



## American Finance Association

---

Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices

Author(s): Lawrence R. Glosten

Source: *The Journal of Finance*, Vol. 42, No. 5 (Dec., 1987), pp. 1293-1307

Published by: Blackwell Publishing for the American Finance Association

Stable URL: <http://www.jstor.org/stable/2328528>

Accessed: 26/11/2009 08:14

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=black>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



Blackwell Publishing and American Finance Association are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Finance*.

<http://www.jstor.org>

## Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices

LAWRENCE R. GLOSTEN\*

### ABSTRACT

The bid-ask spread can be decomposed into two parts: one part due to asymmetric information and the other part due to other factors such as monopoly power. The part due to asymmetric information attenuates statistical biases in mean return, variance, and serial covariance. Thus, using spread data to adjust for biases in return moments requires knowing not only the spread but the composition of the spread. Furthermore, any spread-estimation procedure using transaction prices must estimate two spread components. On the other hand, the appropriateness of some previously suggested statistical corrections is independent of the spread composition.

THE PURPOSE OF THIS paper is to explore further the relation between bid-ask spreads and transaction price behavior by focusing on two distinct sources of the spread. Following the taxonomy suggested in Bagehot [2], the total bid-ask spread can be split into that part that is due to monopoly power, clearing costs, inventory carrying costs, etc., and that part that arises due to the fact that the specialist believes that he or she may be trading with investors who have superior information. The first source of the spread had been analyzed by Amihud and Mendelson [1] and Ho and Stoll [9]; the second source is examined in papers by Copeland and Galai [6] and Glosten and Milgrom [8].

Others have examined the effect of the spread on transaction prices. Starting with Niederhoffer and Osborne [10] and developed further in Cohen, Maier, Schwartz, and Whitcomb [4] and Roll [11], analysis of transaction prices has suggested that the existence of a bid-ask spread will induce negative serial correlation in measured returns. In addition, Blume and Stambaugh [3] argue that the existence of a spread implies that average simple returns measured from transaction prices overstate mean true returns. Finally, French and Roll [7] argue that the existence of a bid-ask spread induces spurious variance; i.e., the return variance estimated from transaction prices overstates the true return variance. This paper will show that the magnitude of these effects is a function of the relative magnitudes of the two components of the spread as well as the width of the spread. Loosely speaking, the entire spread is not the culprit. Rather,

\* J. L. Kellogg Graduate School of Management, Northwestern University. I would like to thank, without imputing responsibility for errors or omissions, participants in the finance workshops at Columbia, University of Minnesota, and University of Chicago. I would particularly like to thank George Constantinides, Doug Diamond, Wayne Ferson, Ravi Jagannathan, Bob Korajczyk, Rob Stambaugh, and an anonymous referee. Research support from the Institute for Quantitative Research in Finance is gratefully acknowledged. Part of the research for this paper was completed while visiting the Curtis L. Carlson School of Management, University of Minnesota.

the serial correlation, spurious variance, and return bias are due to the portion of the spread arising from inventory costs, monopoly power, clearing costs, etc.

The intuition behind this assertion is relatively straightforward. Marketmakers, facing investors who are better informed than themselves, are presented with an adverse-selection problem; better informed investors will be buying when they set a price too low and will be selling when they set a price too high. However, this suggests that the occurrence of a purchase will lead to an upward revision of the true price and the occurrence of a sale will lead to a downward revision. If the marketmakers set the bid and ask rationally, then they will receive at least their revised price on investor purchases and pay no more than their revised price on investor sales. However, that must mean that a part of the spread, the adverse-selection component, is merely the difference between the two revised prices. Since this part of the spread represents possible revisions of the expectations, it should not lead to any statistical biases.

The fact that certain statistical properties of transaction prices are a function not only of the total spread but the composition of the spread as well suggests, first, that any spread-estimation technique based on transaction prices must estimate two components of the spread. For example, as will be argued below, Roll's [11] simple estimator of the spread is reliable only if there is no adverse-selection component to the spread. As another example, comparing portfolio returns calculated in the usual way with the "buy-and-hold" portfolio returns as calculated by Blume and Stambaugh [3] gives an estimate of the bias due to the bid-ask spread. The magnitude of this bias will give some insight into the magnitude of the spread, again, only if there is no adverse-selection spread component.

A second note of caution concerns the use of bid-ask spread data directly to correct for biases in estimators of transaction return moments. For example, it would be inappropriate to take average bid and ask spreads to adjust mean, variance, and serial covariance estimators without feeling assured that no part of the spread were due to the presence (or believed presence) of traders with superior information. In general, use of bid-ask spread data directly will overcorrect for spread-induced biases.

The following model and discussion will clarify the above assertions. Section I will present a simple model of the bid-ask spread. The model is a slight generalization of the model discussed in Glosten and Milgrom [8] and bears some resemblance to the model in Copeland and Galai [6]. While the environment in which we would expect the model to be completely accurate is very specialized, it does illustrate the simple relation between true price dynamics and transaction price dynamics.

It is shown that the spread due to adverse selection and the spread due to other factors are not independent. Furthermore, while we can write the transaction price as the true price plus a random "spread" term (the usual representation of the relation between transaction prices and true prices), the presence of an adverse-selection source of the spread implies that this random term and the true price are correlated. The magnitude of the covariance is related to the size of the part of the spread due to asymmetric information.

Section II investigates the statistical properties of transaction prices and “transaction returns” (returns calculated from transaction prices). It is shown that the bias in simple returns is a function of the proportion of the spread due to factors other than adverse selection, while, under some stationarity assumptions and regardless of the source of the spread, continuously compounded return averages are unbiased estimates of true continuously compounded returns. It is also shown, as in Glosten and Milgrom [8], that the negative serial covariance induced by the bid-ask spread is a function of the composition of the spread. In spite of this result, the variance adjustment suggested by French and Roll [7] is robust to the source of the spread. The paper ends with a summary and some concluding remarks regarding the direction of further research.

### I. Simple Model of Transaction Prices and Bid-Ask Spreads

The goal of this section is to describe a decomposition of the spread into the adverse-selection component, which is due to the presence of traders with private information, and the remainder, which might be termed the gross-profit component. There are two ways to view the adverse-selection spread: an *ex ante* view and an *ex post* view. The modeling to follow takes the *ex ante* view in that the adverse-selection spread is described as the anticipated change in expectations that will occur in response to a buy or a sell. The *ex post* view begins with the observation that, in the end (i.e., *ex post*), the marketmakers will lose to traders with private information. The adverse-selection spread represents the profit from liquidity traders that compensates for the loss to informed traders.<sup>1</sup>

Obviously, these are consistent views, and the connection between them can be seen in the following. Though the change in expectations in response to a trade is on average correct, an omniscient observer would note that the price overreacts to a liquidity trade and underreacts to an information trade. However, this is just another way of saying that the marketmakers profit from liquidity traders and lose to informed traders. Since the reaction is on average correct, the profits and losses balance.

We start by considering an environment in which, at any point in time, all investors with only the common-knowledge information (i.e., investors without private information) agree on the true price ( $p$ ). Then we can describe the market bid and ask prices,  $B$  and  $A$ , respectively, by

$$\begin{aligned} B &= p - Z_B - C_B, \\ A &= p + Z_A + C_A, \end{aligned} \tag{1}$$

where  $Z_A + Z_B$  is the part of the spread due to the believed presence of informed traders (the adverse-selection component) and  $C_A + C_B$  represents an addition to the spread to cover marketmaker transaction costs, inventory costs, and a normal rate of return to being a marketmaker (the gross-profit component).

<sup>1</sup> The analysis of Glosten and Milgrom [8] largely takes the *ex ante* view, while the analysis of Copeland and Galai [6] largely takes the *ex post* view.

The two parts of the adverse-selection component of the spread,  $Z_A$  and  $Z_B$ , are defined to be the magnitude of the adjustments in the (common-knowledge) true price in response to an investor purchase at the ask and a sale at the bid, respectively. That is, if  $p$  is the true price prior to a transaction and an investor purchases at the ask, then all uninformed market participants who know that this purchase has occurred (but do not know who the initiator of the transaction was) will agree that the new true price is  $p + Z_A$ . Similarly, if the transaction was an investor sale at the bid, the revised true price is  $p - Z_B$ .

The determination of  $C_A$  and  $C_B$  is not of interest for our purposes. In fact,  $C_A$  and  $C_B$  will often be treated as being given exogenously.<sup>2</sup> It is merely assumed that the market (or inside) bid and ask prices are the result of competition among all those making a market: specialist, floor traders, and limit-order submitters. It should be noted that there is no assumption that the bid and ask represent quotes by the same individual.

The above is a fairly general characterization of the adverse-selection component, but here is value in imposing somewhat more structure on the model and defining the  $Z$ 's more rigorously. To that end, let  $p^*$  be the true, full-information value of a share of the stock. That is,  $p^*$  is what the true price would be if everyone had access to the private information. Further, suppose that the private information is information about an unpriced risk. In this environment and at any point in time, the true price based on the common-knowledge information,  $H$ , is given by

$$p = E[p^* | H]. \quad (2)$$

Now, assume that all potential marketmakers have only the common-knowledge information and define the functions  $a(\cdot)$  and  $b(\cdot)$  by

$$\begin{aligned} a(x) &= E[p^* | H, \text{"investor buys at } x\text{"}], \\ b(y) &= E[p^* | H, \text{"investor sells at } y\text{"}]. \end{aligned} \quad (3)$$

The functions  $a(\cdot)$  and  $b(\cdot)$  describe how common knowledge expectations are updated in response to transactions at various possible ask and bid prices. Put another way, the functions  $a(\cdot)$  and  $b(\cdot)$  describe the inference problem that a typical marketmaker with rational expectations and the common-knowledge information will solve when determining the ask and bid prices. It is now clear what  $Z_A$  and  $Z_B$  are. If  $A$  and  $B$  are the ask and bid prices, respectively, then  $Z_A$  and  $Z_B$  are given by

$$Z_A = a(A) - p; \quad Z_B = p - b(B). \quad (4)$$

The following structure is imposed on the "expectation functions"  $a(\cdot)$  and  $b(\cdot)$ :

- (A1) Both  $a(x)$  and  $b(x)$  are increasing in  $x$ ;
- (A2)  $a(0) = b(+\infty) = p = E[p^* | H]$ ;
- (A3) For some non-negative  $c$ , there exist  $x_a < \infty$  and  $x_b > 0$  satisfying  $x_a = a(x_a) + c$ ,  $x_b = b(x_b) - c$ .

<sup>2</sup> The determination of the gross-profit component of the spread is dealt with in papers by Amihud and Mendelson [1], Cohen, Maier, Schwartz, and Whitcomb [5], and Ho and Stoll [9].

The first assumption states that the willingness of an investor to transact at a more extreme price leads to a larger revision in expectations. This is a reasonable assumption if there are informed traders and, loosely speaking, “good news” causes them to buy and “bad news” causes them to sell.<sup>3</sup> One would not expect assumption (A1) to hold, on the other hand, if there were no private information. In this case, the occurrence of a trade would not lead to any revision in expectations, and both  $a(\cdot)$  and  $b(\cdot)$  would be constant functions.

Assumption (A2) states that the willingness of an investor to buy at a price of zero or sell at an arbitrarily high price conveys no information about  $p^*$ . Assumption (A3) states that there exist bid and ask prices that allow the marketmaker to at least break even. If (A3) did not hold for some non-negative  $c$ , then assumption (A2) implies that, no matter what ask or bid was set, the revised true price in response to an investor buy would exceed the ask and the revised true price in response to an investor sell would be less than the bid; i.e., the marketmaker would expect to lose money on any transaction.<sup>4</sup>

Given the above discussion, one can imagine the process of determining the market bid and ask. All potential marketmakers perform the mental experiment embodied in the expectation functions  $a(\cdot)$  and  $b(\cdot)$  given in (3). The outcome of a game between the potential marketmakers determines a profit level (gross of transaction and inventory costs and required rate of return) on sells and buys, which sets  $C_A$  and  $C_B$ , respectively. The bid and ask are set to yield this expected profit:

$$\begin{aligned} A &= a(A) + C_A = p + (a(A) - p) + C_A = p + Z_a + C_A \\ B &= b(B) - C_B = p - (p - b(B)) - C_B = p - Z_b - C_B. \end{aligned} \quad (5)$$

Figure 1 illustrates this derivation of the bid and ask prices.<sup>6</sup>

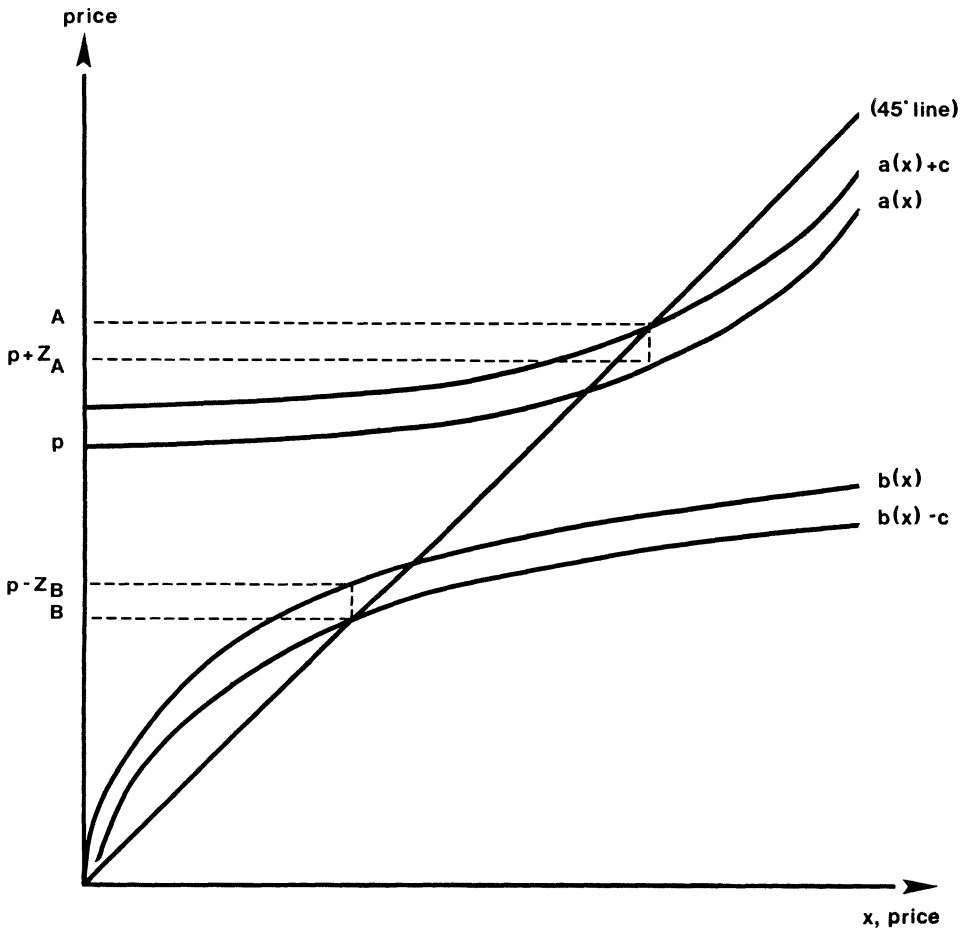
<sup>3</sup> A complete and rigorous analysis of the functions  $a(\cdot)$  and  $b(\cdot)$  will take us somewhat far afield. There is, however, a fairly easily interpretable sufficient condition that will make (A1) true. Intuitively, this condition is that, conditional on  $p^*$ , the probability of an investor purchase (sale) at  $x + s$  relative to the probability of an investor purchase (sale) at  $x$  is increasing in  $p^*$ . More formally, let the events  $A_x$  and  $B_x$  be defined by  $A_x =$  “investor buys at  $x$ ” and  $B_x =$  “investor sells at  $x$ ”, and let  $P(A_x | p^*)$  and  $P(B_x | p^*)$  be, respectively, the probability of an investor purchase at  $x$  given  $p^*$  and all public information and the probability of an investor sale at  $x$  given  $p^*$  and all public information. If  $P(A_{x+s} | p^*)/P(A_x | p^*)$  and  $P(B_{x+s} | p^*)/P(B_x | p^*)$  are both increasing in  $p^*$  for all  $s > 0$ , then  $a(\cdot)$  and  $b(\cdot)$  will be strictly increasing. These conditions are similar to monotone likelihood-ratio properties.

<sup>4</sup> This may occur at times. The rational response of competitive marketmakers would be to refuse to make a market; i.e., trading would be halted.

This assumption is not critical, but it makes interpretation of the results somewhat easier. In any event, one would expect this condition to hold “on average”, for otherwise marketmakers would be driven out of business or the market would never open.

<sup>5</sup> There may be several solutions to the first equality. A reasonable assumption is that the market ask and bid are, respectively, the smallest  $A$  and largest  $B$  satisfying  $A = a(A) + C_A$ ,  $B = b(B) - C_B$ .

<sup>6</sup> Figure 1 and the preceding discussion might suggest that the decomposition works for only a single marketmaker. This is not true as long as all potential marketmakers have the same information. Suppose one views an ask,  $A$ , and a bid,  $B$ , and these may or may not be the quotes of two distinct individuals. In principle, at any point in time, an investor with the common-knowledge information knows the (common-knowledge) true price,  $p$ , and the expectation functions  $a(\cdot)$  and  $b(\cdot)$ . Then, the adverse-selection component is  $a(A) - b(B) = Z_A + Z_B$ , and the gross-profit component is the remainder. However, if marketmakers have different information, the functions  $a(\cdot)$  and  $b(\cdot)$  will depend upon who made the quote.



**Figure 1.** Determination of bid and ask prices.  $p$  is the true price prior to a trade,  $c$  is the exogenously given level of gross profit, and  $a(\cdot)$  and  $b(\cdot)$  are the ask and bid “expectation functions”—e.g.,  $a(x)$  is the revised true price if there is an investor purchase at an ask price of  $x$ .  $A$  and  $B$  are the determined ask and bid prices, respectively. If a trade takes place at the ask, the new true price will be  $p + Z_A$  and, if at the bid,  $p - Z_B$ .

To relate this analysis to other discussions of the spread, note that, while the total quoted spread is  $Z_A + Z_B + C_A + C_B$ , by definition only  $C_A + C_B$ , the gross-profit component, offers compensation to marketmakers for the services they supply. Thus, to use the terminology of Stoll [12], the quoted spread is larger than the effective spread, the difference being the adverse-selection spread. This observation is also relevant for the empirical investigation of Roll [11], an issue that will be discussed below.

Figure 1 and the above discussion might suggest that the adverse-selection component and the gross-profit component interact. This is true in the sense that, if we interpret  $C_A$  and  $C_B$  as being given, then  $Z_A$  and  $Z_B$  are, respectively,

functions of  $C_A$  and  $C_B$ . Specifically, an increase in the  $C$ 's will increase the spread directly but, in addition, will increase the spread indirectly by increasing the  $Z$ 's. This is shown formally in the following proposition.

**PROPOSITION 1:** *Let  $Z_A$  and  $Z_B$  make up the adverse-selection component of the spread, and let  $C_A$  and  $C_B$  make up the gross-profit component. If the expectation functions satisfy (A1) through (A3), then an exogenous increase in  $C_A$  will increase  $Z_A$  and an exogenous increase in  $C_B$  will increase  $Z_B$ .*

*Proof:* See the Appendix.

This result can be understood from either the ex ante view or the ex post view. If the  $C$ 's are increased, then the incidence of a transaction will suggest to the specialist that the informed have even more extreme information. This is because the odds that a trader has extreme information are increased at the larger spread. This will lead to an even greater revision of expectations. However, the possible revision in expectations is precisely what is measured by the  $Z$ 's. Alternatively, as the  $C$ 's decrease, more liquidity traders and informed agents with less extreme information will decide to trade. Thus, the expected loss to informed traders can be spread among a probabilistically greater number of liquidity traders, and, hence, the adverse-selection spread decreases.

The dynamics of the transaction prices can now be described and related to the dynamics of true prices. Let  $\hat{p}_n$  be the  $n$ th transaction price. Suppose the  $n$ th transaction takes place when the ask and bid prices are  $A$  and  $B$ , respectively. Define  $TA$  and  $TB$  to be, respectively, the events "transaction at  $A$ " and "transaction at  $B$ ". Then the transaction price is given by

$$\hat{p}_n = AI_{TA} + BI_{TB}, \quad (6)$$

where  $I$  is the indicator function of the described event; that is, it is one if the event occurs and zero otherwise. Using (4) and (5), this can be expanded to

$$\hat{p}_n = E[p^* | H, TA]I_{TA} + E[p^* | H, TB]I_{TB} + C_A I_{TA} - C_B I_{TB}, \quad (7)$$

where  $H$  is the common-knowledge information prior to the trade and  $p^*$  is the (unobserved) full-information price.

If we let  $H_n$  be the common-knowledge information up to and including the  $n$ th transaction and assume that transactions are public information, then the first two terms on the right represent  $E[p^* | H_n]$ , but this is just the true price subsequent to the  $n$ th trade,  $p_n$ . Define the second two terms on the right to be  $C_n Q_n$ , where  $Q_n$  is +1 if the transaction was initiated by a buyer, -1 otherwise, and  $C_n$  is  $C_A$  if the transaction was initiated by a buyer,  $C_B$  otherwise. Then the transaction price  $\hat{p}_n$  is given by

$$\hat{p}_n = p_n + C_n Q_n. \quad (8)$$

This is the representation commonly used in the literature, and it is interesting that the  $Z$ 's do not appear explicitly in this formula. However,  $p_n$  and  $Q_n$  are, in general, correlated, and the amount of dependence is measured by the size of the adverse-selection component.

**PROPOSITION 2:** Let  $p$  be the true price prevailing just before the  $n$ th transaction, and let  $p_n$  be the true price after the transaction. Then,  $\text{cov}(p_n, Q_n | p) = E[Z | p]$ , where  $Z = Z_A$  if  $Q_n = 1$  and  $Z = Z_B$  if  $Q_n = -1$ .

*Proof:*

$$\begin{aligned}\text{cov}(p_n, Q_n | p) &= E[(p + ZQ_n)Q_n | p] - E[p + ZQ_n | p]E[Q_n | p] \\ &= E[ZQ_n^2 | p] + pE[Q_n | p] - pE[Q_n | p] \\ &\quad - E[ZQ_n | p]E[Q_n | p].\end{aligned}$$

Note that  $Q_n^2 = 1$ ,  $ZQ_n$  is a revision in expectations, and, hence,  $E[ZQ_n | p] = 0$ . Thus,  $\text{cov}(p_n, Q_n | p) = E[Z | p]$ . Q.E.D.

Intuitively, if  $Q_n$  is  $+1$ , then  $p_n$  must have resulted from an upward revision of expectations, and, if  $Q_n$  is  $-1$ , then  $p_n$  must have resulted from a downward revision. Hence,  $p_n$  and  $Q_n$  are correlated, and the amount of covariance is measured by how much revision in expectations is expected, i.e., the expected absolute change in expectations. Thus, the common assumption that  $p_n$  and  $C_n Q_n$  are uncorrelated is implicitly an assumption that no part of the spread is due to adverse selection.

The different roles of the adverse-selection spread and the gross-profit spread can be seen more clearly by examining the first differences of prices. Define  $\varepsilon_{n+1}$  to be revisions in the true price due to the passage of time and the arrival of public information between trade  $n$  and trade  $n + 1$ . The true price subsequent to the  $n + 1$  trade is thus  $p_{n+1} = p_n + \varepsilon_{n+1} + Z_{n+1}Q_{n+1}$ , and the  $n + 1$  transaction price is  $\hat{p}_{n+1} = p_{n+1} + C_{n+1}Q_{n+1}$ . Taking first differences,

$$\hat{p}_{n+1} - \hat{p}_n = C_{n+1}Q_{n+1} - C_n Q_n + Z_{n+1}Q_{n+1} + \varepsilon_{n+1}. \quad (9)$$

One way to interpret the difference between the spread components is to observe from (9) that the gross-profit part of the spread leads to transaction price changes that tend to reverse, whereas the adverse-selection component leads to transaction price changes that are, on average, permanent.

Expressions (8) and (9) might suggest that the uninformed only pay a part of the adverse-selection spread. For example, if the  $C$ 's and  $Z$ 's are constant and no information arrives between trades, then the round-trip experience of an investor buying and then immediately selling is (from (9))  $-2C - Z$ . Since the total spread is  $2(Z + C)$ , this investor appears to have paid only half of the adverse-selection spread. However, when buying at  $\hat{p}_n$ , this investor knows that he or she has no information, whereas the marketmaker does not. Hence, the  $p_n$  established after his or her first trade is not his or her valuation of the stock; it is too high by the adjustment due to his or her trade. Thus, the trader pays the entire spread but earns a profit on his or her information equal to half the adverse-selection spread. As a transaction cost, then, the adverse-selection part of the spread is as important as the gross-profit component.

## II. Statistical Properties of Transaction Prices

This section investigates the relation between the components of the spread and the statistical properties of transaction prices. The objective of the section is

twofold: (a) to show that, if one wanted to use spread information to correct for various biases, one would need to know not only the spread but also the composition of the spread and (b) to show what corrections are possible without such information.

The results of the section may be potentially useful for empirical work, but some care is required in their application. While the general model is rich enough to encompass a variety of environments, in order to derive expressions that are easily interpretable, some fairly strong assumptions are made. In particular, it is assumed that returns (simple and continuously compounded) are serially uncorrelated and that quoted prices are unrestricted. Despite the lack of justification for these assumptions (there is no theoretical reason why returns should be uncorrelated, and certainly quoted prices are restricted to a discrete set), they are often made in empirical work as approximations. So, too, should the specific results of this section be viewed as approximations.

Since most statistical analyses are performed upon periodic observations of transaction prices, notation is introduced to describe the relation between observed holding-period returns, transaction prices, and true prices. Let  $\hat{R}_k$  and  $R_k$  be, respectively, the observed and true simple gross returns over the  $k$ th holding period. It is presumed that the transaction return is calculated from transaction prices; i.e., every day has at least one trade. Define  $\hat{P}_k$  to be the last transaction price of period  $k$  and  $P_k$  to be the true price subsequent to this transaction. Furthermore, any variable with a subscript  $k$  will be understood to be the value of that variable at the last transaction of period  $k$ .

The next proposition shows that the spread-induced bias in average simple returns (noted, for example, by Blume and Stambaugh [3]) is attenuated by the adverse-selection spread. Thus, the bias in average simple returns is a function of not only the spread but the components of the spread. At one extreme, if there is no adverse-selection component, the bias is as described by Blume and Stambaugh [3]. At the other extreme, if the spread is entirely due to adverse selection, then there is no bias.

The intuition for this result is embodied in the fact that transaction price variations due to an adverse-selection spread reflect variations in the true price, whereas variations due to the gross-profit spread represent fluctuations about the true price. It is these latter fluctuations that cause the bias, and the importance of fluctuations about the true price relative to fluctuations in the true price is what determines the magnitude of the bias.

The intuition presented above holds quite generally, but, in order to present a result that is notationally simple, a number of assumptions are made. In particular, it is assumed that the spread is symmetric about the true price just prior to a trade and that the gross-profit component of the spread does not lead to a conditional drift in prices (note that, by definition, the adverse-selection component does not cause drift); i.e.,  $C_{Ak} + Z_{Ak} = C_{Bk} + Z_{Bk}$  and  $E_{k-1}[C_k Q_k] = 0$ , where  $E_{k-1}[\cdot]$  indicates expectation conditioned on information available after the last trade in period  $k - 1$ . Finally, assume that true returns are independent of all past history, and let the expected per-period gross return be denoted  $R$ .<sup>7</sup>

<sup>7</sup> The last transaction of the period need not occur at the end of the period, and, in fact, one can show that this introduces a further bias in average return calculation. The magnitude of this bias,

**PROPOSITION 3:** *Under the assumptions specified in the paragraph above, the expected average per-period return calculated from transaction prices is given by*

$$E[\hat{R}_k] = R(1 + \pi B), \quad B = 0.25s^2/(1 - 0.25s^2),$$

where  $R$  is the true expected per-period return,  $\pi$  is the proportion of the spread due to factors other than adverse selection,  $\pi = C/(C + Z)$ ,  $B$  is the bias when there is no adverse-selection spread, and  $s$  is the proportional spread,  $s = (A - B)/(A + B)/2$ .

*Proof:* See the Appendix.

It should be clear from the proof of the above proposition that the portfolio methods suggested by Blume and Stambaugh [3] will take care of the portfolio-return bias induced by the spread (as long as  $E_{k-1}[C_k Q_k] = 0$ ) no matter what the source of the spread. Unfortunately, the result of Proposition 3 suggests that adjusting an individual security return directly for the bias introduced by the spread requires more information than one is likely to have—not only the spread, but the composition of the spread. The next proposition provides a possible solution.

Proposition 4 shows that, under suitable stationarity assumptions and regardless of the composition of the spread, the average of continuously compounded returns provides unbiased estimates of true mean continuously compounded returns. This result, which has been proven in a symmetric-information environment by Blume and Stambaugh [3], is not vitiated by the presence of an adverse-selection spread. This should not be surprising given the discussion preceding Proposition 3.

**PROPOSITION 4:** *If the transaction price as a percentage of the true price subsequent to the transaction, i.e.,  $(P_k + C_k Q_k)/P_k$ , follows a stationary distribution, then the average of continuously compounded transaction returns is an unbiased estimator of mean continuously compounded true return.<sup>8</sup>*

*Proof:*

$$\begin{aligned} E[\log(\hat{P}_k/\hat{P}_{k-1})] &= E[\log(P_k/P_{k-1})] + E[\log(\hat{P}_k/P_k)] \\ &\quad - E[\log(\hat{P}_{k-1}/P_{k-1})] \\ &= E[\log(P_k/P_{k-1})]. \quad \text{Q.E.D.} \end{aligned}$$

The remainder of the analysis deals with the second moments of continuously compounded returns. Propositions 3 and 4 suggest that, if working with continuously compounded returns can be justified on economic grounds, then the effects

---

however, is trivial even relative to the possibly small bias introduced by the spread. For this reason, this bias is not considered here, and it is assumed that the last trade of the day is at the end of the day.

<sup>8</sup> A sufficient condition for this is that both components of the spread be proportional to the true price prior to the trade, for note that the expression in the proposition will be  $(1 + (z_k + c_k)Q_k)/(1 + z_k Q_k)$ , where  $z$  and  $c$  are the proportional spread components.

of the bias introduced by the spread can be minimized. The questions are what other biases there are and what can be done about them without having to know the composition of the spread.

We turn first to the nature of the serial correlation in transaction returns. One would guess that the composition of the spread will play a role. As has already been pointed out, the adverse-selection spread leads to fluctuations in the true price, and, hence, we would not expect this component of the spread to contribute to any spurious serial correlation. This is approximately true, and the demonstration is in the proof of the following proposition. As with Proposition 3, we make several assumptions so that we can obtain a simple formula for the serial covariance.

**PROPOSITION 5:** *Assume that true continuously compounded returns are uncorrelated, that both components of the spread are proportional to the ex ante (i.e., before trade) true price, and that the buying spread and the selling spread are the same. Then, the serial covariance of continuously compounded returns is approximately given by*

$$scov = -0.25\pi s^2,$$

where  $s$  is the total proportional spread and  $\pi$  is the proportion of the spread due to factors other than adverse selection.

*Proof:* See the Appendix.

The proposition implies that the covariance estimator of the spread suggested by Roll [11] given by  $2\sqrt{-scov}$  may be a downwardly biased estimator of the quoted spread since  $2\sqrt{-scov} = s\sqrt{\pi} \leq s$ . However, the proposition also suggests that Roll's estimator may provide upwardly biased estimates of the effective spread (i.e., the gross-profit component) since  $s\sqrt{\pi} \geq s\pi$ . Of course, these observations are true in a very special environment. Discreteness of prices and correlation of true returns will also affect the accuracy of the Roll estimator.

The observation that the spread may be related to serial correlation in continuously compounded returns has also been used to suggest an estimator of the variance of continuously compounded returns (for example, in French and Roll [7]). Specifically, it has been shown that, if the spread has no adverse-selection component (and under some suitable stationarity and independence assumptions), then adding twice the serial covariance of continuously compounded transaction returns to the variance of these returns produces a consistent estimator of the variance of true continuously compounded returns. The next proposition shows that, despite the fact that the serial covariance is a function of both the spread and the composition of the spread, the appropriateness of the procedure is independent of the spread composition.

We can write the log of the last transaction price in period  $k$  as

$$\begin{aligned} \log(\tilde{P}_k) &= \log(P_k) + \log((P_k + C_k Q_k)/P_k). \\ &= \log(P_k) + \log(1 + [c_k Q_k/(1 + z_k Q_k)]) \\ &= \log(P_k) + e_k, \end{aligned} \tag{10}$$

where  $c_k$  and  $z_k$  are the proportional spread components at date  $k$  and, in general,  $e_k$  and  $P_k$  are not independent. True continuously compounded returns and transaction returns are related by

$$\begin{aligned}\hat{r}_k &= r_k + (e_k - e_{k-1}), \\ r_k &= \log(P_k/P_{k-1}), \\ \hat{r}_k &= \log(\hat{P}_k/\hat{P}_{k-1}).\end{aligned}\tag{11}$$

**PROPOSITION 6:** Assume that (a) true returns,  $\{r_k\}$ , are serially uncorrelated, (b) the transaction noise at date  $k$ ,  $e_k$ , is uncorrelated with the return noise during period  $k - 1$ ,  $e_{k-1} - e_{k-2}$ , (c) the true return during period  $k$ ,  $r_k$ , is uncorrelated with the transaction noise at dates  $k + 1$  and  $k - 2$ ,  $e_{k+1}$  and  $e_{k-2}$ , respectively, and (d)  $\{r_k\}$  and  $\{e_k\}$  follow stationary distributions. Then,

$$\text{var}(\hat{r}_k) = \text{var}(r_k) - 2 \text{cov}(\hat{r}_k, \hat{r}_{k-1}).\tag{12}$$

*Proof:* See the Appendix.

The proof of the proposition provides no insight into why the procedure works regardless of the composition of the spread. Essentially, the intuition is that, while the variance of transaction returns is an increasing function of the total spread, only the gross-profit component adds to the variance spuriously. Thus, the spurious variance and the serial correlation are functions of the spread and the composition of the spread, and the serial covariance adjustment removes only the spurious part of the spread.

As a final note, Propositions 5 and 6 can also be proven for price differences as well as for log price differences. With proportional spreads reinterpreted as dollar spreads and continuously compounded returns reinterpreted as price differences (not returns) and with the analogous assumptions, the derived expressions will be identical.

### III. Conclusion

The basic point of this paper has been to show that the part of the spread due to asymmetric information and the part of the spread due to other factors affect the properties of the transaction-price process differently. In particular, it is the part of the spread that leads to specialist gross profit (rather than the total spread), that induces biases in our measurement of mean (simple) return and variance of the return, and that induces negative serial covariance in the measured returns. The simple intuition is that the gross-profit component leads to a fluctuation in transaction prices about the true price while the adverse-selection component leads only to fluctuations in the true price. In spite of these results, the estimator of the true variance of continuously compounded returns (observed variance plus twice the serial covariance) is correct whether or not there is asymmetric information. In fact, this works even if trades are not limited to unit amounts, as long as the specialist is allowed to quote a price schedule for various quantities. All it really relies on is that true returns and the spread components be uncorrelated.

Since statistical properties of transaction prices are typically a function of both the spread and the composition of the spread, there is no obvious candidate for a simple spread-estimation procedure based on the moments of transaction returns. In particular, the estimator of the spread proposed by Roll [11] estimates the total spread only when there is no adverse-selection spread. Furthermore, the results suggest that any attempt to estimate the spread from transaction prices should estimate two components.

While the general model is consistent with prices limited to a discrete set, the specific results are probably not. An interesting avenue of research would be to examine what effect discreteness has on the time series of bid-ask quotes and, hence, how the propositions of Section II might be made more realistic. Also, the model treats limit-order submitters as competing marketmakers. This is reasonable as long as informed agents do not submit limit orders. Another area of research suggested by this model is an investigation of what the optimal order-placement strategy of informed traders will be.

## Appendix

*Proof of Proposition 1:* The proof presented is for the ask side only. The proof for the bid side is similar. Let  $c_1$  and  $c_2$  be two exogenously specified gross-profit components of the ask side of the spread, with  $c_1 > c_2$ . Let  $A(c)$  be the ask when the gross profit spread is  $c$ . By assumptions (A1) and (A2),  $A(c)$  is the smallest  $A$  satisfying  $A = a(A) + c$ . We need to show that  $a(A(c_1)) > a(A(c_2))$ . Since  $a(\cdot)$  is strictly increasing, it is enough to show that  $A(c_1) > A(c_2)$ . Suppose the contrary, i.e.,  $A(c_1) \leq A(c_2)$ . By the definition of  $A(c_2)$ , for  $A \leq A(c_2)$ ,  $a(A) + c_2 \geq A$ . In particular,  $a(A(c_1)) + c_2 \geq A(c_1) = a(A(c_1)) + c_1$ , and, hence,  $c_2 \geq c_1$ . The contradiction establishes  $A(c_1) > A(c_2)$ . Q.E.D.

*Proof of Proposition 3:* Let  $E_{k-1}[\cdot]$  indicate expectation conditional on information available just prior to the  $k$ th trade. Then,

$$\begin{aligned} E[\hat{R}_k] &= E[\hat{P}_k/\hat{P}_{k-1}] = E[(P_k/\hat{P}_{k-1})] + E[E_k[C_k Q_k/\hat{P}_{k-1}]] \\ &= RE[P_{k-1}/\hat{P}_{k-1}]. \end{aligned}$$

For notational simplicity, denote by  $p$  the midpoint of the spread at the last trade in period  $k-1$ , and note that the conditions of the proposition imply that  $Z_A = Z_B = Z$ ,  $C_A = C_B = C$ . Then,  $E[\hat{R}_k] = RE[(p + ZQ_{k-1})/(p + (C + Z)Q_{k-1})] = (R/2)\{(p + Z)/(p + C + Z) + (p - Z)/(p - C - Z)\}$ . Letting  $\pi$  be the proportion of the spread due to factors other than adverse selection,  $\pi = C/(Z + C)$ , and letting  $s$  be the proportional spread,  $s = 2(C + Z)/p$ , we get  $E[\hat{R}_k] = R\{1 + [0.25\pi s^2/(1 - 0.25s^2)]\}$ . Q.E.D.

*Proof of Proposition 5:* Using  $Q_k$  to indicate the trade of the last investor in the  $k$ th period, we can write  $P_k$  as  $P_k = P_{k-1}(1 + \varepsilon_k)(1 + zQ_k)$ , and  $\hat{P}_k = P_{k-1}(1 + \varepsilon_k)(1 + (c + z)Q_k)$ , where  $\varepsilon_k$  represents the random fluctuations in the price due to the passage of time and the arrival of public information, and  $c$  and  $z$  are the proportional gross-profit and adverse-selection components, respectively.

Then,

$$\begin{aligned}\hat{r}_k &= \log(\hat{P}_k/\hat{P}_{k-1}) \\ &= \log((1 + zQ_{k-1})(1 + \varepsilon_k)(1 + (c + z)Q_k)/(1 + (c + z)Q_{k-1})) \\ &= \log(1 + \varepsilon_k) + \log(1 + (c + z)Q_k) \\ &\quad - \log((1 + (c + z)Q_{k-1})/(1 + zQ_{k-1})).\end{aligned}$$

Assuming that true returns are uncorrelated and that  $Q_k$  and  $Q_{k-1}$  are uncorrelated leads to

$$\begin{aligned}scov &= \text{cov}(\hat{r}_{k-1}, \hat{r}_k) \\ &= -\text{cov}(\log(1 + (c + z)Q_k), \log((1 + (c + z)Q_k)/(1 + zQ_k))).\end{aligned}$$

Working this out and assuming that  $Q_k$  is either one or minus one with equal probability we get

$$\begin{aligned}scov &= -0.25 \log((1 + c + z)/(1 - c - z)) \\ &\quad \times \log((1 + c - z(c + z))/(1 - c - z(c + z)))\end{aligned}$$

Approximating  $\log(1 + x)$  by  $x$ , we get

$$scov = 0.25 \pi s^2,$$

where, as above,  $\pi$  is the proportion of the spread due to factors other than adverse selection and  $s$  is the total proportional spread. Q.E.D.

*Proof of Proposition 6:* To simplify the notation, set  $Y_k = \log(\hat{P}_k)$ ,  $X_k = \log(P_k)$ . Then,

$$\begin{aligned}\text{cov}(Y_{k+1} - Y_k, Y_k - Y_{k-1}) &= \text{cov}(X_{k+1} - X_k + e_{k+1} - e_k, X_k - X_{k-1} + e_k - e_{k-1}) \\ &= \text{cov}(X_{k+1} - X_k, e_k) - \text{cov}(X_k - X_{k-1}, e_k) \\ &\quad - \text{var}(e_k) + 2 \text{cov}(e_{k+1}, e_k) - \text{cov}(e_{k+1}, e_{k-1}) \\ &= -\text{cov}(X_{k+1} - X_k, e_{k+1} - e_k) - \text{var}(e_k) + \text{cov}(e_{k+1}, e_k)\end{aligned}$$

by stationarity and the hypothesis of the proposition. Also,

$$\begin{aligned}\text{var}(Y_{k+1} - Y_k) &= \text{var}(X_{k+1} - X_k - e_{k+1} - e_k) \\ &= \text{var}(X_{k+1} - X_k) + 2 \text{var}(e_k) - 2 \text{cov}(e_{k+1}, e_k) \\ &\quad + 2 \text{cov}(X_{k+1} - X_k, e_{k+1} - e_k).\end{aligned}$$

Thus,  $\text{var}(Y_{k+1} - Y_k) = \text{var}(X_{k+1} - X_k) - 2 \text{cov}(Y_{k+1} - Y_k, Y_k - Y_{k-1})$ . Q.E.D.

## REFERENCES

1. Yakov Amihud and Haim Mendelson. "Dealership Market." *Journal of Financial Economics* 8 (March 1980), 31-53.
2. Walter Bagehot. "The Only Game in Town." *Financial Analysts Journal* 22 (March/April 1971), 12-14.
3. Marshall Blume and Robert F. Stambaugh. "Biases in Computed Returns." *Journal of Financial Economics* 12 (November 1983), 387-404.

4. Kalman J. Cohen, Steven F. Maier, Robert A. Schwartz, and David K. Whitcomb. "On the Existence of Serial Correlation in an Efficient Securities Market." In E. J. Elton and M. J. Gruber (eds.), *Portfolio Theory, 25 Years After: Essays in Honor of Harry Markowitz*. New York: Elsevier North-Holland, 1979, 151-68.
5. ———. "Transaction Costs, Order Placement Strategy and Existence of the Bid-Ask Spread." *Journal of Political Economy* 89 (April 1981), 287-305.
6. Thomas Copeland and Dan Galai. "Information Effects on the Bid-Ask Spread." *Journal of Finance* 38 (December 1983), 1457-69.
7. Kenneth R. French and Richard Roll. "Stock Return Variances: The Arrival of Information and the Reaction of Traders." *Journal of Financial Economics* 17 (September 1986), 5-26.
8. Lawrence R. Glosten and Paul R. Milgrom. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics* 14 (March 1985), 71-100.
9. Thomas Ho and Hans R. Stoll. "Optimal Dealer Pricing under Transactions and Return Uncertainty." *Journal of Financial Economics* 9 (March 1981), 47-73.
10. Victor Niederhoffer and M. F. M. Osborne. "Market Making and Reversals on the Stock Exchange." *Journal of the American Statistical Association* 61 (December 1966), 897-916.
11. Richard Roll. "A Simple Measure of the Effective Bid-Ask Spread in an Efficient Market." *Journal of Finance* 39 (September 1984), 1127-39.
12. Hans R. Stoll. *The Stock Exchange Specialist System: An Economic Analysis*. New York: Salomon Brothers Center, New York University, 1985.