In this paper's formulation, this future inventory value *must* be determined endogenously. Unlike the analyses of Ho and Stoll [10] and Ahmihud and Mendelson [1], we do not assume any fixed, intrinsic value for the stock. Instead, the market maker's expectation of the future stock price is impounded in the value function. It is shown in the next section that the dealer's pricing strategy depends on his or her judgment of this future price.

All transactions settle at the end of each trading day. Gains and losses accumulate during the day but no shares or cash change hands until close. Then, at day's end, the market maker delivers securities sold and takes cash, and accepts securities bought and pays cash. All deliveries, in and out, are at the trading prices agreed upon when the trades were executed. If more securities are bought than sold, inventory grows; if sales exceed purchases, inventory declines. A negative inventory simply means that the market maker is short.

The daily settlement of trades means that inventory is important not only for its effect on tomorrow's operation but also because it affects the current cash flow. If the market maker winds up short at close, he or she must borrow securities overnight to deliver to buyers. Similarly, a positive closing inventory may be lent to brokers overnight. This overnight market is modeled as a competitive repurchase market in which shorts borrow securities and lend cash and longs lend securities and borrow cash. This overnight market establishes a price, \tilde{p} , that is the collateral value of stock loans.⁸ This repurchase price, for which loans are reversed the next day, is not necessarily the expected open price the next day. It is simply a forward price that clears the overnight loan market. If the market maker is short (i.e., $I_n < 0$) he or she pays $r\tilde{p}(I_n)$, where r is the interest rate, or

⁷ Note that $V(0)$ need not be zero. V represents the discounted value of future expected profits. This value is given by the solution $V(\cdot)$ to the fundamental equation of dynamic programming

$$V(I_n) = \max E \left[U \left(\sum_{t=1}^n (\tilde{\pi}_{j+1,t}) + V(\tilde{I}_{j+1,n}) \right) \right].$$

The existence of a solution to this problem follows from standard arguments (see [2]). If the right to be a market maker, the exchange privilege, is valuable in itself, $V(0) > 0$.

⁸ There currently exists such an overnight stock market in which shorts and longs trade cash and shares to balance out their positions. The shares are generally provided by large institutional investors or brokerage houses that have the shares as part of their portfolios. For example, Merrill Lynch had \$800.6 million in securities out on loan at the end of 1982. In the overnight market, a single price prevails, in much the same way that a block trade is executed at a negotiated price. Because this involves only brokerage, a single price can be agreed upon. The daily market, however, involves numerous trades of small volume and the dealer also provides immediacy by standing ready to trade on either side of the market. Because of this, the daily and overnight markets involve different services and different pricing mechanisms. (For more discussion of the overnight market, see "Inquiries Stirred By Stock Lending," *New York Times*, October 31, 1983.)

broker loan rate, on borrowing overnight; if long, he or she receives $r\tilde{p}(I_n)$.⁹ Both r and \tilde{p} are assumed exogenous to the dealer.

The next section analyzes how the market maker determines bid and ask prices and thus the spread. Because no specific form is imposed on the dealer's preferences, the solution is in the form of a general functional equation, the actual value of which depends upon the specific properties of the randomness and the market maker's utility function. This general expression indicates the factors that influence the market maker's optimal strategy and illustrates the complexity of the market making operation. A closed form solution is then determined for the market maker's problem when his or her preferences exhibit constant absolute risk aversion.

III. Optimal Bid and Ask Prices

The market maker's last decision in a trading day is to set a_n and b_n . These prices affect the volume of trades in period n and influence the level of inventory held overnight. The market maker faces a constrained optimization problem because limit orders must be positive

$$\begin{aligned}
 (6) \quad \max_{\{a_n, b_n\}} E \left[U \left(\sum_{t=1}^{n-1} \pi_t + a_n (\alpha - a_n \gamma + \tilde{w}_n) - b_n (\beta + b_n \phi + \tilde{\varepsilon}_n) \right. \right. \\
 \left. \left. + r\tilde{p}(I_{n-1} + \beta + b_n \phi + \tilde{\varepsilon}_n - \alpha + a_n \gamma - \tilde{w}_n) \right) \right. \\
 \left. + V(I_{n-1} + \beta + b_n \phi + \tilde{\varepsilon}_n - \alpha + a_n \gamma - \tilde{w}_n) \right]; \\
 \text{subject to } \alpha^L - a_n \gamma^L \geq 0; \beta^L + b_n \phi^L \geq 0.
 \end{aligned}$$

We define $U' \equiv \partial U / \partial \pi_n$ and $V' \equiv \partial V / \partial I_n$. When the constraints are satisfied as strict inequalities, the first order conditions lead to solutions for the ask and bid prices quotes at period n 's start¹⁰

$$\begin{aligned}
 (7) \quad a_n &= \alpha / 2\gamma + E(U' \tilde{w}_n) / E(U') 2\gamma \\
 &\quad + rE(U' \tilde{p}) / 2E(U') + E(V') / 2E(U'), \\
 b_n &= -\beta / 2\phi - E(U' \tilde{\varepsilon}_n) / E(U') 2\phi \\
 &\quad + rE(U' \tilde{p}) / 2E(U') + E(V') / 2E(U').
 \end{aligned}$$

⁹ The effect on the market maker's cash flow is $(1+r)\tilde{p}(I_n)$ and is reversed the next day for $\tilde{p}(I_n)$. Hence, the effect on the market maker's trading profit is $r\tilde{p}(I_n)$, which is paid (or received) at the end of the trading day. Since the $\tilde{p}(I_n)$ term nets out, it is not included in the trading profit term.

¹⁰ The Kuhn-Tucker conditions for nonbinding constraints mean the Lagrange multipliers are zero, so they do not appear in the first-order conditions or the price equations. The first-order conditions with slack variables h_n and k_n give an alternative expression for prices in which the first terms in (7) are replaced by $(\alpha^L + h_n)/\gamma$ in the a_n equation, and $-(\beta^L + k_n)/\phi$ in the b_n equation. The slack variables h_n and k_n must be negative. They measure the losses avoided because traders must buy at ask and sell at bid prices.

These expressions are not explicit solutions for a_n and b_n because both U' and V' are functions of these prices, and U' and V' appear on the right sides of the equations. Nevertheless, each part of the expressions for a_n and b_n in equations (7) has direct interpretation. The first terms are the solutions for a_n and b_n that derive from the known limit orders and the expected market orders. The second terms give the risk adjustments for the variability in market orders. This variability is impounded in the covariance between U' and \tilde{w}_n or $\tilde{\varepsilon}_n$. For a risk-averse market maker, $\text{Cov}(U', \tilde{w}_n) < 0$ and $\text{Cov}(U', \tilde{\varepsilon}_n) > 0$.¹¹ Because this latter term enters (7) negatively, holding everything else constant, the covariance terms shift both a_n and b_n down.

The last two terms give the direct price level adjustment induced by inventory changes. Inventory affects the market maker immediately through cash flow and, in the future, through the position taken into the next trading day. A nonzero inventory results in lending or borrowing at $r\tilde{p}$, and the third term in (7) incorporates this cash flow effect. The last term shows the effect of carrying this inventory into the future. A nonzero inventory affects the dealer's future profitability because the future stock price is unknown. The last term incorporates the market maker's expectation of this future price and, as is evident from (7), this adjustment affects both a_n and b_n equally.

The market maker's spread may be computed by subtracting b_n from a_n . The optimal spread is

$$(8) \quad a_n - b_n = (\alpha\phi + \beta\gamma)/2\phi\gamma + (\phi E(\tilde{w}_n) + \gamma E(\tilde{\varepsilon}_n))/2\gamma\phi \\ + (\phi \text{Cov}(U', \tilde{w}_n) + \gamma \text{Cov}(U', \tilde{\varepsilon}_n))/2\phi\gamma E(U').$$

Equation (8) gives some interesting insights into the determination of the bid-ask spread. The first two terms reflect the market maker's total expected supply and demand during trading interval n . This is the spread charged by a risk-neutral market maker. The market maker also must contend with market order uncertainty. This is captured in the third term of equation (8). If the market maker is risk neutral, market order uncertainty has no effect since U' is constant and both covariance terms are zero. If the market maker is risk averse, however, market order variability causes an adjustment to the spread. It can be shown that $\text{Cov}(U', \tilde{w}_n) < 0$ and $\text{Cov}(U', \tilde{\varepsilon}_n) > 0$. This means that a risk-averse market maker's spread can widen or narrow in response to numerous factors including risk preferences and the demand and supply parameters' magnitudes. This adjustment process is analyzed more fully below.

The final term in equation (8) also indicates that overnight inventory affects a risk-averse market maker's spread. Although the direct influence cancels when b_n is subtracted from a_n , an indirect linkage exists through U' . More precisely, U' is a function of both I_{n-1} and a_n and b_n . But from equation (7), a_n and b_n are also functions of inventory, I_{n-1} . Since U' appears in the third term of equation (8), overnight inventory may influence the size of this term indirectly.

¹¹ To see this, note that \tilde{w}_n enters $u(\cdot)$ as $(a_n - r\tilde{p})\tilde{w}_n$, where $(a_n - r\tilde{p}) > 0$. The larger is w_n , the larger is $u(\cdot)$ and, hence, the smaller is $u'(\cdot)$. Thus $\text{Cov}(u', \tilde{w}_n) < 0$. The argument is symmetric for $\tilde{\varepsilon}_n$ with $\text{Cov}(u', \tilde{\varepsilon}_n) > 0$.

The market maker has a similar decision problem in period $n-1$. Given the final period optimal prices a_n^* and b_n^* , the market maker must set prices for period $n-1$, a_{n-1} and b_{n-1} . An intertemporal linkage exists between a_n^* , b_n^* and a_{n-1} and b_{n-1} , because a_n^* and b_n^* depend upon π_{n-1} and I_{n-1} and they, in turn, are influenced by a_{n-1} and b_{n-1} . This creates a very complex analytic problem. However, the period $n-1$ solution has exactly the same form as the period n spread; the market maker's spread in periods n and $n-1$ is thus given by the same functional relationship.¹²

We wish to analyze further the market maker's strategy under reasonable assumptions regarding risk preferences. As constant absolute risk aversion is extensively employed in the literature, this preference structure is utilized to analyze the market maker's behavior. The market maker's problem can be represented as

$$(9) \quad \max E \left[-\exp \left(-c \sum_{t=1}^n \tilde{\pi}_t \right) - \exp(-d\tilde{p} \tilde{I}_n) \right],$$

where the parameters c and d are the market maker's coefficients of absolute risk aversion associated with trading profits and the overnight value of inventory, respectively.¹³ It is assumed here that the market maker values the inventory at the overnight repurchase price. This permits calculation of the value to the market maker of carrying a positive or negative inventory position into the next trading day.

If trading profits and inventory value are jointly normally distributed, maximizing expected utility is equivalent to maximizing a function that is linear in expected values and variances. Market orders are assumed to arrive in a manner

¹² In period $n-1$, the dealer's maximization problem is

$$\begin{aligned} \max_{\{a_{n-1}, b_{n-1}\}} E & \left[U \left(\sum_{t=1}^{n-2} \pi_t + a_{n-1}(\alpha - a_{n-1}\gamma + \tilde{w}_{n-1}) - b_{n-1}(\beta + b_{n-1}\phi + \tilde{\epsilon}_{n-1}) \right. \right. \\ & + a_n^*(\alpha - a_n^*\gamma + \tilde{w}_n) - b_n^*(\beta + b_n^*\phi + \tilde{\epsilon}_n) \\ & + r\tilde{p}(I_{n-2} + \beta + b_{n-1}\phi + \tilde{\epsilon}_{n-1} - \alpha + a_{n-1}\gamma + \tilde{w}_{n-1} \\ & \left. \left. + \beta + b_n^*\phi + \tilde{\epsilon}_n - \alpha + a_n^*\gamma - \tilde{w}_n) \right) \right. \\ & \left. + V(I_{n-2} + \beta + b_{n-1}\phi + \tilde{\epsilon}_{n-1} - \alpha + a_{n-1}\gamma - \tilde{w}_{n-1} \right. \\ & \left. + \beta + b_n^*\phi + \tilde{\epsilon}_n - \alpha + a_n^*\gamma - \tilde{w}_n) \right], \\ \text{subject to } & \beta^L + b_{n-1}\phi^L \geq 0; \alpha^L - a_{n-1}\gamma^L \geq 0. \end{aligned}$$

Assuming interior solutions, the first-order conditions can be solved for the spread in period $n-1$ as a_{n-1} minus b_{n-1} ,

$$\begin{aligned} a_{n-1} - b_{n-1} = & (\alpha\phi + \beta\gamma)/2\phi\gamma + (\phi E(\tilde{w}_{n-1}) + \gamma E(\tilde{\epsilon}_{n-1}))/2\gamma\phi \\ & + (\phi \text{Cov}(U', \tilde{w}_{n-1}) + \gamma \text{Cov}(U', \tilde{\epsilon}_{n-1}))/2\gamma\phi E(U'). \end{aligned}$$

This spread equation is in the same form as the interval n spread in equation (8).

¹³ It is assumed that the market maker's value function is also a negative exponential function.

that allows the transformation to a linear equation.¹⁴ With this transformation, the market maker's final period problem is

$$(10) \quad \max_{\{a_n, b_n\}} \sum_{t=1}^{n-1} (\pi_t) + E(\tilde{\pi}_n) - \frac{c}{2} \text{Var}(\tilde{\pi}_n) + E(\tilde{p} \tilde{I}_n) - \left(\frac{d}{2}\right) \text{Var}(\tilde{p} \tilde{I}_n);$$

subject to $\alpha^L - a_n \gamma^L \geq 0; \beta^L + b_n \phi^L \geq 0$.

Expected values (variables with a bar) and variances (denoted $\text{Var}(\cdot)$) are defined in the Appendix.

To obtain insight into this problem, it is first useful to consider how inventory price variability and market order variability each, separately, affect the market maker's pricing policy. If the market maker faces only price variability and not also market order variability, solving the maximization problem shows that the spread is merely the risk-neutral spread. The individual bid and ask prices, however, do *not* equal their risk-neutral values. Since with no market order variability the supply and demand are known, the market maker can determine his or her optimal inventory level. The market maker then incorporates uncertainty about the inventory's future value by moving the bid and ask prices symmetrically.

Alternatively, suppose that the market maker faces only market order variability (with the inventory price fixed at \bar{p}). Solving the maximization problem then yields

$$(11) \quad a_n - b_n = \frac{\alpha\phi + \beta\gamma}{2\gamma\phi} + \frac{\bar{w}_n\phi + \bar{\epsilon}_n\gamma}{2\gamma\phi} + \frac{c}{2\gamma\phi} \left[\gamma \text{Var}(\tilde{\epsilon}_n) \left(\frac{-\beta - \epsilon_n + (1+r)\bar{p}\phi - cr\bar{p} \text{Var}(\tilde{\epsilon}_n)}{2\phi + c \text{Var}(\tilde{\epsilon}_n)} + r\bar{p} \right) - \phi \text{Var}(\tilde{w}_n) \left(\frac{\alpha + \bar{w}_n + (1+r)\bar{p}\gamma - cr\bar{p} \text{Var}(\tilde{w}_n)}{2\gamma + c \text{Var}(\tilde{w}_n)} + r\bar{p} \right) \right].$$

The first two terms are the risk-neutral spread while the third term is the risk adjustment factor. Thus, risk adjustments can be either *positive or negative* depending upon the relative magnitudes of the parameters involved. For example, suppose there exists symmetric market order variability, and the absolute values

¹⁴ Profit in a period is random due to the difference in earnings on buy and sell market orders. It is assumed this difference approaches normality rapidly enough for the periods to be quite short. For example, suppose market orders for round lot trades arrive during a trading period of length t with probabilities qt for a buy order and st for a sell order. The characteristic function for this process (number of shares divided by 100) is

$$e^{\Phi(u)} = e^{-t + t(qe^{iu} + se^{-iu})}.$$

Define $iu = \theta$ and expand e^θ and $e^{-\theta}$ with Maclaurin series. The limiting expression is

$$e^{\Phi(u)} = e^{iut(q-s) - (ut^2/2)},$$

which is the characteristic function for a normal process. See [8], pp. 59, 567, for additional information.

of the slopes of the total order flow are equal (i.e., $\gamma \text{Var}(\tilde{\epsilon}) = \phi \text{Var}(\tilde{w}) \equiv H > 0$). Then the spread becomes

$$(12) \quad a_n - b_n = \frac{\alpha\phi + \beta\gamma}{(2\gamma\phi + cH)} + \frac{\bar{w}_n\phi + \bar{\epsilon}_n\gamma}{(2\gamma\phi + cH)}.$$

For a risk-averse market maker, the risk-aversion coefficient, c , is greater than zero. As c increases, the spread narrows compared to the risk-neutral solution. This occurs because both bid and ask prices decline, but a_n declines more than b_n . The reason a risk-averse market maker behaves like this is because this pricing strategy reduces the variance of profit in each trading interval.¹⁵ The reduced profit variance more than offsets reduced expected profit and increases the market maker's expected utility. This narrower spread implies that a risk-averse market maker cannot always be dominated by a risk-neutral specialist in this analytic framework. Equations (12) and (13) also indicate that, in the absence of price variability, the level of inventory does not affect the spread. The independence of spread and inventory position thus holds true in the model if the market maker's inventory position can always be settled at a known price \bar{p} .

In general, however, the market maker faces both simultaneous market order and inventory price uncertainty. To analyze how this multiple variability affects the market maker's spread, the maximization problem is solved in the presence of both market order and inventory price uncertainty. The solution is given in the Appendix. As expected, the multiple uncertainty greatly complicates the market maker's pricing strategy. In particular, the risk adjustment term in the spread includes the amount of inventory brought into the period, (I_{n-1}) . This means that the market maker's inventory affects the magnitude of the spread. It can also be demonstrated that the inventory affects the placement of the individual prices a_n and b_n . The interaction of market order and inventory price variability thus results in inventory affecting *both* the placement and magnitude of the spread. This inventory effect would be even more important if the randomness of limit orders over time was incorporated into the analysis. If, as in the work of Zabel [14] and Bradfield [4], this paper's model included this dual order variability, the added randomness would make the interaction between inventory price uncertainty and order flow uncertainty more pronounced.

This analysis suggests some interesting insights into the market maker's price formation policy. The size of the market maker's optimal spread depends on the parameters of the supply and demand functions. If the supply and demand are known, the optimizing spread can be set much as in the manner suggested by

¹⁵ The effect of the market maker's risk aversion on prices is given by

$$\frac{\partial a_n}{\partial c} = \frac{-((\alpha + \bar{w}) + \bar{p}\gamma(1+r) + 2\gamma r\bar{p})}{(2\gamma + c \text{Var}(\tilde{w}))^2} < 0$$

$$\text{and } \frac{\partial b_n}{\partial c} = \frac{-(-(\beta + \bar{\epsilon}) + \bar{p}\phi(1+r) + 2\phi r\bar{p})}{(2\phi + c \text{Var}(\tilde{\epsilon}))^2} < 0.$$

It can easily be shown that $\partial a_n / \partial c < \partial b_n / \partial c$. The reduced variance follows from

$$\frac{\partial \text{Var}(\tilde{\pi}_n)}{\partial c} = \left(2a_n \frac{\partial a_n}{\partial c} + 2r\bar{p} \frac{\partial a_n}{\partial c} \right) \text{Var}(\tilde{w}) + \left(2b_n \frac{\partial b_n}{\partial c} + 2r\bar{p} \frac{\partial b_n}{\partial c} \right) \text{Var}(\tilde{\epsilon}).$$

Demsetz [7]. With market order variability, however, the supply and demand are not known and the risk-averse market maker adjusts the spread to compensate for what could be large divergences in either supply or demand. The placement of the spread depends both on today's supply and demand and on the future expected value of the security. If the security price is variable, the market maker adjusts bid and ask prices to compensate for what could be large movements in tomorrow's security prices.¹⁶

If risk-neutral prices do prevail (either because of a risk-neutral dealer or dealers or a computerized clearing process), there exists an interesting relationship between these quoted trading prices and the expected overnight repurchasing price. Substituting the risk-neutral prices into the supply and demand equations (4) and computing expected values yields

$$(13) \quad \begin{aligned} E(\tilde{A}_t^*) &= (\alpha + \bar{w}_t - (1+r)\bar{p}\gamma)/2 \\ E(\tilde{B}_t^*) &= (\beta + \bar{\varepsilon}_t + (1+r)\bar{p}\phi)/2. \end{aligned}$$

In this notation, $E(\tilde{A}_t^*)$ is the expected number of shares sold at the optimal ask price for period t , and $E(\tilde{B}_t^*)$ is the expected number of shares bought at the bid price. Suppose the market is in a stable situation such that, for the optimal bid and ask prices quoted, expected purchases equal expected sales during the trading interval. Then $E(\tilde{A}_t^*) = E(\tilde{B}_t^*)$, and the expected inventory change is zero. From equations (13), another way to write this condition is

$$(14) \quad (1+r)\bar{p} = (\gamma a_t + \phi b_t)/(\gamma + \phi).$$

In words, $(1+r)\bar{p}$ is a weighted sum of the ask and bid prices quoted for the trading interval. Moreover, since both γ and ϕ are positive, the weights sum to unity and $(1+r)\bar{p}$ is bracketed by a_t and b_t . If the demand and supply schedules have the same absolute slope, $\gamma = \phi$, then the pricing relationship is

$$(15) \quad (1+r)\bar{p} = (a_t + b_t)/2.$$

The expected inventory repurchase price (adjusted for interest) equals the arithmetic average of the ask and bid prices. In summary, if risk-neutral trading prices are set for a trading period such that expected inventory change is zero for the period, overnight speculation is not a trading motive. Optimal bid and ask prices bracket the expected overnight price in a stable trading situation.

IV. Summary

Any depiction of the market making process must specify a few essential market characteristics. These include the nature of the order flow, the duration of the service, and the rules for market clearing. This paper presents an analysis of

¹⁶ Zabel [14] and Bradfield [4] also demonstrate that in the absence of price variability the placement of the spread depends on the specialist's book and the time of day. This paper's model suggests that the time of day is also important but, as the specialist's book is not a state variable in the model, the book's effect on the spread is not considered.

one set of market making arrangements. Within a dynamic framework, the market maker's pricing policy is analyzed in the presence of transaction and price uncertainty. It is shown that risk aversion may affect both the spread and the individual bid and ask prices, and that this risk-averse spread may be less than the risk-neutral spread. It is also demonstrated that inventory has a pervasive influence on the market maker's pricing policy, influencing both the size and placement of the spread. Finally, it is found that if risk-neutral prices do prevail, a stable trading situation implies that the expected overnight price lies between the quoted bid and ask prices. This seems quite reasonable because a stable situation means that expected inventory change is zero. In short, market conditions preclude anticipated inventory changes for speculation.

This work also gives a starting point for understanding the process of price formation. As the analysis demonstrates, even in the simplest markets in which price decisions are made by one agent, the process of price formation is extremely complex. As a result, prices incorporate not only current supply and demand conditions but also expectations of future conditions, risk-preference adjustments, and dynamic considerations.

Appendix

The Determination of the n th Period Spread

The market maker's n th period problem is

$$\max_{\{a_n, b_n\}} \sum_{t=1}^{n-1} (\pi_t) + E(\tilde{\pi}_n) - \frac{c}{2} \text{Var}(\tilde{\pi}_n) + E(\tilde{p}\tilde{I}_n) - \frac{d}{2} \text{Var}(\tilde{p}\tilde{I}_n),$$

$$\text{subject to } \alpha^L - a_n \gamma^L \geq 0,$$

$$\beta^L + b_n \phi^L \geq 0,$$

$$\begin{aligned} \text{where } E(\tilde{\pi}_n) &= a_n(\alpha - a_n \gamma + \bar{w}_n) - b_n(\beta + b_n \phi + \bar{\epsilon}_n) \\ &\quad + r\bar{p}(I_{n-1} + \beta + b_n \phi - \alpha + a_n \gamma) + r\bar{p}(\bar{\epsilon}_n - \bar{w}_n), \\ \text{Var}(\tilde{\pi}_n) &= (a_n^2 + 2a_n r\bar{p}) \text{Var}(\tilde{w}_n) + (b_n^2 + 2b_n r\bar{p}) \text{Var}(\tilde{\epsilon}_n) \\ &\quad + (r^2 J^2 + 2r^2 J(\bar{\epsilon}_n + \bar{w}_n)) \text{Var}(\tilde{p}) + r^2 \text{Var}(\tilde{p}\tilde{\epsilon}_n) \\ &\quad + r^2 \text{Var}(\tilde{p}\tilde{w}_n), \end{aligned}$$

$$\text{where } J \equiv (I_{n-1} + \beta + b_n \phi - \alpha + a_n \gamma),$$

$$E(\tilde{p}\tilde{I}_n) = J\bar{p} + E(\tilde{p}\tilde{\epsilon}_n) + E(\tilde{p}\tilde{w}_n),$$

$$\text{Var}(\tilde{p}\tilde{I}_n) = (J^2 + J(\bar{\epsilon}_n + \bar{w}_n)) \text{Var}(\tilde{p}) + \text{Var}(\tilde{p}\tilde{\epsilon}_n) + \text{Var}(\tilde{p}\tilde{w}_n).$$

The first-order conditions can be solved for the spread

$$\begin{aligned}
 a_n - b_n = & \frac{\alpha\phi + \beta\gamma}{2\gamma\phi} + \frac{\bar{w}_n\phi + \bar{\epsilon}_n\gamma}{2\gamma\phi} + \frac{c}{2\gamma\phi} \left[\gamma \text{Var}(\tilde{\epsilon}_n)(r\bar{p} + \Omega) \right. \\
 & - \phi \text{Var}(\tilde{w}_n)(r\bar{p} + \{\alpha + \bar{w}_n + (1+r)\bar{p}\gamma - cr\bar{p} \text{Var}(\tilde{w}_n) \\
 & - (cr^2 + d/2)(\bar{\epsilon}_n + \bar{w}_n) \text{Var}(\tilde{p}) \\
 & - (cr^2 + d)\gamma(I_{n-1} + \beta - \alpha) \text{Var}(\tilde{p}) \\
 & \left. - (cr^2 + d)\gamma\phi\Omega \text{Var}(\tilde{p})\} \right] \\
 & / \left[(2\gamma + c \text{Var}(\tilde{w}_n) + (cr^2 + d)\gamma^2 \text{Var}(\tilde{p})) \right],
 \end{aligned}$$

where

$$\begin{aligned}
 \Omega = & \left\{ (2\gamma + c \text{Var}(\tilde{w}_n) + (cr^2 + d)\gamma^2 \text{Var}(\tilde{p}))(-\beta - \bar{\epsilon}_n + (1+r)\bar{p}\phi \right. \\
 & - cr\bar{p} \text{Var}(\tilde{\epsilon}_n) - (cr^2 + d)\phi(\bar{\epsilon}_n + \bar{w}_n) \text{Var}(\tilde{p}) \\
 & - (cr^2 + d)\phi(I_{n-1} + \beta - \alpha) \text{Var}(\tilde{p})) \\
 & - (cr^2 + d)\phi\gamma(\alpha + \bar{w}_n + (1+r)\bar{p}\gamma - cr\bar{p} \text{Var}(\tilde{w}_n) \\
 & - (cr^2 + d/2)\gamma(\bar{\epsilon}_n + \bar{w}_n) \text{Var}(\tilde{p}) \\
 & \left. - (cr^2 + d)\gamma(I_{n-1} + \beta - \alpha) \text{Var}(\tilde{p})) \right\} / R,
 \end{aligned}$$

$$\begin{aligned}
 \text{and } R = & \left[(2\gamma + c \text{Var}(\tilde{w}_n) + (cr^2 + d)\gamma^2 \text{Var}(\tilde{p}))(2\phi + c \text{Var}(\tilde{\epsilon}_n) \right. \\
 & \left. + (cr^2 + d)\phi^2 \text{Var}(\tilde{p})) + (cr^2 + d)\gamma\phi \text{Var}(\tilde{p}) \right].
 \end{aligned}$$

References

- [1] Amihud, Y., and H. Mendelson. "Dealership Market: Market Making with Uncertainty." *Journal of Financial Economics*, 8 (March 1980), 31-54.
- [2] Blume, L.; D. Easley; and M. O'Hara. "Characterization of Optimal Plans for Stochastic Dynamic Programs." *Journal of Economic Theory*, 28 (Dec. 1982), 221-234.
- [3] Bradfield, J. "A Formal Dynamic Model of Market Making." *Journal of Financial and Quantitative Analysis*, 14 (June 1979), 275-291.
- [4] ———. "Optimal Dynamic Behavior of a Market-Maker in Choosing Bid and Asked Prices." *Journal of Economics and Business*, 34 (1982), 303-316.
- [5] Burdett, K., and M. O'Hara. "Building Blocks: An Introduction to Block Trading." Working Paper, Cornell Univ. (Oct. 1983).
- [6] Cohen, K.; S. Maier; R. Schwartz; and D. Whitcomb. "Transactions Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread." *Journal of Political Economy*, 89 (April 1981), 287-305.
- [7] Demsetz, H. "The Cost of Transacting." *Quarterly Journal of Economics*, 82 (Feb. 1968), 33-53.
- [8] Feller, W. *An Introduction to Probability Theory and Its Applications*. Vol. II, 2nd Ed., New York: John Wiley & Sons (1971).
- [9] Garman, M. B. "Market Microstructure." *Journal of Financial Economics*, 3 (June 1976), 257-275.
- [10] Ho, T., and H. Stoll. "Optimal Dealer Pricing under Transactions and Return Uncertainty." *Journal of Financial Economics*, 9 (March 1981), 47-73.
- [11] Smith, V. "Bidding Theory and the Treasury Bill Auction: Does Price Discrimination Increase Bill Prices?" *Review of Economics and Statistics*, 48 (May 1966), 141-146.
- [12] Smidt, S. "Continuous vs. Intermittent Trading on Auction Markets." *Journal of Financial and Quantitative Analysis*, 14 (Nov. 1979), 837-866.
- [13] Tinic, S., and R. West. "The Securities Industry under Negotiated Brokerage Commissions: Changes in the Structure and Performance of New York Stock Exchange Member Firms." *Bell Journal of Economics*, 11 (Spring 1980), 29-41.
- [14] Zabel, E. "Competitive Price Adjustment without Market Clearing." *Econometrica*, 49 (Sept. 1981), 1201-1221.