14

TAPMI

Working Paper Series

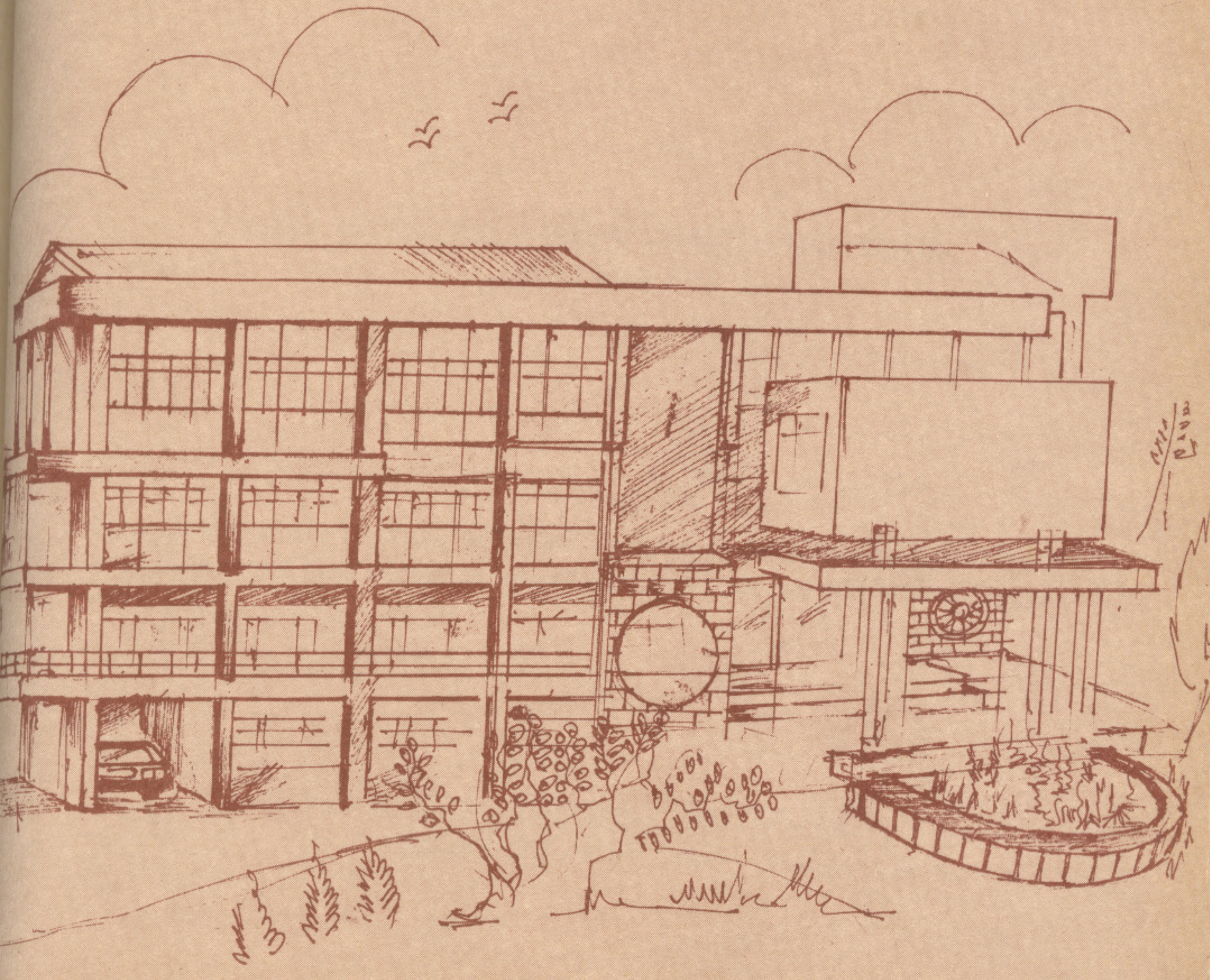
AN ORDER PROCESSING MODEL OF THE BID-ASK SPREAD: THE INDIAN EVIDENCE

By

SUNIL K.PARAMESWARAN
Professor

and

TUHIN S. BANERJEE
Faculty Associate



AN ORDER PROCESSING MODEL OF THE BID-ASK SPREAD: THE INDIAN EVIDENCE

By

SUNIL K.PARAMESWARAN
Professor

and

TUHIN S. BANERJEE
Faculty Associate

TAPMI WORKING PAPER SERIES NO. 98/04
March 1998

The objective of TAPMI Working Paper series is to help faculty members of TAPMI to test out their research ideas/findings at the pre-publication stage.

T. A. PAI MANAGEMENT INSTITUTE
MANIPAL - 576 119

An Order Processing Model of the Bid-Ask Spread:

The Indian Evidence

Sunil K. Parameswaran

and

Tuhin S. Banerjee

Introduction:

The quoted bid-ask spread is defined as the difference between the price at which a dealer is willing to sell the security, the 'ask', and the price at which he is willing to buy it, the 'bid'. There are three major theories which purport to explain the existence of the spread. These are, order processing cost models, inventory cost models, and adverse selection models. Stoll (1989) develops a unified framework for studying the relative importance of these, and shows that the differences between the models lie in their predictions regarding the probability of a price reversal, and the magnitude of the price reversal.[1] Stoll also makes a distinction between the quoted spread, and the effective or realized spread. The realized spread is defined as the difference betweeen the price at which the dealer sells a security, and the price at which he buys at an earlier

[1] Foster and Viswanathan (1990) also study the three components of the spread simultaneously. The major focus of their paper is the adverse selection cost.

instant. The order processing model, which is the focus of this paper, predicts that the two spreads are equal.

The simplest of the three models are the order processing, or fixed cost models. One of the most popular models in this class was developed by Roll (1984), and later extended by Choi, Salandro and Shastri (1988), hereafter referred to as CSS. Its premise is that the spread compensates the dealer for the costs incurred in processing buy and sell transactions. In this framework, transactions at the bid are offset by transactions at the ask, and the difference between the two is used by the dealer to defray his expenses.

The literature on the estimation of the realized spread is fairly diverse. Roll (1984) estimates realized spreads in an order processing framework using daily and weekly data and attempts to relate them to firm size. CSS (1988) use transaction data for stock options to estimate the spread, and the conditional probability of a price continuation. Stoll (1989) estimates the relative components of the spread using transaction prices and bid-ask quotes. Hasbrouck and Ho (1987) do not estimate spreads, but use transaction data and intraday quotes to show that observed stock returns may best be described by an ARMA(2,2) process. In addition to the spread, their model incorporates a delayed adjustment mechanism for observed prices. The partial adjustment model for the evolution of asset prices is also developed by Amihud and Mendelson (1987), who study the impact of the trading mechanism on the return generating process.

Parameswaran (1991) studied the Roll model as well as the partial adjustment model developed by Hasbrouck and Ho using intraday transaction prices and bid-ask quotes.

The objective of this paper is to study the order processing model using data from the Indian stock market. We carry out this exercise under the assumption that observed prices instantly adjust to fully reflect new information. This is the setting in which Roll develops his model. By focusing solely on the fixed cost model, we are ignoring the other two components of the spread, namely inventory and adverse selection costs. However, this does not imply that these costs are insignificant, although empirical research has shown that the fixed cost component constitutes the bulk of the spread.[2]

Our objective is to test and see whether a pure order processing model is an adequate first approximation as a theory of the spread. The order processing model is often used by researchers who are forced to take the bid-ask spread into account as a 'nuisance' parameter, although they are not interested in the spread per se. If we were to find that the pure fixed cost model is an appropriate first approximation, then, it will be a vindication of the approach taken by these researchers.

We use daily closing prices from the Mumbai stock exchange, for the thirty companies constituting the Sensex, for the period January 1996 to October 1997. The data for each company is divided into twenty two monthly subsamples.

[2] See Foster and Viswanathan (1990).

In the second section, we discuss the order processing cost model. The notation used for the Roll model is identical to that used by Harris (1989b). In this context, we also briefly examine an extension of the original model, as proposed by CSS.

The order processing model was developed under the following set of assumptions.

1. Security prices at any given time, fully incorporate all relevant available information. What this means is that the price that would be observed in the absence of market imperfections accurately reflects the intrinsic value of the asset.

2. Buy or sell orders are equally probable at any instant in time, on an unconditional basis, and order flows are serially independent.

3. The underlying price of the asset is independent of the order flow in the market. This implicitly rules out any adverse information effects.

4. The true value of the security is bracketed by the bid-ask spread and the spread is assumed to be constant and symmetric[3], at least for the time period for which the analysis is conducted.

Mathematically, the model may be stated as follows.

$$P_t = P^*_t + \frac{s}{2} Q_t$$

---

[3] Roll (1984) has shown that the size of the spread can be allowed to depend on the magnitude of the price changes, without changing the structure of the model. If so, the estimated realized spread is the average spread over the sample period.

and,

$$P^*_t = P^*_{t-1} + \mu + \varepsilon_t ,$$

where P_t is the observed price at time t, P^*_t is the intrinsic value of the asset at time t, which would have been observed in the absence of the spread, and Q_t is the indicator of the transaction type. $Q_t = 1$, if the observed price is at the ask, and -1, if it is at the bid. The value innovations, $\{\varepsilon\}$ are assumed to be serially uncorrelated, and have a zero mean and constant variance σ^2_ε.[4] The drift in the true prices μ is assumed to be constant over time, as is the spread s.

With these assumptions, the covariance of the observed price changes equals minus the square of half the bid-ask spread, i.e., $Cov(\Delta P_t , \Delta P_{t-1}) = -s^2 / 4$, and the estimates of the covariance can be transformed to get the spreads.[5]

The CSS Extension:

CSS generalize the model by partially relaxing the second assumption made above. In their model, although the unconditional probability of a bid or an ask is the same at a particular point in time, conditional on a bid, the probability of a consecutive bid is greater than a half, and likewise for the ask. In other words, transaction types are positively serially correlated. A major difference between the two models, is that in the CSS framework, the interval at which successive transactions are measured is

---

[4] Researchers often assume that the $\{\varepsilon\}$s are i.i.d. This, however, is not necessary for our study.

[5] The resulting estimate is subject to underestimation due to Jensen's inequality, because the transformation uses a concave function.

important, while it is irrelevant for the Roll model. If stocks are assumed to trade at fixed intervals, it can be shown that the longer the measurement interval, the better the fit of the Roll model, which assumes that transaction types are serially uncorrelated.[6]

There are two reasons as to why transaction types may be serially correlated in practice. Firstly, large market orders may be split up by floor brokers, thereby ensuring that successive transactions take place on the same side of the market. Similarly, a movement in the price will cause limit orders on the same side of the order book to be executed in succession.

A Survey of Empirical Tests:

Roll (1984) uses daily and weekly data for all actively traded stocks on the CRSP data base for the years 1963-82. Serial covariances are calculated for every year in this period, and the estimates are transformed to get the realized spreads. Not having access to the actual quotes, he attempts to indirectly verify the model by regressing the realized spread on the log of firm size. The underlying hypothesis is that firm size is positively correlated with share volume, which in turn is negatively correlated with the size of the spread. His results support this hypothesis. However, Roll finds some disconcerting features. First, the serial covariances computed from daily returns are often positive, and, in many cases, the same is true for covariances computed from weekly returns. Second, the average of the spread computed from daily returns is much smaller

---

[6] See Appendix A. These results were derived by Parameswaran (1991).

than the mean estimate obtained from weekly returns, which violates one of the predictions of the model.

CSS (1988) use transaction data and intraday quotes for options traded on the CBOE for the period August 1977 to August 1978. The seial covariances obtained by them are negative for all the option contracts they study. The estimated spreads are then regressed on the average quoted spreads, and are found to have significant explanatory power. However, the means of the estimated spreads are lower than those of the quoted spreads.

Using NASDAQ data for three periods of one month each in 1984, Stoll computes the serial covariances of prices and bid-ask quotes. He is then able to estimate the two parameters of interest, the magnitude and the probability of a price reversal, and thereby infer the relative components of the spread. Stoll's results indicate that the order processing component is about 47% of the quoted spread.

Harris (1989b) derives the population moments of the finite sample estimates of the variance and serial covariance of stock price changes. His results show that the serial covariance is negatively biased for small sample sizes and large observation intervals, i.e. periods for which the variance of price innovations is large. This negative bias can lead to overestimation of the spread. Moreover, the variance of the serial covariance estimate is so large in small samples that it alone can account for the large frequencies of positive covariances found by Roll in daily and weekly data. The size of the variance can also cause underestimation of the spread due to Jensen's inequality.

Parameswaran (1991) tests the Roll as well as the Hasbrouck and Ho models, using the Generalized Method of Moments estimation procedure, with data consisting of intraday transaction prices and bid-ask quotations obtained from the Institute for the study of Securities Markets (ISSM), for the year 1988, for the thirty companies constituting the Dow Jones Industrial Average. Using data measured at half hourly intervals he finds that the Roll model is rejected for 26.67% of the total 360 sample months, whereas the Hasbrouck and Ho model is rejected for 12.5% of the sample months.

Results:

The model is rejected for 374 out of 660 sample months, because the covariances are positive. The rejection rate is 56.67%. The average spread for the other months varies from Rs. 0.54 for SAIL to Rs. 30.31 for ACC. Table 1 shows the results companywise. It depicts the number of months for which the model is rejected, the minimum realized spread, the maximum realized spread, and the average realized spread.

From these results we conclude that the Roll model does not fit the data very well. Not only are a large number of covariances positive, the realized spread displays significant variation from month to month.

Avenues for Further Research:

Harris (1989b) demonstrates that for studies which use daily or weekly price observations to infer spreads, the small samples used ensure that the serial covariance estimates from which the spreads are computed are negatively biased and have large standard errors. Depending on the relative magnitudes of the spread and the variance of the innovations in the underlying price series, these problems can cause the estimates of the spread to be either severely biased upward or downward. He therefore concludes that for empirical studies of this nature, it is imperative that transaction data be used. The advantage with transaction data is that the overall sample can be split into many subsamples without having to invoke the stationarity assumptions over long periods of time. In addition to the issues discussed earlier, there are other major problems when daily data is used to estimate serial covariances. Amihud and Mendelson (1987), and Smith (1989), find that the covariance of daily returns is dependent on the time of the day at which returns are measured. They show that returns measured at the beginning of the day are negatively correlated, while those measured at the end of the day are often positively correlated. Also Harris (1989a) and others have found that closing prices tend to be at the ask quote more frequently than at the bid. Since most samples of daily data use closing prices, it is less likely that one can detect a significant bid-ask bounce. It will be interesting to conduct similar studies for Indian markets.

Hence we conclude that the Roll model can be better tested using intraday prices and bid-ask quotes. Extensions of the study could include testing of the partial adjustment/ overreaction model, and the testing of models which account for all the components of the spread simultaneously.

References

1. Amihud Y. and Mendelson H., 1987, Trading mechanisms and stock returns: An empirical investigation, Journal of Finance, 533-553.

2. Choi J.Y., Salandro D. and Shastri K., 1988, On the estimation of bid-ask spreads: Theory and evidence, Journal of Financial and Quantitative Analysis, 219-230.

3. Foster D.F. and Viswanathan S., 1990, Variations in volumes, variances and trading costs, Unpublished Manuscript, Duke University.

4. Harris L., 1989a, A day-end transaction price anomaly, Journal of Financial and Quantitative Analysis.

5. Harris L., 1989b, The power of the Roll serial covariance bid-ask spread estimator, Unpublished Manuscript.

6. Hasbrouck J. and Ho T.S.Y., 1987, Order arrival, quote behaviour, and the return generating process, Journal of Finance, 1035-1048.

7. Parameswaran S.K., 1991, Testing order processing models of the bid-ask spread: A method of moments approach, Unpublished Manuscript, Duke University.

8. Roll R., 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, Journal of Finance, 1127-1139.

9. Smith T., 1989, Financial models and market microstructure effects, Unpublished Manuscript, Duke University.

10. Stoll H., Inferring the components of the bid-ask spread: Theory and empirical tests, Journal of Finance.

COMPANY	NUMBER OF MONTHS REJECTED	MINIMUM REALIZED SPREAD	MAXIMUM REALIZED SPREAD	AVERAGE REALIZED SPREAD
ARVIND	14	0.84	4.99	2.79
ACC	11	11.41	54.47	30.31
BSES	9	1.20	10.21	3.40
BAJ. AUTO	15	6.65	11.77	8.67
BHEL	15	0.94	4.32	3.03
COL.PALM.	12	1.89	6.97	3.71
GLAXO	11	0.94	7.23	2.87
GRASIM	14	2.22	11.55	6.29
GE. SHIP.	8	0.08	1.52	0.58
G.A.CEM.	15	3.08	9.44	5.07
HINDALCO	13	2.72	21.38	9.49
HLL	16	6.90	25.52	11.75
HPCL	15	1.99	15.04	6.63
ITC	10	1.89	15.87	7.48
IND.HOTEL	18	5.06	17.95	10.56
IPCL	8	0.70	4.35	2.64
ICICI	10	0.82	2.90	1.93
IDBI	11	0.56	3.27	1.58
L&T	11	1.71	9.06	5.05
MTNL	10	1.33	4.60	2.90
M&M	18	4.01	7.15	5.31
NESTLE	13	1.94	8.28	3.88
RANBAXY	13	2.38	13.21	6.41
RIL	9	1.00	7.65	4.35
SBI	13	2.09	9.84	6.31
SAIL	12	0.24	0.95	0.54
TAT.CHEM	16	1.18	2.58	1.84
TELCO	13	2.33	10.38	6.66
TISCO	11	0.70	6.56	4.03
TAT.POW	12	1.03	7.50	3.18
TOTAL	374			

Appendix I

Serial Covariance of Price Changes with Serially Correlated Transaction Types

We derive an expression for the spread, when 'δ', the conditional probability of a bid following a bid, or of an ask following an ask, is not equal to 1/2. The covariance, in this case, depends on the length of the measurement interval. More specifically, it depends on 'N', where 'N' is the number of transactions that have occurred between t-1 and t. For the purpose of this analysis, we will assume that transactions occur at fixed intervals in time, i.e., N is fixed for a given measurement interval.

a) N = 1: In this case, the joint probability distribution can be shown to be,

		ΔP_t		
		-s	0	s
ΔP_{t-1}	-s	0	$\delta(1-\delta)/2$	$(1-\delta)^2/2$
	0	$\delta(1-\delta)/2$	δ^2	$\delta(1-\delta)/2$
	s	$(1-\delta)^2/2$	$\delta(1-\delta)/2$	0

$$Cov(\Delta P_t, \Delta P_{t-1}) = -s^2 (1 - \delta)^2 .$$

From the above distribution, two main features of the covariance are obvious.

1) The only two entries which affect the covariance are (s,-s) and (-s,s).

2) Because of the symmetric nature of the model, the two probabilities are equal. Hence, we need only focus on the probability of the sequence bid-ask-bid, in order to derive the covariance.

In general,

$$\mathrm{Cov}(\Delta P_t, \Delta P_{t-1}) = -s^2 \left[\sum_{k=1}^{N} C^N_{2k-1} \delta^{N-(2k-1)} (1-\delta)^{2k-1}\right]^2,$$

where $C^a_b = 0$ if $a<b$, else $C^a_b = \dfrac{a!}{(a-b)!b!}$.

For a given value of $\mathrm{Cov}(\Delta P_t, \Delta P_{t-1})$, we will now compare the values of the spread obtained using Roll's formula and the exact formula.

	$\delta = .60$	
N	Roll	Exact
1	$2(-c)^{1/2}$	$2.5(-c)^{1/2}$
2	$2(-c)^{1/2}$	$2.083(-c)^{1/2}$
3	$2(-c)^{1/2}$	$2.016(-c)^{1/2}$
4	$2(-c)^{1/2}$	$2.003(-c)^{1/2}$
5	$2(-c)^{1/2}$	$2.001(-c)^{1/2}$
10	$2(-c)^{1/2}$	$2.000(-c)^{1/2}$

.	$\delta = .75$	
N	Roll	Exact
1	$2(-c)^{1/2}$	$4(-c)^{1/2}$
2	$2(-c)^{1/2}$	$2.667(-c)^{1/2}$
3	$2(-c)^{1/2}$	$2.286(-c)^{1/2}$
4	$2(-c)^{1/2}$	$2.133(-c)^{1/2}$
5	$2(-c)^{1/2}$	$2.065(-c)^{1/2}$
10	$2(-c)^{1/2}$	$2.002(-c)^{1/2}$
.	$\delta = .90$	
N	Roll	Exact
1	$2(-c)^{1/2}$	$10(-c)^{1/2}$
2	$2(-c)^{1/2}$	$5.556(-c)^{1/2}$
3	$2(-c)^{1/2}$	$4.098(-c)^{1/2}$
4	$2(-c)^{1/2}$	$3.388(-c)^{1/2}$
5	$2(-c)^{1/2}$	$2.445(-c)^{1/2}$
10	$2(-c)^{1/2}$	$2.241(-c)^{1/2}$

As we can see, the exact formula converges to Roll's formula as N increases. Thus, even for a high value of δ, like .75, if the stock trades every minute, Roll's formula will be fairly accurate if 5 minute price changes are used.