

A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market

RICHARD ROLL*

ABSTRACT

In an efficient market, the fundamental value of a security fluctuates randomly. However, trading costs induce negative serial dependence in successive observed market price changes. In fact, given market efficiency, the effective bid-ask spread can be measured by

$$\text{Spread} = 2\sqrt{-\text{cov}}$$

where "cov" is the first-order serial covariance of price changes. This implicit measure of the bid-ask spread is derived formally and is shown empirically to be closely related to firm size.

FINANCIAL SCHOLARS AND PRACTITIONERS are interested in transaction costs for obvious reasons: the net gains to investments are affected by such costs and market equilibrium returns are likely to be influenced by cross-sectional differences in costs.

For the practical investor, the measurement of trading costs is painful but direct. (They appear on his monthly statement of account.) For the empirical researcher, trading cost measurement can itself be costly and subject to considerable error. For example, brokerage commissions are negotiated and thus depend on a number of hard-to-quantify factors such as the size of transaction, the amount of business done by that investor, and the time of day or year. The other blade of trading costs, the bid-ask spread, is perhaps even more fraught with measurement problems. The quoted spread is published for a few markets but the actual trading is done mostly within the quotes.

This paper presents a method for inferring the effective bid-ask spread directly from a time series of market prices. The method requires no data other than the prices themselves; so it is very cheap. It does, however, require two major assumptions:

- 1) The asset is traded in an informationally efficient market.
- 2) The probability distribution of observed price changes is stationary (at least for short intervals of, say, two months).

Given these assumptions, an implicit bid-ask spread measure is derived in Section I. It is investigated empirically in Section II.

* Graduate School of Management, University of California at Los Angeles. I am grateful for the thoughtful and constructive comments of Gordon Alexander, Eugene Fama, Dan Galai, Jon Ingersoll, Eduardo Lemgruber, Ron Masulis, Mark Rubinstein, and the referee.

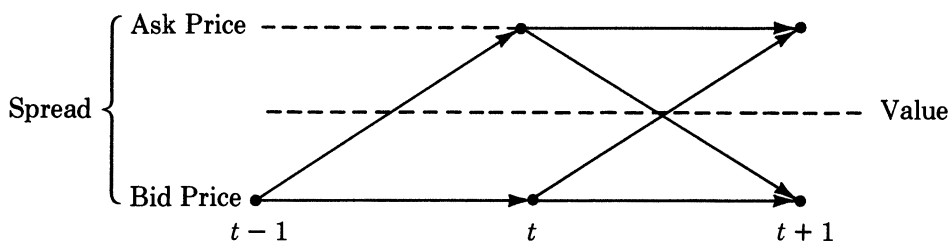
I. The Implicit Bid-Ask Spread

If the market is informationally efficient, and trading costs are zero, the observed market price contains all relevant information.¹ A change in price will occur if and only if unanticipated information is received by market participants. There will be no serial dependence in successive price changes (aside from that generated by serial dependence in expected returns).

When transactions are costly to effectuate, a market maker (or dealer) must be compensated; the usual compensation arrangement includes a bid-ask spread, a small region of price which brackets the underlying value of the asset. The market is still informationally efficient if the underlying value fluctuates randomly. We might think of "value" as being the center of the spread. When news arrives, both the bid and the ask prices move to different levels such that their average is the new equilibrium value. Thus, the bid-ask *average* fluctuates randomly in an efficient market.

Observed market price changes, however, are no longer independent because recorded transactions occur at either the bid or the ask, not at the average. As pointed out by Niederhoffer and Osborne [7], negative serial dependence in observed price changes should be anticipated when a market maker is involved in transactions. To see why, assume for simplicity of illustration that all transactions are with the market maker and that his spread is held constant over time at a dollar amount s . Given no new information about the security, it is reasonable to assume further that successive transactions are equally likely to be a purchase or a sale by the market maker as traders arrive randomly on both sides of the market for exogenous reasons of their own.

The schematic below illustrates possible paths of observed market price between successive time periods, given that the price at time $t - 1$ was a sale to the market maker, at his bid, and given that no new information arrives in the market.



Each path is equally likely. There is a similar but opposite asymmetric pattern if the price at $t - 1$ happened to be a purchase from the market maker, at his ask price.

Thus, the joint probability of successive price changes ($\Delta p_t \equiv p_t - p_{t-1}$) in trades *initiated other than by new information* depends upon whether the last transaction was at the bid or at the ask. This probability distribution (conditional on no new information) consists of two parts.

¹ Cf., Samuelson [9] and Fama [4]; but see also Grossman and Stiglitz [6] for proof that "strong-form" efficiency will not usually obtain.

| | | p_{t-1} is at the bid | | p_{t-1} is at the ask | | | |
|--------------------|------|-------------------------|---------------|-------------------------|---------------|--------------------|--|
| | | Δp_t | | Δp_t | | | |
| | | 0 | $+s$ | $-s$ | 0 | | |
| $\Delta p_{t+1} =$ | $-s$ | 0 | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ | | |
| | 0 | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | | |
| | $+s$ | $\frac{1}{4}$ | 0 | $\frac{1}{4}$ | 0 | | |
| | | | | | | $= \Delta p_{t+1}$ | |

Notice that if the transaction at $t - 1$ is at the bid (ask) price, the next price change cannot be negative (positive) because there is no new information. Similarly, there is no probability of two successive price increases (or declines).

Since a bid or an ask transaction at $t - 1$ is equally likely, the combined joint distribution of successive price changes is

| | | Δp_t | | |
|------------------|------|---------------|---------------|---------------|
| | | $-s$ | 0 | $+s$ |
| Δp_{t+1} | $-s$ | 0 | $\frac{1}{8}$ | $\frac{1}{8}$ |
| | 0 | $\frac{1}{8}$ | $\frac{1}{4}$ | $\frac{1}{8}$ |
| | s | $\frac{1}{8}$ | $\frac{1}{8}$ | 0 |

To compute the covariance between successive price changes, note that the means of Δp_t and Δp_{t+1} are zero; so the middle row and column can be ignored and the covariance is simply

$$\text{Cov}(\Delta p_t, \Delta p_{t+1}) = \frac{1}{8}(-s^2 - s^2) = -s^2/4. \quad (1)$$

The covariance is minus the square of one-half the bid-ask spread. Similarly, the variance of Δp is $s^2/2$ and the autocorrelation coefficient is $-1/2$.

The magnitude of this autocorrelation coefficient might appear to be implausible because much smaller (in absolute value) autocorrelations are invariably found in asset returns; cf., Fama [3], the original and classic article on the subject. But observed autocorrelation coefficients may be small because the covariance is divided by the sample variance of unconditional price changes. The *variance* of observed price changes is liked to be dominated by new information, whereas the *covariance* between successive price changes cannot be due to new information if markets are efficient.² The large new information component in the observed sample variance results in small observed serial correlation coefficients. Thus, in attempting to measure the bid-ask spread, we would be well-advised to work only with serial covariances, not with autocorrelations or with variances since these latter statistics are polluted (for present purposes) by news.

There are several aspects of this analysis which should be pointed out before going to the data. First, note that s is not necessarily the quoted spread. Successive price changes are recorded from actual transactions—so the s in the probability table above and in Equation (1) is the effective spread, i.e., the spread faced by the dollar-weighted average investor who actually trades at the observed prices.

² A formal proof of this statement is provided in Appendix A, Part (A).

In other words, the illustrative assumption above that all trades are with the market maker is innocuous. Even though many trades on organized exchanges are not with the market maker,³ the probability distribution above still applies, but s is the average absolute value of the price change when the price *does* change and yet no information has arrived.

Second, the *expected value* of the spread-induced serial covariance is independent of the time interval chosen for collecting successive prices.⁴ This is implied by the fact that the serial covariance depends only on whether successive sampled transactions are at the bid or the ask, not on whether any news arrives between the sample observations. Of course, in the interest of efficient estimation, the more frequent the observations the better—because nonstationarity is less likely to affect the results and because the larger sample size means that the spread will be buried in relatively less noise.

II. Empirical Estimation of the Implicit Bid-Ask Spread and Verification by Its Relation to Firm Size

The first-order serial covariance in price changes is inversely related to the effective bid-ask spread (Equation (1) above). This implies that the spread can be inferred from the sequence of price changes simply by computing and transforming the serial covariance. If percentage returns, rather than first differences of prices, are used in these calculations, we will obtain an estimate of the *percentage* bid-ask spread.⁵ (This is a more relevant measure for comparing spreads across firms.)

To verify directly that the resulting estimates of spreads are valid, it would be necessary to collect bid-ask spreads from market data (a costly procedure we are attempting to avoid). But the results can be validated indirectly by relating the measured implicit spread to firm size. Since firm size is positively related to volume (another variable for which comprehensive data are not available), and volume is negatively related to spread (see Demsetz [2] and Copeland and

³ For instance, on the New York Stock Exchange, about 12 percent of the transactions are with the specialist and about 15 percent are with other Exchange members for their own accounts; cf., *NYSE Fact Book* [8, p. 12].

⁴ A formal proof of this statement is provided in Appendix A, Part (B).

⁵ Actually, this is only approximately true. Using arithmetic returns rather than price first differences introduces a slight bias if the spread is fixed in dollar amount. This is due to the denominator of the return being either the bid or the ask which causes the expected return not to be exactly zero. It is straightforward to show that the first-order serial covariance of *returns* is exactly

$$-s_R^2/4 - s_R^4/16 = \text{Cov}(R_{t+1}, R_t)$$

where R_t , the return, is $\Delta p_t/p_{t-1}$ and the percentage spread s_R is taken with respect to the geometric mean of bid and ask prices, i.e., it is

$$s_R \equiv s/\sqrt{p_A p_B}$$

where s is the dollar spread and p_A and p_B are the ask and bid prices, respectively.

Since s_R is typically quite small, say one to three percent, the term $s_R^4/16$ can be safely ignored; its order of magnitude is 0.0000000625 to 0.0000050625. For example, if the true percentage spread s_R is three percent, ignoring the second-order term in estimating the spread from the covariance will result in an estimate of 3.00033 percent instead of exactly three percent.

Galai [1] for two different reasons), we should find a strong negative cross-sectional relation between measured spread and measured size.

Evidence for this cross-sectional relation was developed as follows. For each whole year in the CRSP⁶ daily sample, 1963–82, the serial covariance of returns was calculated for every stock which (a) had a sufficient number of observations during that year and (b) was present with a price on the last day of the previous year. Size was calculated as closing price times number of outstanding shares at the end of the preceding year.

Let $\hat{c}_{j,t}$ be the estimated serial covariance of returns of stock j in year t ; then, according to our previous analysis

$$\hat{s}_{j,t} = 200\sqrt{-\hat{c}_{j,t}} \quad (2)$$

is an estimate⁷ of the percentage bid-ask spread for the stock. (The constant 200 instead of 2.0 converts the units to percent). Two estimates of serial covariance were made for each stock, one estimate using daily returns and one estimate using weekly returns. A “sufficient number of observations” was arbitrarily chosen to be one month (21 trading days) for calculations with daily returns and 21 weeks for calculations with weekly returns.

Table I reports year-by-year cross-sectional regressions of $\hat{s}_{j,t}$ on the log of size and the predicted strong negative relation is confirmed. Indeed, the significance levels are high except for daily returns in one aberrant year, 1968. During the last half of 1968, the exchanges were closed on Wednesdays (because of a paperwork backlog). Perhaps this has something to do with the 1968 daily results in Table I being so atypical; but if it does, I certainly do not understand the mechanism.

Because of conceivable misspecifications in this parametric linear regression, a cross-sectional rank correlation is also reported. It gives much the same inference. Finally, since the estimated errors in serial covariance are probably cross-sectionally correlated, thereby biasing the t -statistics but not the estimated coefficients, the 20 yearly coefficients were used in a time-series test of significance, which is reported in the last row of the table. Although the t -statistics of the time-series mean coefficients are lower than most of the cross-sectional t -statistics, they are nevertheless large in absolute value, confirming a strong and negative relation between estimated spread and size.

The differences in the regression results between daily and weekly returns are quite minor in most years and the mean values of the cross-sectional slope coefficients are similar in size and in significance. Weekly returns produce somewhat more significant slopes and rank correlations on average.

In contrast to the cross-sectional regressions, there is a large difference in the mean values of the estimated spreads calculated from daily versus weekly serial covariances. The mean spreads derived from weekly data are larger in every year and are about six times as large on average as those derived from daily data. Notice too that the weekly-derived means are more stable over time. They are

⁶ The Center for Research in Securities Prices, Graduate School of Business, University of Chicago, equities data base. It consists of daily data for stocks listed on the New York and American Exchanges since July, 1962.

⁷ This estimator is downward biased but Appendix B shows that the bias is immaterial.

Table I
Estimated Bid-Ask Spread and Size,^a AMEX and NYSE Listed Stocks, 1963–82, One-Day (Daily)
and Five-Day (Weekly) Returns

| Year | Sample Size | | Cross-Sectional Mean Spread $\bar{s}(\%)$ (<i>t</i> -statistic) | | Cross-Sectional Regression $s_{it} = a + b \log_e(\text{Size}_{i,t-1})$ <i>b</i> (<i>t</i> -statistic) | | Cross-Sectional Rank Correlation of s_{it} and $\text{Size}_{i,t-1}$ | |
|------|-------------|----------|---|-----------------|--|-------------------|--|-------------------|
| | One-Day | Five-Day | One-Day | Five-Day | One-Day | Five-Day | One-Day | Five-Day |
| 1963 | 2043 | 2008 | 1.07 (20.1) | 2.18 (26.0) | -0.480 (-17.6) | -0.709 (-16.2) | -0.390 (-19.1) | -0.397 (-19.3) |
| 1964 | 2061 | 2036 | 0.829 (16.5) | 2.06 (24.6) | -0.521 (-20.9) | -0.731 (-17.1) | -0.447 (-22.7) | -0.377 (-18.4) |
| 1965 | 2113 | 2085 | 0.423 (8.93) | 1.89 (22.2) | -0.410 (-16.9) | -0.666 (-15.1) | -0.321 (-15.6) | -0.351 (-17.1) |
| 1966 | 2150 | 2116 | 0.193 (4.29) | 2.08 (23.7) | -0.387 (-16.2) | -0.568 (-11.9) | -0.288 (-14.0) | -0.276 (-13.1) |
| 1967 | 2181 | 2146 | -0.0726 (-1.79) | 1.69 (18.5) | -0.186 (-8.25) | -0.556 (-11.1) | -0.0881 (-4.13) | -0.259 (-12.4) |
| 1968 | 2173 | 2127 | -0.605 (-18.0) | 1.74 (19.1) | +0.0270 (1.28) | -0.555 (-9.89) | +0.121 (5.67) | -0.258 (-12.3) |
| 1969 | 2194 | 2154 | -0.468 (-14.8) | 1.58 (19.5) | -0.0597 (-2.84) | -0.234 (-4.35) | -0.00169 (-0.0791) | -0.126 (-5.91) |
| 1970 | 2297 | 2269 | -0.343 (-7.83) | 3.05 (28.2) | -0.372 (-14.6) | -0.534 (-8.26) | -0.224 (-11.0) | -0.206 (-10.0) |
| 1971 | 2406 | 2373 | -0.123 (-3.22) | 0.838 (9.98) | -0.297 (-14.2) | -0.434 (-9.22) | -0.191 (-9.55) | -0.196 (-9.76) |
| 1972 | 2540 | 2516 | 0.0553 (1.43) | 1.60 (20.7) | -0.374 (-19.1) | -0.573 (-13.3) | -0.313 (-16.6) | -0.298 (-15.6) |
| 1973 | 2678 | 2642 | 0.355 (6.78) | 1.25 (12.3) | -0.742 (-28.0) | -0.522 (-9.05) | -0.459 (-26.7) | -0.175 (-9.14) |
| 1974 | 2707 | 2669 | 1.22 (19.3) | 3.32 (30.0) | -0.961 (-34.6) | -0.769 (-13.6) | -0.564 (-35.5) | -0.260 (-13.8) |
| 1975 | 2657 | 2624 | 0.742 (13.4) | 2.21 (19.9) | -0.714 (-28.3) | -0.530 (-9.37) | -0.460 (-26.7) | -0.193 (-10.1) |

| | | | | | | | | |
|-------|-----------------------|-----------------------|--------------------|-----------------|---------------------|---------------------|---------------------|---------------------|
| 1976 | 2610 | 2572 | 0.688 (15.6) | 1.65 (19.0) | -0.533 (-26.1) | -0.471 (-10.6) | -0.443 (-25.2) | -0.220 (-11.4) |
| 1977 | 2570 | 2542 | 0.834 (20.2) | 2.03 (27.8) | -0.565 (-30.0) | -0.518 (-13.9) | -0.498 (-29.1) | -0.290 (-15.3) |
| 1978 | 2515 | 2468 | 0.221 (5.23) | 0.729 (8.42) | -0.368 (-16.9) | -0.410 (-8.83) | -0.255 (-13.2) | -0.160 (-8.07) |
| 1979 | 2439 | 2394 | 0.315 (7.69) | 1.00 (12.1) | -0.417 (-19.7) | -0.483 (-10.8) | -0.311 (-16.1) | -0.214 (-10.7) |
| 1980 | 2375 | 2328 | 0.265 (5.62) | 1.11 (12.5) | -0.551 (-23.4) | -0.250 (-5.10) | -0.390 (-20.6) | -0.0939 (-4.55) |
| 1981 | 2348 | 2294 | 0.170 (2.90) | 1.38 (16.7) | -0.504 (-16.8) | -0.238 (-5.34) | -0.358 (-18.5) | -0.111 (-5.34) |
| 1982 | 2339 | 2277 | -0.0492 (-1.07) | 1.41 (14.8) | -0.418 (-17.4) | -0.141 (-2.67) | -0.309 (-15.7) | -0.0565 (-2.70) |
| All | 47414 | 46658 | 0.298 | 1.74 | -0.442 ^c | -0.495 ^c | -0.310 ^c | -0.226 ^c |
| Years | (24358 ^b) | (16225 ^b) | (27.9) | (84.3) | (-8.60) | (-12.6) | (-7.93) | (-10.7) |

^a The variables used are $s_{jt} = 200 \sqrt{-\hat{c}_{jt}}$, the estimated bid-ask spread, where \hat{c}_{jt} is the serial-covariance of returns in year t on stock j . (The sign of the covariance was preserved after taking the square root.) Size $_{j,t-1}$ is the market capitalization in \$ millions (number of shares times the price) on the last trading day of year $t - 1$. Stocks were discarded from the sample in year t unless they had at least 21 observations (21 trading days or approximately one month of data for daily returns and 21 weeks for five-day returns).

^b Number of negative spread estimates.

^c Means and t -statistics based on 20 time series observations of cross-sectional coefficients.

positive in every year whereas the daily-derived estimates have negative means in six years out of 20.

Using daily data, the average value of the implicit bid-ask spread across all stocks and time periods was only 0.298 percent. This is an estimate of the average *effective* spread and should be smaller than the quoted spread; but the minimum quoted spread is 1/8th of a dollar, which would be about 0.3 percent of a stock selling for 41%. This may not be too far from the average price of a NYSE issue but it seems too high for an AMEX stock. The average implicit bid-ask spread estimated from weekly returns was 1.74 percent, which is certainly in a more believable range for the average over all issues on both exchanges.

The difference between spreads estimated from daily and weekly data is too large to be attributed to small sample bias in the smaller sample sizes used for the weekly calculations (see Appendix B). The difference is statistically significant. This is verified by performing a paired *t*-test of the difference in the two estimates; i.e., the difference $d_{jt} = \hat{s}_{5jt} - \hat{s}_{1jt}$ was calculated for stock *j* and year *t* between the spread estimated from five-day (\hat{s}_{5jt}), and one-day (\hat{s}_{1jt}), returns. The cross-sectional mean of $d_{j,t}$ for year *t* was tested for significance from zero using a standard *t*-statistic. The minimum *t* value over the 20 years was 5.94 and the average over the 20 years was 16.6. Out of 46658 values of $d_{j,t}$, 29611, or 63.5 percent, were positive.

Since the spreads inferred from any observation interval must be equal when markets are informationally efficient, these results cast doubt on the contention that the New York and American Exchanges really are in fact perfectly efficient. The degree of inefficiency may be economically insignificant and too small to exploit profitably, and yet still be large enough to cause estimation problems for the spread. Apparently, the serial dependence is less positive for weekly than for daily returns. Perhaps daily returns have inefficiency-induced positive dependence. Perhaps weekly returns have negative dependence.

Another possibility is that mean returns are nonstationary. Positive dependence in observed daily returns could be induced by short-term fluctuations in expected returns which dampen out over a period as long as a week, thus leaving less dependence in observed weekly returns. Nonstationarity in the spread itself, caused by the reactions of dealers to stochastic information arrival, is less likely to be an explanation.⁸ But some more complex type of nonstationarity could be present. Further work will have to decide whether market inefficiency or nonstationarity, or both, is the problem.

III. Summary

The effective bid-ask spread can be inferred from the first-order serial covariance of price changes, provided that the market is informationally efficient. The implicit percentage spread is given by

$$s_j = 200\sqrt{-\text{cov}_j}$$

where s_j is the spread and cov_j is the serial covariance of returns for asset *j*.

This implicit measure of trading costs was estimated annually from daily and

⁸ See the discussion in Appendix A, Part (C).

weekly returns of stocks listed on the New York and American Exchanges. The resulting estimates were strongly negatively related to firm size, thus supporting the measure of being related to trading costs (which are negatively related also to firm size). However, a sizeable difference was detected between spreads estimated from daily and weekly data. This implies informational inefficiency, (although not necessarily profit opportunities) or else very short-term nonstationarity in expected returns.

Appendix A

The following are proofs that: (A) the covariance between successive price changes cannot be due to new information if markets are informationally efficient, (B) the implicit spread measure is independent of the observation interval if markets are efficient, and (C) even if the spread changes in reaction to news, the serial covariance will still be $-s^2/4$ where s^2 is the average squared spread in the sample.

(A) If markets are efficient, the effective spread brackets the “value” of the assets. Denote this true but unobserved value p^* . The observed price change in period t consists then, of two parts, a change in p^* caused by new information and a component determined entirely by whether the transaction at the end of t was initiated by the same side of the market; i.e., the observed price change, $\Delta\hat{p}_t$, is given by

$$\Delta\hat{p}_t = \Delta p_t^* + \Delta p_t$$

where Δp_t is the transaction cost component whose probability distribution is given by the table in the text on page 1129.

If markets are informationally efficient, we must have

$$\text{Cov}(\Delta p_t^*, \Delta p_{t-j}^*) = 0 \quad j \neq 0 \quad (\text{A1})$$

and we also require

$$\text{Cov}(\Delta p_t^*, \Delta p_{t-j}) = 0. \quad (\text{A2})$$

The first (A1) covariances are zero because changes in value are surprises in efficient markets. The second (A2) covariances are zero because movements between bid and ask prices cannot be predicted by, nor be predictors of, changes in value. For $j \neq 0$ in (A1) and $j \geq 0$ in (A2) the zero values follow directly from the proposition that Δp_t^* is unrelated to *all* other preceding variables (including its own value in earlier periods). The value of zero for covariances $j < 0$ in (A2) might seem to require the added assumption that the current information surprise does not affect the spread (but I shall argue in (C) below that only a weaker assumption is actually necessary).

Using (A1) and (A2), we obtain

$$\begin{aligned} \text{Cov}(\Delta\hat{p}_t, \Delta\hat{p}_{t-1}) &= \text{Cov}(\Delta p_t^* + \Delta p_t, \Delta p_{t-1}^* + \Delta p_{t-1}) \\ &= \text{Cov}(\Delta p_t, \Delta p_{t-1}) = -s^2/4. \end{aligned}$$

Thus, the covariance between successive price changes is not due to new information (but only to the spread).

(B) Now consider calculating the covariance over a longer interval. The price change over some longer interval, say N periods, is simply the sum of N successive changes; i.e., define

$$\Delta \hat{p}_T \equiv \sum_{(T-1)N+1}^{TN} \Delta \hat{p}_t$$

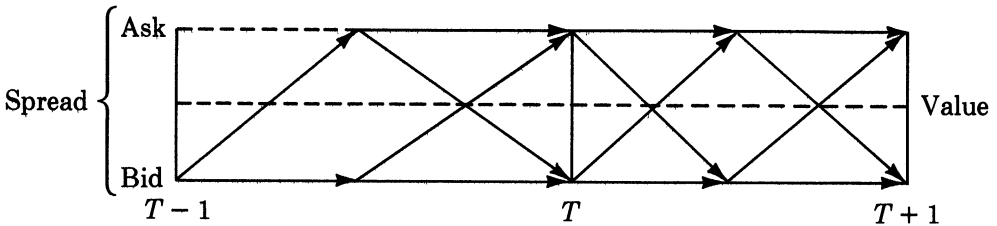
where T is an index of the longer interval.

Thus,

$$\Delta \hat{p}_T = \Delta p_T^* + \Delta p_T$$

where Δp_T^* and Δp_T are sums of, respectively, the new information components and the spread components over the N periods.

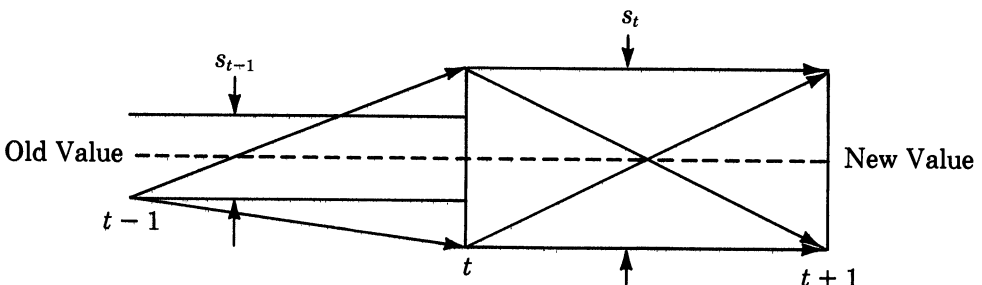
The sum of the spread components Δp_T has exactly the same distribution as an individual component, because, although the price now bounces back and forth between bid and ask up to a maximum of N times during interval T , it still comes to rest at one end or the other of the spread. For example, a diagram analogous to the simpler one of the text for $N = 2$, is



In general, there are 2^N possible paths between $T - 1$ and T (given that the price at $T - 1$ is at the bid), and there are 2^{N+1} paths possible from T to $T + 1$. All paths are equally likely but the diagram proves that exactly half of the paths produce the same value of Δp . Thus, $\text{Cov}(\Delta p_T, \Delta p_{T-1}) = -s^2/4$.

For nonoverlapping intervals, it is straightforward to apply conditions (A1) and (A2) to the sums and obtain $\text{Cov}(\Delta p_T^*, \Delta p_{T-1}^*) = \text{Cov}(\Delta p_T^*, \Delta p_{T-j}^*) = 0$ for all j . Thus, $\text{Cov}(\Delta \hat{p}_T, \Delta \hat{p}_{T-1}) = -s^2/4$, which is independent of N , the number of periods within the measurement interval.

(C) Now consider the possibility that the spread is affected by information arrival. It would seem sensible that the spread might widen the larger the absolute value of the price change from $t - 1$ to t , and that this widening would occur whether the information inducing the price change were good or bad. If the spread does react symmetrically, without loss of generality the schematic below can be used to model the process.



Value has changed from $t - 1$ to t and this induces a change in the spread $\Delta s = s_t - s_{t-1}$ (which can be either positive or negative). The schematic lines up the new and old values. Notice that the possible price paths (indicated by arrows) now permit price increases or decreases in both periods, in contrast to the situation in which the spread is constant.

By virtue of the symmetry assumption, that the change in spread is independent of the algebraic sign of the change in value (although it does depend on the absolute magnitude of that change), we can ignore Δp^* ; i.e., the symmetry assumption implies

$$\text{Cov}(\Delta p_{t-1}^*, \Delta p_t) = \text{Cov}(\Delta p_t^*, \Delta p_{t-1}) = 0.$$

Since market efficiency guarantees also that $\text{Cov}(\Delta p_t^*, \Delta p_{t-1}^*) = 0$, it is still the spread alone which determines the *observed* serial covariance. Now, however, the probability distribution is more complex,

| | | Δp_t | | | |
|------------------|--------|-------------------------|---------------|--------------|------------------------|
| | | $-s_{t-1} - \Delta s/2$ | $-\Delta s/2$ | $\Delta s/2$ | $s_{t-1} + \Delta s/2$ |
| Δp_{t+1} | $-s_t$ | 0 | 0 | $1/8$ | $1/8$ |
| | 0 | $1/8$ | $1/8$ | $1/8$ | $1/8$ |
| | s_t | $1/8$ | $1/8$ | 0 | 0 |

where $\Delta s = s_t - s_{t-1}$. However, it is straightforward to verify that $\text{Cov}(\Delta p_t, \Delta p_{t+1}) = -s_t^2/4$!

Surprisingly, the only thing that matters is the *new* spread; and the formula for the serial covariance is identical to that derived under the assumption of a constant spread. The spread can change over time in reaction to news but the observed serial covariance will still be related by the same simple formula to the spread.

For the sample cross-product computed from $t - 1$ to t and t to $t + 1$,

$$E[\Delta p_t, \Delta p_{t+1}] = -s_t^2/4.$$

The expected value of the covariance for the entire sample is thus

$$E[\sum_t (\Delta p_t, \Delta p_{t+1})/\tau] = -\bar{s}^2/4, \quad t = 1, \dots, \tau.$$

where \bar{s}^2 is the average squared spread during the sample of length τ .

Finally, note that changes in the spread can conceivably be induced by *any* alteration in the distribution of new information. For example, a nonstationary variance of returns would be a likely source of a nonstationary spread. However, regardless of the source of nonstationarity, so long as the spread is not related to the *algebraic* values of past price changes, our simple formula will still be valid.

Appendix B

The Bias in the Sample Estimator of Spread

Due to Jensen's inequality, the sample estimator (2) of spread will be biased in small samples. This appendix derives the approximate size of the bias. The

bias arises because the true autocovariance of price changes in (1) is solved for the value of the spread in order to obtain (2). But since only a sample estimate of the autocovariance is known, the square root transformation induces the well-known Jensen's inequality problem, i.e., although

$$E(\hat{c}) = c = -s^2/4,$$

where $\hat{\cdot}$ indicates a sample estimate and c denotes the first-order serial covariance

$$E(\hat{s}) = E[2\sqrt{-\hat{c}}] < 2\sqrt{-E(\hat{c})} = s,$$

the sample estimator \hat{s} is biased downward.

To obtain the approximate size of the bias, we first expand the function $E(\hat{s})$ in a Taylor series. Define $\epsilon \equiv \hat{c} - c$, the sample error in estimating the serial covariance. Expanding $E(\hat{s})$ about $\sqrt{-c}$ and dropping the higher than third-order terms, we obtain

$$\begin{aligned} E(\hat{s}) &= 2E[\sqrt{-c} + \epsilon \cdot \frac{\partial}{\partial c} \sqrt{-c} + \frac{\epsilon^2}{2} \frac{\partial^2}{\partial c^2} \sqrt{-c} + \dots] \\ &= s - 2\sigma_\epsilon^2/s^3 \end{aligned} \quad (\text{B1})$$

where σ_ϵ^2 is the estimation error variance of the serial covariance.

An expression for σ_ϵ^2 can be obtained from asymptotic formulae such as those given in Fuller [5, Section 6.2, pp. 236–44]. Using formulae (6.2.2) and theorem 6.2.2 in Fuller,

$$\lim_{n \rightarrow \infty} (n-1)\sigma_\epsilon^2 = (\eta-3)c^2 + \sum_{p=-\infty}^{\infty} [\gamma^2(p) + \gamma(p+1)\gamma(p-1)]$$

where n is the time series sample size, η is the kurtosis of the price changes (it is equal to 3 if they are normally distributed), and $\gamma(p)$ is the (true) serial covariance of order p . Note that $\gamma(p) = 0$ in the present case except for $-1 \leq p \leq 1$; while $\gamma(-1) = \gamma(1) = c = -s^2/4$ and $\gamma(0) = s^2/2$. Thus, we can estimate σ_ϵ^2 by

$$\begin{aligned} \sigma_\epsilon^2 &= \frac{1}{n-1} [(\eta-3)c^2 + \gamma^2(-1) + \gamma^2(0) + \gamma^2(1) + \gamma(-1)\gamma(1)] \\ &= \frac{s^4}{16(n-1)} [(\eta-3) + 7]. \end{aligned} \quad (\text{B2})$$

Using the estimate σ_ϵ^2 from (B2) in (B1), we obtain

$$E(\hat{s}) = s \left[1 - \frac{(\eta-3) + 7}{8(n-1)} \right]. \quad (\text{B3})$$

For normally distributed data, the term $\eta-3$ disappears. Even for very thick-tailed data, say $\eta = 6$, the bias in \hat{s} with, say, $n = 60$ observations is

$$E(\hat{s}) = 0.979s$$

or slightly more than two percent. Most of the stocks in the sample had around 250 observations in the typical covariance estimate with daily data (this is approximately the number of trading days in the year). The bias in such cases

with $\eta = 6$ is about 0.5 percent of s . The minimum sample size was 21 observations and the bias in such a case could be as large as six percent.

REFERENCES

1. Thomas E. Copeland and Dan Galai. "Information Effects on the Bid-Ask Spread." *Journal of Finance* 38 (December 1983), 1457-69.
2. Harold Demsetz. "The Cost of Transacting." *Quarterly Journal of Economics* 82 (February 1968) 33-53.
3. Eugene F. Fama. "The Behavior of Stock Market Prices." *Journal of Business* 38 (January 1965) 34-105.
4. ———. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25 (May 1970) 383-417.
5. Wayne A. Fuller. *Introduction to Statistical Time Series*, New York: John Wiley & Sons, 1976.
6. Stanford J. Grossman and Joseph E. Stiglitz. "Information and Competitive Price Systems." *American Economic Review* 66 (May 1976), 246-53.
7. Victor Niederhoffer and M. F. M. Osborne. "Market Making and Reversal on the Stock Exchange." *Journal of the American Statistical Association* 61 (December 1966), 897-916.
8. *NYSE Fact Book 1982*. New York: New York Stock Exchange, 1982.
9. Paul A. Samuelson. "Proof that Properly Anticipated Prices Fluctuate Randomly." *Industrial Management Review* 6 (Spring 1965) 41-49.