

An Extended Model of Serial Covariance Bid-Ask Spreads

Dar-Hsin Chen*

Department of Banking and Finance, Tamkang University, Taiwan

Lloyd P. Blenman

*Department of Finance and Business Law, University of
North Carolina-Charlotte, U.S.A.*

Abstract

This paper presents a generalized serial covariance spread pricing model that unifies and improves existing spread models. We analyze three cost components of spread: order processing, adverse information, and inventory holding costs. We modify Stoll's (1989) model by incorporating a two-period conditional probability trading model to derive a new spread estimator. We propose a methodology to estimate the input parameters. We then show this extended model potentially avoids some of the limitations associated with earlier models.

Key words: bid-ask spread; implicit spread; tick test

JEL classification: G10; D80

1. Introduction

We derive the spread in the serial covariance model from the statistical properties of price changes. We first estimate the serial covariance of price changes and then use that estimate to infer the spread. This indirect approach differs from the usual methods of directly calculating the spread. Roll (1984) derives an implicit spread estimator in the equity market. He observes that the quoted spread is not necessarily equal to the effective spread, and sometimes the quoted spread overstates the real transaction costs faced by traders. He thus uses a relationship between transaction price changes to estimate indirectly the effective spread in an efficient market, and his method basically only requires the transaction prices themselves.

Stoll's (1989) work builds on that of Roll (1984). Traditional bid-ask spread models use the inventory liquidation mechanism, transaction costs, or informational asymmetry to explore the spreads, but in this paper we will identify these different cost components of the spread by studying the transaction types and the statistical properties of transaction price changes. We will first introduce the notion of Roll's

Received December 3, 2002, revised March 17, 2003, accepted March 27, 2003.

*Correspondence to: Department of Banking and Finance, Tamkang University, Tamsui 251, Taiwan.
E-mail: dhchen@mail.tku.edu.tw.

simple bid-ask spread estimator and then extend the more complicated estimators which have been developed by Stoll (1989) and others in order to improve their potential applications.

Recently there has been concern over the performance of the Roll (1984) and Stoll (1989) estimators. George et al. (1991) allow for time-varying returns and implement their model only requiring one auto-covariance. Brooks and Masson (1996) report that the estimator of the realized spread suggested by Stoll suffers from difficulty in estimating quotes and price change covariances. Moreover, upon applying the Stoll's technique, they find that the estimated spread ranges from 14% to 257% of actually quoted spreads in a sample of NYSE stocks. Lesmond et al. (1999) report similar problems in the George et al. (1991) model as in that of Stoll (1989). It is clear that these estimators have serious problems and there is scope for an improved model.

Our new model allows for the existence of the order processing costs as well as inventory holding costs and asymmetric information costs. This permits us to get rid of the assumption of the independence of successive trades and allows for the presence of serially correlated effects. As such, our model easily accommodates price reversal effects that can be brought about by asymmetric information and inventory holding costs. Such price reversal effects are absent in the models we generalize.

We organize our paper as follows. Section 2 presents the theory and proposes a more generalized serial covariance spread model, which not only reconciles the above models but also considers the different cost components of the spread. In Section 3, we discuss how to estimate the parameters in our model. Section 4 proposes new ways to apply this new model in empirical research. We then conclude and summarize our results.

2. Background and Model Development

Under Roll's ideal market assumptions, the possible paths of successive transaction price changes are restricted. Since transaction prices can only bounce either at the ask price or at the bid price, he derives the spread estimator by computing the combined joint distribution of successive price changes. His spread estimator results from solving: $COV(\Delta P_t, \Delta P_{t+1}) = E(\Delta P_t)(\Delta P_{t+1}) - E(\Delta P_t)E(\Delta P_{t+1}) = E(\Delta P_t)(\Delta P_{t+1}) = (1/8)(-S^2 - S^2) = -S^2/4$. His effective (implied) spread estimator is:

$$S = 2\sqrt{-COV(\Delta P_t, \Delta P_{t+1})}. \quad (1)$$

However, earlier studies such as Garbade and Lieber (1977) report that ask (bid) orders are more likely to follow ask (bid) orders and are therefore serially dependent. If future prices are not independent of current prices, then Roll's technique is no longer appropriate.

Choi et al. (1988) modify Roll's estimator and introduce an unequal and serially-correlated assumption regarding order arrivals. They allow for the conditional probability of a transaction at the bid (ask) following one at the bid (ask) to be dif-

ferent from 0.5. They derive their new spread estimators by computing the combined joint distribution of successive price changes.

Their model implies that $COV(\Delta P_t, \Delta P_{t+1}) = -S^2(1-\delta)^2$, and the effective spread, S , can be derived as:

$$S = \frac{\sqrt{-COV(\Delta P_t, \Delta P_{t+1})}}{1-\delta}, \quad (2)$$

where δ is the conditional probability that the transaction at time $t+1$ is at the bid (ask) price given that the transaction at time t is at the bid (ask) price.

Chu et al. (1996) develop another generalized model by further considering the possibility of a transaction type conditional upon two preceding transaction periods rather than one preceding period as that in Choi et al. (1988). They incorporate not only δ as the conditional probability of a subsequent transaction being the same type as that of the current transaction but also (1) α as the conditional probability of the next transaction being the same type as the current type but different from the previous type and (2) β as the conditional probability of the next transaction being the same type as both the current type and the previous type.

If the initial transaction is a bid transaction, the paths are similar by symmetry. This implies that in their model $COV(\Delta P_t, \Delta P_{t+1}) = -S^2[(1-\alpha)(1-\delta)/2] - S^2[(1-\alpha)(1-\delta)/2] = -S^2(1-\alpha)(1-\delta)$, and rearranging S to the left-hand-side results in:

$$S = \frac{\sqrt{-COV(\Delta P_t, \Delta P_{t+1})}}{(1-\delta)(1-\alpha)}. \quad (3)$$

However, the original Roll model and the modified models above still focus on the existence of the order processing costs while the other two critical costs usually mentioned in the literature—namely, the inventory holding costs and asymmetric information costs—are ignored. Thus no inventory adjustment is considered and independence of trade flows is assumed. Their models generally focus on modifying the serial correlation of transaction types while the price reversal effects caused by holding inventory and facing asymmetric information simply disappear.

Stoll (1989) decomposes the possible reasons for security price changes (ΔT) into three possible changes: the changes due to the basic security characteristics (SC), the price changes due to the bid-ask bounce (ΔP), and the price changes due to new information (e). The total price changes are as follows:

$$\Delta T_t = SC + \Delta P_t + e_t. \quad (4)$$

The most important assumption is that in an informationally efficient market the price changes resulting from new information are independent and also are uncorrelated with lagged and leading values of the price changes due to the spread, while e_t is assumed to be white noise with a zero mean and a constant variance. Updating $\Delta T_{t+1} = SC + \Delta P_{t+1} + e_{t+1}$ implies that the serial covariance of price changes

becomes:

$$\begin{aligned} COV(\Delta T_t, \Delta T_{t+1}) &= COV(\Delta P_t, \Delta P_{t+1}) + COV(\Delta P_t, \Delta e_{t+1}) + \\ &\quad COV(e_t, \Delta P_{t+1}) + COV(e_t, \Delta e_{t+1}) \\ &= COV(\Delta P_t, \Delta P_{t+1}). \end{aligned} \quad (5)$$

Thus, in an informationally efficient market, the spread is the only possible cause for the serial covariance of price changes in this model. If a transaction starts at the bid price, it can only take on two possible values: (1) $A_{t+1} - B_t = (1 - \lambda)S$ with probability π or (2) $(B_{t+1} - B_t) = -\lambda S$ with probability $1 - \pi$, where A is the ask transaction, B is the bid transaction, π is the probability of a price reversal, and $1 - \lambda$ represents the magnitude of a price reversal as a fraction of the spread, S . Similarly, if the preceding transaction is at the ask price, then ΔP can take on only two values: (1) $B_{t+1} - A_t = -(1 - \lambda)S$ with probability π or (2) $A_{t+1} - A_t = \lambda S$ with probability $1 - \pi$. A price reversal is defined as a trade occurring at the bid (ask) and the next trade occurring at the ask (bid).

Stoll shows that the dealer's compensation for order processing and inventory costs—that is, the expected revenue earned on a round-trip trade—is equal to $2S(\pi - \lambda)$ and that $S[1 - 2(\pi - \lambda)]$ reflects the asymmetric information component of the spread. The order processing and inventory costs can be further decomposed into $2S(\pi - 0.5)$ and $S(1 - 2\lambda)$, respectively, with the former being the inventory costs and the latter reflecting the compensation for order processing costs. Thus, by using the serial covariance of buy and sell transaction prices, π and λ can be estimated.

Based on these arguments, Stoll derives the serial covariance of transaction price changes as:

$$COV(\Delta P_t, \Delta P_{t-1}) = S^2 [\lambda^2(1 - 2\pi) - \pi^2(1 - 2\lambda)]. \quad (6)$$

Stoll models the probability of price reversal depending only on the preceding period. Thus, his one-period conditional probability is:

$$\begin{aligned} P(A_t | A_{t-1}) &= P(A_{t+1} | A_t) = P(B_t | B_{t-1}) = P(B_{t+1} | B_t) = 1 - \pi \\ P(A_t | B_{t-1}) &= P(A_{t+1} | B_t) = P(B_t | A_{t-1}) = P(B_{t+1} | A_t) = \pi. \end{aligned}$$

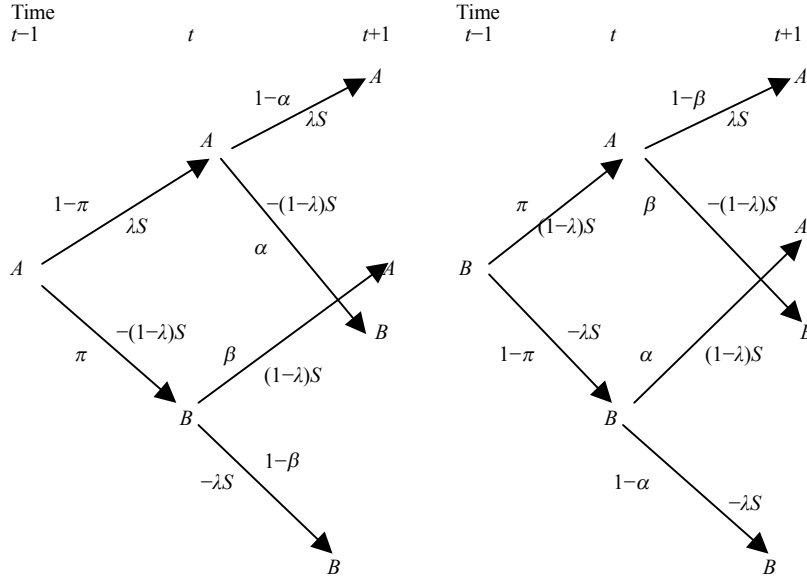
In our study we allow for transaction reversals to follow a two-period conditional probability distribution. In particular, the transaction type does not merely depend on the previous period but also on the period before that. Figure 1 shows the bid-ask probability and reversal determination diagram under the new probability assumptions.

Statistically, the new probability assumption in Figure 1 can be summarized as:

$$\begin{aligned} P(A_{t+1} | A_{t-1}, A_t) &= P(B_{t+1} | B_{t-1}, B_t) = 1 - \alpha \\ P(A_{t+1} | B_{t-1}, B_t) &= P(B_{t+1} | A_{t-1}, A_t) = \alpha \\ P(A_{t+1} | A_{t-1}, B_t) &= P(B_{t+1} | B_{t-1}, A_t) = \beta \end{aligned}$$

$$P(A_{t+1} | B_{t-1}, A_t) = P(B_{t+1} | A_{t-1}, B_t) = 1 - \beta.$$

Figure 1. The Probability and Reversal Structure of Trade



Note: A is the ask transaction; B is the bid transaction; $1-\lambda$ is the magnitude of a price reversal as a fraction of the spread, S ; and π , α , and β are conditional probabilities of price reversal under different conditions. (The probability of transaction type at time $t-1$ is assumed to be equally likely at bid or ask.)

In this model we assume that the size of spread is constant at each point in time. Based on the transaction process depicted by Figure 1, a more generalized model can be obtained. Assuming the initial trade at $t-1$ occurs with equal probability at either the bid or the ask, the new joint probability distribution of successive transaction price changes is shown in Table 1. The serial covariance of transaction price changes is: $COV(\Delta P_t, \Delta P_{t-1}) = \lambda^2 S^2 (1-\pi)(1-\alpha) - \lambda S(1-\lambda)S(1-\pi)\alpha + \lambda S(1-\lambda)S\pi(1-\beta) - (1-\lambda)^2 S^2 \pi\beta = S^2(\lambda^2(1-2\pi) + \lambda\pi(\alpha + \beta) + \lambda(\pi - \alpha) - \pi\beta)$.

The effective spread, S , can be derived as:

$$S = \sqrt{\frac{COV(\Delta P_t, \Delta P_{t+1})}{\lambda^2(1-2\pi) + \lambda\pi(\alpha + \beta) + (\lambda\pi - \lambda\alpha - \lambda\beta)}} \tag{7}$$

If the price reversal always equals the spread, i.e., $\lambda = 0$ and $\pi = \alpha = \beta = 0.5$, then (7) simply reduces to the Roll's serial covariance model. If $\pi = \alpha = \beta$, it reduces to Stoll's model. If there is no price reversal effect ($\lambda = 0$), then it reduces to the estimator developed by Chu et al., if there is no price reversal effect and $\beta = \pi$, then it reduces to Choi et al.'s model. Therefore, our proposed model will include Roll's and other more recent estimators as special cases. Our model always generates a spread estimator that is positive. This is not the case for models that rely on negative covariances to generate a positive estimator.

Table 1. The New Combined Joint Distribution of Successive Price Changes

		ΔP_t			
		Initial Trade at Bid		Initial Trade at Ask	
		$-\lambda S(BB)$	$(1-\lambda)S(BA)$	$\lambda S(AA)$	$-(1-\lambda)S(AB)$
ΔP_{t+1}	$-\lambda S(BB)$	$(1-\pi)(1-\alpha)$	0	0	$\pi(1-\beta)$
	$(1-\lambda)S(BA)$	$(1-\pi)\alpha$	0	0	$\pi\beta$
	$\lambda S(AA)$	0	$\pi(1-\beta)$	$(1-\pi)(1-\alpha)$	0
	$-(1-\lambda)S(AB)$	0	$\pi\beta$	$(1-\pi)\alpha$	0

3. Estimation Procedure

Under the serial covariance model, only the transaction price data are employed instead of the quoted bid-ask prices. Therefore, we need a classification procedure to identify each time series transaction prices as either a bid or an ask. The tick test discussed by Lee and Ready (1991) as well as Aitken and Frino (1996) can be used for this purpose. The tick test classifies a trade as an ask (market buyer initiated) if it occurs on an uptick or zero-uptick; otherwise it is identified as a bid (market seller initiated). An uptick (downtick) is defined as if the current price is higher (lower) than the price of the previous transaction. When the current price is the same as the previous trade (a zero tick), if the last price change was an uptick (downtick), then the trade is a zero-uptick (zero-downtick). Lee and Ready provide the empirical support by claiming a high accuracy rate as 90%. Aitken and Frino report a lower overall accuracy rate as 74%.

In addition to the trade type, from (7), there are also several input parameters we need to estimate before computing the serial covariance spread. These parameters can be estimated by the maximum likelihood estimation technique. For the estimation of the conditional probability π , the likelihood function can be specified as:

$$L(\pi) = \prod_{t=1}^n [\pi I_t + (1-\pi)(1-I_t)], \quad (8)$$

where n is the total number of observations used, $I_t = 1$ if a trade in time t is different from the trade type in time $t-1$, and $I_t = 0$ otherwise. The maximum likelihood estimator (MLE) of π is derived by finding the value of that maximizes the following log likelihood function:

$$\begin{aligned} \ln L(\pi) &= \ln \left\{ \prod_{t=1}^n [\pi I_t + (1-\pi)(1-I_t)] \right\} \\ &= \ln[\pi I_1 + (1-\pi)(1-I_1)] + \dots + \ln[\pi I_n + (1-\pi)(1-I_n)]. \end{aligned}$$

Therefore, the resulting MLE estimator of π is:

$$\pi = \frac{\sum_{t=1}^n I_t}{n}. \quad (9)$$

Similarly, the conditional probabilities α and β are estimated by the MLE as shown by Chu et al.:

$$\begin{aligned} \hat{\alpha} &= \frac{N(A_{t-1}, A_t, B_{t+1}) + N(B_{t-1}, B_t, A_{t+1})}{N(A_{t-1}, A_t, B_{t-1}) + N(A_{t-1}, A_t, A_{t+1}) + N(B_{t-1}, B_t, A_{t+1}) + N(B_{t-1}, B_t, B_{t+1})}, \\ \hat{\beta} &= \frac{N(A_{t-1}, B_t, A_{t+1}) + N(B_{t-1}, A_t, B_{t+1})}{N(A_{t-1}, B_t, A_{t+1}) + N(A_{t-1}, B_t, B_{t+1}) + N(B_{t-1}, A_t, B_{t+1}) + N(B_{t-1}, A_t, A_{t+1})}, \end{aligned} \quad (10)$$

where N is the number of occurrences with respect to a particular event described in the parenthesis. From the properties of the maximum likelihood function, $\hat{\pi}$, $\hat{\alpha}$, and $\hat{\beta}$ are unbiased, efficient, and consistent estimators of π , α , and β , respectively.

Since this study produces an improved version of the previous serial covariance spread estimators, in addition to being a more accurate measure of the serial correlation relationship among trade types, the magnitude of price reversal, $1-\lambda$, plays a very important role in order to infer the different cost components in the bid-ask spread. Recall that if either the inventory adjusting costs or asymmetric information costs are present, the magnitude of price reversal is no longer equal to the spread, S , and thus λ is not constrained to zero.

Here the magnitude of price reversal, $1-\lambda$, is expressed as a percentage of the quoted spread and thus the magnitude of price continuation, λ , must be estimated first. Mathematically λ can be expressed as:

$$\lambda = \frac{\Delta P_{t+1} | P_t = A_t, P_{t+1} = A_{t+1}}{S} = \frac{\Delta P_{t+1} | P_t = B_t, P_{t+1} = B_{t+1}}{S}, \quad (11)$$

where P_{t+1} denotes the transaction price at time $t+1$ and thus $\Delta P_{t+1} = P_{t+1} - P_t$ is the price difference between two periods conditioning upon whether the transaction prices at times t and $t+1$ are of the same transaction type.

So we can use the following formula to estimate $\hat{\lambda}$:

$$\hat{\lambda} = \frac{\sum_{n=1}^h \Delta P_n - \sum_{n=1}^k \Delta P_n}{h(P_t = A_t, P_{t+1} = A_{t+1}) + k(P_t = B_t, P_{t+1} = B_{t+1})} \times \frac{1}{S}, \quad (12)$$

where h (k) is the number of occurrences when the transactions at both time $t+1$ and time t are an ask (a bid); ΔP denotes the corresponding price change. To obtain S , we need the quoted bid and ask price data to calculate it as the average of absolute quoted spreads.

4. Conclusions

In this paper we present a more generalized dealer spread pricing model in which the existing spread models are unified and three cost components of spread—order processing, adverse information, and inventory holding costs—are considered. In particular, this model is extended in two ways from the previous models to estimate the bid-ask spread. The first extension relies on the fact that only considering the serial correlation of transaction types cannot capture all components of the spread. The second extension improves the Stoll's approach, which is the first to mathematically decompose the three components of the spread by taking into account not only the magnitude of price reversal but also a two-period serial correlation of transaction types.

This new statistical spread estimator can be applied to (1) more accurately estimate the effective spread [by using the formula in (7)] and (2) improve the results of Stoll among others in decomposing three cost components of the spread. Many other applications can also be found in a number of ways, such as to determine the sources of spread variations during the day, to compare between dealer and auction markets, to compare across trade sizes, to compare markets with different geographical locations, or to assess the effect of a particular event on the spread component.

References

- Aitken, M. and A. Frino, (1996), "The Accuracy of the Tick Test: Evidence from the Australian Stock Exchange," *Journal of Banking and Finance*, 20, 1715-1729.
- Brooks, R. and J. Masson, (1996), "Performance of Stoll's Spread Component Estimator: Evidence from Simulations, Time Series and Cross-Sectional Data," *Journal of Financial Research*, 19, 459-476.
- Choi, J. Y., D. Salandro, and K. Shastri, (1988), "On the Estimation of Bid-Ask Spread: Theory and Evidence," *Journal of Financial and Quantitative Analysis*, 23, 219-229.
- Chu, Q. C., D. K. Ding, and C. C. Pyun, (1996), "Bid-Ask Bounce and Spread in the Foreign Exchange Futures Market," *Review of Quantitative Finance and Accounting*, 6, 19-37.
- Garbade, K. and Z. Lieber, (1977), "On the Independence of Transactions on the New York Stock Exchange," *Journal of Banking and Finance*, 1, 151-172.
- George, T., G. Kaul, and M. Nimalendran, (1991), "Estimating Components of the Bid-Ask Spread and Its Components: A New Approach," *Review of Financial Studies*, 4, 622-656.
- Lee, C. M. C. and M. J. Ready, (1991), "Inferring Trade Direction from Intraday Data," *Journal of Finance*, 46, 733-746.
- Lesmond, D. A., J. P. Ogden, and C. A. Trzcinka, (1999), "A New Estimate of Transactions Costs," *Review of Financial Studies*, 12, 1113-1141.
- Roll, R., (1984), "A Simple Implicit Measure of the Effective Bid-Ask Spread in an

- Efficient Market,” *Journal of Finance*, 39, 1127-1139.
- Stoll, H. R., (1989), “Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests,” *Journal of Finance*, 44, 115-134.