The Dynamics of Dealer Markets Under Competition
Author(s): Thomas S. Y. Ho and Hans R. Stoll
Source: *The Journal of Finance*, Sep., 1983, Vol. 38, No. 4 (Sep., 1983), pp. 1053-1074
Published by: Wiley for the American Finance Association
Stable URL: https://www.jstor.org/stable/2328011

THE JOURNAL OF FINANCE • VOL. XXXVIII, NO. 4 • SEPTEMBER 1983

The Dynamics of Dealer Markets Under Competition

THOMAS S. Y. HO and HANS R. STOLL*

ABSTRACT

The behavior of competing dealers in securities markets is analyzed. Securities are characterized by stochastic returns and stochastic transactions. Reservation bid and ask prices of dealers are derived under alternative assumptions about the degree to which transactions are correlated across stocks at a given time and over time in a given stock. The conditions for interdealer trading are specified, and the equilibrium distribution of dealer inventories and the equilibrium market spread are derived. Implications for the structure of securities markets are examined.

IN THIS PAPER the behavior of competing dealers in security markets is examined. Much of the theoretical work on dealers (Demsetz [6], Tinic [18], Garman [8], Stoll [16], Amihud and Mendelson [1], Ho and Stoll [11], Copeland and Galai [3], Mildenstein and Schleef [13]) has recognized that dealers may face competition from other dealers or investors placing limit orders, but nonetheless has analyzed only a single (representative) dealer. This approach is quite reasonable for the New York Stock Exchange specialist who has a quasi-monopoly position, but it is less applicable when considering other markets such as the over-the-counter market where there are several dealers with equal access to the market. Similarly the empirical studies of dealer bid-ask spreads (Demsetz [6], Tinic [18], Tinic and West [19], Benston and Hagerman [2], Stoll [17], Smidt [15]) have either been based on models of a single dealer or have lacked a theoretical foundation based on the microeconomics of the dealer.

This paper develops a theoretical model of equilibrium in a market with competing dealers and provides a basis for empirical work that would distinguish competing and monopolistic dealer markets. The paper is concerned with the behavior and interaction of individual competing dealers and with the determination of the market bid-ask spread. Markets with several dealers, several stocks and several periods are considered. Dealers bear risk arising not only from uncertainty about the returns on their inventories but also from uncertainty about the arrival of transactions. Each dealer also recognizes that his welfare depends on the actions of other dealers and each sets bid and ask prices to maximize his own expected utility of terminal wealth. A recent paper by Cohen, Maier, Schwartz and Whitcomb [5] examines similar issues in the context of an auction market in which the market spread is determined by limit orders. However, unlike the model of this paper, their analysis is not based as clearly on a model of individual traders' maximizing behaviors nor are the costs of placing

* Financial support of the Dean's Fund for Faculty Research at the Owen Graduate School of Management is gratefully acknowledged.

limit orders explicitly derived. In an interesting paper, Garbade and Silber [7] examine some of the broader issues of structuring a securities market in terms of the effect of dealers on the frequency with which markets clear. However, they do not examine the transaction-by-transaction dynamics of a market, something with which we deal.

Our approach is to apply a model of an individual dealer operating under return and transactions uncertainty (Ho and Stoll [11]) to the case of more than one dealer and the problem of determining the equilibrium market bid-ask spread.[1] The paper is organized as follows. The dynamics of the individual dealer and the dynamic programming problem of the dealer in the presence of a competitor are specified in Section I. In Section II dealer reservation buying and selling fees under a one period horizon are derived. The interaction of dealers and of incoming orders is examined in Section III, and the equilibrium distribution of inventories across dealers and the equilibrium market spread are derived under the assumption of homogeneous risk preferences. In Section IV we show that the equilibrium market spread is not affected by the introduction of heterogeneous opinions about the true price of the stock. The effect on dealer fees of extending the time horizon and of permitting serial correlation in transactions is examined in Section V. The paper concludes with an examination of the implications of the model for the structure of securities markets and with a section summarizing the major conclusions of the paper.

I. The Individual Dealer

As a simplifying assumption, the formal analysis in this paper is restricted to two dealers, A and B, each making a market in two stocks. Unless specified otherwise, the dealers have homogeneous opinions about the true value of each stock. At various points, the analysis is extended to many stocks, to many dealers, and to heterogeneous opinions. In this section the return dynamics and transaction dynamics affecting each component of dealer A's wealth are presented, and the dealer's maximization problem is specified. The dynamics for dealer B are the same except for a change of notation. Our notational convention is to denote variables of dealer B by the superscript "^{0}".

A. Dynamics of a Dealer

Each dealer's wealth consists of three components: inventory of each of the stocks, cash and base wealth.

1. Inventory. The dollar values of the dealer's inventories of his two stocks are given by

$$M_{t-1} = (1 + r_M)[M_t + q_M(Q, -Q)] + [M_t + q_M(Q, -Q)]Z_M \tag{1}$$

and

$$N_{t-1} = (1 + r_N)[N_t + q_N(Q, -Q)] + [N_t + q_N(Q, -Q)]Z_N \tag{2}$$

[1] The formal model is changed slightly in that we use discrete time stochastic processes in this paper rather than the continuous time stochastic processes in Ho and Stoll [11]. Some preliminary results under competition—particularly the idea of second best pricing and a preliminary formulation of the condition for interdealer trading are in Ho and Stoll [10].

where

M, N Dollar value of dealer's inventory of stock M or N. As a subscript also serves as an identifier of the stock.

t Subscript giving the number of periods remaining to the horizon date.

$r_i, i = M, N$ Dealer's expected per period rate of return in stock i in the absence of bid or ask fees.

$\dot{Z}_i \backsim N(0, \sigma_i^2), i = M, N$ Stochastic component in the return of stock i.

Q Dollar transaction size in each stock.[2]

$q_i(Q, -Q)$ Realization of the stochastic transaction process in stock i, defined as follows:

$$q_i(Q, -Q) \begin{cases} \xrightarrow{\lambda_i} & Q \quad \text{if} \quad b_i < b_i^0, \text{ 0 otherwise} \\ \xrightarrow{\lambda_i} & -Q \quad \text{if} \quad a_i < a_i^0, \text{ 0 otherwise} \\ \xrightarrow{1 - 2\lambda_i} & 0 \end{cases} \tag{3}$$

$\lambda_i, i = M, N$ Probability of a public sale (dealer purchase) of Q dollars or of a public purchase (dealer sale) of $-Q$ dollars in each period. Unless indicated otherwise we assume that transactions in a given stock or in different stocks are independent and therefore that the probability of two transactions in a period—λ^2—is small enough to be ignored.

$a_i, b_i, i = M, N$ Dealer's proportional reservation selling fee and proportional reservation buying fee, respectively. (The superscript "^{0}" indicates the fee of dealer B.) The reservation fee is the minimum fee such that the dealer's expected utility of terminal wealth would not be lowered were he to buy at a bid price, $p(1 - b_i)$ or sell at an asking price, $p(1 + a_i)$, where p is the dealer's opinion of the true price of the stock. Inventories (M, N) and transactions (Q) are valued by the dealer at his true price; and the return, $r_i, i = M, N$, is calculated on the basis of his opinion of the true price. For simplicity this price is the same for each stock, something that can be accomplished by appropriate stock splits.

The transaction process (3), facing dealer A, depends on his reservation fee relative to the fee of his competitor, B. (However, the fee which dealer A will set is not generally equal to his reservation fee. This issue is discussed below.) Although the probability of a public sale (dealer purchase) in stock i is λ_i, the

[2] Contrary to actual practice, we assume a fixed transaction value rather than a fixed number of shares per transaction. This is because the transaction value is the relevant variable in determining dealer fees. Since one can pick any value of Q, the assumption does not constrain our conclusions in any way.

probability that dealer A makes the transaction is λ_i only if $b_i < b_i^0$. In words, (3) says that, conditional on having the lower reservation fee, the process $q_i(Q, -Q)$ for dealer A takes on the value Q with probability λ_i and $-Q$ with the same probability, λ_i; and no transaction occurs with probability $1 - 2\lambda_i$. The probability of a transaction is zero if B has the lower reservation fee.

The interpretation of (1) or (2) is now as follows. The dealer enters period t with an inventory M_t, N_t that is changed according to the stochastic transaction process, $q_i(Q, -Q)$. The new inventory is then subject to return dynamics consisting of two components: a deterministic return, r_i, and a stochastic return, Z_i. Thus both transaction uncertainty and return uncertainty affect the position of the dealer as he moves from period t to period $t - 1$.

2. Cash. Cash is accumulated when the dealer sells securities (short selling is permitted) and paid out when the dealer buys securities. Any balance in the cash account earns (or pays) the risk-free rate, r. Then the value of the cash account is

$$F_{t-1} = (1 + r)[F_t + q_M(-Q + b_M Q, Q + a_M Q) + q_N(-Q + b_N Q, Q + a_N Q)] \quad (4)$$

where the transaction arrival process, q_M or q_N is the same as (3) with q taking on the values $-Q + bQ$, $Q + aQ$ or 0; i.e., the dollar cash flows are opposite in sign from the value of shares traded and different by the amount of the dealer's fee.

3. Base wealth. Base wealth is an efficient portfolio of the dealer's assets other than cash and share inventory generated by his dealership activity. Base wealth is used as collateral for the borrowing of money or shares. The dollar value of the dealer's base wealth, Y, follows the following return dynamics:

$$Y_{t-1} = (1 + r_Y)Y_t + Y_t Z_Y \quad (5)$$

where r_Y = expected return on base wealth and $Z_Y \backsim N(0, \sigma_Y^2)$.

B. Competitive Equilibrium and the Dealer's Problem

Analogous to a progressive auction, investors are assumed to interrogate dealers to elicit the maximum buying price and minimum selling price each dealer is willing to post at each point in time. Trading is assumed to take place only at the bid or ask prices posted by the dealers.

The dealer determines his buying and selling fee so as to maximize expected utility of terminal wealth for every state of the world in which the dealer may find himself. The dealer must set bid and ask prices not only in anticipation of random transaction arrivals and returns, but also in anticipation of what his competitor may do. Indeed only if the dealer believes he can make transactions at fees in excess of his reservation fees, is it worthwhile to continue as a dealer; and he must determine if his competitor's position will allow him this freedom.

Dealer A's elementary utility function over terminal wealth (at $t = 0$) is $U(W_0)$ where

$$W_0 = F_0 + Y_0 + M_0 + N_0 \quad (6)$$

Viewed from earlier periods ($t > 0$), terminal wealth is uncertain and dependent on dealer actions in the intervening periods. At each earlier period, dealer A acts on the basis of a derived utility function $J(\ )$ defined as

$$J(t, F, M, N, Y, F^0, M^0, N^0, Y^0) = \max_{\substack{a_M, b_M \\ a_N, b_N}} EU(W_0 \mid t, F, M, N, Y, F^0, M^0, N^0, Y^0) \quad (7)$$

The function $J(\ )$ is the solution to the maximization problem. It is the derived utility when the dealer follows the optimal strategy for setting a_i, b_i at each point in time; and it depends on the state of the world described by the variables in the function; namely, time remaining and the portfolio characteristics of both dealers.[3]

We have elsewhere solved an analogous function $J(\ )$ in a continuous time framework for the case of a single monopoly dealer (Ho and Stoll [11]). The procedure for two dealers trading two stocks is the same except that the optimal strategies of the two dealers must be simultaneously determined. Many insights may be achieved, however, without solving this complicated dynamic programming problem. We consider first the determinates of dealer reservation fees in a simple one period environment.

II. Dealer Reservation Fees Under a One Period Horizon

Only when there is one period remaining does the reservation fee of each dealer not depend on the inventory position of the other dealer. This is because all positions are assumed to be liquidated in the next period at next period's true price. As a result, the probability of trading next period and the price to be received are independent of the other dealer's actions.

Proposition 1: With independent transactions in M and N and with one period left to the horizon date, a dealer with inventories of M and N has reservation buying and selling fees for stock M that are approximately given by

$$b_M = \frac{1}{2}\,\sigma_M^2 R(Q + 2I_M) \quad (8a)$$

$$a_M = \frac{1}{2}\,\sigma_M^2 R(Q - 2I_M) \quad (8b)$$

where

$I_M = M + \beta_{NM} N$

$\beta_{NM} = \sigma_{NM}/\sigma_M^2$

σ_M^2 = per period variance of return of stock M.

σ_{NM} = per period covariance of return between stock M and N.

$R = \dfrac{-U''(W)}{(1 + r)U'(W)}$, a discounted coefficient of absolute risk aversion.

[3] It may not be necessary to know the portfolio characteristics of the other dealer directly since they may be inferred from pricing actions.

The same result, with appropriate changes in notation, holds for stock N.

Proof: When the return dynamics are given as in (5), the expected utility, using the Pratt-Arrow approximation, is given by

$$EU(Y_0) = U(Y_1) + r_Y Y_1 U'(Y_1) + \frac{1}{2} U''(Y_1)\sigma_Y^2 Y_1^2 \tag{9}$$

where $U'(\ )$ and $U''(\ )$ are the first and second derivatives of the elementary utility function. In an analogous fashion, for the return dynamics of the dealer's total portfolio (which is $W_1 = M_1 + N_1 + F_1 + Y_1$), the expected utility without any transaction is given by

$$EU(W_0) = U(W) + (r_Y Y + r_M M + r_N N + rF)U'(W) + \frac{1}{2} U''(W)(Y^2\sigma_Y^2 + M^2\sigma_M^2 + N^2\sigma_N^2 + 2MN\sigma_{MN} + 2MY\sigma_{MY} + 2NY\sigma_{NY}) \tag{10}$$

where the time subscripts on the right hand side of (10) are omitted, and it is understood that one period remains. Immediately after a sale of stock M (such that, in (10), M becomes $M - Q$ and F becomes $F + Q + \pi Q$, π being the dealer's proportional fee) the dealer's expected utility of his new uncertain terminal wealth position, $W_0^{\#}$ is

$$EU(W_0^{\#}) = EU(W_0) + U'(W)(1 + r)\pi Q + \frac{1}{2} U''(W)(\sigma_M^2 Q^2 - 2\ \sigma_M^2 MQ - 2\sigma_{MN} NQ) \tag{11}$$

where we assume stock M is in the efficient portfolio Y and, therefore, meets the standard first order condition:

$$r_M - r = \frac{-U''(W)}{U'(W)} Y\sigma_{MY}. \tag{12}$$

If expected utility is not to be lowered after the transaction, we require that $EU(W_0^{\#}) \geq EU(W_0)$. By the definition of the reservation selling fee a_M being the minimum of the selling fees π meeting this condition, Equation (8b) follows directly from (11) and the condition that $EU(W_0^{\#}) = EU(W_0)$. The reservation buying fee for stock M and the reservation fees for stock N can be calculated by the same procedure. Q.E.D.

Inspection of (8) shows that the dealer's fee depends on the stock's risk as measured by the variance of return, on his attitude toward risk, R, on the transaction size, Q, and on the current inventory of the dealer in his stocks, M and N. The appropriate measure of risk is σ_M^2, not σ_{MY}. The optimal holding of stock M, which is part of Y, has a risk measured by σ_{MY}, but this "systematic" risk is offset by the excess return of the stock, $r_M - r$, and therefore does not appear in (8). Because any holding of the stock in the dealer's inventory is non-optimal, σ_M^2 is the relevant measure of risk with respect to that amount by which inventory exceeds the optimal.

The impact of a transaction, Q, depends on the dealer's initial inventory. For example, a purchase by a dealer with a short position would be risk reducing and would result in a negative buying fee. In measuring the initial inventory for the purpose of evaluating the impact of a transaction in stock i, all stocks, j, correlated with stock i are aggregated, using $\beta_{ji} = \sigma_{ji}/\sigma_i^2$ as a weight. In other words, the size of the dealer's initial inventory is measured in terms of equivalent risk units of stock i, where equivalence is determined from the covariance of returns between other stocks and stock i. An implication of this result is that a change in the holding of any stock changes the bid and ask fees of every other stock with which it is correlated, the fee change depending on the size of the change in holdings and the degree of equivalence between the two stocks as measured by β_{ji}.

Corollary: From the definition of the spread as $s = a + b$, it follows immediately from (8) that the reservation spread in stock i is given by

$$s_i = \sigma_i^2 RQ \tag{13}$$

In a one period world, while the dealer's inventory affects a or b alone, the reservation spread is independent of inventory. A positive inventory raises the reservation buying fee but reduces the reservation selling fee by a corresponding amount. Thus, in adjusting to changes in inventory, the dealer raises or lowers both the reservation bid and ask price relative to the true price without changing the distance between them.

It is often suggested that diversification of the dealer's inventory beyond one or two stocks would reduce return risk and thereby reduce the reservation spread. However, since inventory does not appear in (13), it follows that the degree of diversification of the dealer's inventory has no effect on his reservation spread. This is because bid and ask prices are adjusted to the dealer's inventory with the result that the spread reflects only the risk of the incremental transaction.[4] Thus under the assumption of independent transactions processes for stocks, the spread is independent of the number of stocks in which the dealer makes a market. Furthermore the dealer is indifferent as to whether the next transaction is a purchase or sale because bid and ask fees are prices of contingent services determined such that dealer welfare is unaffected by the outcome of the transaction process.

When transaction processes in different stocks are not independent, dealer reservation fees and the reservation spread are affected by the degree to which transactions in the dealer's stocks are correlated. Intuitively, if returns are positively correlated, positively correlated transactions in two stocks have the same effect as increasing the transaction size in a single stock; and we know dealer fees increase with Q. Conversely offsetting transactions in two stocks with positively correlated returns have the same effect as reducing the transaction size and the fee in a single stock. The dealer is thus concerned to diversify his set of stocks against transaction uncertainty (rather than return uncertainty), and he will attempt to choose stocks in which transactions are offsetting or independent. The reservation fee in the case of correlated transactions is derived in the appendix.

[4] Base wealth of the dealer is optimally diversified and the issue is whether there is any benefit to diversification of that portion of total wealth which is the dealer's trading account. Since R is a function of wealth, the spread depends on dealer wealth.

We now turn to the process by which the market bid-ask spread is determined, and we shall assume in the following section that transactions are uncorrelated. If transactions are correlated the results would be modified in a straight forward way along the lines discussed in the appendix.

III. The Market Bid-Ask Spread Under a One Period Horizon

A. Dealer Pricing Strategy and Conditions on the Market Spread

While reservation bid and ask prices may differ across dealers, there is at each moment only one market bid price and one market ask price. The competitive market equilibrium we assume reduces the profits that would accrue to a single dealer (as in Ho and Stoll [11]), but it does not necessarily cause each dealer to quote his reservation price. Indeed, if that were the case, no incentive to be a dealer would exist. A dealer is able to quote a fee above his reservation fee because he is in a temporarily advantageous position (in terms of inventory and/or attitude toward risk) with respect to his competitor, and thereby earns a temporary producer surplus. Thus, which dealer makes the next transaction and what market fee above the reservation fee can be charged depends on the relative positions of the dealers.

Consider a market with any number of competing dealers. Rank dealers in ascending order of their reservation buying fees and their reservation selling fees. The first dealer is the dealer with the lowest reservation fee. The second dealer is the dealer with the next lowest reservation fee, which may be the same fee. Reservation fees of different dealers are identified by different superscripts as follows:

First seller	a	b
Second seller	a^0	b^0
Second buyer	a'	b'
First buyer	a^*	b^*

where

$$a \le a^0 \le a' \le a^* \quad \text{and} \quad b^* \le b' \le b^0 \le b$$

Then the market bid-ask spread, s, is

$$s = \frac{p(1 + a^0) - p(1 - b')}{p} = a^0 + b'. \tag{14}$$

The dealer with the lowest reservation fee has no incentive to quote that fee. He will instead quote his nearest competitor's reservation fee (less a small amount). In other words the second best dealers set the market spread. This result requires knowledge of competing dealers' inventory positions, which might be inferred from observed transactions, and knowledge of competing dealers' risk attitudes.[5]

[5] Alternatively, one can view the process of setting public bid-ask quotes as a progressive auction in which all dealers (except one) raise bid prices and lower ask prices to their reservation prices and thereby reveal their reservation prices. The theory of such auctions is discussed in Vickrey [20].

The first seller and the first buyer are determined by their risk attitudes and inventories. Given identical inventories for all dealers, the dealer with the lowest value of R is both first buyer and first seller. However, as he acquires (disposes of) inventory, he changes his fee so that he loses his position as first buyer (seller). Thus the least risk averse dealer does not have a natural monopoly.

Given identical R, the dealer with most inventory of a stock (measured in equivalent risk units of all his stocks) is the first seller. If a dealer trades two stocks, M and N, the inventory of the first seller in equivalent risk units of stock M is said to exceed that of the second seller if

$$I_M = M + \beta_{NM}N > I_M^0 = M^0 + \beta_{NM}N^0,$$

a condition that can be satisfied even if $M < M^0$. The dealer with the least inventory of a stock (measured in equivalent risk units of all stocks) is the first buyer.

B. Inter-dealer Trading and the Equilibrium Distribution of Inventories Under Homogeneous Risk Preferences

Since inventory differences are responsible for differences in reservation fees, one is immediately led to ask whether inter-dealer trading would eliminate such differences and thereby would eliminate the producer surplus dealers periodically earn. Inter-dealer trading arises if a dealer prefers to trade immediately with another dealer rather than wait to trade with the next incoming order with probability λ. We modify the model to allow an instant of time for inter-dealer trading prior to the public order flow and the period of price uncertainty. Consider a market with two dealers trading two stocks and examine the problem from the point of view of dealer A and stock M.

Proposition 2: Assume dealer A is the first seller and has a larger inventory of stock M than does dealer B the second seller; and assume the dealers have the same R. Then A will choose option 1—wait to sell Q of stock M to the next incoming market order if

$$\pi_M > R\sigma_M^2[I_M - \tfrac{1}{2}Q - Q(\lambda_M + \lambda_N\beta_{NM})]. \tag{15}$$

Otherwise A will choose option 2—sell Q of stock M immediately, and pay a fee of π_M.

Proof: Under option 1 (to wait), A's expected utility of terminal wealth can be shown to be

$$EU(W_0) + (1 + r)U'(W_1)Q[\lambda_M(a_M^0 - a_M) + \lambda_N(a_N^0 - a_N)] \tag{16}$$

where $EU(W_0)$ is given by (10) and a_M and a_N are given by (8). Equation (16) says that the expected increment to A's expected utility of waiting for an incoming market order depends on the difference between the market fees (a_M^0, a_N^0) and A's reservation fees (a_M, a_N) weighted by the probabilities (λ_M, λ_N) of receiving the differences. Note that marginal utility is being evaluated just prior to a transaction. Thus, the probability of each transaction enters. There is no possibility that A will make the next buy transaction in either stock, and buying fees do not therefore appear. Under option 2 (to trade now and pay a fee of π_M), A's expected utility of terminal wealth is

$$EU(W_0^{\#}) + (1 + r)U'(W)Q[\lambda_M(a_M^0 - a_{M\#}) + \lambda_N(a_N^0 - a_{N\#})]$$

where $EU(W_0^{\#})$ is given by (11) and represents the dealer's expected utility after a transaction of size Q, and $a_{M\#}$ and $a_{N\#}$ represent the reservation ask fees of dealer A immediately after the inter-dealer transaction in stocks M and N and before realization of the random transactions and return processes. (We assume A continues to be the first seller.) For option 1 to exceed option 2 requires:

$$U'(W)(1 + r)\pi_M Q - \tfrac{1}{2}U''(W)(\sigma_M^2 Q^2 - 2\sigma_M^2 MQ - 2\sigma_{MN} NQ) + U'(W)Q[\lambda_M(a_{M\#} - a_M) + \lambda_N(a_{N\#} - a_N)] > 0$$

Using (8) and simplifying yields the desired result, (15). Q.E.D.

Thus, whenever the fee, π_M, that A must pay B to trade immediately meets condition (15), A would prefer to set an ask fee and bear the uncertainties of waiting to trade with an incoming market order. The decision of dealer A is thus like the decision of any investor coming to the market considering placing a limit order. The investor will place a limit order and trade with the next market order if the fee charged in the market satisfies (15). Otherwise the investor will trade at the market quote.

From (14) we know that dealer A pays the market bid price set by the second buyer. In the case of two dealers, A himself is the second buyer; and therefore $\pi_M = b_M$. Substituting the expression for b_M (Equation 8) for π_M in (15) shows that (15) is always satisfied under the reasonable assumption that $(1 + \lambda_M + \lambda_N \beta_{NM}) > 0$.[6] The reason is simple. Even if λ, the probability of selling to a market order, were zero, A would not sell to B and *pay* b_M to reduce inventory he is willing to add to at b_M. If $\lambda > 0$, A has the additional possibility of selling to a market order and earning a fee.

Proposition 2 holds for a market with any number of dealers since B can be chosen to be any dealer other than A. However, when there are more than two dealers, the market buying fee may be set by a dealer other than A. In this case, inter-dealer trading may arise. As noted above let a prime (′) refer to the second dealer on the buy side who sets the market buying fee. Then $\tau_M = b'_M$ in (15). Applying (8) and simplifying yields the condition for A to wait for an incoming order rather than to sell at b'_M:

$$I_M - I'_M < Q[1 + \lambda_M + \lambda_N \beta_{NM}] \tag{17}$$

Since inter-dealer trading will not occur if condition (17) is met, (17) imposes a limit on the divergence of inventories. We examine this first in the single stock case. Reference to Table 1 is helpful in understanding our notation and the allowable inventory divergence. In the case of one stock (17) becomes

$$M - M' < Q(1 + \lambda_M) \tag{17.a}$$

[6] The substitution yields

$$R\sigma_M^2 Q > -R\sigma_M^2 Q(\lambda_M + \lambda_N \beta_{NM})$$

which is always satisfied when $(1 + \lambda_M + \lambda_N \beta_{NM}) > 0$.

Since $0 < \lambda_M < 1$, the largest divergence between the first seller (dealer A) and the second buyer (dealer E) is $M - M' = Q$. Similarly, the largest divergence between the first buyer (dealer F) and the second seller (dealer B) is $M^0 - M^* = Q$. This implies that the largest divergence of inventories between first seller and first buyer is $M - M^* = 2Q$. In Table 1 inventory patterns one and two are not equlibria. In each case there is an incentive for interdealer trading because market bid and ask prices determined by the inventories of the second dealers on each side of the market are "cheap" relative to the inventory holding costs of some of the dealers.

In inventory patterns three and four, all pairs of dealers meet the conditions for no interdealer trading, (17.a). In pattern three, there are no ties of inventory positions at the margin so that the second dealers have no probability of trading with an incoming order. As a result (17.a) applied to second dealers yields $M^0 - M' < Q$, which implies that all dealers other than first dealers have identical inventories. If only three dealers make a market, two will be first dealers; and there can only be one second dealer. In that case it is necessarily true that $M^0 - M' < 0$ since the second seller and second buyer are the same dealer.

In pattern four, a first and a second dealer have the same inventory and would therefore quote the same price. We assume that the second dealer as well as the first has a non-zero probability of trading with the next incoming order. In that case λ_M in (17.a) is not zero when applied to second dealers, and $M^0 - M' = Q$. To the extent that dealers differ on dimensions other than those considered in this paper—such as clerical efficiency—the likelihood of ties at the margin is reduced and with it the likelihood that second dealers have different inventories.

If one makes the reasonable assumption that $0 < \lambda_M + \lambda_N\beta_{NM} < 1$, everything that has been said about the individual stock also applies to the dealer's inventory stated in risk units of stock M, I_M. The difference in inventories of first dealers is $I_M - I_M^* < 2Q$. The difference in inventories of any other pairs of dealers is Q, or less, directly from (17). If no second dealer has the same inventory as a first dealer, the difference in inventories of second dealers is 0 from (17); and, in this case, the probability of trading is zero for a second dealer.

In the case of two stocks, the allowable divergence of inventories in stock M is dependent on the divergence in stock N:

$$M - M' < -\beta_{NM}(N - N') + Q(1 + \lambda_M + \lambda_N) \tag{18}$$

This follows directly from (17) by writing I_M and I'_M in terms of individual stocks as in proposition 1. When returns are correlated, the inventory divergence in M is dependent on the divergence in N. Assuming $\beta_{NM} > 0$, a divergence in excess of Q in M can be offset by an opposite divergence in N.

Corollary: Interdealer trading will not be observed if all transactions of a stock are of a fixed size, Q.

Proof: Once an equilibrium distribution of inventories is established that satisfies (17) for all pairs of dealers, an incoming order cannot disturb inventories in a way that induces interdealer trading. If dealers have identical inventories, an incoming transaction can only produce an inventory difference of Q, which is insufficient to induce interdealer trading. If dealer inventories differ, an incoming transaction is carried out by the dealer whose inventory is most divergent, and

Table 1

Example of Alternative Inventory Patterns in a Market for One Stock Made by Six Dealers

Dealers Ranked on Inventory	Dealer Description	Reservation Fee	Disequilibria			Equilibrium, No Ties at Margin		Equilibrium With Ties at Margin	
			Prob. of a Trade	Pattern One	Pattern Two	Prob. of a Trade	Pattern Three	Prob. of a Trade	Pattern Four
A	First Seller	a	λ	$M = 3Q$	$M = 2Q$	λ	$M = Q$	$\frac{1}{2}\lambda$	$M = Q$
B	Second Seller	a^0	0	$M^0 = 2Q$	$M^0 = Q$	0	$M^0 = 0$	$\frac{1}{2}\lambda$	$M^0 = Q$
C		0	0	0	0	0	0	0	
D		0	0	0	0	0	0	0	
E	Second Buyer	b'	0	$M' = -Q$	$M' = -Q$	0	$M' = 0$	$\frac{1}{2}\lambda$	$M' = 0$
F	First Buyer	b^*	λ	$M^* = -2Q$	$M^* = -Q$	λ	$M^* = -Q$	$\frac{1}{2}\lambda$	$M^* = 0$

who is therefore able to offer the lowest fee. As a result the transaction can only narrow inventory differences. Q.E.D.

C. Equilibrium Market Spread

The "second best" pricing rule specified above and the conditions for inter-dealer trading given in proposition 2, along with the resulting equilibrium distribution of inventories implied by (17) yield limits on the value of the equilibrium market spread when dealers have identical coefficients of absolute risk aversion and identical opinions of the true price of the stock. The limits on the spread depend on the number of dealers making a market in the stock.

Proposition 3: Under homogeneous preferences and opinions, the equilibrium market bid ask spread in a stock satisfies the following conditions, depending on the number of dealers making a market in the stock.

Two dealers: $$s \geq R\sigma^2 Q \tag{19.a}$$

Three dealers: $$s = R\sigma^2 Q \tag{19.b}$$

More than three dealers $$0 \leq s \leq R\sigma^2 Q \tag{19.c}$$

where $R\sigma^2 Q$ is the reservation spread of any dealer.

Proof: From (14), the market bid-ask spread is $s = a^0 + b'$. From (8) and the assumption of identical R, $s = R\sigma^2[Q - (I^0 - I')]$. Thus the equilibrium market spread depends on the equilibrium inventory difference between the second seller and the second buyer which is specified by proposition 2 and by (17). For two dealers, the nearest competitor to the first (and only) seller is the first (and only) buyer. Thus I^0 is to be interpreted as the first buyer's inventory and I' is to be interpreted as the first seller's inventory. By definition of first buyer and first seller, $I^0 \leq I'$ which is sufficient to prove (19.a).[7] For three dealers, we have shown $I^0 - I' = 0$, which is sufficient to prove (19.b). For four or more dealers, we have shown $0 \leq I^0 - I' \leq Q$, which is sufficient to prove (19.c). Q.E.D.

While the market spread may be less than $R\sigma^2 Q$ when there are more than three dealers, this would be relatively rare since it depends on the existence of ties on each side of the market and for an inventory difference between second dealers of Q. In the absence of ties, second dealers would have the same inventory and the market spread would be $R\sigma^2 Q$.[8] Furthermore, should $s = 0$, there is a tendency for inflowing orders to cause this spread to jump to its upper bound of $R\sigma^2 Q$. A public sell order will be traded by the first buyer who will then lower his bid price. Correspondingly a public buy order will be traded by the first seller who will then raise his ask price.

With two dealers there is a corresponding tendency of inflowing orders to push

[7] In the case of one dealer, which we have considered in Ho and Stoll [11], there is no limit on the spread from the supply side. There is a limit determined by the demand for dealer services, which is modeled in the cited paper. With only two competing dealers, demand conditions could affect the desirability of setting a price permitted by the competitor's inventory position. However, the flow of incoming orders keeps inventories from diverging by an amount that would make this an important consideration.

[8] This could, for example, be broken by slight differences in risk aversion or clerical costs.

the spread to its lower bound of $R\sigma^2Q$. A public sell order will be purchased by the dealer with the smaller inventory. This process narrows inventory differences and narrows the market spread. Under homogeneous preferences there is thus a tendency for the equilibrium market spread to be the reservation spread of any dealer.

IV. Equilibrium Market Spread Under Heterogeneous Opinions

In this section the effect of disagreement about the true price of the stock on the pattern of dealer inventories and on the equilibrium market spread is analyzed. We assume there are more than three dealers each with the same R, and that all dealers initially agree and have $I = 0$. Now suppose new information generates disagreement about the true current price of the stock without changing opinions of σ^2 or the underlying market clearing price. Disagreement induces trading until the inventory risk of the shares acquired by optimistic dealers tends to offset the optimists' more favorable opinion, and conversely such that the inventory risk of the short position of pessimistic dealers tend to offset the pessimists' unfavorable opinion.

Under heterogeneous expectations, (14) may be written as

$$s = \frac{p^0(1 + a^0) - p'(1 - b')}{1}$$

where subscripts for the stock are suppressed and where

1 underlying market clearing price in the absence of any dealer trading costs—an average of dealer's opinions of the true price. The price is scaled to one for convenience.[9]

$p^0 = (1 + \varepsilon^0)$ opinion of stock's true price of second seller who sets ask price; $\varepsilon^0 < 0$ is the deviation of his opinion from the average.

$p' = (1 + \varepsilon')$ opinion of true price of second buyer who sets bid price; $\varepsilon' > 0$ is the deviation of his opinion from the average.

We can write the market spread as

$$s = (a^0 + b') - (\varepsilon' - \varepsilon^0) \tag{20}$$

(where we have assumed $a^0\varepsilon^0 = -b'\varepsilon'$), and using (8) this becomes

$$s = \tfrac{1}{2} R\sigma^2(Q^0 + Q') - R\sigma^2(I^0 - I') - (\varepsilon' - \varepsilon^0) \tag{21}$$

where Q^0 and Q' are the transaction values at the personal valuations of each of the dealers.

Under homogeneous opinions, $Q^0 = Q' = Q$, $e^0 = \varepsilon' = 0$, and $I' - I^0 \leq Q$ in equilibrium. Under heterogeneous opinions, differences of opinion generate inventory differences such that the last two terms of (21) tend to be offsetting.

[9] When dealers disagree it is not clear how one should define the proportional market spread. We have simply chosen to divide by the average "true" price. In fact, otherwise identical dealers will, simply because they have different opinions of the true price, define different proportional spreads. This causes some unnecessary complexities that do not invalidate the basic points. Thus we have simply assumed that proportional spreads are the same for different dealers.

However the effect may not be exact just as in the case of homogeneous expectations $(I' - I^0)$ is not necessarily zero. Define the net unwanted inventory as

$$\Omega^0 = I^0 - \varepsilon^0/R\sigma^2 \tag{22.a}$$

$$\Omega' = I' - \varepsilon'/R\sigma^2 \tag{22.b}$$

Suppose, for example, that the second seller is a pessimist so that $\varepsilon^0 < 0$. He will sell shares to acquire a short inventory position. To the extent his short position exceeds or falls short of his desired inventory determined by ε^0 and $R\sigma^2$, he has unwanted inventory of Ω^0, The market spread can now be written as

$$s = R\sigma^2 Q - R\sigma^2(\Omega^0 - \Omega') \tag{23}$$

where we let $Q = \frac{1}{2}Q^0 + \frac{1}{2}Q'$, which is the transaction value at the average price. It can be shown that the propositions about the equilibrium distribution of inventories under homogeneous opinions also apply to the net unwanted inventories, Ω. And since the market equilibrium spread depends directly on Ω^0 and Ω', it can be shown that proposition 3 specifying the equilibrium market spread holds under heterogeneous opinions.

The key concept, as before, is that if the unwanted inventory difference between second dealers exceeds Q, interdealer trading is induced. In equilibrium, pessimistic dealers have acquired a (desired) short position and attendant inventory risk which is reflected in a high bid and ask price relative to their opinion of the true price. Optimistic dealers have acquired a (desired) long position and attendant inventory risk which is reflected in a low bid and ask price relative to their opinion of the true price. In equilibrium, the reservation bid and ask prices of all dealers are in the neighborhood of the market clearing price. Under proposition 3, reservation bid or ask prices may differ at most by the fee on a transaction of size Q.

The reservation spread of each dealer is independent of his inventory. Therefore, under the assumption of identical R, all dealers have the same reservation spread under heterogeneous as well as homogeneous opinions. Furthermore, as in the case of homogeneous opinions, there is a tendency for the observed market spread to be the reservation spread of any dealer. This should be of some comfort to those undertaking empirical studies of the spread.

V. Effect on Dealers' Reservation Fees of Extending Horizon to Two Periods

The dynamic programming character of a dealer's problem and the dependence of a dealer's fee on the position of his competitors can be shown in a two period case. In this context the effect of serially correlated transactions can also be examined.

In this section a dealer—A—operating in a market with several other competing dealers is analyzed. For simplicity the analysis is restricted to one stock—stock M—and all dealers are assumed to have identical absolute risk aversion and to start with zero inventory of the stock. The problem is considered from the perspective of what selling fee to set with two periods remaining. In the dynamic programming framework discussed earlier that selling fee is set such that expected derived utility of wealth (the J function) next period, under optimal pricing strategy in later periods, is maximized.

Proposition 4. Assume that two periods remain ($t = 2$) and initial inventory of each dealer of the one stock being traded is $M = 0$. Since dealers are assumed to have identical R, this implies that no dealer earns a producer surplus at $t = 2$. Assume that, given a purchase at $t = 2$ by A, the following conditional probabilities apply with respect to transactions in the next period ($t = 1$):

Event at $t = 1$ Given a sale by A at $t = 2$	Market's Conditional Probability	Dealer A's Conditional Probability
No trade	θ	$\theta + \gamma$
Purchase	μ	μ
Sale	γ	0

Then A's reservation selling fee at $t = 2$ is

$$a_2 = \frac{1}{2(1 + r)} \sigma_M^2 RQ(1 - 2\mu) \tag{24}$$

Proof: When one period remains, dealer A's derived utility, given a sale at $t =$ 2 is

$$J(1) = E_1 U(W_0^+) + (1 + r)U'(W_1)Q\mu(b_1^0 - b_1) \tag{25}$$

where $E_1 U(W_0^+)$ is given by (10) letting $N = 0$ and $M = -Q$, and the subscript on the expectation operator indicates the point in time at which the expectation is taken. The last term in (25) represents the marginal utility of the expected producer surplus earned by A at $t = 1$. Since A is assumed to be the seller at $t = 2$, he will be the first buyer at $t = 1$ earning $b_1^0 - b_1$, where b^0 is the reservation buying fee of the second buyer. From the principle of dynamic programming we know that $J(2) = E_2 J(1)$; for if a change in derived utility were expected, the dealer would choose some other strategy for selecting buying and selling fees.

At $t = 2$, A set a selling fee, a_2, such that the derived utility conditional on a sale, $J(2 \mid \text{sale})$, equals the derived utility conditional on no transaction in the remaining two periods. The derived utility when there is no trading for two periods is the two period analog of (10) where we have made the additional assumption that $M = N = F = 0$:

$$J(2 \mid \text{no trade}) = E_2 E_1 U(W_0)$$

The derived utility if A sells Q shares at $t = 2$ is

$$\begin{aligned} J(2 \mid \text{sale}) &= J(2, F_2 + Q, M_2 - Q, Y_2, \bullet) + J_F(2, F_2 + Q, M_2 - Q, Y_2, \bullet)aQ \\ &= E_2 J(1) + U'(W_0)aQ \\ &= E_2 E_1 U(W_0^+) + E_2(1 + r)U'(W_1)Q\mu(b_1^0 - b_1) + U'(W_0)aQ \end{aligned}$$

Setting $J(2 \mid \text{no trade}) = J(2 \mid \text{sale})$, noting that

$$E_2E_1U(W_0^+) - E_2E_1U(W_0)$$
$$= E_2[(-r_MQ + rQ)U'(W_1) + \tfrac{1}{2}U''(W_1)(Q^2\sigma_M^2 - 2QY\sigma_{MY})],$$

and dropping certain small cross product terms gives

$$a = \frac{1}{1+r}\left[\frac{1}{2}\,\sigma_M^2RQ - \mu(b_1^0 - b_1)\right]$$

Using (8) and noting the assumption that M_1^0 0, $M_1 = -Q$ and $N_1^0 = N_1 = 0$, yields the desired result. Q.E.D.

The implications of (24) are quite straight forward. The greater the probability of a reverse transaction in the next trading interval the lower the reservation fee today. Indeed, the fee will be lower than if the dealer were able to cover his short position in the next trading interval at the true price (as in the one period case). This is due to the fact that he is able to collect a producer surplus next period by buying at a price below the true price. Since we have assumed that all dealers start from identical positions, each would have a lower reservation fee; and as a result the market spread is lower in a two period framework. Since actively traded stocks have a larger μ, it also follows directly that actively traded stocks have a lower spread than infrequently traded stocks.

It should be noted that even though purchases and sales have the same probability of arrival, dealer competition introduces negative serial correlation in transactions as far as the inventory of a single dealer is concerned. Dealer A, having sold, is unlikely to sell again. He dominates the bid side of the market in the second period and his probability of a purchase next period is the probability of a public sale, as reflected in (24).

Were the assumption of a single transaction of size Q to be modified, A would no longer dominate the bid side in the second period. If several dealers each purchase Q at $t = 2$, competition would deny them a producer surplus in period 2 and the fee at $t = 2$ would be the same as in the one period case. This is immediately evident from (25).

VI. Implications: Exchange Markets Versus Dealer Markets

This paper has implications for two strands of research that have been subsumed under the heading of "the microstructure of securities markets." First, it provides a basis for analyzing the transaction-by-transaction price behavior of individual stocks and specifies the role of dealers in causing the observed transaction price to deviate from the underlying true price. Research in this area is frequently concerned with statistical modeling of the time series of security returns. For a review of this strand of research see Cohen, Hawawini, Maier, Schwartz and Whitcomb [4]. Second, the paper provides a framework for analyzing alternative structures of securities markets. First we show that the model is applicable to the analysis of auction markets in which investors place limit orders; and, second, we consider the relationship between spreads in a dealer market and an auction market such as the NYSE.

A. Limit Orders

In an auction market, investors (represented by brokers) trade directly with each other either by meeting at the same time or, more likely, by trading against a previously entered limit order. In a dealer market, investors (represented by brokers) trade at the quoted bid or ask price of dealers. While our analysis has been framed in terms of a dealer market, the framework is equally applicable to an auction market.[10] The decision of an investor in an auction market to leave a limit order or to trade immediately with an existing limit order is exactly analogous to the decision of a dealer to post his price and wait or to trade immediately with another dealer.

In an auction market the depth and liquidity of the market is sometimes measured by the thickness of the book of limit orders (i.e., the number of orders at each price). However, under the assumption of our model one would expect a relatively thin book regardless of the activity in the stock. We have shown that second dealers (or investors), who have no chance of trading with the next incoming order, would rather trade immediately than place an order and wait to trade. Anyone with a reservation bid price below the market bid price would sell stock at the market bid or place a limit order to sell inside the market ask rather than placing a buy order at a price which has no probability of being executed against the next incoming order. Analogously, anyone with a reservation ask price above the market ask would buy at the market ask or place a limit order to buy slightly inside the market bid rather than placing a limit sell order which has no probability of being met in the next trading interval. Since all orders except the one at the margin in each side of the market (and those tied with it at the same price) will be executed, a thin book results.

If monitoring costs are high for investors or if investors face short sales constraints, they may choose to place limit orders outside the market spread in anticipation of a hoped-for change in the market price. In the absence of continuous monitoring such limit orders offer an option to traders continuously in the market to trade at known prices should new information justify a change in the underlying market price. On the New York Stock Exchange the option is particularly valuable to the specialist because he has knowledge of the book of limit orders.[11]

B. Bid-ask Spread on the New York Stock Exchange (NYSE) *and on the Over-the-Counter* (OTC) *Market*

The NYSE combines aspects of auction and dealer markets in that there is a single dealer—the specialist—who is responsible for quoting bid and ask prices while there is also a book of limit orders, maintained by the specialist. In the OTC markets in stocks or bonds, several dealers compete by simultaneously

[10] Cohen, Maier, Schwartz & Whitcomb [5] analyze an auction market.

[11] Copeland and Galai [3] have treated the deviation of the bid or ask price from the true market price as reflecting the premium on an option to trade at a known price. This assumes that even the dealer cannot monitor the market continuously and is unable to change bid and ask prices rapidly in response to new information. While we are not willing to accept this assumption, we are willing to assume their model applies to investors who face monitoring costs.

quoting bid and ask prices. An important issue in the development of securities markets is whether an evolving national market should be structured along the lines of the NYSE or of the OTC.

It has been argued elsewhere, both on theoretical and empirical grounds, that the specialist system is a monopoly franchise and as such sets a monopoly spread.[12] Furthermore, markets in certain stocks may be poor simply because of excessive risk aversion or inadequate wealth of the particular specialist making a market in the stock. From this perspective, a market of competing dealers is generally judged to be preferable. However, empirical work by Newton and Quandt [12] indicates that the OTC market spread appears to be higher than the specialist's spread on the NYSE when similar stocks are compared.[13] The analysis of this paper provides a framework for examining this result and for comparing the two types of markets.

An important characteristic of the specialist system is that each order is funneled through the single specialist, thereby enforcing the transaction size limit, Q, for which the specialist's quote applies. The specialist then has the opportunity to reset his quotes prior to the next transaction. On the OTC it is possible for an investor simultaneously to make several transactions each of size Q with the different dealers. In effect the dealer market offers greater liquidity at each spread and, therefore, could be expected to have a larger spread.

To see the effect more precisely, imagine two ways of organizing trading in a single stock. One in which there is a single specialist with wealth of W and one—an OTC market—in which there are 10 dealers each with wealth $W/10$. Now suppose a sell order of $10Q$ is to be traded. On the OTC market that order could be split into 10 orders and sold at posted bid prices. On the exchange only Q could automatically be traded at the posted bid, and successive sales would take place at lower bid prices. Because the coefficient of absolute risk aversion declines with wealth, each of 10 small dealers sets a higher spread for a transaction of size Q than a large dealer with the same total wealth. Assume the coefficient of relative risk aversion $Z = RW$ is constant. Then, from (13), the reservation spread of the specialist is $\sigma_i^2 ZQ/W$; and the reservation spread of each of the 10 dealers is $\sigma_i^2 10ZQ/W$. We have shown that the market spread tends to the reservation spread of any dealer. Thus the market spread can be expected to be higher in the OTC market than on the NYSE.

VII. Summary and Conclusions

The behavior of competing dealers in securities markets is analyzed in this paper. Securities are characterized by stochastic returns and stochastic transactions of a fixed size, Q. Reservation bid and ask fees of dealers are derived under alternative assumptions about the degree to which transactions are correlated across stocks at a given time and over time in a given stock. In a multi-period

[12] See for example Smidt [14], Stoll [16] and Ho and Stoll [11].

[13] Newton and Quandt calculate the market spread for OTC stocks traded on NASDAQ (an automated quotation system) by taking the difference between the lowest ask price of any dealer and the higher bid price of any dealer—the inside quote.

framework, reservation fees of a dealer depend on the anticipated actions of other dealers. The dynamic programming formulation of this problem is presented for the case of two dealers in two stocks; but, except for a two period case, solutions are not attempted.

The interaction of any number of dealers and the determination of the market spread are examined under the assumption that dealers have a one-period horizon. The market spread is shown to be determined by the second best dealers. An equilibrium market spread is defined to exist when the distribution of inventories across dealers is such that no dealer wishes (at that moment) to trade with any other dealer. The conditions under which the dealer would trade immediately with another dealer rather than wait for an incoming order is specified. We show that the equilibrium market spread must be non-negative. Under homogeneous risk preferences and with at least four competing dealers, the equilibrium market spread may be zero; but such a spread is not stable over time and would tend to be driven to the reservation spread of an individual dealer by the flow of incoming orders. That the equilibrium market spread tends to the reservation spread of any dealer holds true under heterogeneous as well as homogeneous opinions about the true price of the stock (so long as dealers have homogeneous risk preferences and agree on the variance of the stock's return).

The model is shown to be applicable to the placement of limit orders by investors as well as the placement of bid and ask prices by dealers, and it can therefore be used in comparing spreads between dealer markets such as the OTC and exchange markets such as the NYSE. We show that spreads on comparable stocks may be expected to be higher in an OTC market with many dealers than in an exchange market with a single specialist because a market with many dealers stands ready to trade more shares at quoted prices.

The model potentially has a much wider applicability than to the securities markets. Any sequential bidding problem under uncertainty in which a bidder's price depends on his position relative to that of his competitors may be analyzed in our framework. The problem of bidding on securities is perhaps simpler than others since the stochastic processes for returns and transactions can be specified in a realistic way.

Appendix: Dealer Reservation Fee When Trading in Two Stocks is Correlated

When the two stocks arrive together, the dealer can only determine fees for buying or selling the combination. The fee for each stock is indeterminate without some additional assumptions. To investigate the effect of correlated transaction arrivals in stock M and N, we assume the stocks have the same risk—$\sigma_M^2 = \sigma_N^2 = \sigma^2$—and the inventory of the stocks is $M = N = 0$. Under independent transactions arrival this implies $b_M = b_N = a_N$ by proposition 1. We assume the fees are equal under correlated transactions arrival and investigate the level of the common fee as compared with independent arrival.

Proposition A: Assume that one period remains to the horizon date, that $\sigma_M^2 = \sigma_N^2 = \sigma^2$, that $M = N = 0$, that the joint probability of transactions in stock

M and N is symmetric as given in Table A. Then the common bid or ask fee, π_i, in each stock, i, is given by

$$\pi_i = \frac{1}{2} R\sigma^2 Q\left(1 + 2\beta_{ji} \frac{\mu - \gamma}{\lambda_i + \mu + \gamma}\right), \qquad i = M, N \tag{A.1}$$

where $\beta_{ji} = \sigma_{ji}/\sigma^2$.

Proof: We sketch the proof from the perspective of a dealer setting the bid price in stock M but because of our assumptions any other perspective would give a corresponding result. Define the change in the dealer's expected utility of terminal wealth, ϕ_k, for each possible event, k, under which M may be purchased:

Buy M, no trade in N	$\phi_1 = EU(W^{\#}_{01}) - EU(W_0)$
Buy M, Buy N	$\phi_2 = EU(W^{\#}_{02}) - EU(W_0)$
Buy M, Sell N	$\phi_3 = EU(W^{\#}_{03}) - EU(W_0)$

The values of $EU(W^{\#}_{0k})$ are calculated in the manner of (11) for the alternative uncertain terminal wealth positions implied by each event. We place the dealer just an instant before the transaction and assume one period of return uncertainty remains after the realization of the event. Then the expected change in expected utility of terminal wealth conditional on a purchase of stock i, ψ_i, is

$$\psi_i = \frac{\theta}{\lambda_i}\phi_1 + \frac{\mu}{\lambda_i}\phi_2 \frac{\gamma}{\lambda_i}\phi_3$$

Substituting for ϕ_k in the above equation and requiring $\psi_i = 0$ then yields (A.1) Q.E.D.

We see from a comparison of (A.1) and (8) (under the assumption of $M = N = 0$) that the effect of correlated transactions is to modify the transaction size, Q. The condition under which (A.1) becomes (8) are (a) $\beta_{ji} = 0$. If returns are not correlated, correlated transactions do not increase return risk. (b) $\mu = \gamma$. If the probability of parallel transactions equals the probability of offsetting transactions, there is no increase in expected return risk (c) If $\mu = \gamma = 0$, we have the assumption leading to (8). Since the spread is the sum of the bid and ask fees, the spread is changed in proportion to the change in the effective transaction size.

Table A

Probabilities of Transactions in Stocks M and N

Transaction Amount in Stock M	Transaction Amount in Stock N: Q	$-Q$	0	Marginal Probabilities of Transactions in Stock M
Q	μ	γ	θ	λ_M
$-Q$	γ	μ	θ	λ_M
0	θ	θ		$1 - 2\lambda_M$
Marginal Probabilities of Transactions in Stock N	λ_N	λ_N	$1 - 2\lambda_N$	

The practical effect of proposition A is that the dealer is concerned about the stocks in which he makes a market when transactions are correlated. One can imagine two sets of stocks for which return variances of all stocks are the same. Yet the dealer handling the set of stocks for which transactions arrive in parallel would set a higher spread for each stock than the dealer for which transactions are offsetting or independent.

REFERENCES

1. Y. Amihud and Y. Mendelson. "Dealership Market: Market Making with Inventory." *Journal of Financial Economics* 8 (March 1980), 31–53.
2. G. Benston and R. Hagerman. "Determinants of Bid-Asked Spreads in the Over-the-Counter Market." *Journal of Financial Economics* 1 (March 1974), 353–364.
3. T. Copeland and D. Galai. "Information Effects on the Bid-Ask Spread" UCLA Working Paper, (October, 1982).
4. K. Cohen, G. Hawawini, S. Maier, R. Schwartz and D. Whitcomb. "Implications of Microstructure Theory for Empirical Research on Stock Price Behavior." *Journal of Finance*, 35 (May, 1980), 249–257.
5. K. Cohen, S. Maier, R. Schwartz and D. Whitcomb. "Transaction Costs, Order Placement Strategy, and the Existence of the Bid-Ask Spread." *Journal of Political Economy* 89 (April, 1981), 287–305.
6. Harold Demsetz. "The Cost of Transacting." *Quarterly Journal of Economics* 82 (February, 1968), 33–53.
7. K. Garbade and W. Silber. "Structural Organization of Secondary Markets Clearing Frequency, Dealer Activity and Liquidity Risk," *Journal of Finance* 34 (June 1979), 577–593.
8. M. Garman. "Market Microstructure," *Journal of Financial Economics* 3 (June, 1976), 257–275.
9. J. Hamilton. "Marketplace Fragmentation, Competition and the Effieciency of the Stock Market." *Journal of Finance* 34 (March, 1979), 171–187.
10. T. Ho and H. Stoll. "On Dealers Markets Under Competition," *Journal of Finance*, 35 (May, 1980), 259–267.
11. T. Ho and H. Stoll. "Optimal Dealer Pricing Under Transactions and Return Uncertainty." *Journal of Financial Economics*, 9 (March, 1981), 47–73.
12. W. Newton and R. Quandt. "An Empiricial Study of Spreads," Research Memorandum No. 30, Financial Research Center, Princeton University (January, 1979).
13. E. Mildenstein and H. Schleef. "The Optimal Pricing Policy of Monopolistic Marketmaker in the Equity Market." Working Paper, University of Oregon, Eugene, Oregon (1980).
14. S. Smidt. "Which Road to an Efficient Market: Free Competition vs. Regulated Monopoly." *Financial Analysts Journal* 27 (September-October 1971), 18ff.
15. S. Smidt. "Continuous Versus Intermittent Trading on Auction Markets." *Journal of Financial and Quantitative Analysis* 14 (November, 1979), 837–866.
16. H. Stoll. "The Supply of Dealer Services in Securities Markets," *Journal of Finance*, 33 (September, 1978), 1133–1151.
17. H. Stoll. "Pricing of Security Dealer Services: An Empirical Study of NASDAQ Stocks." *Journal of Finance* 33 (September, 1978), 1153–1172.
18. S. Tinic. "The Economics of Liqudity Services," *Quarterly Journal of Economics* 86 (February, 1972), 79–83.
19. S. Tinic and R. West. "Competition and the Pricing of Dealer Services in the Over-the-Counter Market," *Journal of Financial and Quantitative Analysis*, 7 (June, 1971), 1707–1727.
20. W. Vickrey. "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, 16 (March, 1961), 8–37.