Estimation of Bid-Ask Spread and its Components in Indian Stock Market Using Trade Prices

Priyanka Singh[1]

Ajay Pandey[2]

Abstract

In the absence of order-book data and limited information on quoted bid-ask spreads in the Indian stock market, this paper attempts to analyze the bid-ask spread in Indian market by estimating bid-ask spreads and its components from trade prices. The sample consists of tick-by-tick data for the time period January 2002 through to October 2008 of 160 stocks traded on the National Stock Exchange of India. We estimate implied bid-ask spreads and its components (adverse selection costs; combined inventory and order processing costs) using theoretical models. We find that all the models used in the study produce consistent estimates of bid-ask spreads and its components. In the Indian Stock Market, we find that the adverse selection cost and the combined order-processing and inventory-holding cost each account for approximately 50 percent of the bid-ask spread. We also find that the estimated bid-ask spreads are approximately 80 percent of the quoted bid-ask spreads. In our sample period, we find that the relative bid-ask spreads have decreased over the years.

Keywords: Bid-ask spread, adverse selection costs, inventory holding costs, order processing costs, order driven market

[1] Fellow, Finance & Accounting, Indian Institute of Management Ahmedabad, email: priyankas@iimahd.ernet.in
[2] Professor, Finance & Accounting, Indian Institute of Management Ahmedabad, email: apandey@iimahd.ernet.in

Estimation of Bid-Ask Spread and its Components in Indian Stock Market using Trade Prices

1. Introduction

Post 2008 crisis, the role of liquidity in financial markets has attracted renewed attention from regulators, exchanges and investors. Liquidity is the ability to access funds when needed at a minimal cost. For institutions, it is the ability to raise funds easily for normal activities and for investors; it is the ease with which an asset can be traded at an exchange. The former is called *funding liquidity* and the latter is called *market liquidity*. Liquidity reflects the financial health of any market. In practice, management of liquidity is an important concern in portfolio management, risk management, trading strategies etc. Understanding the liquidity features of a market has been an important strand of research for a long time. The research has tried to understand the aspects of liquidity and various transaction costs, which drive the liquidity of an asset market. Bid-ask spread is one of the measures of market liquidity (Bervas, 2006).The bid-ask spread is also an important indicator of the financial stability of the stock market. Its sudden widening is a cause of concern and may generate systemic crisis. As bid-ask spread reduces the return on investment, all investors would like to face minimum bid-ask spread. For the firms, the cost of capital depends on the transaction costs, of which bid-ask spread is an important component. As bid-ask spread represents the market liquidity, the determinants of bid-ask spread and its components have been examined extensively theoretically as well as empirically. Bid-ask spread arises as the compensation for various costs borne by the liquidity providers. These costs are identified to be- order processing, inventory holding and adverse selection.

Markets are typically organized either as an order-driven market or as a quote-driven market. Quote driven market has a third party or a dealer providing liquidity, whereas the order-driven market does not have one. Order-driven markets are the markets where liquidity is provided by the limit orders. Microstructure literature suggests that there are two economic forces that drive trading: liquidity events and information events (Bagehot, 1971). A liquidity event is an exogenous random event leading to trading by the liquidity trader and is specific to that trader. On the other hand, an information event is a result of some new news that affects the asset value. This information can be public or private. However, often the traders do not receive the news simultaneously, particularly if the information is private, and therefore face the risk of trading with a better informed trader. This adverse selection risk is borne by market-makers in the quote-driven market and limit-order traders in the order driven market (Handa and Schwartz, 1996b). In addition, the trader also faces non execution risk of his order in an order-driven market.

Although there are numerous studies, which estimate the bid-ask spread in quote-driven markets especially the markets in US, yet relatively little is known about the behavior of bid-ask spread in an emerging order-driven market. Indian Stock Markets, dominated by the National Stock Exchange and the Bombay Stock Exchange, are order-driven markets. This study examines the behavior of the bid-ask spread as well as its components in the Indian Stock market using data from the National Stock Exchange. The purpose of this study to examine and estimate the bid-ask spreads and its components for the liquid stocks traded on Indian stock market. Like some other markets, Indian stock exchanges did not store order-book and quoted bid-ask spread. In such a case, the bid-ask spread, has to be inferred or estimated from the trade prices. We estimate bid-ask spreads using tick-by-tick transaction data as suggested by various estimation models proposed and compare the estimates. We find that the bid-ask spread of liquid stocks in our sample is of the order of 0.4-0.7% over the years. This is comparable to markets around the world. The proportion of adverse selection costs in the bid-ask spread is approximately 50 %, which is consistent with the findings elsewhere for the order driven market (Ahn et al., 2002). We also find that over the years, the relative spread has decreased, rising only in the year 2008. 2008 being the year of financial crisis, the rise in the relative spread is not completely unexpected. The paper contributes to the existing literature in several ways. Firstly, we consider high frequency data for a reasonably period of six years and ten months (sample period: Jan. 2002 to Oct. 2008) unlike most other studies, which are based on shorter sample periods. Secondly, we do the estimation of bid-ask spread and its components in the context of an emerging order-driven market like India, where such estimation on a large data set has not yet been undertaken. Thirdly, we compare five different models of bid-ask spread and in the process check not only the robustness of our results but the consistency of models suggested for estimating bid-ask spread and its components .

The rest of the paper is organized as follows. Section 2 presents the review of literature on the models proposed for estimating bid-ask spread as well as it components from trade prices. This section also explains the methodology used by us to estimate the bid-ask spread and its components. The institutional details on the National Stock Exchange and on the data and its descriptive statistics are provided in section 3. Findings of the study are given in section 4. Last section concludes the paper.

2. Estimating Bid-Ask Spreads and its Components

In this section, we briefly outline the various components of bid-ask spread. Subsequently, we describe some of the models that have been proposed for estimating bid-ask spread and its components. Finally, we discuss the empirical findings on estimated bid-ask spread as well as its three components in other markets.

2.1 Components of the Bid-Ask Spread

The bid-ask spread has been argued to have three components: order processing, inventory holding costs, and adverse selection costs. Out of these, the inventory holding costs are expected to be minimal in an order-driven market as discussed later. The modeling of bid-ask spread started with focus on order processing costs (Tinic, 1972; Roll, 1984), followed by incorporating the inventory holding costs (Ho and Stoll, 1983; Ho and Macris, 1984) and finally, adverse selection costs (Copeland and Galai, 1983; Admati and Pfleiderer, 1988).

2.1.1 Order Processing Costs

Order processing costs are those costs that are directly related to the trading. In case of the dealer markets, these will be the costs involved in making the market like exchange seat, floor space rent, computer costs, labor costs, opportunity of time involved, etc. These costs are largely fixed. In an order-driven market, these costs will involve the computer costs, time involved (limit order traders' waiting time), etc. Additionally a major component of order processing costs will be order execution costs. Limit order traders in an order-driven market face the risk of non-execution of the order and incur associated costs.

2.1.2 Inventory Holding Costs

In a quote-driven market, the market maker's work is to act as counterparty to the incoming trade in the dealer market. In light of this role, he is faced with the uncertainty of the order flow, both in terms of quantity of the order and the time of the order. This forces him to hold a portfolio, which is not optimal, which he could have otherwise held as an investor. He bears some cost in maintaining this sub-optimal inventory which is classified as the inventory holding costs. The dealer will factor in this cost while quoting bid and ask prices. There are several models like Ho and Stoll, which analyses this problem faced by the dealer. They show that dealer tackles this problem through two variables, 'spread' and 'price adjustment factor'. Dealer controls the flow by adjusting these two variables. If he has positive inventory, he would want to encourage sales and hence will lower the selling price. On the other hand, if the inventory is negative, he will encourage purchase and so will increase the bid price. He does so through the spread and demands higher spread in the case of an order imbalance.

2.1.3 Adverse Selection Costs

There are traders in the market with different information levels. Different traders receive information at different points in time and at various levels of precision and consequently some are better informed than the others. Dealers in the quote driven market and the limit order traders in the order driven market are faced with the risk of trading with a better informed trader. A better informed trader will buy when the stock is undervalued and sell when it is overvalued. In both cases, the liquidity providers lose and faced with this risk, they demand some compensation. The adverse selection component in the bid-ask spread is the expected asymmetric information costs. This type of costs arises only in case of an informational event and was first suggested by Bagehot (1971). The first attempt to formalize a model of bid-ask spread incorporating this cost was Copeland and Galai (1983).

The effect of order processing costs and inventory holding costs on price is transitory in nature as it not information driven. These costs cause the price to mean revert or price changes to be negatively serially correlated. However, the expected adverse selection costs are relatively permanent in nature. The effect on prices does not revert in future. This is due to the informational nature of the adverse selection costs. The models used in this study, which are based on trade data, split the bid-ask spread into only two components: adverse selection costs and combined order processing and inventory holding costs. However, inventory holding costs is expected to be minimal for order-driven markets and therefore, not too important in an order driven market like NSE. Even though recently many new markets have adopted the order driven system, the literature on modeling behavior of bid-ask spreads is relatively sparse on such markets. Empirically, we expect to see negligible inventory holding costs in an order driven market. Instead, the order processing costs should be the major part of the trading costs. Due to the greater transparency of the order book and more information available to the investors, adverse selection costs should also be less than in quote driven markets.

2.2 Models for Estimating the Bid-Ask Spread

In many cases, quote data on bid-ask spread is not available and has been estimated based on the transaction prices in empirical research. In most cases, the transaction prices used are tick-by-tick trade data. In this section, five models estimating the bid-ask spread using the trade data are discussed. These models are Roll (1984), Glosten and Milgrom (1985), Glosten and Harris (1988), De Jong, Nijman and Roell (1996) and Madhavan, Richardson and Roomans (1997). Models for estimating bid-ask spread based on trade data are based on certain basic assumptions. Firstly, it is assumed that the true unobserved price depends only on information and not noise in the market. The observed trade price depends on the true unobserved price and other frictions in the market. This can also be viewed in a different manner. True unobservable prices depend only on permanent effects but the observed price depend on temporary effects also. Secondly, as is known, bid-ask spread comprises of order processing cost, inventory holding cost, and adverse selection cost. Of these costs, adverse selection cost is permanent, i.e., its effect will not reverse over time. Hence, it will affect the true unobservable price. The other two, order processing cost and inventory holding cost do not have a permanent effect on the prices, as these are noises or frictions of the market and do not carry any fundamental information. Therefore, these costs will affect the observed trade price. This is the basic structure of the trade based models for estimating bid-ask spread. In addition to this, different models use different variables to extract the components of the bid-ask spread.

Model 1: Roll (1984)

Roll's model is the simplest model to estimate the bid ask spread using transaction prices. It assumes that information is homogenous; orders are executed at the best bid price or at the best ask price, trading process itself has no impact on the equilibrium price, the probabilities of buying and selling are equal and probability of continuation is same as reversal. Essentially, Roll's estimator assumes only the presence of order processing cost in the spread. Since, it is a non-information costs, it will not have a permanent impact on the prices. Therefore this cost will lead to price reversal and negative covariance between successive price changes. Roll employs this property to estimate the bid-ask spread.

In presence of transaction costs, assuming a constant spread S, whose midpoint is assumed to be equal to the equilibrium price m_t , the transaction price can be written as:

$$P_t = m_t + \frac{S}{2} Q_t \qquad (1)$$

$$m_t = m_{t-1} + U_t \qquad (2)$$

where Q_t is a dummy variable whose value is + 1 for a buyer initiated transaction and -1 for a seller initiated transaction and U_t is the public information.

Using the above equations, the auto-covariance of the price change is given by

$$Cov(\Delta P_t, \Delta P_{t-1}) = -\frac{S^2}{4} < 0 \qquad (3)$$

Hence,

$$S_{roll} = 2\sqrt{-Cov(\Delta P_t, \Delta P_{t-1})} \qquad (4)$$

Model 2: Glosten and Milgrom (1985)

In the absence of private information, prices follow a random walk independent of the trading process. However, in the presence of asymmetric information, trading process has an impact on the equilibrium price. In other words, unobservable true price will also have information asymmetry cost. As is known, the bid-ask spread has both permanent cost (information asymmetry) and temporary cost (order processing and inventory holding cost), the part that will affect true price is the permanent cost. Glosten and Milgrom (GM) gave a theoretical model that incorporates the asymmetric information effect. Their model is as follows:

$$m_t = m_{t-1} + (1-\pi)\frac{S}{2}Q_t + U_t \qquad (5)$$

where $(1-\pi)$ is the adverse selection cost and π is the order processing cost as a fraction of total costs. This equilibrium price includes the public information (U_t) as well as the private information revealed by Q_t .

The transaction price P_t reflects the order processing cost also. Hence, it comprises unobservable true price as well as the order processing cost.

$$P_t = m_t + \pi\frac{S}{2}Q_t \qquad (6)$$

Taking the first difference of P_t from the above equation we get the following equation:

$$\Delta P_t = \Delta m_t + \pi\frac{S}{2}\Delta Q_t$$

$$= (1-\pi)\frac{S}{2}Q_t + \pi\frac{S}{2}\Delta Q_t + U_t \qquad (7)$$

The parameters S and π can be estimated from the regression coefficients in a regression of ΔP_t on Q_t and ΔQ_t. The variable Q_t estimates the permanent effect or adverse selection component and the variable ΔQ_t estimates the temporary effect.

Model 3: Glosten and Harris (1988)

Glosten and Harris (GH) also model the spread incorporating the asymmetric information by extending the theoretical model of GM. They make both the adverse selection cost as well as combined order processing & inventory cost to be a function of trade size or order size. This is based on other theoretical models, like Glosten (1987a), Easley and O'Hara (1987) and Kyle (1985), which propose that trade size carries information. In GM model, information asymmetry and order processing costs were inferred from Q_t variable only.

It is important to reiterate here that true prices will have component only from the information cost and not from the inventory and order processing cost. However, the observed transaction price will have both. In their model, spread comprises of two components: transitory (order processing and inventory) and permanent (adverse selection).

$$m_t = m_{t-1} + U_t + Q_t Z_t \quad (8)$$

$$P_t = m_t + Q_t C_t \quad (9)$$

$$Z_t = z_0 + z_1 q_t \quad (10)$$

$$C_t = c_0 + c_1 q_t \quad (11)$$

$$e_t \sim iid \;\; Normal$$

$$\Delta P_t = c_0(Q_t - Q_{t-1}) + c_1(Q_t q_t - Q_{t-1} q_{t-1}) + z_0 Q_t + z_1 Q_t q_t + U_t + r_t - r_{t-1} \quad (12)$$

Effective spread: $2C_t + Z_t$

Quoted spread: $2C_t + 2Z_t$

where,

q_t is observed number of shares traded in transaction t, Z_t is the unobservable adverse selection spread component at transaction t and C_t is unobservable transitory spread component at transaction t.

Model 4: De Jong et al. (1996)

The GM and GH models, although used for an order driven market too, were motivated in the context of a quote driven market. However Glosten (1994) presented a theoretical model of price evolution in an electronic order driven market, which is later modified by De Jong et al. (DJ) by explicitly including order processing costs.

The transaction price equation is as follows:

$$p_t = m_t + (R_0 + R_1 q_1)Q_t + v_t \quad (13)$$

where,

p_t is the logarithm of the transaction price, v_t is the random pricing error, R_0 and R_1

The price revision can also be modeled by the change in the expected value of m_t as follows:

$$m_{t+1} = m_t + (e_0 + e_1 q_t)Q_t + u_{t+1} \quad (14)$$

where, u_{t+1} is the publicly observed information that comes in between transaction *t* and *t+1*, but unrelated to the current transaction.

Substituting transaction price equation into expected value equation, we get the following:

$$\Delta p_{t+1} = \Delta(R_0 + R_1 q_{t+1})Q_{t+1} + (e_0 + e_1 q_1)Q_t + e_{t+1}, \tag{15}$$

where, $e_t = u_t + \Delta u_t$.

Rewriting the above equation, we get the following:

$$\Delta p_{t+1} = c + R_0 \Delta Q_{t+1} + R_1 \Delta(q_{t+1} Q_{t+1}) + e_0 Q_t + e_1 q_t Q_t + e_{t+1} \tag{16}$$

the order processing costs: $2[c_0 + c_1 q]$

where $c_0 = R_0 - e_0 - e_1 \alpha$

$$c_1 = R_1 - \frac{1}{2} e_1$$

α is estimated by the median of the transaction size distribution, divided by ln2.

The adverse selection costs: $2[(R_0 - c_0) + (R_1 - c_1)q]$

Model 5: Madhavan et al. (1997)

The previous models considered the effect of public and private information. Madhavan, Richardson and Romans (MRR) modeled the bid-ask spread including both the effect of public/private information as well as market microstructure effects. What is meant by the market microstructure effect is that the trading process itself generates changes in bid-ask spread/ prices in addition to revealing information. MRR develop a structural model of prices capturing the frictions in a unified setting. They estimate two components of the spread: adverse selection and inventory cost & order processing. In the quote driven markets, an order may be executed within the quotes, and hence MRR incorporate this feature in the model. However, in an order-driven market, trades cannot happen in between quotes, and hence this factor is not incorporated here (Ahn et al., 2002). In MRR model, it is assumed that the prices change due to the new public information, which are not associated with trading and order flow but which may provide noisy signal about the future values. MRR argue that indicator variable gives more information than trade size as the informed traders tend to break their order. The breaking of orders leads to the same type of orders which results in order flow autocorrelation and this is more informative as compared to the trade size. Hence, the information asymmetry cost is inferred from autocorrelation of trades rather than signed trade size.

One can use the MRR[3] model to decompose the spread assuming that the following three parameters affect the spread:

(1) θ, the asymmetric information parameter

(2) ϕ, the cost of supplying liquidity

(3) ρ, the autocorrelation of order flow

$$p_t - p_{t-1} = (\phi + \theta)Q_t - (\phi + \rho\theta)Q_{t-1} + \varepsilon_t + \xi_t - \xi_{t-1} \qquad (17)$$

Madhavan, et al. used generalized method of moments (GMM) to identify the parameters $\beta = (\theta, \phi, \rho)$ and a constant drift α implied by the model:

$$E\begin{pmatrix} Q_t Q_{t-1} - Q_t^2 \rho \\ u_t - \alpha \\ (u_t - \alpha)Q_t \\ (u_t - \alpha)Q_{t-1} \end{pmatrix} = 0 \qquad (18)$$

[3] Original MRR model has four parameters; we have not used λ, the probability of trades happening in between the quotes.

The first equation defines the autocorrelation in trade initiation, the second equation defines the drift term, and the last two equations are OLS normal equations.

2.3 Implied and Quoted Bid-Ask Spread

Quoted spread is actual posted spread by the dealer in the quote driven market and the difference between the best limit orders in the order driven market. Roll (1984) argued that in the continuous market setting, there is difference between the quoted and the implied spread. This is because the actual trading is done mostly within the quotes. This price improvement is mainly due to the intervention by a specialist, hidden limit orders and floor traders and matching market orders (Peterson and Fialkowski, 1994). Specialist can improve the price in two ways. Firstly, he can 'stop and order' and secondly by taking order at prices better than the standing order. The standing order could be due to public limit order. It was reported by the fact book of New York Stock Exchange (NYSE) that specialist is involved only in 20 percent of the trade volume. Public limit orders provide a significant amount of trading. The specialist is required to post a representative quotes and not necessarily the best quotes. Some limit orders are not the standing orders and if they are better than the standing orders quoted, they get executed and result in price improvement. These are called hidden limit orders. Floor traders present at the exchange also provide better price to market orders. They also function as hidden limit order traders. Instead of trading through the formal limit order book, the trade can be executed with them. They compete with the specialist by improving the posted quotes. Implied spread is therefore more relevant trading costs. If the reasons given by Peterson and Fialkowski (1994) are correct, then the significant difference between the quoted spread and implied spread should be less in an order-driven market as compared to a quote driven market. The correlation between the posted and implied spread would also be higher in case of an order-driven market.

Implied and quoted bid-ask spreads were compared by Stoll (1989) using the transaction data for National Market System (NMS) securities on the NASDAQ systems for the months October, November, and December 1984. He found that average implied spread was approximately 57 percent of the quoted spread. Lee, Mucklow, and Ready (1993) studied the sample of 230 stocks selected for 1988 from Institute for the Study of Security Markets (ISSM) transaction tape. It was found that mean effective spread is 72 percent of the mean of the quoted spread. Peterson and Fialkowski (1994) examined the posted and effective spread using two days, April 25th and August 20th, 1991 for stocks listed on NYSE. They found that for most orders, the effective spread averages half the quoted spread. In addition, if the quoted spread widens, only 10 to 22 percent of the increase gets transmitted to the implied spread.

Wang (1995) examined the stocks that comprised of the Composite Stock Price Average (CSPA) of Taiwan Stock Exchange for the period January 5 to April 29, 1994. It was found that implied spread is 88 percent of quoted spread. Also the correlation was very high (81 percent) between the two. Ahn, Cai, Hamao, and Ho (2002) found that implied spread was much lesser than quoted spread using data from Nikkei Economic Electronic Database System (NEEDS) for January 5, 1998 to July 31, 1998, straddling April 13, 1998. Declerck (2000) studied the stocks of CAC 40 traded at Paris Bourse for data of January 1998 through June 1998. Opening prices were neglected as the opening prices came through a call auction. He found that traded spread accounted for 85 percent of the posted spread.

In this section, we discussed the relationship between actual quoted spread and implied spread arrived from bid-ask spread models. The relationship is important if traders/researchers want to use implied spread as a proxy of quoted spread. In the next section, we will discuss the empirical findings of the components which make the implied spread.

2.4 Empirical Findings on the Components of the Bid-Ask Spread

The behavior of bid-ask spread is mainly driven by its three components. Different markets have different proportion of these componets depending on the type of market, nature of investors, reforms and maturity of the market, etc. In this section, we summarize some of the empirical findings on components of the bid-ask spread.

Brockman and Chung (1999) provided the components of the bid-ask spread from the Stock Exchange of Hong Kong. Their data set included intraday data for 345 companies covering the period May 1, 1996 to August 1997. The median adverse selection cost was 33 percent of spread and median order processing was 45 percent of the spread. Ness, Ness and Warr (2001) applied four models, GH, GKN, HS, MRR, to the data from NYSE for three months- April, May and June 1999. The adverse selection costs as a proportion of spread was found to be 0.389 (GH), 0.476 (GKN), 0.180 (HS) and 0.666 (MRR). McInish and Ness (2002) examined the NYSE stocks and decomposed the bid-ask spread into order processing and adverse selection using GKN and MRR models. The sample included 30 Dow Jones Industrial Average from October 1997 through September 1998. The order processing cost was 59.78 percent and adverse selection cost was 40.22 percent.

Menyah and Paudyal (2000) estimated the components of bid-ask spreads on the London Stock Exchange (LSE) using data for one year spanning from Jan., to Dec., 1995 of 819 stocks using Stoll's method. They found that the asymmetric information component was approximately 47%. They also employed DJ and Kim & Ogden (K&O, 1996) models for robustness. GKN (1991) showed asymmetric information component to be 21% and K&O model showed 38%. Hanousek and Podpiera (2003) examined the components of the bid-ask spread in an emerging market, namely, Czech equity market (dealer market) by studying ten stocks for the period March 1999 to Dec. 2001. They found that only 17% of the bid-ask spread was explained by informed trading.

Ahn, Cai, Hamao, Ho (2002) examined the stocks of Tokyo Stock Exchange. The data included 204 firms and period January 5, 2000 to March 31, 2000. The adverse selection cost was 49.8% of the bid-ask spread. Huang (2004) compared the components of the bid-ask spread of Taiwan Futures Exchange (TAIFEX), which is a call order driven market and Singapore Exchange Derivatives Trading Limited (SGX-DT). The sample period was from January 1, 2001 to September 30, 2001. Only near month futures were used in the study. Adverse selection cost was lower in the TAIFEX as compared to SGX-DT.Ahn et al. (2005) examined the adverse selection costs of bid-ask spread for Tokyo Stock Exchange (TSE) using 930 stocks for the period July 1 to Oct. 31, 2001. They used MRR, DJ and GH models to estimate the components of the bid-ask spread. They sampled the stocks into three groups according to their tick sizes of 1, 5 and 10 Yen. The adverse selection cost as a proportion of the total spread was found to be 56.47% for the DJ model versus 60.68% for the GH model. As per the MRR model the adverse selection component was found to be 69.56%. Frey and Grammig (2006) estimated the adverse selection component of the bid-ask spread using Xetra open limit order system employed at Frankfurt Stock Exchange using 30 stocks for the period Jan., 2nd to March, 31st 2004. They modified the Glosten/Sandas models by giving away the parametric assumption of the market order size and employing the nonparametric approach. Additionally they replaced the assumption of marginal by average profit conditions. They found that adverse selection was more severe for smaller capitalization stocks.

Frijns, Gilbert and Tourani (2008) investigated the relationship between insider trading and the bid-ask spread for the period Jan, 2001 to Aug., 2004 using 70 most actively traded stocks listed on exchange. They decomposed the spread before and after the enactment of insider trading rules in New Zealand. They found that information asymmetry component of the spread decreased after enactment, mainly observed in illiquid stocks. They used MRR and GH models for decomposing the spread components. Though both the models showed decline in the information asymmetry component but they showed different values. According to the MRR model, they found that information asymmetry decreased from 58.93% to 55.19%. As per GH model information asymmetry decreased from 29.20% to 27.76%.

Giouvris and Philippatos (2008) examined the components of the bid-ask spread under different trading regimes using FTSE100 and FTSE250 stocks. LSE changed from quote driven market to order driven market and hybrid market. They took the sample period of two months before and after the introduction of SETS (order book) for FTSE100 and two months before and after the introduction of SETSMM for FTSE250. They used GKN and K&O models. They found that the asymmetric component of the spread for FTSE100 stocks was higher when the market was quote driven. However, the difference in asymmetric information components was not statistically different between the two regimes- quote driven and hybrid system. Results obtained for FTSE250 stocks showed that the asymmetric information component of the spread did not reduce when the market changed from quote driven to hybrid.

Angelidis and Benos (2009) estimated the components of bid-ask spread for the Athens Stock Exchange (ASE) using stocks of FTASE-20 and FTASE-40 for the period 4th Feb. to 30th Dec., 2002. They applied MRR and GH to estimate the theoretical components of the bid-ask spread. Both these models gave consistent results. They found that the adverse selection component on an average comprised of 56.74% of the bid-ask spread. For robustness, they applied the models for another period, namely June to August 1999 and found that order handling costs comprised of 98% of the implied spread. They argued that at that time of year most investors participating in the market had no information of any stock and trading was due to herd behavior.

Strother, Wansley and Daves (2009) compared the components of the bid-ask spread for trades originating in Electronic Communication Networks (ECNs) and Nasdaq Market Makers using 63 stocks for the period March 1, 1999 to September 30, 1999. They decomposed the bid-ask spread into its theoretical components using Stoll, LSB, GKN and MRR methods. In general they found adverse selection costs to be lower when ECNs was alone at both the bid and ask. According to the Stoll's model, the adverse selection comprised of 60.8% of the bid-ask spread; according to LSB, it was 11.05% and according to GKN, it was 69.8%.

Ahn, Kang and Ryu (2010) investigated the information effect of trade size and trade direction for KOSPI 200 options for the period Jan., 1 to Dec., 31, 2002. They used modified MRR model (size dependent, (SDM), and dummy variable model, (DVM)). They found that for call options, the adverse selection component of the bid-ask spread was around 41.4 % and for put options, the adverse selection component of the bid-ask spread was around 45.3%. They also found that for call options, buyer (seller) initiated trades had around 39.8% (40.1%) adverse selection costs of the bid-ask. For put options, the adverse selection component in the buyer initiated trades was around 40% and seller initiated trades had around 38%.

We find that there is dearth of empirical research in order driven markets in an emerging stock market context including Indian stock market. This study tries to bridge this gap by taking a case of a large order driven market in an emerging context.

3. The Institutional Setting, Data and Descriptive Statistics

3.1 The National Stock Exchange

In November 1992, leading financial institutions, banks, insurance companies and other financial intermediaries set up the National Stock Exchange (NSE). National Stock Exchange is one of the first de-mutualized stock exchanges in the country. NSE provides for nation- wide, anonymous, order driven, screen based trading. The trading platform is known as National Exchange for Automated Trading (NEAT) system. NSE is the first exchange in the world to use satellite communication technology for trading and also has the largest VSAT based trading network in the world. The traders enter the details of their order like price and quantity of the stocks they want to trade in the NEAT system. All the orders are matched based on price/time priority. This system allows large number of participants, improving the liquidity and depth of the market. A single consolidated order book for each stock displays the buy and sell limit orders. The trading facility is made available by NSE not only from computer terminals but also via personal computers. SEBI has also allowed trading through wireless medium or Wireless Application Protocol platform. The SEBI (Prohibition of Insider Trading Regulations, 1992) prohibits insider trading.

The NEAT has four types of market: normal market, odd lot market, auction market and retail debt market. Normal market is for all orders of regular lot size and in dematerialized mode, the market lot size is one. All orders less than regular lot size are traded in the odd lot market. For a trade to take place in this market, both the price and quantity should match exactly with the opposite side. It is used as Limited Physical Market as per the SEBI directions. In an auction market, auctions are initiated by the Exchange for settlement related reasons on behalf of trading members. The retail debt market on the NEAT system is used for trading in retail debt in the same manner as in equities. NSE provides six types of order books: regular lot book, special terms book, negotiated trade book, stop-loss book, odd lot book, and auction book. Regular lot book contains all the regular lot orders that don't have the options 'all or none', 'minimum fill' or 'stop loss'. The Special Terms book contains all the orders that have the features 'all or none' or 'minimum fill'. Negotiated order book has all negotiated order entries matched by the system before they have matched against their counterparty trade entries. Stop-loss order book contains the stop loss orders that trigger only at a pre-specified price. When this order is triggered, the order is released in the regular lot book. Odd lot book has all the odd lot orders i.e. orders less than the marketable lot in the system. Auction book contains all the orders entered in the auctions.

The market currently[4] opens at 9:00 am with a call auction concluded at 9:12 am. Then it starts as a continuous order driven market from 9:15 am throughout the trading period except for closing price. NSE has a tick size of Rs. 0.05 which implies that NSE encourages tight bid-ask spreads

We use the real time National Stock Exchange (NSE) historical tick data supplied by the NSE. The database is time stamped to the nearest second and hence contains full trade data. For quotes, NSE provides data only at four snapshot times: 11:00 hrs, 12:00 hrs, 13:00 hr and 14:00 hrs. The sample period for the study is from January 2002 through October 2008. The choice of beginning from 2002 was made due to the fact that *badla system* in India was abolished in July 2001. Hence, to avoid two different trading systems our sample period starts from 2002. For a stock to be included in the sample for this study, it had to be traded on 80% of the trading days every year throughout the sample period of January 2002 to October, 2008. Following this we arrive at a final sample of 160 stocks. The names of these 160 stocks are given in Appendix A. We also report the total number of days a particular stock traded in the whole period. The anomalous data is deleted from the sample. Firstly for the snapshot quote data, any bid or ask quote which is less than or equal to zero is deleted. We also delete the quote data which has a corresponding negative or zero depth. Additionally, we remove any snapshot data that results into negative bid-ask spread. Any snapshot data that results in the relative spread greater than 20 % is also deleted. For the trade data, we delete the transaction if the price or volume at that price is zero or less than zero assuming an error in recording. The data that is recorded before 9:55 am or after 15:30 pm (for the period we have taken) is also removed. After all these filters, we put one final filter based on price change. Any change in prices which is more than 20 percent in a day on either side is removed.

[4] During the sample period, market opened without a call at 9:55 AM.

3.1 Classification of Buy and Sell Transactions

In the database provided by the National Stock Exchange of India, there are no flags for classifying whether a trade was buyer or seller initiated. In the microstructure literature, most of the models require this classification in their model. This variable is called trade indicator (Q_t). Hence, it is important to impute this variable from the trade data. Lee and Ready (1991) came up with an algorithm to obtain the trade indicator variable from the tick by tick data. Popularly known as "tick test", this test compares the trade price to preceding trades. This test classifies each trade into five categories: an uptick, a downtick, a zero uptick, a zero-downtick and zero-zero tick. If the price is higher than the previous price then it is called uptick and if the price is lower than the previous price then it is called downtick. When the current price is same as the previous price then it is zero tick. In this case, we check the previous trade category. If the zero tick follows an uptick then it is classified as zero uptick. Alternatively, we get a zero tick following a downtick it is classified as zero downtick. A zero tick following a zero tick is classified as zero zero tick. The uptick trade is classified as ask trade or buyer initiated trade (buy) and the downtick trade is classified as bid trade or seller initiated trade (sell). The zero zero tick is not classified as either ask or bid trade. The value of Q_t is +1 if is it a buy order and the value of Q_t is -1 if it is a sell order. Lee and Ready using their methodology on 150 New York Stock Exchange (NYSE) stocks for the year 1988 showed that their tick test provides accurate directional inferences. They also argued that tick test is better than quote test (based on quote and trade data combined) as quotes recorded at the time of trade are not the quotes at which trade happened. Quotes that resulted into trade are not recorded. Order book shows any point of time quotes that are outstanding and executed trades. Although this problem cannot be resolved to a great extent by matching current trades with quotes 5 second before the time stamp.

Before starting to classify the data into buy and sell order, we re-examined the data to remove the effect of standing orders. Hasbrouck (1988) points out that when there is a large order on one side, it gets executed with a series of market orders. These executions will happen at the same price and classifying trade indicator variable based on this will lead to series of zeros in the database implying trade within the spread. Therefore, as suggested by Hasbrouck first and later widely used in the microstructure literature, trades at same timestamp and same price are merged together to form one trade. The trade sizes are added together to get the final trade size at that timestamp and price. We also first merge the trades at same price and time and then classify the trades into buy/sell based on Lee and Ready's tick test. After merging the data and classifying it according to the tick test, 11% data is still with zero zero tick. Such trades may happen in a quote driven market where improvement in the spread may happen due to hidden orders (hidden limit orders). The trades in such cases will happen in between the posted bid and ask quotes. However, in an order driven market like the NSE all the trades happen through the posted limit order prices and hence ruling out the possibility of use of zero zero tick. This is the limitation of the Lee and Ready algorithm used here that we land with these zero zero ticks. Therefore like work elsewhere (Ahn. et al., 2002), these trades are deleted.

4. Estimated Bid-Ask Spread as well as its Components based on Transaction Prices

We report the estimates of the spread as well as its components year wise for the sample period (Jan., 2002- Oct., 2008). Table 1 gives the year wise implied as well as the quoted spread. As can be seen, the estimates by various models are consistent with each other. This is in contrast to the findings of Ness, Ness and Warr (2001) for NYSE, where they find that the estimates are inconsistent with each other. Similar result has been found by Ahn et al. (2002), wherein they find that various models produce consistent results. We find that model estimated spreads are less than the quoted spreads in each case. The spread is highest in the year 2008. Implied spread is approximately 80 percent of the quoted spread on an average across the models. From Table 2, it can be seen that the correlations are also very high across the spread estimates. Except for the correlation between the Roll's spread and spreads estimated using other methods as well as the quoted spread, all others exhibit correlations of almost 1. Correlation between Roll's spread and other spreads is of the order of 0.93-0.94. Figure 1 plots the spread calculated using different methods. As we can see, the spread is almost of the same magnitude till the year 2005 in the sample period beginning 2002 and then increases in the year 2006 followed by a reduction in the year 2007. The average spread in the year 2002 for our sample is of the order Rs. 1.01 to 1.76. In the year 2008, the average spread is of the order Rs. 1.75 to 2.58. Differences across the years are not significantly different.

Table 1: Year-wise Average Estimates of the Spread

This table presents the year wise estimates of the spreads. Both the implied spreads as well as spread calculated from quotes are given. Implied spread is estimated using five methods: Roll; Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks. Standard errors are given in the parenthesis. The unit of spread is Rs.

Year	Roll	GM	GH	DJ	MRR	Quoted
2002	1.01 (3.69)	1.35 (5.04)	1.35 (5.13)	1.36 (5.28)	1.23 (4.95)	1.76 (7.40)
2003	1.00 (3.34)	1.33 (4.56)	1.31 (4.55)	1.33 (4.61)	1.17 (4.24)	1.72 (6.20)
2004	1.18 (3.34)	1.39 (4.05)	1.39 (4.07)	1.38 (4.03)	1.15 (3.34)	1.70 (4.99)
2005	1.07 (2.45)	1.29 (2.59)	1.27 (2.59)	1.27 (2.60)	1.08 (2.39)	1.52 (3.46)

2006	1.53 (3.66)	1.75 (4.00)	1.74 (4.04)	1.74 (4.04)	1.47 (3.61)	2.13 (5.15)
2007	1.43 (3.46)	1.60 (3.53)	1.60 (3.58)	1.60 (3.52)	1.33 (2.99)	1.96 (4.31)
2008	1.75 (5.01)	2.13 (6.19)	2.11 (6.11)	2.13 (6.14)	1.73 (4.79)	2.58 (7.03)

* significant at 1 % level

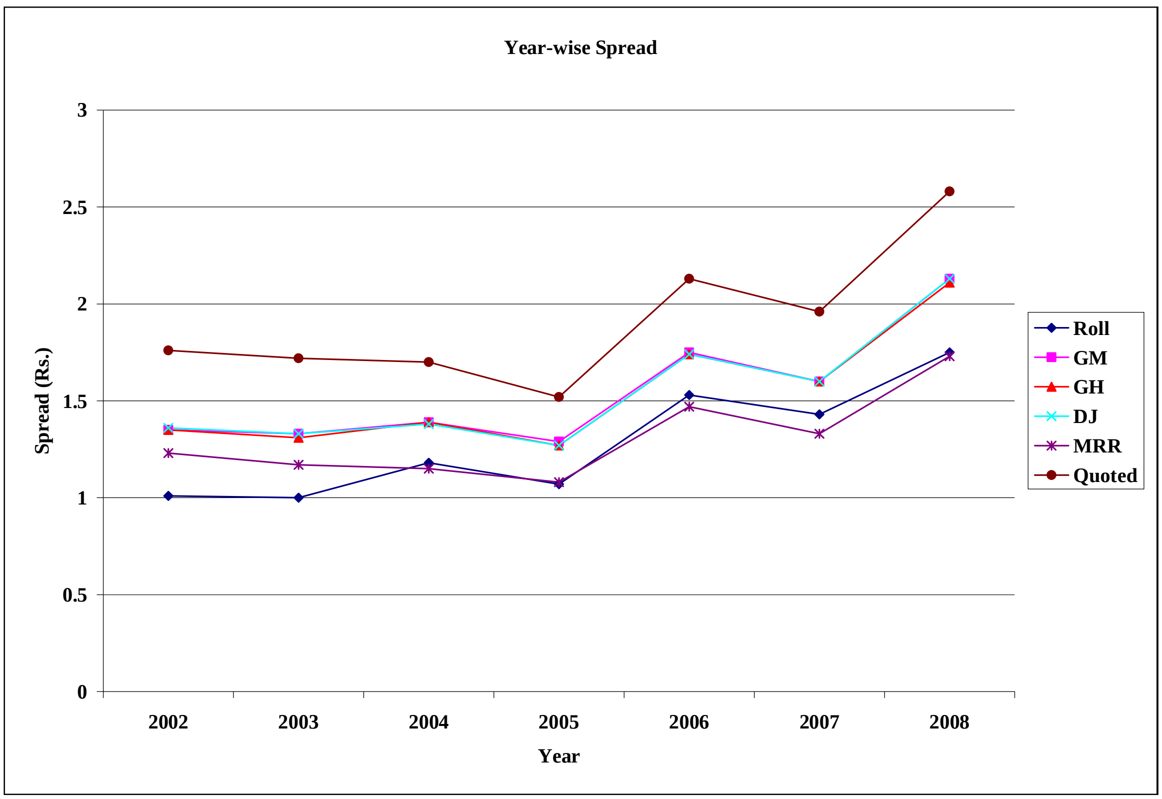

Figure 1: Year-wise Spread calculated using different models

Table 2: Correlations between Spread

This table presents the correlations between various spread measures. The spread is calculated using quotes as well as the models: Roll; Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks.

Correlations	Spread	Roll	GM	GH	DJ	MRR
Spread	1.00	0.94	1.00	1.00	1.00	1.00
Roll		1.00	0.94	0.93	0.94	0.93
GM			1.00	1.00	1.00	1.00
GH				1.00	1.00	1.00
DJ					1.00	1.00
MRR						1.00

It is more appropriate to compare relative bid-ask spread than bid-ask spread when dealing with different stocks as spread is a direct function of price. Bid-ask spread is limited by the price of the stock i.e. it cannot be greater than the price of the stock. Relative spread is defined as the absolute spread in Rupee divided by the price. Hence, we compute and report the relative spread in Table 3. From Table 3, it can be seen that the relative spread estimates using quoted spread as well as spread calculated through models also produce consistent values just like absolute spreads. The relative spread has reduced over time and increased only in the year of financial crisis (2008). In the year 2002, the relative spread on an average was of the order 0.55 % to 0.78% and it became 0.39% to 0.55% in the year 2008. Figure 2 gives the plot of the relative spread over years. As can be seen the relative spread reduced till the year 2005 (~0.29%) and then there was slight increase in the year 2006 (~0.30%) and it remained at the same level for the year 2007 (~0.30%). In 2008, the relative spread became approx. 0.48 %. The reduction is significant for the year 2003 from 2002 and for the year 2004 and 2005. Reduction of relative spread from the year 2002 to 2008 is also significant. The increase in the relative spread in the year 2008 from the year 2007 is significant too. The correlation between the relative spread estimated using different models as well as the quoted relative spread is very high except for the relative spread using Roll's estimator (Table 4).

Table 3: Year-wise Average Estimates of the Relative Spread

This table presents the year wise estimates of the relative spreads. Relative spread is spread by price. Both the implied relative spreads as well as relative spread calculated from quotes are given. Implied relative spread is estimated using five methods: Roll (R); Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks. The relative spreads are in percentages. Standard errors are given in the parenthesis. Statistical significance is given for difference between consecutive years as well as the difference between the last year and first year.

Year	R/P	GM/P	GH/P	DJ/P	MRR/P	Relative Spread
2002	0.55* (0.75)	0.74* (0.65)	0.70* (0.59)	0.70* (0.59)	0.63* (0.57)	0.78* (0.72)
2003	0.36 (0.36)	0.48 (0.39)	0.48 (0.38)	0.48 (0.38)	0.42 (0.35)	0.56 (0.46)
2004	0.41 (1.20)	0.43* (0.36)	0.40* (0.33)	0.40* (0.33)	0.34* (0.30)	0.48* (0.45)
2005	0.25 (0.47)	0.29 (0.21)	0.28 (0.20)	0.29 (0.20)	0.24 (0.18)	0.32 (0.25)
2006	0.30 (0.49)	0.32 (0.26)	0.32 (0.26)	0.32 (0.26)	0.27 (0.23)	0.37 (0.32)
2007	0.27*	0.32*	0.32*	0.32*	0.27*	0.37*

	(0.33)	(0.26)	(0.26)	(0.26)	(0.23)	(0.30)
	0.39	0.48	0.48	0.48	0.40	0.55
2008	(0.39)	(0.45)	(0.45)	(0.45)	(0.40)	(0.54)

* significant at 1 % level

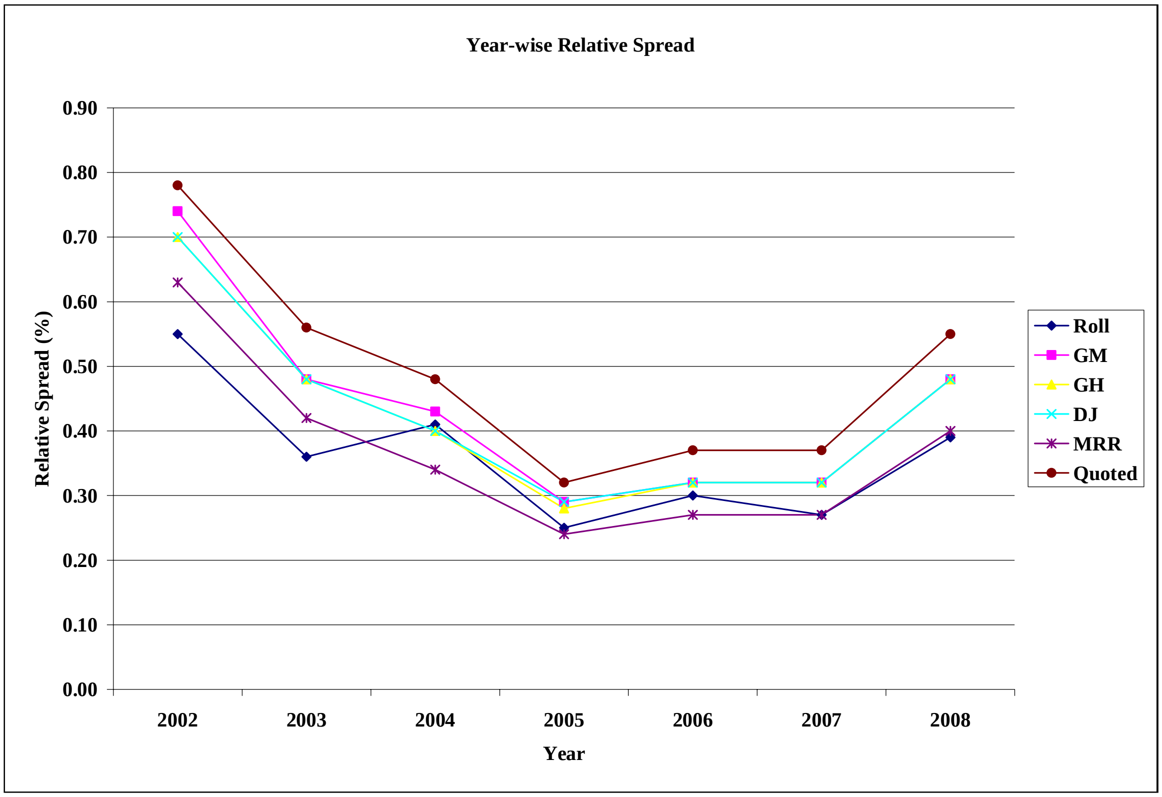

Figure 2: Year-wise Relative Spread calculated using different models

Table 4: Correlations between Relative Spread

This table presents the correlations between various relative spread measures. Relative spread is spread by price. This relative spread is calculated using quotes as well as the models: Roll; Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks.

Correlations	Spread	Roll	GM	GH	DJ	MRR
Spread	1.00	0.58	0.96	0.98	0.98	0.97
Roll		1.00	0.55	0.56	0.56	0.48
GM			1.00	0.99	0.99	0.98
GH				1.00	1.00	0.99
DJ					1.00	0.99
MRR						1.00

Table 5: Comparison of Indian Market's Relative Spread with Other Markets*

Market	Year	Relative Spread	Source
NSE	2002	0.78%	NSE paper[5]
NSE	2003	0.56%	NSE paper
NSE	2004	0.48%	NSE paper
TSE	2002-2004	0.49%	Ahn, Cai and Hamao (2005)
AuSE	2002-2004	0.66%	Popular press

[5] http://www.nse-india.com/content/research/comppaper128.pdf

Taiwan	2002-2004	0.53%	Huang (2004)
KSE	2002-2004	0.39%	Ahn, Kang and Ryu (2010)
SEHK	2002-2004	1.08%	Popular press
ASE	2002-2004	0.80%	Angelidis and Benos (2009)
NYSE	2002	0.39%	Jiang, Kim and Wood (2009)
NYSE	2003	0.26%	Jiang, Kim and Wood (2009)
LSE	2002-2004	1.39%	Giouvris and Philippatos (2008)
Euronext Paris	2002-2004	0.92%	Popular press
Deutsche Boerse	2002-2004	0.82%	Popular press

In this section, we also present the mean values of adverse selection costs (Table 6) as well as adverse selection costs by price (Table 7). As can be seen from the Table 6, adverse selection costs increases with time till the year 2006 and then it reduces in the year 2007 and again increases in the year 2008 when it is the highest. We also see that an adverse selection cost is usually highest for the MRR model (though the difference is insignificant statistically). Table 7 presents the adverse selection costs divided by the price. It can be seen that it reduces till the year 2005 and thereafter, it remains almost constant till the year 2007 after which it increases in the year 2008.

Table 6: Year-wise Average Estimates of the Adverse Selection Costs

This table presents the year-wise averages of the adverse selection costs. It is estimated using four methods: Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks. Standard errors are given in the parenthesis. The unit of Adverse Selection Costs is Rs.

Year	GM	GH	DJ	MRR
2002	0.63 (2.14)	0.64 (2.25)	0.66 (2.56)	0.68 (2.56)
2003	0.66 (2.30)	0.66 (2.31)	0.65 (2.29)	0.70 (2.53)
2004	0.69 (2.04)	0.68 (2.04)	0.68 (2.05)	0.70 (2.00)
2005	0.63 (1.24)	0.62 (1.22)	0.61 (1.21)	0.66 (1.38)
2006	0.84 (1.85)	0.84 (1.89)	0.84 (1.88)	0.88 (2.01)
2007	0.78 (1.71)	0.78 (1.78)	0.77 (1.71)	0.80 (1.73)
2008	1.06 (3.25)	1.06 (3.22)	1.06 (3.28)	1.04 (2.92)

* significant at 1 % level

Table 7: Year-wise Average Estimates of the Adverse Selection Costs by Price

This table presents the year-wise averages of the adverse selection costs by price. It is estimated using four methods: Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks. The adverse selection costs by price are in percentages. Standard errors are given in the parenthesis. Statistical significance is given for difference between consecutive years as well as the difference between the last year and first year.

Year	GM/P	GH/P	DJ/P	MRR/P
2002	0.37* (0.33)	0.35* (0.29)	0.35* (0.29)	0.36* (0.31)
2003	0.24 (0.20)	0.24 (0.19)	0.24 (0.19)	0.25 (0.20)
2004	0.21* (0.18)	0.20* (0.16)	0.20* (0.16)	0.20* (0.17)
2005	0.15 (0.11)	0.14 (0.10)	0.14 (0.10)	0.14 (0.11)
2006	0.16 (0.13)	0.16 (0.13)	0.16 (0.12)	0.16 (0.14)
2007	0.16* (0.13)	0.16* (0.13)	0.16* (0.12)	0.16* (0.14)
2008	0.24 (0.22)	0.24 (0.22)	0.24 (0.22)	0.24 (0.23)

* significant at 1 % level

Table 8 gives the mean values of the order processing costs estimated using all the models. Order processing cost values from the models GM, GH and DJ are consistent with each other. However, the order processing costs estimated using the MRR model and lower than the other three models. Order processing component of the spread follows a similar pattern over time as the adverse selection component of the spread. Order processing cost divided by the price is given in the Table 9.

Table 8: Year-wise Average Estimates of the Order Processing Costs

This table presents the year-wise averages of the order processing costs. It is estimated using four methods: Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks. Standard errors are given in the parenthesis. The unit of Order Processing Costs is Rs.

Year	GM	GH	DJ	MRR
2002	0.72 (2.91)	0.71 (2.91)	0.70 (2.72)	0.55 (2.39)
2003	0.67 (2.26)	0.66 (2.24)	0.67 (2.32)	0.47 (1.75)
2004	0.71 (2.01)	0.70 (2.04)	0.70 (1.98)	0.45 (1.34)
2005	0.66 (1.35)	0.66 (1.38)	0.66 (1.39)	0.42 (1.01)
2006	0.91 (2.15)	0.90 (2.15)	0.90 (2.16)	0.59 (1.61)
2007	0.83 (1.82)	0.82 (1.81)	0.83 (1.82)	0.53 (1.26)
2008	1.07 (2.95)	1.06 (2.91)	1.07 (2.88)	0.69 (1.89)

* significant at 1 % level

Table 9: Year-wise Average Estimates of the Order Processing Costs by Price

This table presents the year-wise averages of the order processing costs by price. It is estimated using four methods: Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks. The order processing costs by price are in percentages. Standard errors are given in the parenthesis. Statistical significance is given for difference between consecutive years as well as the difference between the last year and first year.

Year	GM/P	GH/P	DJ/P	MRR/P
2002	0.37* (0.33)	0.35* (0.30)	0.35* (0.30)	0.27* (0.27)
2003	0.24 (0.20)	0.24 (0.19)	0.24 (0.19)	0.17 (0.16)
2004	0.21* (0.18)	0.20* (0.17)	0.21* (0.17)	0.13* (0.13)
2005	0.15 (0.11)	0.14 (0.11)	0.15 (0.11)	0.09 (0.08)
2006	0.16 (0.14)	0.17 (0.14)	0.17 (0.14)	0.11 (0.11)
2007	0.16* (0.14)	0.16* (0.13)	0.17* (0.14)	0.11* (0.10)
2008	0.24 (0.23)	0.24 (0.23)	0.25 (0.24)	0.16 (0.17)

* significant at 1 % level

In Table 10, we report the average year-wise proportion of adverse selection costs to spread. Proportion of adverse selection costs by the spread is calculated by taking the ratio of adverse selection component of the spread to bid-ask spread. The proportions using all the four methods are given. As can be seen from the table, according to the models, GM, GH and DJ the average proportion of adverse selection to spread is 50 % with insignificant variation across stocks and over the years. This is similar to the findings from other order driven markets like Tokyo Exchange, where Ahn et al. (2002) found approximately 53% adverse selection component in the spread. We find higher proportion of adverse selection to spread using the MRR model because as mentioned above, it gives higher values for adverse selection component and lower values for order processing component of the spread as compared to the other three models.

Table 10: Year-wise Proportion of Adverse Selection Costs to Spread

This table presents the year-wise averages of the proportion of adverse selection costs to spread. It is estimated using four methods: Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and (4) Madhavan et al. (MRR). This is based on 160 stocks. Standard errors are given in the parenthesis.

Year	GM	GH	DJ	MRR
2002	0.51 (0.03)	0.50* (0.03)	0.50 (0.03)	0.60 (0.07)
2003	0.51 (0.02)	0.51* (0.03)	0.50 (0.03)	0.62 (0.07)
2004	0.51* (0.03)	0.50 (0.03)	0.50 (0.03)	0.62 (0.08)
2005	0.50 (0.03)	0.50 (0.03)	0.50 (0.03)	0.62 (0.08)
2006	0.50 (0.02)	0.50 (0.03)	0.50 (0.03)	0.61 (0.07)
2007	0.50*	0.50*	0.50	0.60

	(0.02)	(0.02)	(0.02)	(0.06)
2008	0.51 (0.02)	0.51* (0.02)	0.50 (0.02)	0.61 (0.06)

* significant at 1 % level

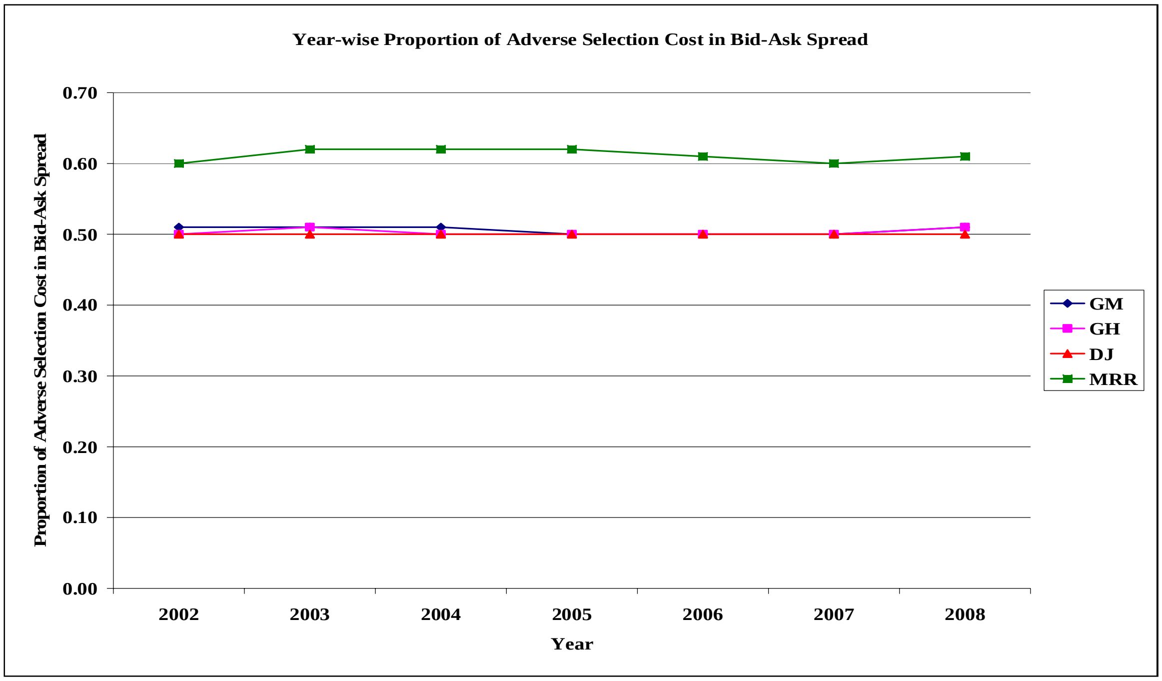

Figure 3: Year-wise Proportion of Adverse Selection Cost in Bid-Ask Spread

Table 11: Comparison of Proportion of Adverse Selection Costs in the Bid-Ask Spread with Other Markets

Market	Proportion of adverse selection costs
NSE	50%

TSE	50%
ASE	56.74%
NYSE, AMEX	9.60%
LSE	21%

5. Conclusions

In this paper, we report the estimated bid-ask spread (and the relative spread) as well as its components- adverse selection costs (and the adverse selection costs divided by price) and order processing costs (and the order processing costs divided by price), using five/ four different models. These five models are- Roll; Glosten and Milgrom (GM); Glosten and Harris (GH); De Jong et al. (DJ); and Madhavan et al. (MRR). These estimated bid-ask spreads are benchmarked with the spreads (relative spreads) calculated using four snapshots based quotes.

We find that all the models produce consistent results with each other for the absolute bid-ask spreads as well as the relative spreads. For the adverse selection component (adverse selection costs by price) and the order processing component (order processing costs by price) of the bid-ask spread, similar values for the models, GM, GH and DJ were found. We get higher values for adverse selection and lower values for order processing component of the spread using the MRR model. However, the difference is not statistically significant. The correlation is also very high among the bid-ask spreads and relative spreads estimated using different models and estimates using four snapshot of actual quotes (except spread and relative spread calculated using the Roll's method). We find that over years, the spread has increased but relative spread has decreased, rising only in the year 2008. 2008 being the year of financial crisis, the rise in the relative spread is as expected. The proportion of adverse selection costs in the bid-ask spread is approximately 50 %, which is consistent with the findings elsewhere for the order driven market (Ahn et al., 2002).

References

Admati, A., and Pfleiderer, P., (1988) "A Theory of Intraday Patterns: Volume and Price Variability", *Review of Financial Studies*, Vol. 1, 3-40

Ahn, H., Cai, J., Hamao, Y., and Ho, R, (2002) "The components of the bid-ask spread in a limit-order market: evidence from the Tokyo Stock Exchange", *Journal of Empirical Finance*, Vol. 9, 399-430

Ahn, H., Cai, J., Hamao, Y., and Ho., R., (2005) "Adverse selection, brokerage coverage, and trading activity on the Tokyo Stock Exchange", *Journal of Banking & Finance*, Vol. 29, 1483-1508

Ahn, H., Kang, J., and Ryu, D., (2010) "Information Effects of Trade Size and Trade Direction: Evidence from the KOSPI 200 Index Options Market", *Asia-Pacific Journal of Financial Studies*, Vol., 39, 301-339

Angelidis, T., and Benos, A., (2009) "The components of the Bid-Ask Spread: The Case of the Athens Stock Exchange", *European Financial Management*, Vol. 15, No. 1, 112-144

Bagehot, W., (1971) "The Only Game in Town", *Financial Analysts Journal*, Vol.8, 31-53

Bervas, A., (2006) "Market liquidity and its incorporation into risk management", http://www.banquedefrance.net/gb/publications/telechar/rsf/2006/etud2_0506.pdf

Brockman, P., Chung, D., (1999), "An analysis of depth behavior in an electronic, order-driven environment", *Journal of Banking and Finance*, Vol. 23, 1861–1886

Copeland, T., and Galai, D., (1983) "Information Effects on the Bid-Ask Spreads", *Journal of Finance*, Vol. 33, 1457-69

De Jong, F., Nijman, T., and Roell, A (1996) "Price effects of trading and components of the bid-ask spread on the Paris Bourse", *Journal of Empirical Finance*, Vol. 3, 193-213

Declerck, F., (2002) “Le prix de l’immédiateté: Le cas de la Bourse de Paris”, *Banque et Marchés*, Vol. 57, 31–45

De Jong, F., Nijman, T. and Roell, A. (1996), “Price effects of trading and components of the bid-ask spread on the Paris Bourse”, *Journal of Empirical Finance*, Vol. 3, 193-213

Easley, D., and O’Hara, M., (1987) “Price, Trade Size, and Information in Securities Markets”, *Journal of Financial Economics*, Vo. 19, 69-90

Frey, S. and Gramming, J., (2006) “Liquidity supply and adverse selection in a pure limit order book market”, *Empirical Economics*, Vol. 30, No. 4, 1007-1033

Frijns, B., Gilbert, A., and Tourani-Rad, A., (2008) “Insider Trading, Regulation, and the Components of the Bid-Ask Spread”, *Journal of Financial Research*, Vol. 31, No., 3

George, T., Kaul, G., and Nimalendran, M., (1991) “Estimation of the bid-ask spreads and its components: A new approach”, *Review of Financial Studies*, Vol. 4, 623-656

Giouvris, E., and Philippatos, G., (2008) “Determinants of the Components of the Bid-Ask Spreads on the London Stock Exchange: The Case of Changes in Trading Regimes”, *Journal of Money, Investment and Banking*, Issue 1, 49-60

Glosten, L., and Milgrom, P., (1985) “Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders”, *Journal of Financial Economics*, Vol. 14, 71-100

Glosten, L. (1987a) “Components of the bid/ask spread and the statistical properties of transaction prices”, *Journal of Finance*, Vol. 42, 1293-1308

Glosten, L. and Harris, L., (1988) “Estimating the components of the bid/ask spread”, *Journal of Financial Economics*, Vol. 21, 123-142

Glosten, L. (1994) “Is the electronic open limit order book inevitable?”, *Journal of Finance*, Vol. 49, 1127-1161

Handa and Schwatrz, (1996) “Limit Order Trading”, Journal of Finance, Vol. 51, 1835-1861

Hanousek, J., and Podpiera, R., (2003) “Informed trading and the bid-ask spread: Evidence from an emerging market”, *Journal of Comparative Economics*, Vol. 31, 275-296

Hasbrouck, J., (1988), “Trades, quotes, inventories and information”, Journal of Financial Economics, Vol. 22, 229-252

Ho., T., and Stoll, H., (1983) “The dynamics of dealer markets under competition”, *Journal of Finance*, Vol. 38, 1053-1074

Ho, T., and Macris, R., (1984) “Dealer bid-ask quotes and transaction prices: An empirical study of some AMEX options”, *Journal of Finance*, Vol. 39, 23-45

Huang, Y., (2004) “The market microstructure and relative performance of Taiwan stock index futures: A comparison of the Singapore exchange and the Taiwan futures exchange”, *Journal of Financial Markets*, Vol. 7, 335-350

Jiang, C., Kim, J. and Wood, R., (2009), “Adverse selection costs for NASDAQ and NYSE after decimalization”, *International Review of Financial Analysis*, Vol. 18, 205-211

Kyle, A., (1985) “Continuous auctions and insider trading”, *Econometrica*, Vol. 53, 1315-1336

Lee, C., and Ready, M.., (1991) “Inferring Trade Direction from Intraday Data”, *Journal of Finance*, Vol. XLVI, 733-746

Lee, C., Mucklow, B., and Ready, M., (1993) “Spreads, depths, and the impact of earnings information: an intraday analysis”, *Review of Financial Studies*, Vol. 6, 345-374

Madhavan, A., Richardson, M., and Roomans, M., (1997) "Why do security prices fluctuate? A transaction level analysis of NYSE stocks", *Review of Financial Studies,* Vol. 10, 1035-1064

McInish, T., and Ness, B., (2002) "An intraday examination of the components of the bid-ask spread", *The Financial Review*, Vol. 37, 507-524

Menyah, K and Paudyal, K., (2000) "The components of bid-ask spreads on the London Stock Exchange", *Journal of Banking & Finance*, Vol., 24, Issue 11, 1767-1785

Ness, R., Ness, B., and Warr, R., (2001) "Determinants of adverse selection on the NYSE and NASDAQ: A paired comparison", *Working Paper*, SSRN

Petersen, M., and Fialkowski, D., (1994) "Posted versus effective spreads: Good prices or bad quotes?", *Journal of Financial Economics*, Vol. 35, 269-292

Roll, R., (1984) "A simple implicit measure of the effective bid-ask spread in an efficient market", *Journal of Finance*, Vol. 39, 1127-1139

Sandas, P., (2001) "Adverse selection and competitive market making: Empirical evidence from a limit order market", *Review of Financial Studies*, Vol. 14, 705-734

Stoll, H., (1989) "Inferring the components of the bid-ask spread: Theory and Empirical tests", *Journal of Finance*, Vol. 44, 115-134

Strother, T., Wansley, J., and Daves, P., (2009) "Electronic Communication Networks, Market Makers, and the Components of the Bid-Ask Spread", International *Journal of Managerial Finance*, Vol. 5, Iss 1, 81-109

Tinic, S., (1972) "The economics of liquidity services", *Quarterly Journal of Economics*, Vol. 86, 79-93

Wang, C., (1995) "Bid-ask spread and stock return behavior: An empirical investigation of the Taiwan stock market", *University of Rhode Island*, 140 pgs

Appendix 1

Table A: List of Stocks in the Sample Set

This table gives the name of the stock along with their symbol. After filtering the stocks, there are a total of 160 stocks in the final sample. We also report here the number of trading days for a particular stock throughout the sample period: Jan. 2002 to Oct., 2008.

Company Name	N
Aban Offshore Ltd.	1675
Abb Limited	1680
Acc Limited	1678
Adani Enterprises Limited	1671
Ador Welding Limited	1672
Alfa Laval (India) Ltd.	1672
Alok Industries Limited	1671
Ambuja Cements Ltd	1671
Apollo Hospitals Enterprise Ltd	1671
Apollo Tyres Ltd	1671
Arvind Limited	1675

Ashok Leyland Ltd	1671
Asian Hotels Ltd	1678
Asian Paints Limited	1674
Astrazeneca Pharma India Limited	1602
Atul Ltd.	1640
Axis Bank Limited	1671
Bajaj Auto Limited	1627
Balrampur Chini Mills Ltd	1671
Bank Of Baroda	1671
Basf India Ltd	1678
Bata India Ltd	1671
Beml Limited	1678
Berger Paints (I) Ltd	1670
Bharat Electronics Ltd	1680
Bharat Forge Ltd	1671
Bharat Heavy Electricals Ltd	1673
Bank Of India	1671
Bombay Dyeing & Mfg Co. Ltd	1671
Bongaigaon Refinery & Petrochemicals Ltd	1671
Bharat Petroleum Corpn. Ltd	1677
Britannia Industries Ltd	1671

Cadila Healthcare Limited	1651
Carborundum Universal Ltd	1668
Century Enka Ltd	1671
Chemplast Sanmar Ltd.	1662
Cipla Ltd.	1677
Cosmo Films Ltd	1674
Chennai Petroleum Corporation Limited	1674
Crisil Limited	1660
Crompton Greaves Ltd	1674
Cubex Tubings Ltd.	1680
Cummins India Ltd.	1674
Dabur India Ltd	1677
Dcm Shriram Consolidated Ltd	1670
Dcw Ltd	1680
Dena Bank	1671
Dr. Reddy's Laboratories Ltd.	1653
Eid Parry India Ltd.	1671
Escorts Ltd	1674
Finolex Industries Ltd.	1674
First Leasing Company Of India Ltd	1672
Garden Silk Mills Ltd.	1647

Ghcl Limited	1678
Gujarat Industries Power Co. Ltd	1677
Glaxosmithkline Consumer Healthcare Limited	1671
Glaxosmithkline Pharmaceuticals Limited	1574
Glenmark Pharmaceuticals Ltd.	1674
Gujarat Mineral Development Corporation Limited	1674
Gujarat Narmada Valley Fertilizer Co. Ltd.	1678
Godfrey Phillips India Ltd	1651
Grasim Industries Ltd.	1652
Gujarat State Fertilizers & Chemicals Ltd.	1657
Gujarat Gas Co. Ltd	1659
Hcl Technologies Ltd	1651
Housing Development Finance Corporation Ltd.	1656
Hdfc Bank Ltd	1671
Hero Honda Motors Ltd.	1651
Hindalco Industries Ltd.	1671
Hindustan Unilever Limited	1646
Hindustan Motors Ltd.	1637
Hindustan Oil Exploration Co. Ltd	1674
Hindustan Petroleum Corporation Ltd.	1672
Hindustan Sanitaryware And Industries Ltd	1667

Icici Bank Ltd.	1681
Idbi Bank Limited	1677
Indian Hotels Company Limited	1671
India Cements Limited	1663
Indraprastha Medical Corporation Ltd.	1631
Infosys Technologies Ltd.	1650
Infotech Enterprises Ltd	1634
Ing Vysya Bank Limited	1669
Indian Overseas Bank	1678
Indian Oil Corporation Ltd	1680
Iol Netcom Limited	1680
Itc Ltd.	1678
Iti Ltd.	1596
Jayshree Tea & Industries Ltd	1661
Jb Chemicals & Pharmaceuticals Ltd.	1647
Jammu & Kashmir Bank Ltd.	1674
Karur Vysya Bank Ltd	1645
Kirloskar Oil Engines Ltd.	1665
Ksb Pumps Ltd.	1664
Lic Housing Finance Ltd	1671
Larsen & Toubro Ltd.	1654

Madras Cements Ltd.	1660
Maharashtra Seamless Ltd	1672
Mastek Ltd	1656
Bosch Limited	1667
Mirc Electronics Ltd.	1645
Mahindra & Mahindra Ltd.	1676
Moser-Baer (I) Ltd	1673
Mrf Ltd.	1677
Mahanagar Telephone Nigam Ltd.	1677
Heidelbergcement India Limited	1631
Nagarjuna Fertilizer & Chemicals Ltd.	1668
Nahar Spinning Mills Ltd.	1629
National Aluminium Co. Ltd.	1671
Piramal Healthcare Limited	1668
Niit Ltd.	1643
Nirma Ltd.	1673
Nocil Ltd.	1580
Oil & Natural Gas Corpn. Ltd.	1676
Oriental Bank Of Commerce	1644
Paper Products Ltd.	1667
Pfizer Ltd.	1671

Polaris Software Lab	1669
Polyplex Corporation Ltd.	1561
Punjab Tractors Ltd. [Merged]	1627
Ranbaxy Laboratories Ltd.	1667
Raymond Ltd.	1671
Rashtriya Chemicals & Fertilizers Ltd.	1676
Reliance Industries Ltd.	1667
Reliance Industrial Infrastructure Ltd.	1672
Reliance Infrastructure Ltd.	1679
Steel Authority Of India Ltd.	1674
Satyam Computer Services Ltd.	1667
State Bank Of India	1674
Shipping Corpn. Of India Ltd.	1677
Sesa Goa Ltd.	1671
Shanthi Gears Ltd.	1659
Siemens Ltd.	1671
South Indian Bank Ltd.	1668
S R F Ltd.	1675
Sun Pharmaceutical Inds. Ltd.	1670
Supreme Petrochem Ltd.	1667
Supreme Industries Ltd.	1642

Su-Raj Diamonds & Jewellery Ltd.	1643
Surya Roshni Ltd.	1664
Swaraj Engines Ltd.	1631
Tata Motors Ltd.	1672
Tata Power Co. Ltd.	1669
Tata Sponge Iron Ltd.	1669
Tata Steel Ltd.	1671
Tata Tea Ltd.	1671
Thermax Ltd.	1671
Thomas Cook (India) Ltd.	1643
Tamil Nadu Newsprint & Papers Ltd.	1633
Ucal Fuel Systems Ltd.	1667
Unitech Ltd.	1657
Vardhman Textiles Ltd.	1671
Vesuvius India Ltd.	1665
V I P Industries Ltd.	1559
Voltas Ltd.	1672
Tata Communications Ltd.	1675
V S T Industries Ltd.	1625
Wipro Ltd.	1676
Zandu Pharmaceutical Works Ltd.	1613

Zee Entertainment Enterprises Ltd.	1666
Zuari industries	1667