

OPTIMAL DEALER PRICING UNDER  
TRANSACTIONS AND RETURN UNCERTAINTY

by

Thomas Ho and Hans Stoll

Working Paper No. 27-79

RODNEY L. WHITE CENTER  
FOR FINANCIAL RESEARCH

University of Pennsylvania  
The Wharton School  
Philadelphia, PA 19104

OPTIMAL DEALER PRICING UNDER  
TRANSACTIONS AND RETURN UNCERTAINTY

Thomas Ho and Hans Stoll

Introduction

Dealers in securities markets stand ready immediately to trade fixed amounts of securities at stated bid or ask prices. In this paper the optimal bid and ask prices are determined such that the dealer's expected utility of terminal wealth is maximized. Our concern is primarily with the risk assumed by the dealer, which in this case arises not only from uncertainty about the return on his inventory but also from the uncertainty about when future transactions will occur (which affects how long he must bear return uncertainty). We consider a single dealer trading a single stock and facing a downward sloping stochastic demand for the service of immediacy provided by the dealer.

This paper is related to an earlier paper by Stoll (1978a); indeed the economic setting of the problem is much the same. What is different is the nature of uncertainty (primarily the introduction of transactions uncertainty), explicit treatment of a multiperiod strategy for the dealer, and the introduction of the demand side. The paper should also be placed in the context of the growing literature on the micro-structure of securities markets. This literature is concerned with the difference between transactions prices and "true" underlying prices as determined by the available information set. One approach to this issue is to propose price or return generating processes that are consistent with our understanding of market structure but do not necessarily derive from any

principles of individual economic behavior. The end product of such research is the relationship between the observed distribution of returns and the "true" underlying distribution. The work of Clark (1973); Blattberg and Gonedes (1974); Westerfield (1977); Oldfield, Rogalski and Jarrow (1977) Schwartz and Whitcomb (1977), and Goldman and Beja (1979) is in this spirit. The latter paper is most closely related to our paper since it is concerned with the impact of the specialist on the process by which transactions prices adjust to the evolving "true" price of the stock. However, they do not build an explicit micro-economic model of the dealer.

An alternative approach is to specify the demand for and supply of the services of intermediaries -- i.e. dealers -- whose compensation is reflected in a transaction price different from the "true" price. For an extensive review of the literature, see Cohen, Maier, Schwartz and Witcomb [CMSW] (1979). On the supply side, work began with the paper by Demsetz (1968), and additional theoretical contributions were made by Tinic (1972) and Stoll (1978a). There are also a host of empirical papers on bid-ask spreads that have arisen out of this or closely related approaches. In addition to Demsetz and Tinic, see Tinic and West (1972), Benston and Hagerman (1974), Stoll (1978b), and Smidt (1979), for example. Papers by Copeland (1976); Epps (1976); CMSW (1978); and Goldman and Sosin (1979) can be viewed as the basis for the demand for dealer services because they specify why investors trade and how trading affects prices. However, the role of intermediaries is not explicitly considered.

In an excellent paper, Garman (1976) does consider the role of the intermediation process, but he is more concerned with the equilibrium price of securities rather than the equilibrium price of the services of intermediaries. We take the "true" price of the stock to be exogenously

determined by the information set and ask how the dealer prices relative to the "true" price whereas Garman is concerned with the necessary behavior of transactions prices to meet an exogenous inventory objective of the dealer. We adopt Garman's concept of stochastic supply of and demand for securities but view both the demand to sell securities and the demand to buy securities as demands for dealer services. Unlike Garman and recent papers by Amihud and Mendelson (1979) and Schleef and Miltenstein (1979), based on Garman, we do not require our dealer to maximize expected profits. Indeed our key concern is with the risk the dealer faces, and how this affects his willingness to supply dealer services. In addition we introduce stochastic processes for the "true" return on the dealer's stock and the other components of his portfolio. In the face of stochastic transactions and stochastic returns, the dealer maximizes expected utility of terminal wealth by adjusting bid and ask prices through time. We present a dynamic programming solution to this more general problem.

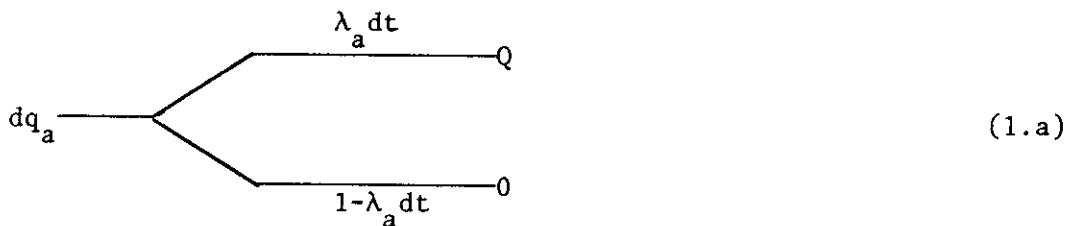
This paper is limited to the behavior of a single dealer making a market in a single stock. The generalization to many stocks is not difficult but adds needless complexity and is not considered in this paper. The determination of market bid and ask prices when there are several dealers (or, for that matter, individuals placing limit orders) is also left to another paper.

The next section of this paper discusses the assumptions underlying our model. The dynamic programming problem is then stated and a solution presented. The model is used to analyze the dealer's optimal bid-ask spread and price adjustment strategy under symmetric as well as asymmetric demand conditions. A numerical illustration is also presented. The model is then used to analyze the decision to become a dealer in a particular stock. Finally the paper is summarized.

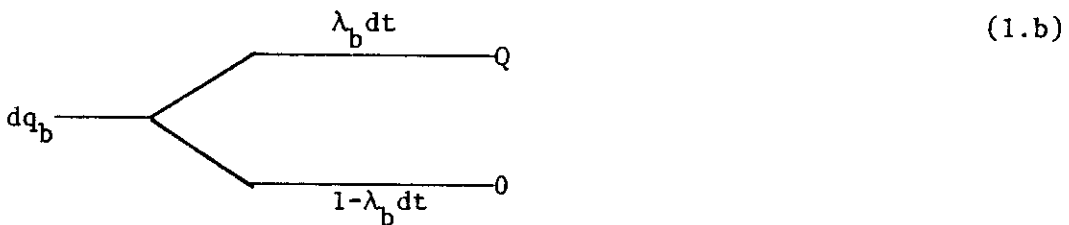
Assumptions

Transactions are assumed to evolve as a stationary continuous time stochastic jump process as in Garman (1976). A different process is allowed for purchases by the dealer and sales by the dealer: the process may be represented as follows.

Dealer sales (public purchases):



Dealer purchases (public sales):



where

$Q$  = jump size = number of shares in a transaction.

$\lambda_a, \lambda_b$  = intensities of processes. These represent the average number of public purchase or public sale transactions respectively per unit time.

The use of a Poisson jump process to model transactions is quite natural. In actual markets, dealers are obligated to trade a minimum quantity (i.e. 100 shares), and transactions occur at discrete time intervals.

The dealer determines a price of immediacy,  $b$ , should a public sale (dealer purchase) order arrive and a price,  $a$ , should a public purchase

(dealer sale) order arrive. The dealer does not directly quote  $b$ ,  $a$ . Rather he quotes bid and ask prices which are defined as,

$$p_b = p - b \quad (2.a)$$

$$p_a = p + a \quad (2.b)$$

where  $p$  is the dealer's opinion of the true price of the stock at the time he sets the bid-ask quotation. The public's opinion of the true price is denoted by  $p^*$ , and the two need not be the same, although in an informationally efficient market one would expect them to be the same. Both  $p$  and  $p^*$  are assumed to be exogenous.

The dealer is assumed to trade only with the public in a passive way. That is, he sets prices and must wait until a transaction occurs. He cannot immediately dispose of inventory by trading with a dealer of last resort. The arrival rate of dealer purchases  $\lambda_b$ , and dealer sales,  $\lambda_a$ , depends on the dealer's buying fee and selling fee respectively. Linear demand curves are shown in Figure 1. By raising  $b$ , the dealer can reduce the probability of further public sales and by lowering  $a$ , he raises the probability of public purchases. Of course public purchases and sales are also influenced by the difference between  $p$  and  $p^*$ . This difference is assumed to be parameterized in the functions  $\lambda_a$  and  $\lambda_b$ .

In addition to uncertainty about the timing of subsequent transactions, the dealer faces uncertainty about the return on his existing portfolio. In the absence of a transaction, portfolio growth,  $\tilde{dX}$ , is represented by a stochastic differential equation of the following form:

$$d\tilde{X} = r_X \tilde{X} dt + \tilde{X} dZ_X \quad (3)$$

where  $r_X$  is the mean return per unit time;  $dZ_X$  is a (non standardized) Wiener process with mean zero and instantaneous variance,  $\sigma_X^2$ .

The dealer's wealth is divided into three components:

1. 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 of interest,  $r$ . The change in the value of the cash account  $\tilde{F}$ , is

$$d\tilde{F} = r\tilde{F}dt - (p-b)dq_b + (p+a)dq_a \quad (4)$$

Uncertainty in the cash account is due to uncertainty about transactions not to any uncertainty about the interest to be earned.

2. Inventory

The dealer's inventory consists of shares of the one stock in which he makes a market. The change in the value of the inventory account,  $\tilde{I}$ , is,

$$d\tilde{I} = r_I\tilde{I}dt + pdq_b - pdq_a + \tilde{I}dZ_I \quad (5)$$

where notation is as in (3) and (4). Shares enters inventory at  $p$  whereas the associated cash flow in the cash account is  $(p-b)$  or  $(p+a)$  per share. The difference reflects the dealer's profit. Uncertainty arises both from transactions uncertainty reflected in  $dq_a$  and  $dq_b$  and from return uncertainty reflected in the Wiener process  $dZ_I$ .

Since we are concerned with dealer pricing relative to the true price, we adopt the convention that  $p$  is constant. This may be justified by assuming that a stock dividend is paid such that  $p$  is constant. This means the value of transactions is unchanged, which is desirable for our analysis. The constancy of transaction value could also be achieved directly by interpreting  $Q$  in (1) as value rather than number of shares.

Alternatively, the analysis could be carried out with stochastic  $p$  with drift, but this procedure results in needless complexities. We have made the assumption which is most simple and yet also most illuminating from the point of view of realistically describing dealer behavior.

### 3. Base wealth

The dealer has base wealth  $\tilde{Y}$ . On the day he starts as a dealer  $F_0 = 0$  and  $I_0 = 0$ , and base wealth is his only wealth. Base wealth serves as collateral to back up the borrowing of shares or money that arise in the course of acting as a dealer. The change in base wealth is

$$d\tilde{Y} = r_Y \tilde{Y} dt + \tilde{Y} dZ_Y \quad (6)$$

where notation is as in (3).

Figure 2 is a representation of the value of the dealer's inventory as it moves through time between  $t_0$  and  $T$ . The discrete jump represent transactions of  $Q$  shares at price  $p$ . There is an offsetting jump of slightly smaller size (smaller by the transaction cost,  $a$  or  $b$ ) in the cash account (not shown). The fluctuations in value between jumps are due to return uncertainty represented by the last term in (5), and this return uncertainty is the underlying source of risk facing the dealer.

The objective of the dealer is to maximize the expected utility of his total wealth,  $EU(W_T)$ , at time  $T$ , his horizon, where

$$\tilde{W}_T = \tilde{F}_T + \tilde{I}_T + \tilde{Y}_T \quad (7)$$

Numerous transactions and price changes occur between  $t_0$  and  $T$ . We seek the optimal strategy for choosing the values of  $a$  and  $b$  that maximize the dealer's preference function. This involves weighing the costs and benefits of being a dealer. The costs arise from the nonoptimal portfolio the



dealer stands ready to accept. Not only is his degree of diversification and level of risk not optimal but the presence of transactions uncertainty makes uncertain when and how much his portfolio will be unbalanced. The benefits arise from the fee he is able to charge for his service of immediate trading. Because the model assumes a single dealer in the stock, that dealer has the ability to earn abnormal profits.

The optimal strategy is complicated by the multiperiod framework which permits the dealer to adjust  $a$  and  $b$  as he moves through time -- usually in response to inventory changes. Although stochastic, these inventory changes are in turn influenced by the bid and ask prices. This is known as a closed loop control problem because the optimal values of  $a$  and  $b$  depend on the observed state variables  $(F, I, Y)$  as well as on time. In other words the optimal dealer strategy we seek is a function that specifies the choice of  $a$  and  $b$  for any position (described by  $t, F, I, Y$ ) in which the dealer finds himself. The appropriate procedure for such a problem is dynamic programming. Note that there is only one consumption point at  $T$  and that the dealer's actions before that point in time affect his opportunity set at the terminal point. Under the strategy allowed for the dealer, there is no way to turn the multiperiod effects on the opportunity set into a single period problem by assuming a particular utility function for the dealer (such as the log utility function).

#### Statement of dynamic programming problem

For any state of the world which is described by time and the values of the state variables - dealer's cash, inventory and base wealth - the dealer chooses  $a$  and  $b$  to maximize the expected utility of his terminal wealth. Define the optimal performance function  $J ( )$  as

$$J(t, F, I, Y) = \max_{a,b} [EU(W_T) | t, F, I, Y] \quad (8)$$

The function  $J(\cdot)$  is the solution to the maximization problem (8) for an optimal bid and ask strategy from  $t_0$  to  $T$  and therefore describes the result of optimal dealer behavior conditional on any observed  $t, F, I, Y$ . The function  $J$ , can be viewed as a derived utility function. Since there is no intermediate consumption prior to  $T$ , the fundamental recurrence relation implied by the principle of optimality of dynamic programming is simply that

$$\max_{a,b} dJ(\cdot) = 0, \text{ and } J(T, F, I, Y) = U(W_T) \quad (9)$$

The partial differential equation implied by (9) - Bellman's equation--is

$$\begin{aligned} \max_{a,b} \frac{dJ}{dt} = & J_t + LJ \\ & + \max_{a,b} \{ \lambda_a [J(F+pQ+aQ, I-pQ, Y) - J(F,I,Y)] \\ & + \lambda_b [J(F-pQ+bQ, I+pQ, Y) - J(F,I,Y)] \} = 0 \end{aligned} \quad (10)$$

where  $L$  is defined as the operator that takes the differential with respect to time of the mean returns and Wiener risk components of the function  $J$ .

Thus<sup>1</sup>

$$LJ = J_F r_F + J_I r_I + J_Y r_Y + \frac{1}{2} J_{II} \sigma_I^2 I^2 + \frac{1}{2} J_{YY} \sigma_Y^2 Y^2 + J_{IY} \sigma_{IY} IY \quad (11)$$

where  $\sigma_I^2$  and  $\sigma_Y^2$  are the instantaneous variance of  $dZ_I$  and of  $dZ_Y$  respectively, and  $\sigma_{IY}$  is the covariance between  $dZ_I$  and of  $dZ_Y$ , and where subscripts on a function indicate the variable with respect to which a partial derivative is taken.

The expression (11) results from the application of Ito's Lemma to the Wiener process components of the state variables, I and Y, that J depends on. As is evident in (11) the mean return and Wiener risk do not depend directly on a, b, and therefore are not of immediate concern to the dealer. Of course a, b indirectly affect Y, I and thereby influence the dealer's ultimate return and risk.

More important to the dealer and to our problem are the last two terms of (10) which represent the increments to  $J( )$  resulting from transactions the value of which are  $pQ$  and which occur with probability  $\lambda_a dt$  in the case of dealer sales or  $\lambda_b dt$  in the case of the dealer purchases. The dealer also adds his fee,  $aQ$  or  $bQ$ , to his cash account. We are interested in the maximized value of these increments which do depend on a and b. The independence of the processes,  $dq_a$  and  $dq_b$ , allows us to write the effect as the sum of the two increments. The time argument in  $J( )$  has been suppressed.

It is convenient to restate the problem in terms of time remaining to the horizon date T. Let  $\tau$  denote this amount of time, and replace  $t$  in (8) by  $\tau$ . In other words we now think of J as a derived utility function that depends on the time remaining to the horizon date. Thus when  $\tau = 0$

$$J(0, F, I, Y) = U(W), \quad (12)$$

The derived utility at the horizon data is just the utility of wealth at that point. The transformation of the time variable also implies that

$$J_t = - J_\tau \quad (13)$$

Next consider the last two lines of (10), and since  $p$  is arbitrary let  $p = 1$ . Because  $aQ$  and  $bQ$  are small relative to  $Q$ , we make the following first order approximations:

$$J(F + Q + aQ, I - Q, Y) = J(F + Q, I - Q, Y) + J_F(F + Q, I - Q, Y) aQ \quad (14.a)$$

$$J(F - Q + bQ, I + Q, Y) = J(F - Q, I + Q, Y) + J_F(F - Q, I + Q, Y) bQ \quad (14.b)$$

Note that there is no approximation with respect to the effect of a transaction on  $J$ ; the approximation is only with respect to the fee collected on a transaction. We assume the same linear demand relation for dealer sales and purchases. This assumption of symmetric demand is made for expositional clarity and will be relaxed later.

$$\lambda_a = \lambda(a) = \alpha - \beta a \quad (15.a)$$

$$\lambda_b = \lambda(b) = \alpha - \beta b \quad (15.b)$$

Now define the sell operator,  $S$ , as

$$SJ = S[J(F, I, Y)] = J(F + Q, I - Q, Y), \quad (16.a)$$

and define the buy operator,  $B$ , as

$$BJ = B[J(F, I, Y)] = J(F - Q, I + Q, Y). \quad (16.b)$$

The sell operator acting on  $J$  describes the dealer's derived utility after a sale by the dealer. Utility will decrease if the sale drives the dealer farther away from his desired portfolio by increasing a short inventory position. Utility will increase if the sale reduces a long position. A corresponding result holds on the buy side. If the dealer

starts with  $I > 0$ ,  $BJ - J < 0$ . If the dealer starts with  $I < 0$ ,  $BJ - J > 0$ .

Substituting (13), (14) and (15) into (10) and using (16) allows us to restate (10) as follows:

$$J_t = LJ + \max_{a,b} \{ [\lambda(a)aQSJ_F - \lambda(a)(J-SJ)] + [\lambda(b)bQBJ_F - \lambda(b)(J-BJ)] \} \quad (17)$$

This equation says that the change with respect to time remaining of derived utility depends on the return and risk of the dealer's current wealth (LJ), over which he has no direct control, and the maximized value of two terms that reflect the net contribution to the dealer's derived utility from dealer sales and purchases respectively, over which he has control. The essence of the dealer's problem is contained in these two terms.

Consider the last term of (17), which represents the net benefit from a purchase by the dealer.<sup>2</sup> The expected gross benefit (in utile dimensions) from a dealer purchase is

$$\lambda(b)bQBJ_F$$

where  $\lambda(b)bQ$  is the expected revenue from purchase transactions; and  $BJ_F$  is the marginal utility of cash after a purchase transaction, i.e.  $BJ_F$  converts expected revenues into "utiles". Since  $Q$  and  $BJ_F$  are given and since  $\lambda(b)$  has been assumed to be linear, the expected gross benefit is a quadratic function of  $b$  and is drawn in Figure 3. The curve passes through the origin, is at a maximum when  $b = \frac{\alpha}{2\beta}$  [from (15)], and is zero again at  $b = \frac{\alpha}{\beta}$ . Because of the assumed symmetry in demand conditions and on the assumption that  $BJ_F = SJ_F$  the curve is identical for dealer sale transactions.

What is typically different between an incremented purchase and an incremental sale is the gross "cost" to the dealer (in utile dimensions). Unless the dealer begins with zero inventory, he prefers one type of transaction over the other. Suppose the dealer begins with positive inventory which makes him reluctant to purchase additional shares. The expected decrease in derived utility from a purchase,

$$\lambda(b)(J-BJ),$$

is positive because  $(J-BJ) > 0$ . Since  $\lambda(b)$  is linear and downward sloping this "cost" function is also linear and downward sloping as shown in Figure 3. The larger  $b$  the lower is the probability of a dealer purchase and therefore the lower is the expected "cost" to the dealer of a purchase. The optimum buying fee,  $b^*$ , is set to maximize the difference between the "benefit" and "cost" curve. Performing this maximization yields

$$b^* = \frac{\alpha}{2\beta} + \frac{J-BJ}{2BJ_F Q} \quad (18.a)$$

In other words set  $b$  at the level which maximizes expected profits associated with the stochastic arrival of public sellers plus a risk premium term. The risk premium term ultimately depends on the dealer's fundamental attitude toward risk and the level of return and transactions uncertainty of the stock in question all of which enter into  $J$  and  $BJ$ . On the other hand, when inventory is positive, utility is increased by a sale transaction, i.e.  $\lambda(a)(J-SJ) < 0$  because  $(J-SJ) < 0$ . The new position after a sale,  $SJ$ , would be superior to the old position,  $J$ . Thus the "cost" function for a sale transaction is negative and upward sloping as shown in Figure 3. The larger is  $a$  the lower is the probability of a dealer

sale and therefore the less negative is the expected cost (the smaller is the expected increase in utility). The optimum selling fee,  $a^*$ , is set to maximize the difference between the "benefit" curve and the "cost" curve for sales. Analogously to the case of a purchase this solution is

$$a^* = \frac{\alpha}{2\beta} + \frac{J-SJ}{2S_F JQ} \quad (18.b)$$

As one would expect,  $b^* > a^*$ . The two "cost" functions coincide when dealer inventory is zero, in which case  $a = b$ . As inventory changes, the "cost" functions shift in opposite directions and  $a$  and  $b$  change in opposite directions. For given  $p$ , this implies that both the bid price ( $p_b$ ) and ask price ( $p_a$ ) fall in response to inventory increases and both rise in response to inventory decreases, as shown in Figure 4. We shall see later that the spread, the difference between  $p_a$  and  $p_b$ , is to an important degree invariant with respect to inventory changes.

The heuristic discussion of Figure 3 does not however provide us with an explicit solution for the dealer's problem, which is our objective. In order to understand where an individual dealer sets his bid and ask price, the solution to  $J$  is required. Given a solution for  $J$  we can see explicitly the determinants of  $a$  and  $b$  and of the dealer's spread.

We now proceed to a final statement of the dynamic programming problem by performing the maximization required in (17). Substitution of (18) into (17) and using (15) yields

$$J_T = LJ + \frac{[\alpha S J_F Q - \beta(J-SJ)]^2}{4\beta S J_F Q} + \frac{[\alpha B J_F Q - \beta(J-BJ)]^2}{4\beta B J_F Q} \quad (19)$$

The problem is to solve (18) subject to (18) and the terminal condition, (12). The solution for J can then be substituted into (18) to yield the overall optimal values of a and b as functions of the state variables, F, I, Y.

### Solution

Unfortunately to our knowledge there is no closed form solution for equations of the form (19). However by using Taylor's series expansion we can, in principle, come arbitrarily close to the solution. Indeed this procedure has certain advantages since it allows one to see the nature of the solution as a function of time to the horizon date. Thus write  $J(\tau)$  as a Taylor's expansion around  $\tau = 0$ .

$$J(\tau, F, I, Y) = \phi^0(F, I, Y) + \phi^1(F, I, Y) \tau + \phi^2(F, I, Y) \tau^2 \dots \quad (20)$$

where higher order terms are assumed to be small enough to be neglected. It will be evident from the methodology that follows that higher order terms could be incorporated. However, for the purpose of this paper the increased complexity that results is not offset by a corresponding increase in understanding of the dealer. The interpretation of (20) is that  $J(\tau)$  depends on how much time remains in the dealer's planning horizon. When  $\tau = 0$ , only the first term of (20) is relevant. When  $\tau > 0$  but sufficiently small, only the second term is relevant, and so on.<sup>3</sup>

The solution to (20) is found by finding the different functions  $\phi^0(\cdot)$ ,  $\phi^1(\cdot)$  and  $\phi^2(\cdot)$ . The superscripts on the  $\phi$  are notations



associating each function with the degree of the derivative of  $J$  with respect to  $\tau$ . Equation (20) holds for all values of  $\tau$ . For  $\tau = 0$ ,  $J$  is given by (11). Therefore

$$\phi^0(F, I, Y) = U(W), \quad (21)$$

the utility of terminal wealth. In other words if the dealer is at maturity the optimal performance function is just his utility function.

For  $\tau > 0$ , (19) must be satisfied. Substituting for  $J_\tau$ ,  $J$  and  $J_F$  in (19) from (20) and evaluating the resulting expression at  $\tau = 0$  gives a solution for  $\phi^0$  ( ). Since  $\phi^0$  ( ) is given by (21) a final solution for  $\phi^1$  ( ) in terms of the state variables, risk parameters and the dealer's attitude toward risk is achieved:

$$\phi^1 = LU(W) + 2U'\Pi \quad (22)$$

The operator,  $L$ , has been defined in (11). Its application in this case leads to

$$LU(W) = U' r_W W + U'' \left( \frac{1}{2} \sigma_I^2 I^2 + \sigma_{IY} IY + \frac{1}{2} \sigma_Y^2 Y^2 \right)$$

where

$$r_W = r \frac{F}{W} + r_I \frac{I}{W} + r_Y \frac{Y}{W}$$

and  $U' > 0$  and  $U'' < 0$  are the first and second derivatives of  $U(W)$ . In other words  $LU(W) = L\phi^0$  represents the positive contribution of mean return to utility when there is only a short instance of time remaining plus the

negative contribution of return variance. The term,  $\Pi$  is defined as the maximum expected profits from either sales or purchases (which are the same under symmetric demand assumed here), i.e.<sup>4</sup>

$$\Pi = \frac{\alpha^2 Q}{4\beta} .$$

The term  $2U'\Pi$  thus represents the contribution to dealer welfare of the profits he expects to make from purchase plus sale transactions. Unlike return uncertainty the uncertainty of transactions is reflected only through  $\phi^2$ .

The function  $\phi^2$  is found in a similar way. Substitute for  $J_\tau$ ,  $J$  and  $J_F$  in (19) from (20), take the derivative with respect to time, and evaluate at  $\tau = 0$ . The solution depends on  $\phi^1$ , and involves derivatives and moments of third and fourth order. To simplify the problem, we now assume the dealer's elementary utility function is quadratic so that we may restrict ourselves to a mean-variance world.<sup>5</sup> The solution for  $\phi^2$  is then

$$\phi^2 = \frac{1}{2} L^2 U(W) + U' r \Pi + 2U'' \Pi^2 + 2U'' \Pi r_W W + \frac{1}{2} U'' \frac{\alpha}{2} \sigma_I^2 Q^2 \quad (23)$$

This expression is a negative adjustment to the dealer's derived utility for aspects of return uncertainty and the effect of transactions on portfolio return uncertainty.

The first term of (23) reflects the compound effect of return and return risk resulting from twice applying the operator  $L$  to  $U(W)$ . The term is written out in the appendix and inspection will indicate that it is normally negative. The second term is the increase in utility resulting from the interest earnings on the transaction profit. Because  $U'' < 0$  the remaining terms are always negative. The third term,  $2U''\Pi^2$ , is an

adjustment to the profit term,  $2U'\Pi$ , that appears in  $\phi^1$ . If the dealer restricts himself to a very short horizon, transactions appear very profitable because he sees only the fee he collects immediately. As his horizon is extended, derived utility is adjusted downward to reflect the uncertainty of maintaining that profitability. The fourth term is a similar adjustment for the compound effect of  $\Pi$  and total portfolio return  $r_W W$ .

The last term is an adjustment for transactions uncertainty. Transactions uncertainty is not undesirable per se. It is undesirable only because it affects return uncertainty. Therefore  $\sigma_I^2$  is part of this term. The expression,  $\frac{\alpha}{2}$ , is the expected arrival rate of purchases or sales at the dealer's  $a, b$  which maximizes expected profits.<sup>4</sup> When viewed in the context of a small time interval,  $\frac{\alpha}{2}$  is the probability of a purchase or sale. (i.e.  $\frac{\alpha}{2} dt$  is the probability). Thus for a stock in which the probability of a transaction is large, the probability of incurring return risk (the value of which is  $\sigma_I^2 Q^2$ ) is also large. Since  $U'' < 0$ , this term has a negative effect on  $J$ .

#### Dealer's pricing strategy

Having  $\phi^0$ ,  $\phi^1$  and  $\phi^2$ , we have the solution for  $J$  as given by (20); and we can specify precisely how the dealer's bid and ask prices depend on the return characteristics of the stock, the stochastic demand for dealer services and the dealer's elementary utility function. We describe the dealer's pricing strategy by specifying the dealer's spread,  $s$ , and a price adjustment variable,  $d$ . The spread is defined as  $s = a + b$ . (Since  $p = 1$ , this is the proportional spread). We show first that the spread is largely independent of the dealer's inventory position and depends on fundamental

characteristics of the stock and the dealer. Second, the dealer encourages public customers to trade in the direction he desires by changing the adjustment variable which is defined as  $d = a - b$ , not by changing the spread. With the help of Figure 4 one can see that  $d = 0$  whenever the mid-point of the spread is at  $p$ ;  $d < 0$  whenever the mid-point is less than  $p$  (and the dealer is relatively more anxious to sell than to buy);  $d > 0$  whenever the mid-point is greater than  $p$  (and the dealer is relatively more anxious to buy than to sell).<sup>6</sup> Given  $s$  and  $d$ , the selling fee and buying fee can always be derived according to

$$a = \frac{s+d}{2}, \quad b = \frac{s-d}{2}$$

#### The bid-ask spread.

From (18.a) and (18.b) the dealer's spread is

$$s = \frac{\alpha}{\beta} + \frac{J-SJ}{2S_{FQ}} + \frac{J-BJ}{2B_{FQ}} \quad (24)$$

The first term in (24) is the spread which maximizes expected revenues per share from sale and purchase transactions from the symmetric linear demand functions. The more inelastic the demand for dealer services the larger the spread.<sup>7</sup> The remaining terms are "risk premiums" for a sale and purchase transaction, which assume optimal dealer behavior in the future. The sum of the risk premiums is required because the dealer sets the spread without knowing whether the next transaction will be a purchase or sale.

Substitution of the solution for  $J$  into (24) yields the following optimal dealer spread:

$$s = \frac{\alpha}{\beta} - \frac{1}{2} \frac{U''}{U'} \sigma_I^2 Q \tau - \frac{1}{2} \frac{U''}{U'} \sigma_I^2 Q [(r_I - r + G_I) - \frac{U'''}{U'} W (r_W + \frac{2\pi}{W})] \tau^2 \quad (25)$$

where  $G_I = r_I + \frac{1}{2} \sigma_I^2$ , which is the instantaneous growth in the variance of  $I$ . The spread equation actually omits certain small second order terms. The full equation is given in the appendix.

We can now see what factors determine the spread and how the spread is changed as the horizon of the dealer is lengthened. When  $\tau = 0$ , the dealer does not look beyond the current moment and is concerned only with the fee he can collect from a purchase or sale. Thus only the risk neutral spread  $\frac{\alpha}{\beta}$ , already discussed above, is relevant.

At the time horizon increases, the second term of (25), which is the first order risk adjustment, is relevant; and the spread is higher (since  $U'' < 0$  and all other terms are positive). This term is essentially the same as the term derived in Stoll (1978a) where a one period framework was relied upon. As was the case there, the risk adjustment term does not depend on the dealer's inventory. Rather it depends on factors which are relatively stationary through time i.e.

1. The dealer's attitude toward risk as reflected in the coefficient of absolute risk aversion,  $U''/U'$ .
2. The value of the transaction for which the bid-ask quotation is made,  $Q$ . (Recall  $p = 1$ ). Larger transactions require larger spreads.
3. The risk of the stock as measured by the instantaneous variance,  $\sigma_I^2$ . Note that as far as the spread is concerned "total risk" not just "systematic risk" matters. This is because the dealer is not diversified with respect to the transaction,  $Q$ .

In fact, inventory position does not appear in (25) (except as it affects  $r_w$ ). The reason is that the dealer adjusts the bid and ask prices equally whenever there is an inventory change. Thus when he acquires shares, he lowers his bid price (raises  $b$ ) to discourage additional sales to him; and he lowers his ask price (lowers  $a$ ) to encourage additional purchases from him. In other words he places himself on the margin where he is indifferent to a public sale or purchase because the higher compensation derived from a sale to him is equal in utility to the lower risk (cum smaller compensation) from a purchase from him.

Lengthening the dealer's time horizon makes the last term of (25) relevant and causes the spread to be still larger since the term is positive. For a given time horizon, one might expect the spread to be less under our multiperiod model than under the one period model of Stoll (1978a) since the dealer has an opportunity to adjust prices as his inventory changes. However, this is not the case. In a one period model, one transaction occurs; and the dealer holds the risky position to the horizon. In our multiperiod model, additional transactions as well as dealer price adjustments may occur. The uncertainty of these additional transactions (which increases the uncertainty of the dealer's portfolio return and profits) more than offsets the risk reduction the dealer chooses to achieve by changing his prices. The single period approximation simply ignores these additional risks.

As in the case of the first order risk adjustment, the second order risk adjustment depends on the dealer's attitude toward risk, the transaction value, and the stock's return variance, which are multiplied by the bracketed term. The first term inside the brackets is a risk premium plus

the growth in the variance of  $I$ . The second term is the coefficient of relative risk aversion ( $\frac{U''}{U'}$   $W$ ) multiplied by the return from two sources - the return on his total portfolio ( $r_W$ ) and dealer profits expressed as a fraction of his wealth ( $\frac{2\Pi}{W}$ ). Since the expected returns and growth in variance of  $I$  are likely to be small relative to the dealer's horizon, the principal effect comes from  $\frac{2\Pi}{W}$ : the larger the monopoly profits today the more the dealer stands to lose tomorrow, and therefore the larger the second order risk adjustment.

Probably more surprising than the terms which enter (25) are the terms that do not enter. We had expected the transactions processes to play a more explicit role. They do not enter at all in the first order risk term. They enter the second order risk term through  $\Pi$ , but this is simply an adjustment for the expected growth in dealer wealth from his monopoly profits. There is no term for the variance of the transactions process. The absence of such a term in the spread equations is due to two factors. First, in this model transactions uncertainty is not undesirable per se because there is no direct adverse effect of a below average number of buyers or sellers. We have not assumed a minimum cost of doing business, for example. Transactions uncertainty is undesirable only because it increases the uncertainty of the return on the dealer's portfolio by making uncertain how large an unbalanced position must be held. The effect is seen in (23) in the term  $\frac{\alpha}{2} \sigma_I^2 Q^2$ . Second, the reason that this term does not appear in (25) is analogous to the reason that inventory does not appear. For any  $I$ , the dealer sets his bid and ask price so that he is indifferent on the margin as to whether the next transaction is a purchase or sale. This is true whatever is the probability of the next purchase or

sale even if these probabilities were to differ because of asymmetric demand. As a result the term,  $\frac{\sigma}{2} \sigma_I^2 Q^2$ , does not enter the spread. Just as current inventory does not matter for the spread, the expected change in inventory does not affect the spread.<sup>8</sup>

### Price adjustment

On the assumption of symmetric demand, the price adjustment variable derived from (18.a) and (18.b) is

$$d = \frac{1}{2} \frac{J-SJ}{S J_F Q} - \frac{1}{2} \frac{J-BJ}{B J_F Q} \quad (26)$$

Because of symmetric demand,  $d$  is not needed to regulate an expected imbalance in orders. The principal function of  $d$  is to promote adjustment of the dealer's inventory when it is temporarily thrown out of balance by the random arrival of transactions. Substituting the solution for  $J$  into (26) yields

$$d = \frac{U''}{U'} \sigma_I^2 I \tau + \frac{U''}{U'} \sigma_I^2 I [r_I - r + G_I - \frac{U''W}{U'} (r_W + \frac{2\pi}{W})] \tau^2 \quad (27)$$

where certain small terms have again been neglected. The full equation is in the appendix.

Equation (27) is identical to (25) except for the fact that  $2I$  replaces  $Q$ . Whereas the spread depends on the size of the transaction yet to come ( $Q$ ), the price adjustment depends on the inventory which has been acquired ( $I$ ). Equation (27) is an inventory response equation that specifies the price adjustments variable be negative (positive) when inventory is positive (negative). When  $d < 0$ , both the bid price and ask



price are "low" so as to encourage dealer sales and discourage dealer purchases, thereby reducing the dealer's inventory. When  $I < 0$ ,  $d > 0$  to encourage dealer purchases and discourage dealer sales. The degree of price response to an inventory change depends on the same factors determining the size of the spread -- dealer's risk aversion, variance of the stock; and, in the second order term, future increases in wealth which will be at risk.

#### Asymmetric demand

Asymmetric demand arises if the dealer's perception of the true price,  $p$ , differs from the public's perception,  $p^*$ . For example if  $p^* > p$ , the expected number of dealer sales exceeds the expected number of dealer purchases as in Figure 1. This is represented by different  $\alpha$  in the stochastic demand equations. Suppose the demand curves for dealer sales and purchases differ both in their slope and intercept:

$$\lambda_a = \alpha_a - \beta_a a \quad (15.c)$$

$$\lambda_b = \alpha_b - \beta_b b \quad (15.d)$$

Asymmetric demand alters the spread equation by changing the first term in (24) or (25) from  $\frac{\alpha}{\beta}$  to  $(\frac{\alpha_a}{2\beta_a} + \frac{\alpha_b}{2\beta_b})$ . However, if this term - the risk neutral spread - and expected dealer profits remain the same under symmetric as under asymmetric demand, there is no additional effect on spread through the risk premium terms in (25).<sup>9</sup> For example suppose that  $\alpha_a = \alpha_b = 100$  initially, that  $\alpha_a$  rises to 120 and  $\alpha_b$  falls to 50, and that the appropriate changes occur in  $\beta_a$  and  $\beta_b$  to hold constant expected profits and the risk neutral spread. Then there is no additional effect

through the risk terms despite the fact that the expected number of public sellers now exceeds the expected number of public buyers ( $\frac{\alpha_a}{2} = 60$  vs  $\frac{\alpha_b}{2} = 25$ ). This is because the dealer knows today that he can follow a price adjustment strategy in the future that makes him indifferent to this asymmetry.

What is crucially affected by asymmetric demand is the price adjustment variable,  $d$ . Just as the dealer adjusts prices if current inventory is non-zero so he adjusts prices for expected imbalances in demand which could drive future inventory to undesirable levels and thereby increase the dealer's risk to desirable levels. Let  $d^1$  denote the price adjustment under asymmetric demand. Following the procedure used in deriving (27), except that we now assume (15.c) and (15.d), gives

$$d^1 = \left[ \frac{\alpha_a}{2\beta_a} - \frac{\alpha_b}{2\beta_b} \right] - \frac{1}{2} \frac{U''}{U'} \sigma_I^2 Q \frac{\alpha_a - \alpha_b}{2} \tau^2 + d, \quad (28)$$

where  $d$  is given by (27).

The first term is simply the price adjustment necessary to maximize expected profits. This may require inventory accumulation because the stock is a good buy ( $p > p^*$ ) or inventory reductions because the stock is a poor buy ( $p < p^*$ ).<sup>10</sup>

However an expected imbalance in orders significantly raises the dealer's portfolio risk, and this requires a change in  $d$  by the amount of the second term of (28). The expected dollar inventory change is  $(\frac{\alpha_a - \alpha_b}{2})Q$ , and the inventory risk for which the dealer must adjust his price is therefore  $(\frac{\alpha_a - \alpha_b}{2}) \sigma_I^2 Q$ . If the probability of a sale exceeds that of a purchase ( $\alpha_a > \alpha_b$ ), the dealer raises  $d^1$  to discourage public sellers and encourage public buyers; and the converse is true for if the probability of purchase exceeds that of a sale. As in the case of the adjustment to

current inventory, the magnitude of the change in  $d^1$  depends on the instantaneous variance of the stock and the dealer's attitude toward risk.

These price adjustments do not eliminate inventory accumulation or reduction because the source of this inventory change is the favorable or unfavorable price of the stock in the eyes of the dealer. The price adjustment compensates the dealer for the loss of portfolio diversification and change in portfolio risk resulting from the additions to or reductions from inventory. If the dealer always sets  $p = p^*$ , when  $p^*$  is defined as the price at which the expected demand for dealer sales is equal to the expected demand for dealer purchases,  $\alpha_a = \alpha_b$ ; and the case of asymmetric demand would not arise.<sup>11</sup>

#### Numerical illustration

Consider a dealer with a planning horizon of  $\tau = .04$  of a year, approximately two weeks. For simplicity assume he faces symmetric demand. What will be his spread and price adjustment for the following assumed values of the exogenous variables?

$\alpha = 2000$  per year. Expected number of purchase transactions plus expected number of sale transactions at optimal dealer fee assuming symmetric demand.<sup>12</sup>

$\beta = 100,000$  per year. Determined to give risk neutral spread of  $\alpha/\beta = .02$ .

$$Z = - \frac{U'' \cdot W}{U'} = 2. \text{ Coefficient of relative risk aversion,}$$

assumed constant.

$$\sigma_I^2 = .5 \text{ per year.}^{13}$$

$$r_I = .20 \text{ per year; } r_y = .22 \text{ per year; } r = .10 \text{ per year.}$$

$$\sigma_y^2 = .06 \text{ per year; } \sigma_{Iy} = .05 \text{ per year.}$$

We also assume the dealer begins anew from alternative initial positions on which he has not earned profits so that  $F = -I$  and  $W = Y$ .

On the basis of these assumed values (25) and (27) now become

$$s^1 = \frac{1}{2} \frac{Q}{W} \tau + \frac{1}{2} \frac{Q}{W} \left[ .99 + \frac{.2I}{W} + \frac{40Q}{W} \right] \tau^2 \quad (25.a)$$

$$d = -\frac{I}{W} \tau - \frac{I}{W} \left[ .99 + \frac{.2I}{W} + \frac{40Q}{W} \right] \tau^2 \quad (27.a)$$

where  $s^1 = s - \frac{\alpha}{\beta}$ . The risk neutral spread,  $\frac{\alpha}{\beta} = .02$  in this illustration, is invariant to the variables of interest to us, and we look only at the risk premium terms. It is obvious from the equation that, under the assumption of constant relative risk aversion, only the ratios,  $Q/W$  and  $I/W$  are important. Values of  $s^1$  and  $d$  for alternative values of these ratios are given in Table 1. In the equation, the spread depends on  $I$  because the level of  $I$  affects portfolio return,  $r_w$ , in (25). However, as the table shows, this affect is less than one basis point and our earlier assertion that for practical purposes the spread is independent of inventory position is correct.

It is evident from the Table that for a wealthy individual the risks of making a market in a single stock can be small. For example for  $Q/W = .05$  the risk premium in the spread is only .122 percent (for values of  $I/W$  between  $-.20$  and  $+.20$ ) or \$.6.60 on a \$5000 transaction. This implies a proportional buying or selling fee net of monopoly profits of only .056 percent or \$3.30. We expect that for certain classes of dealers, such as certain NYSE specialists, block positioning firms, or underwriters, the ratio,  $Q/W$ , may be considerably higher. And in such cases the risk

premium term in the spread can increase significantly as the Table shows. An obvious recommendation for improving the quality of securities markets is to increase the amount of wealth devoted to market making either by raising wealth of the dealer or by having more dealers. However one cannot compel an increase in a dealer's wealth, and institutional restrictions (as on the NYSE) as well as clerical and informational economies of scale may limit the number of dealers in a stock.

The values in the Table depend on the time horizon of  $\tau = .04$ . Doubling  $\tau$  to .08 of a year would raise the spread (for  $Q/W = .05$ ) to .248 percent or \$12.40 on a \$5000 transaction. The spread more than doubles because the second order term becomes relatively more important. (Nevertheless its contribution to  $s^1$  is only .048 percent, which implies that the degree of approximation achieved by equation (20) is still adequate). While the variance of inventory return,  $\sigma_I^2 \tau$ , increases as  $\tau$  and the variance of inventory due to transactions,  $(\lambda_a + \lambda_b) \tau$ , increases as  $\tau$ , the combined effect on portfolio risk, which appears in the second order term increases faster than  $\tau$ .

The Table is helpful in understanding the behavior of bid and ask prices implied by our model. Let a prime on a variable indicate the risk premium component of the variable. Now consider a particular transaction size, say  $Q/W = .05$ , where  $s^1 = .00112$ . At  $I = 0$ ,  $d = 0$ ; and  $p_a^1 = p + a^1 = p + \frac{s^1 + d}{2} = 1.00056$ ; and  $p_b^1 = p - b^1 = p - \frac{s^1 - d}{2} = .99944$ . Suppose the dealer acquires \$5000 of inventory so that  $I/W = .05$ . He would lower his prices by setting  $d = -.00224$  so that  $p_a = p + \frac{s^1 + d}{2} = .99944$  and  $p_b = p - \frac{s^1 - d}{2} = .99832$ . The dealer has lowered bid and ask prices by the value of the spread, which causes his selling price,  $p_a$ , to be below his opinion of the stock's true price by  $a^1 = b^1$ . In other words to

encourage investors to buy the stock from him he is willing to pay an amount,  $a^1 = b^1$ , equal to the risk premium he demands to continue holding the stock. Conversely his buying price,  $p_b$ , is lowered so that a public seller would in effect pay not only for the marginal effect of his transaction on the dealer's risk but also for the risk of the inventory position of the dealer.

If, by chance, no additional transaction occurs before the horizon date, the model assumes the dealer liquidates his inventory at  $p$ , thereby yielding an expected fee of  $b$ . If additional transactions occur that cause ending inventory to be zero, the dealer earns no risk premium because in effect he has borne no risk. Of course, prior to the horizon date, there is uncertainty about his ending inventory which increases the dealer's portfolio risk, and this causes him to demand a risk premium today.

Because the dealer has monopoly power it is unlikely that the ask price will in fact fall below  $p$ . Noting that  $s = s^1 + \frac{\alpha}{\beta}$ , and  $\frac{\alpha}{\beta} = .02$ , the actual ask and bid price corresponding to the previous example when  $I/W = .05$  are  $p_a = p + \frac{s+d}{2} = 1.00944$  and  $p_b = p - \frac{s-d}{2} = .98832$ .

#### Demand conditions and the decision to become a dealer

Our model is perhaps most applicable to the NYSE where each stock has only one dealer - the specialist. Consider the question of whether a newly listed stock would always find a specialist willing to make a market in the stock. That decision depends on the demand conditions and on the risk of being a dealer. The demand conditions in turn depend on the fundamental desires of investors to trade the stock and on the alternative sources of immediacy. There may be competing markets which reduce  $\alpha$  and increase  $\beta$  to the specialist. Furthermore the demand for dealer services depends on the

ease with which investors could "make their own immediacy" by using limit orders and borrowing or lending. For example instead of selling immediately to the specialist at the specialist's  $p_b$ , an investor could enter a higher limit price and borrow funds while he waits for his transaction to be executed. The cost to investors of these competing sources of immediacy influence the elasticity of demand for dealer services.

In terms of our model the formal answer to the question is that the dealer will compare his derived utility function,  $J(\tau, F, I, Y)$ , when not a dealer with that function were he to become a dealer. At the time of the decision  $F = I = 0$ . From the solution to (20), derived utility is

$$J(\tau, 0, 0, Y) = U(Y) + (LU)\tau + \left(\frac{L^2U}{2}\right) \tau^2 \quad (29)$$

if the individual does not become a dealer and is

$$J^*(\tau, 0, 0, Y) = U(Y) + \phi^1(0, 0, Y)\tau + \phi^2(0, 0, Y)\tau^2 \quad (30)$$

if he becomes a dealer. To become a dealer it must therefore be the case that

$$J^*(\tau) - J(\tau) = 2\Pi U'(Y)\tau \quad (31)$$

$$+ [\Pi(2U''(Y)r_Y Y + rU''(Y)) + 2\Pi^2 U''(Y) + \frac{\alpha}{4} \sigma_I^2 Q^2 U''(Y)] \tau^2 > 0$$

We are assuming symmetric demand.

Using  $\Pi = \frac{\alpha^2 Q}{4\beta}$  to substitute into (31) and manipulating yields the condition,

$$\frac{\alpha^2}{2\beta} \frac{Q}{Y} \tau + \frac{\beta}{\alpha} \frac{Q}{Y} \sigma_I^2 \tau < \frac{2}{Z} + \left(\frac{r}{Z} - 2r_Y\right)\tau \quad (32)$$

where  $z = \frac{-U''(Y)Y}{U'}$ . This condition can always be satisfied for transaction size,  $Q$ , small enough. But suppose there is some minimum  $Q$  the dealer must be ready to trade. Then the condition can be violated if  $\alpha$ , the transactions arrival rate when the dealer's fee is zero, is sufficiently small. For example  $\frac{Q}{Y} = .1$ ,  $z = 2$ ,  $\sigma_I^2 = .5$  per year,  $r = .10$  per year,  $r_y = .20$  per year,  $\beta = 100,000$ , and  $\tau = .04$ . Then the dealer would not make a market if there were fewer than 200 transactions per year, for in that case expected profits would not offset the risk. It is also clear that regulations to increase the transactions size for which a bid-ask quotation must be honored would reduce the willingness of individuals to become a dealer. Since the NYSE specialist also earns brokerage on limit orders he executes, one may see an individual declare his willingness to be a specialist without at the same time being willing to act as a dealer. In such circumstances the provision of immediacy by the specialist is likely to be particularly poor.

### Summary

We have considered the optimal behavior of a single dealer in a single stock who is faced with a stochastic demand for his services (modeled by a continuous time Poisson jump process) and faces return risk on his stock and on the rest of his portfolio (modeled by diffusion processes). As time unfolds and transactions occur, the dealer is able to set his bid price and ask price relative to his opinion of the "true" price of the stock so as to maximize the expected utility of terminal wealth. This stochastic dynamic programming problem is solved for the dealer's derived utility function; and using this solution, we are able to specify how the dealer sets the bid and ask price at any moment of time as a function of the state in which he



finds himself. Indeed the solution could be programmed on the computer to specify how the dealer should set his bid and ask price in response to any situation in which he finds himself.

We show that the spread, the difference between the ask price and bid price, is given by a risk neutral spread that maximizes expected profits for the given stochastic demand functions plus a risk premium that depends on transaction size, the return variance of the stock and the dealer's attitude toward risk. Consistent with the one period model of Stoll (1978a), the spread does not depend on the dealer's inventory position. However dealer price adjustment does depend on inventory. When inventory increases both bid price and ask price decline, and the converse is true when inventory decreases.

Despite the dealer's ability to adjust prices in response to inventory changes as time unfolds, the dealer's risk is greater than implied by the one period model of Stoll (1978a). This results from the fact that the uncertainty of the demand to trade with the dealer is not totally eliminated or offset by the dealer's pricing strategy.

Other issues investigated are the case of asymmetric demand for trading with the dealer and the conditions under which an individual chooses or refuses to make a market in a stock. In an inactive stock, when the dealer is required to trade a minimum amount, the expected profit from trading may not be enough to offset the risk.

Although couched in terms of a dealer, our model can apply to any individual who is determining his reservation price for a stock as a function of his opinion of the stock's value and the degree to which the stock will cause his portfolio to be unbalanced. Thus the placement of limit orders by individuals could, for example, be modeled by our approach.

## FOOTNOTES

<sup>1</sup>A source on the dynamic programming solution to problem in stochastic control is Davis (1977). Davis also uses the operator L in the sense we do.

<sup>2</sup>Because purchases and sales are independent, we can separately maximize the last term.

<sup>3</sup>As a practical matter the omitted higher order terms are not likely to be large for the short planning horizons over which most dealers operate.

<sup>4</sup>Expected profits per sale transaction of size Q are

$$\lambda aQ = \alpha aQ - \beta a^2 Q$$

The value of a which maximize this expression is  $a = \frac{\alpha}{2\beta}$ . Substituting this in the expected profit expression yields maximum expected profits equal to  $\frac{\alpha^2 Q}{4\beta}$  a value later denoted by  $\Pi_a$ . Finally substituting the value,  $\frac{\alpha}{2\beta}$ , in the demand relation yields the optimal arrival rate at the price which maximizes expected profits, i.e.  $\frac{\alpha}{2}$ . Corresponding terms arise in the case of dealer purchases.

<sup>5</sup>Since we are examining a short term interval, the change in dealer wealth is small so that he will not be drawn into the unreasonable range of the quadratic utility function.

<sup>6</sup>Smidt (1979) has called the mid point of the dealer's spread (p + d), the market price, and he has called our p the equilibrium price. He views the dealer as adjusting his inventory by altering the difference between "market price" and the "equilibrium price", i.e., by altering our d.

<sup>7</sup>The papers of Garman (1976) and Amihud and Mendelson (1979) are concerned with this term or the related expression for profits.

<sup>8</sup>One might ask whether transactions uncertainty would be more critical if one introduced serial dependence in the process for purchases and sales. Positive serial dependence in purchases and in sales would be like increasing the transaction size Q, and would therefore increase the spread. Negative serial dependence would reduce the spread. Whether any additional effects would enter is not at all obvious, and we have not worked out this case.

<sup>9</sup>The term  $2\Pi$  is now defined as  $\frac{\alpha_a^2}{4\beta_a} + \frac{\alpha_b^2}{4\beta_b}$ . See footnote 4. We

consider only changes for which this term remains constant because we are not analyzing changes in the monopoly power of the dealer.

<sup>10</sup>This term may be zero even in the case of asymmetric demand if the demand curves are displaced so that profits continue to be maximized at the same prices. This is the case for a rotation that maintains the elasticity of demand; For example  $\alpha_a = 2\alpha_b$  and  $\beta_a = 2\beta_b$ .

<sup>11</sup>The problem then is how to determine expected demand. One possibility is to base this expectation on recent transactions thereby causing  $p = p^*$  to be based on the evolution of transactions.

<sup>12</sup>Under optimal pricing and symmetric demand the expected arrival of purchases on sales is  $\alpha/2$ . The number chosen is based on Smidt (1979) who found six of twelve NYSE stocks had fewer than 6000 transactions per year and on the assumption that the specialist participated as a dealer in one third of those transactions.

<sup>13</sup>The annual return variance is on the high side in part to capture the special risk of the dealer from informational trading, which is not explicitly modeled here. By standing ready to trade at bid or ask price, the dealer incurs losses when investors have information. Since the dealer's position tends to increase just before a price decline and decrease just before a price increase, his risk is greater than implied by the underlying variance of the stock's true return.

## REFERENCES

- Amihud, Y., and Mendelson, H. 1979. "Dealership Market: Market Making with Inventory". Working paper no.608/79. (May 1979)
- Benston, G., and Hagerman, R. 1974. "Determinants of Bid-Asked Spread in the Over-the-Counter Market". Journal of Financial Economics, 1 (March)
- Blattberg, R., and Gonedes, N., 1974. "A Comparison of the Stable and Student Distribution as Statistical Models for Stock Prices," Journal of Business, 47 (April)
- Clark, Peter, 1973. "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices", Econometrica, 41 (January)
- Cohen, K.J., Maier, S.F., Schwartz, R.A., and Whitcomb, D.K. (1978a) "The Returns Generation Process Returns Variance and the Effect of Thinness in Securities Markets", Journal of Finance, 33 (March)
- Cohen, K. Maier, S., Schwartz, R., and Whitcomb, D. (1978b) "Limit Orders, Market Structure, and the Returns Generation Process," Journal of Finance, 33 (June)
- Cohen, K., Maier, S., Schwartz, R., and Whitcomb, D. (1979) "Market Makers and the Market Spread: A Review of Recent Literature," Working paper, (August)
- Copeland, T. 1976. "A Model of Asset Trading Under the Assumption of Sequential Information Arrival", Journal of Finance, 31 (September)
- Davis, M.A.A. 1977. Linear Estimation and Stochastic Control. London: Chapman and Hall. New York: John Wiley & Sons
- Epps, T. 1976. "The Demand for Broker's Services: The Relation Between Security Trading Volume and Transaction Cost", Bell Journal of Economics, 7 (Spring)
- Garbade, K. and Silber, W. 1979. "Structural Organization of Secondary Markets: Clearing Frequency, Dealer Activity and Liquidity Risk", Journal of Finance, 34 (June)
- Garman, Mark. 1976. "Market Microstructure" Journal of Financial Economics, 3 (June)
- Goldman, M.B. and Beja, A. 1979. "Market Prices vs Equilibrium Prices: Returns Variance, Serial Correlation and the Role of the Specialist", Journal of Finance, 34 (June)

- Goldman, M.B. and Sosin, H.B. 1979. "Information Dissemination, Market Efficiency and the Frequency of Transactions", Journal of Financial Economics 7 (March) pp 29-61
- Oldfield, G., Rogalski, R., Jarrow, R. 1977. "An Autoregressive Jump Process For Common Stock Returns" , Journal of Financial Economics, 5
- Schleef, H. and Mildenstein, E. 1979. "A Dynamic Model of Security Dealer's Bid and Ask Prices", Working Paper (June 18)
- Schwartz, R. and Whitcomb, D. 1977. "The Time-Variance Relationship: Evidence on Autocorrelation in Common Stock Returns", Journal of Finance, 32 (March)
- Smidt, Seymour. 1979. "Continuous vs Intermittent Trading on Auction Markets", Paper presented at the Western Finance Association Meetings. June 22, 1979
- Stoll, Hans R. 1978a "The Supply of Dealer Services in Securities Markets", Journal of Finance, 33 (September)
- Stoll, Hans R. 1978b "The Pricing of Security Dealer Services: An Empirical Study of NASDAQ Stocks", Journal of Finance 33 (September)
- Tinic, Seha. 1972. "The Economics of Liquidity Services", Quarterly Journal of Economics, 86 (February)
- Tinic, Seha, and West, R. 1972. "Competition and the Pricing of Dealer Services in the Over-the-Counter Market", Journal of Financial and Quantitative Analysis, 7 (June)
- Westerfield, R. 1977 "The Distribution of Common Stock Price Changes: An Application of Transactions Time and Subordinated Stochastic Models". Journal of Financial and Quantitative Analysis (December)

TABLE 1

Price Adjustment ( $d$ ) and Risk Premium  
Component of Spread ( $s^1$ ) for Alternative  
Transaction and Inventory Values,  
Assuming  $\tau = .04$ . (in percent)

Q/W	I/W								
	-.40	-.20	-.10	-.05	0.0	.05	.10	.20	.40
0.0									
$s^1$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$d$	1.668	0.830	0.415	0.208	0.0	-.208	-.416	-0.833	-1.668
.05									
$s^1$	.112	.112	.112	.112	.112	.112	.112	.112	.112
$d$	1.786	0.894	.447	.224	0.0	-.224	-.448	-.897	-1.796
.10									
$s^1$	.239	.239	.240	.240	.240	.240	.240	.240	.241
$d$	1.914	.958	.479	.240	0.0	-.240	-.480	-.961	-1.924
.20									
$s^1$	0.542	0.543	.543	.543	.544	.544	.544	.544	.545
$d$	2.170	1.086	.543	.272	0.0	-.272	-.544	-1.089	-2.180
.40									
$s^1$	1.341	1.342	1.343	1.343	1.343	1.344	1.344	1.345	1.346
$d$	2.682	1.342	.671	.336	0.0	-.336	-.672	-1.345	-2.692

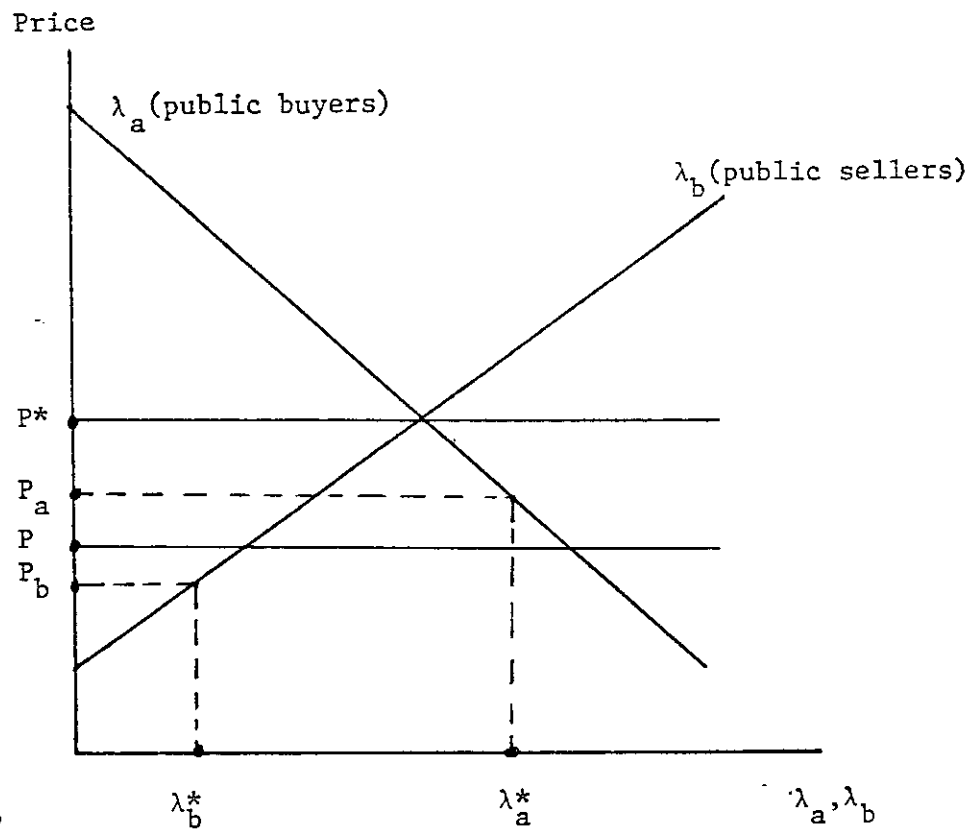


Figure 1

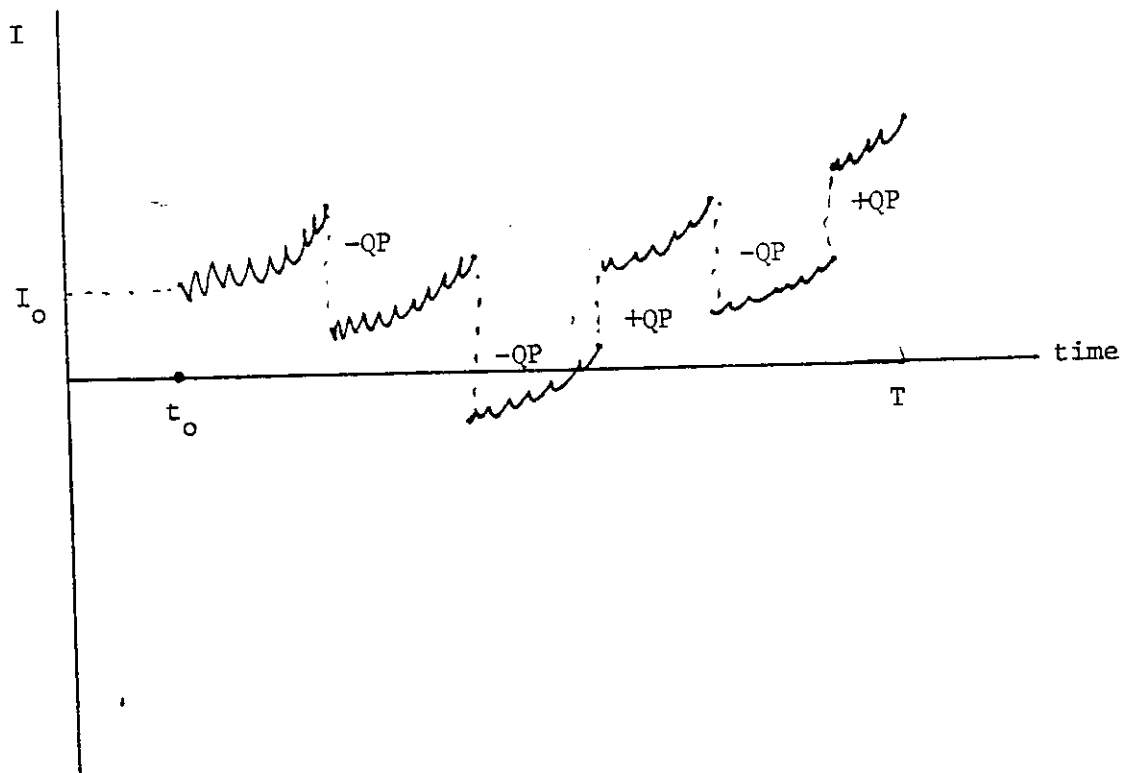


Figure 2



Derived  
Utility

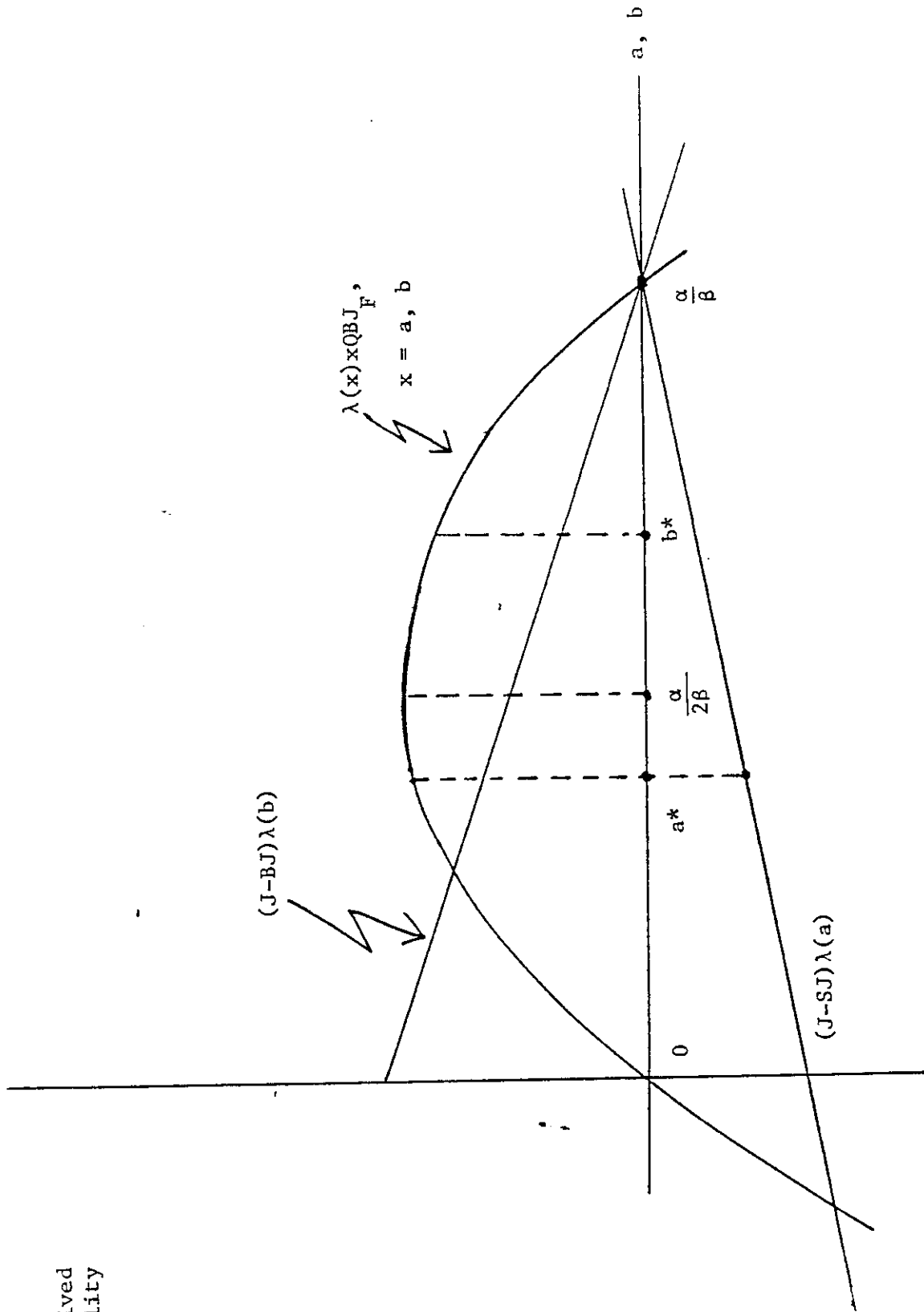


Figure 3

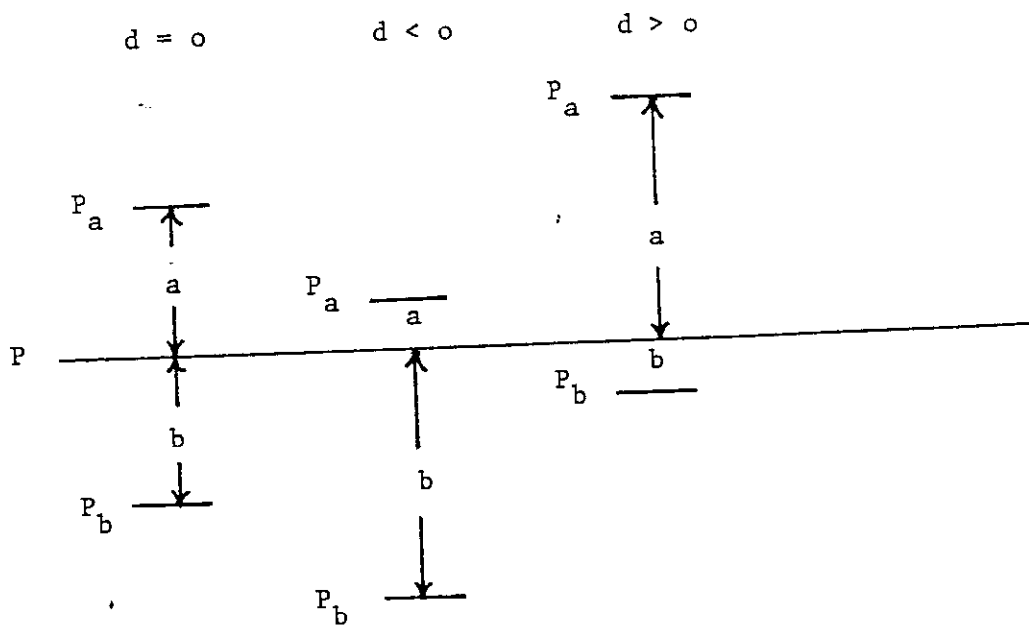


Figure 4

APPENDIX

To determine the bid-price.

$$\text{For } a = \frac{\alpha}{2\beta} + \frac{J-BJ}{2BJ_F Q} \quad (1)$$

$$b = \frac{\alpha}{2\beta} + \frac{J-SJ}{2SJ_F Q} \quad (2)$$

one can show that  $(\alpha-\beta a) (BJ_F a - (J-BJ))$  and  $(\alpha-\beta b)(SJ_F b - (J-SJ))$  are at maximum.

For  $J = \phi^0 + \phi^1 \tau + \phi^2 \tau^2$ , we have

$$b = \frac{\alpha}{2\beta} - \left[ \frac{BL\phi^0 - L\phi^0}{2U'Q} \right] \tau + \left\{ \frac{[BL\phi^0 - L\phi^0]}{U'} [B(L\phi^0)_F + 2U''\Pi] - [B\phi^2 - \phi^2] \right\} \frac{\tau^2}{2U'Q} \quad (3)$$

$$\text{But } \phi^0 = U(W) \quad (4)$$

$$\phi^1 = LU + 2\Pi U' \quad (5)$$

$$\phi^2 = \frac{1}{2}L^2U + \Pi LU' + \Pi(U''r_W W + rU') + 2\Pi^2U'' + \frac{\alpha}{4} \sigma_I^2 Q^2 U'' \quad (6)$$

where

$$\begin{aligned} L^2U &= U''(r_F + r_I I + r_Y Y)^2 + U'(r_F^2 + r_I^2 I + r_Y^2 Y) \\ &+ U''(\sigma_I^2 r_I I^2 + \sigma_{IY} r_I IY + \sigma_{IY} r_Y IY + \sigma_Y^2 r_Y Y^2) \\ &+ U''(r_I \sigma_I^2 I^2 + (r_I + r_Y) \sigma_{IY} IY + r_Y \sigma_Y^2 Y^2) \\ &+ U''(\frac{1}{2}\sigma_I^4 I^2 + \sigma_{IY}^2 IY + \frac{1}{2}\sigma_Y^4 Y^2) \end{aligned}$$

By substituting (4), (5), (6) in (3), we obtain,

$$\begin{aligned}
 b &= \frac{\alpha}{2\beta} - \left[ \varepsilon - Z \sigma_I^2 \frac{I}{W} - \frac{1}{2} Z \sigma_I^2 \frac{Q}{W} \right] \frac{r}{2} \\
 &+ \left\{ \left[ \varepsilon - Z \sigma_I^2 \frac{I}{W} - \frac{1}{2} Z \sigma_I^2 \frac{Q}{W} \right] \left[ -\frac{Z}{2} r_W + \frac{r}{2} - Z \frac{\Pi}{W} - \frac{(r_I - r)}{2} \frac{Q}{W} Z \right] \right. \\
 &- \left[ -\frac{Z}{2} (r_I - r) r_W - \frac{1}{2} Z (r_I - r)^2 \frac{Q}{W} + \frac{1}{2} (r_I^2 - r^2) \right. \\
 &- \left. \frac{Z}{2} (2\sigma_I^2 r_I \frac{I}{W} + \sigma_I^2 r_I \frac{Q}{W} + \sigma_{ZY} r_I \frac{Y}{W} + \sigma_{IY} r_Y \frac{Y}{W}) \right. \\
 &- \left. \frac{Z}{4} (\sigma_I^4 \frac{I}{W} + \frac{1}{2} \sigma_I^4 \frac{Q}{W} + \sigma_{IY}^2 \frac{Y}{W}) \right] + \frac{Z\Pi}{W} (r_I - r) \left. \right\} r^2 \\
 a &= \frac{\alpha}{2\beta} - \left[ -\varepsilon + Z \sigma_I^2 \frac{I}{W} - \frac{1}{2} Z \sigma_I^2 \frac{Q}{W} \right] \frac{r}{2} \\
 &+ \left\{ \left[ -\varepsilon + Z \sigma_I^2 \frac{I}{W} - \frac{1}{2} Z \sigma_I^2 \frac{Q}{W} \right] \left[ -\frac{Z}{2} r_W + \frac{r}{2} - Z \frac{\Pi}{W} + \frac{(r_I - r)}{2} \frac{Q}{W} Z \right] \right. \\
 &- \left[ \frac{Z}{2} (r_I - r) r_W - \frac{1}{2} Z (r_I - r)^2 \frac{Q}{W} - \frac{1}{2} (r_I^2 - r^2) \right. \\
 &- \left. \frac{Z}{2} (-2\sigma_I^2 r_I \frac{I}{W} + \sigma_I^2 r_I \frac{Q}{W} - \sigma_{IY} r_I \frac{Y}{W} - \sigma_{IY} r_Y \frac{Y}{W}) \right] \\
 &- \left. \frac{Z}{4} (-\sigma_I^4 \frac{I}{W} + \frac{1}{2} \sigma_I^4 \frac{Q}{W} - \sigma_{IY}^2 \frac{Y}{W}) \right] - \frac{Z\Pi}{W} (r_I - r) \left. \right\} r^2
 \end{aligned}$$

$$\text{where } \varepsilon = r_I - r - Z \sigma_{IY} \frac{Y}{W}$$

$$Z = - \frac{U''}{U'} W$$

It follows then

$$\begin{aligned} a + b &= \frac{\alpha}{\beta} + \frac{Z}{2} \sigma_I^2 \frac{Q}{W} \tau + \frac{Z}{2} \sigma_I^2 \frac{Q}{W} [(r_I - r + G_I) + Z (r_W + \frac{2\Pi}{W})] \tau^2 \\ &+ [ \frac{Z}{2} \frac{Q}{W} (r_I - r)^2 + Z^2 \sigma_I^2 \frac{I}{W} (r_I - r) \frac{Q}{W} - \varepsilon (r_I - r) \frac{Q}{W} Z ] \tau^2 \end{aligned}$$

If the portfolio of the base wealth of the dealer has the stock being traded, one can assume that  $\varepsilon = r_I - r - Z \frac{Y}{W}$  is small. Moreover,  $(r_I - r)^2$  and  $(r_I - r) \frac{Q}{W} \frac{I}{W}$  can assumed to be small. Hence the last term of the expression can assumed to be small. Similarly, we can compute b-a.

$$\begin{aligned} b-a &= Z \sigma_I^2 \frac{I}{W} \tau + Z \frac{I}{W} \sigma_I^2 (r_I - r + G_I + Z (r_W + \frac{2\Pi}{W})) \tau^2 \\ &+ \{ (r_I - r)(Z (r_W + \frac{2\Pi}{W})) + \frac{1}{2} (r_I - r + \sigma_{IY}) + \frac{Z^2}{2} \sigma_I^2 (\frac{Q}{W})^2 \} \\ &+ \varepsilon (r - Z(r_W + \frac{2\Pi}{W}) + \frac{Y}{W} (r_I + r_Y + \sigma_{IY})) \tau^2 \end{aligned}$$

For the same reason as above, we can assume the last term to be negligible.