Bid/Ask Spread and Volatility in the Corporate Bond Market

Madhu Kalimipalli
Faculty of Management
McGill University

Arthur Warga
Department of Finance, College of Business
University of Houston

Correspondence to:

Arthur Warga
Judge James A. Elkins Professor of Banking and Finance
College of Business Administration
University of Houston
Houston, TX 77204-6282
e-mail: warga@uh.edu
office (713) 743-4779
fax (713) 743-4584

Abstract

Bid/Ask Spread and Volatility in the Corporate Bond Market

This paper examines the relationship between price volatility and bid-ask spreads on individual bonds trading on the NYSE's Automated Bond System. Retail-sized trades and thin volume mandate a data analytic approach that accommodates irregularly spaced transactions and quotes. Latent volatility for each bond is extracted using an Autoregressive Conditional Duration (ACD) model that provides input into an ordered probit model for observed spreads. For the most part we find a significant negative relationship between latent volatility and observed spread among the ten most actively traded bonds on the ABS. The results contrast with earlier findings in the foreign exchange market, and suggest that in some thinly traded markets volatility may proxy for liquidity.

Bid/Ask Spread and Volatility in the Corporate Bond Market

In this paper we provide the first microstructure analysis of the relationship between bid-ask spreads and volatility in the market for corporate bonds. The market we examine is the New York Stock Exchange, which maintains the largest corporate bond exchange[1]. NYSE bonds are traded on the Automated Bond System (ABS), which can be described as a fully automated electronic trading and information system whose schedules of bid and ask prices are fully transparent. Trades on the ABS are typically small in size and average about twenty bonds per trade. Also, like most issues available for trade in the dealer market, the ABS is illiquid. The average time interval between two trades for the most actively traded bond in our sample is about seven minutes while for the tenth most actively traded bond it is about thirty eight minutes. Trading activity declines rapidly for bonds not among the top ten in volume.

The aforementioned features of the ABS market compel us to draw on the statistical literature for techniques dealing with irregularly spaced data. Transforming data to regularly spaced intervals, an approach often employed in microstructure studies of more liquid markets, cannot feasibly be implemented with ABS data. We employ an ordered probit analysis similar to that found in Hausman, Lo, and MacKinlay (1992) and Bollerslev and Melvin (1994) to uncover links between volatility and the magnitude of the bid-ask spread. The analysis is complicated by the fact that volatility cannot be measured in the traditional manner of employing GARCH or other time series methods because the data is too irregular and thin to admit that sort of analysis. To solve this problem we draw on techniques

[1] NYSE traded bonds account for about 10% of the market in these bonds in terms of number of trades, but closer to 1% in terms of the dollar value of those trades. There have been several recent initiatives to start up alternative electronic bond exchanges. Among these are Intervest (www.intervest.com) and Bond Connect (www.bondconnect.com).

developed in Engle and Russell (1998) to estimate proxies for volatility using the time duration between trades.

The dealer market in corporate bonds, which to this date has no central trade or reporting system, accounts for most of the dollar volume of corporate bond transactions. Nevertheless, as documented in Dueweke, Hyland, and Siesel (1992) the ABS system accounts for a significant percentage of the number of transactions in bonds traded there, especially if attention is restricted to the most active issues. If we restrict ourselves to the most actively traded bonds on the ABS, it is unlikely that more frequent transactions or quotes will be available from any dealer. Furthermore, as shown in Hong and Warga (1999), pricing on the ABS is in accord with pricing in the dealer market. ABS is the only corporate bond market capable of carrying out transactions and providing time-stamps for quotes and trades. Transaction-based databases from the dealer market, such as the National Association of Insurance Commissioners' (NAIC) schedule D database, lack time stamps and only record transactions by the day they occur[2]. In this paper we will focus on the top ten most active issues on the ABS, thereby assuring that we examine a stream of quotes representing a significant portion of the appropriate quote universe for those instruments.

In a previous study, Bollerslev and Melvin (1994) examine the nature of the relationship between bid-ask spreads for exchange-rate quotes and the volatility of the underlying exchange-rate process. They develop a model for the foreign exchange market characterized by trading between market makers and two types of traders viz., informed traders and liquidity traders. Expected profit of the market makers is obtained as a sum of their expected profits from trading with the two groups of traders, and is set equal to zero in equilibrium. The authors show that, in equilibrium, spreads are proportional to the conditional volatility of the fundamental value of the exchange rate. In a two step process, Bollerslev and Melvin obtain the GARCH estimates of the underlying volatility of the exchange rate process and then use them as inputs into an ordered probit model. The ordered probit model measures the temporal relationship between observed spreads (having a discrete support) and other predetermined variables (having continuous support) like lagged spreads and volatility estimates of the exchange rates**.** Bollerslev and Melvin find a positive relationship between latent volatility and observed spreads on the Deutschemark/dollar exchange market.

In this paper, we examine the link between volatility and bid-ask spreads in the ABS market. We estimate latent volatility of a bond price process using the time duration between trades (Engle and Russell (1997)). We then examine the relationship between bid and ask quotes and the underlying volatility of the bond price process. Examining the ten most frequently traded bonds on the ABS system, we find results that suggest latent volatility and observed spreads are either significantly negatively related or insignificantly related. While our results say nothing about liquid markets (like foreign exchange and equities), they do raise doubts about the value of volatility as a proxy for asymmetric information or adverse selection cost in markets like the ABS that are characterized by thin trading. We discuss how the lack of liquidity and the presence of a relatively large percentage of uninformed traders might generate these results.

The rest of the paper is developed as follows. Section I discusses the organization and microstructure of the ABS market. Section II presents the analytical framework to understand

---

[2] A small sample of high yield bond transactions (data on 67 bonds) beginning in 1994 is available on NASDAQ's Fixed Income Pricing System (FIPS). See Alexander, Edwards, and Ferri (1999) and Hotchkiss and Ronen (1999) for

the determinants of observed spread in the ABS market. Section III presents the empirical models used in our study. Section IV discusses the data used in the paper. Section V presents the empirical implementation and results from the study. Section VI concludes the paper.

I. The Automated Bond System

The NYSE maintains a fully automated electronic trading and information system[3] for bonds known as the Automated Bond System (ABS). Unlike its counterpart stock market, there is no specialist in the NYSE bond market. Instead, there are brokers who are subscribing members of the ABS. There are 58 ABS member brokers operating on about 210 terminals. The member brokers usually trade on behalf of their customers, though at times they could trade for their own account. Member brokers receive limit orders from the public and enter the corresponding bid-ask quotes and the respective quantities into the automated system. They also enter their own quotes into the system. Liquidity to the ABS market is therefore jointly supplied by public limit orders and dealers' own quotes. The ABS matches the orders automatically and informs the member brokers once an order is executed. The ABS is thus a limit order market with a strict price-time priority. The ABS market is also very transparent. All subscribers to the ABS market have full access to the complete order schedule, which they can divulge to investors upon request.

The ABS market is an outlet for retail trades and odd lot trading with an average trade size close to 20 bonds[4], and a median trade size of ten bonds. Most institutional trading in corporate bonds occurs in the dealer market. This means that trades initiated by uninformed

a description and analysis of this data. Only hourly reports are currently available for analysis.

[3] Further details of the ABS market are described in Hong and Warga (1999).

[4] The " nine-bond rule" obligates member brokers to direct trades having a size of 9 or less bonds to the ABS market. With few exceptions, corporate bonds traded on the ABS have par values of $1000.

(liquidity) traders are likely to be concentrated on the ABS. These "uninformed" trades are less likely than other trades to cause any revision of expectations among dealers.

The average time interval between two transactions for the most actively traded bond in our sample is seven minutes, while for the tenth most actively traded bond it is thirty-eight minutes. The distribution of the time intervals between trades (also referred to as durations[5]) is skewed to its right, i.e. long durations are more likely than short durations. Transaction and quote data on the ABS market are also irregularly spaced. We are also likely to observe long (short) trading intervals followed by long (short) trading intervals, a phenomenon referred to as duration clustering.

II. Transactions, Volume and Volatility

Empirical evidence, such as Bollerslev and Melvin (1994) indicates a positive relationship between volume and volatility. A possible explanation could be that trading is generated due to asymmetric information that results from either differential information or differences in opinion. The size of trades or volume reflects the extent of disagreement among traders about a security's value. The larger the disagreement the larger will be the price changes and higher will be the price volatility and trading volume. This forms the basis for adage that "it takes volume to move prices".

Another explanation for a positive relationship between volume and volatility is that they are both jointly determined by the number of information events, as postulated by the Mixture of Distributions Hypothesis (MDH). According to the MDH, volume and volatility are positively correlated only because they are positively related to the number of information events that serves as a mixing variable. The MDH implies that volume does not have any

[5] The term duration in this context should not to be confused with Macaulay's duration, which is the negative of the price elasticity of a bond with respect to a change in interest rates.

explanatory power in volatility beyond what is contained in the mixing variable, a result that is verified by Jones, Kaul and Lipson (1994).

In this paper, we study the relationship between volatility and spreads for individual bonds. Hence, we would like to ensure that our volatility measure captures price movements only and does not capture volume. Volatility, in our paper, is measured as an inverse function of the time interval between two successive trades (details in section III.A.). When the time interval between two successive trades is short, it implies that there is new information that is flowing into the market rapidly. This implies a high volatility. Conversely, when the time interval between two successive trades is long, there is no new information that is being released into the market. This, in turn, implies a low volatility. Our volatility measure, therefore, is related to the occurrence of trade as described by the MDH. Volatility goes up in response to new market information and not in response to a higher trading volume. As discussed in the results section below, our tests for volume reveal that it has no explanatory value in the ABS market, and is uncorrelated with our volatility measure. This makes sense in the ABS market because only relatively low volume trades occur in the first place, with most of the volume (although not necessarily frequency) taking place in the dealer market.

III. Models for Irregularly Spaced Data

A. The Autoregressive Conditional Duration Model

Quotes in the ABS market do not arrive in equal intervals. This raises the question of how we estimate latent volatility for unevenly spaced data. We can aggregate the data over a fixed time interval and use a GARCH model as in Bollerslev and Melvin (1994). If we choose too short an interval there may be many intervals with no new information and we may introduce some form of heteroscedasticity in the data. If we, however, choose too long an interval then features of the data will be smoothed and potentially hidden and microstructure aspects of the data will be lost. These problems are certain to be severe in the ABS market, at least relative to the markets where aggregating data over a fixed interval of time has been successfully applied (e.g. foreign exchange or equities).

We mitigate the problem by using an alternative method to obtain latent volatility of the bond price process. Recently, Engle and Russell (1997,1998) introduced the Autoregressive Conditional Duration (ACD) model to analyze irregularly spaced data. The ACD model focuses on the intertemporal correlations of the durations, where "durations" refer to the time interval between arrivals. Instead of aggregating the data to some fixed interval, the ACD model treats the arrival times of the data as a point process with an intensity defined conditional on past activity. The ACD model corrects for duration clustering in the data[6]. Duration clustering refers to the phenomenon where long (short) durations are followed by long (short) durations.

Let the sequence of time arrival of successive quotes be represented as {t_i for i = 1,2… n.}, where t_i refers to the arrival time of the ith quote. The stochastic process for {t_i for i = 1,2… n.} is called a point process. Corresponding to the point process is the counting process {N(t), t>0}, which is the number of events that have occurred by time t. The conditional intensity of the point process (or the hazard rate), refers to the instantaneous probability of an event conditional on past information and is defined as follows.

$$l(t; N(t), t_1, \ldots\ldots, t_{N(t)}) = \lim_{\Delta t \to 0} \frac{P(N(t+\Delta t) > N(t) | N(t), t_1, \ldots\ldots, t_{N(t)})}{\Delta t} \quad (1)$$

[6] The GARCH model corrects for volatility clustering in the data. Considering that durations are inversely related to price volatility, the ACD model also indirectly corrects for volatility clustering.

The conditional intensity (hazard rate) uniquely defines a point process. Engle and Russell in their ACD model introduce a new family of point processes and corresponding conditional intensities.

Let x_i refer to the duration between i th and i-1 th quote. i.e $x_i = t_i - t_{i-1}$. According the ACD(m,q) model, we have

$$x_i = \psi_i \varepsilon_i$$

$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{t-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} \text{ for } \alpha_j, \beta_j \geq 0, \omega > 0, \text{for all i, } i = 1,.......N \tag{2}$$

where ψ_i is the conditional expected duration given past information and parameter set θ i.e.

$$\psi_i = E[x_i \mid x_{i-1,}........., x_1; \theta] \tag{3}$$

and where m and q refer to the orders of the lags used and { ε_i } is the error term that is i.i.d and follows a specified distribution. Under an exponential distribution for the error term, long durations are as likely to occur as short durations. However, in our bond data we find that long durations occur more often than short durations, a phenomenon that can be captured using a Weibull distribution.

Under a Weibull distribution, the likelihood function of the ACD(1,1)[7] model is defined as

[7] The conditional intensity of the ACD model i.e. $\lambda(t, N(t), t_1, \ldots.. t_{N(t)})$ depends upon the distribution of the error term and specification of ψ_i.

$$L(q) = \sum_{i=1}^{N(T)} \ln\left(\frac{g}{x_i}\right) + g\ln\left(\frac{\Gamma(1+\frac{1}{g})\, x_i}{y_{\,i}}\right) - \left(\frac{\Gamma(1+\frac{1}{g})\, x_i}{y_{\,i}}\right)^{g}$$

where

$y_{\,i} = w + a x_i + b y_{\,i-1}$, for i > 1

$y_{\,i} = \dfrac{w}{(1-b)}$ for i = 1

$q = (g, w, a, b,),\ a, b \geq 0, w > 0,\ a + b < 1$

Γ is a Gamma distribution

(4)

The log likelihood function in (4) is maximized iteratively with respect to parameters in θ and subject to the non-negativity and stationarity conditions above[8].

The ACD model gives us the latent conditional duration of the quotes. Following Engle and Russell the conditional duration can be in turn used to construct latent volatility as follows:

- Define a new price process as an average of bid-ask prices. The new price process does not have the bid-ask bounce in it.
- Next, 'thin' the price process by keeping a quote only if its average price differs from the average price of the previous quote by at least a constant c. For each bond, the constant c is set equal to the spread with the highest frequency for that bond. This turns out to be 1/8 for all the bonds. The purpose of excluding quotes whose average prices have moved within ± 1/8 is to exclude possible noisy quotes and to include only those quotes that have significant information embedded in them. We refer to the new price process as the thinned price process.

[8] Maximum likelihood estimation is done using the BFGS algorithm in GAUSS.

- Deseasonalize the thinned price durations for any possible seasonal (time-of-the day and day-of-the week) patterns.
- Choose an ACD(1,1) model with Weibull distribution (WACD(1,1)) as the benchmark model for the data[9]. Fit a WACD(1,1) model for the thinned price process and obtain the conditional expected price durations (ψ_i) .
- The ψ_i {i = 1,2,….,N} are a measure of time per unit price change. The inverse of ψ_i represents the price change per unit of time, and this is a measure of volatility. Engle and Russell (1998) show that volatility can be estimated as

$$s_i^2 = \frac{c^2}{y_i} \qquad (5)$$

B. Ordered Probit Model

Spreads in the ABS market are in increments of eighths, and therefore do not have continuous support. We, therefore, employ an ordered probit model (Hausman et. al. (1992) and Bollerslev and Melvin (1994)) to study the relationship between spreads and volatility in the ABS market.

As we shall explain later in section IV, spreads observed in the bond market seem to belong to a few distinct groups. The observed spreads take on a fixed number of discrete values and can be explained by a set of predetermined values that have a continuous support. The observed spread is a function of a latent spread that is continuous and is itself a function of lagged volatility and lagged spread. Specifically, our ordered probit model can be written as

---

[9] WACD(1,1,) is a fairly general specification and a good starting point for any specification search (Engle and Russell (1997)).

$$k_t^* = d_0 + d_1 h_{t-1} + d_2 k_{t-1} + e_t$$
$$e_t \sim N(0,1) \qquad (6)$$

(all the variables are expressed in log form) where the conditional volatility of the bond price (h_{t-1}) and lagged value of observed spread (k_{t-1}) jointly explain the conditional mean and variance of k^*_t, the unobserved spread. Furthermore, the ordered probit model requires that there is a one to one mapping between the discrete observed spread (k_t) and continuous unobserved spread (k^*_t)

i.e.
$$k_t = a_j \text{ iff } k_t^* \in A_j,\ j = 1,2,........,m \qquad (7)$$

where the a_j's are a set of discrete values that k_t can take on and where A_j's are a set of m ordered disjoint intervals that k^*_t can be partitioned into. We will see below in the data section that an adequate and parsimonious number of partitions "m" will be equal to 4 for all bonds with the exception of one, for which we will choose m=3.

IV. Data

The bid-ask quote data on corporate bonds covers a period of approximately five months extending from 10/01/96 to 02/21/97. The data contains bonds that are traded on the NYSE from 9.30 a.m. to 4:00 p.m. every trading day. We present a sample of quotes for one bond (issued by Penn Traffic Co.) in Table I.

Insert Table I here

Next, we determine the bonds with the highest number of trades and quotes. We sort all bonds by the total number of transactions and separately by the number of quotes. The top ten bonds in the intersection of these two sortings are presented in Table II and a description of each bond in Table III. As discussed above, transaction frequency drops off markedly as we move beyond the top ten traded bonds, and so we restrict our attention to these instruments. All of the bonds examined fall in the non-investment grade category (below Baa in rating).[10]

Insert Tables II, III here

For the ordered probit analysis, the observed spread has to be grouped into a finite number of ordered categories. The first two columns in Table IV present the frequency distribution of spreads for the top ten bonds. The next two columns present the groups and the respective spread intervals we adopt for each bond.

V. Results

We adopt the following scheme in implementing our study. For each of the ten bonds,

1. Obtain price durations from the quotes.
2. Obtain thinned price durations.
3. Deseasonalize price durations for daily and weekly effects. We employ 3 dummy variables[11]to account for time-of- the day effects and 5 dummy variables to account for day of the week effects. In all we have 15 dummy variables (3 time- of- the day dummies times 5 day-of-the week dummies). We deseasonalize by a) regressing the thinned price durations on 15 dummy variables and b) dividing the thinned price durations by the regression estimate from step a.
4. Fit a WACD(1,1) model to the deseasonalized thinned price durations.
5. Extract latent conditional expected durations and duration based annual price volatilities using the WACD(1,1) estimates.
6. Fit an ordered probit model using the groups specified in last two columns of Table IV.

---

[10] The bond labeled "unrated" was originally issued with a high yield rating as well.

[11] The first dummy variable indicates the morning interval from 9.30 a.m.-12 a.m., the second dummy variable indicates the interval from 12 a.m. - 2 p.m. and the third dummy variable indicates the interval from 2.pm to 4 p.m. on

These steps were described in section III above. Table V presents the WACD(1,1) estimates. Table VI presents ordered probit estimates corresponding to the Table V inputs.

Insert Tables V, VI here

We have the following results from Table V:

- The WACD(1,1) estimates for all bonds are highly significant at 5% level[12]. This indicates that there is a significant presence of duration clustering in the data.
- The estimate for γ for all the bonds is significantly lower than one. This indicates that long durations are more likely than short durations.

We have the following results from Table VI:

Examining the p-values for mean coefficients in the unobserved spread equation, we have three groups of bonds. Our main interest is in the sign and statistical significance of the parameter δ_1, which determines the relationship between bid-ask spreads and latent volatility[13]:

Group A:

This group includes bonds where we find a negative and statistically significant relationship between the bid-ask spread and latent volatility of the bond price process. The six bonds in this group are PNF 05, CLAR 02, BBY 00, PCS 03, AGY 04, and SME 01.

Group B:

This group includes bonds where we find a negative but not statistically significant relationship between the bid-ask spread and latent volatility of the bond price process. The two bonds in this group are SME 04 and STO 01.

---

a week day. NYSE bonds have a U shaped trading pattern over the day. Our three dummy variables correspond to the three time intervals of the U shape.

[12] For some bonds, the constant term in the conditional duration equation is significant at 10% level.

[13] The partition coefficients for the ordered probit model (not reported) are all highly significant.

Group C:

This group includes bonds where we find a positive but not statistically significant relationship between the bid-ask spread and latent volatility of the bond price process. Bonds in this group are HDS 03 and WHX 03.

As mentioned in a previous section, we also want to examine the relationship between volume, volatility and spreads using the ordered probit model for individual bonds. One of our arguments for employing the ACD model for volatility rests on the fact that it isn't proxying for volume. A major qualification is that tests for volume can only be conducted on a subset of the quote data corresponding to actual transactions. The identification procedure allowing us to associate a particular transaction with a particular quote is inexact, and this leads to a further loss of data. Sample sizes per bond decline by between fifty and seventy percent. In results not reported here, we find that volume has no significant explanatory power for spreads. More importantly, we find no significant correlation between trading volume and volatility[14].

Microstructure theory decomposes the bid-ask spread into adverse selection costs, inventory, and order processing costs. In the present context, brokerage fees that typically accompany retail-sized orders from brokerage customers can hide order processing costs and inventory costs. It is also quite possible that adverse selection costs are minimal because the thinness of the market does not attract market makers wishing to trade on the basis of information flow. The bid-ask spread narrowing in reaction to increased volatility observed here may be a result of the fact that the market we are observing exists under the "shadow" of the larger dealer market. Since it is likely that liquidity-based trading is dominating the ABS market, liquidity-based effects not captured in the classical bid-ask spread decomposition may lead to the result that higher volatility simply reflects increased liquidity for a bond.

Our results suggest that there are circumstances when employing volatility as a proxy for adverse selection costs can be a mistake. For example, Krishnaswamy and Subramaniam (1999) and Krishnaswamy et al (1999) both use residual volatility in daily stock returns as a proxy for information asymmetry for each firm. The residual volatility is defined as the standard deviation of residuals obtained from market adjusted daily stock returns for a given firm. If the firm's managers and investors are equally well informed about the economy wide factors influencing the firm's value, then the residual volatility in a firm's stock returns captures the information asymmetry between managers and investors about the specific firm. In other words, residual volatility captures firm specific uncertainty that remains after removing the uncertainty that is common to the firm's managers and the investors from total uncertainty. Firms with higher information asymmetry about their cash flows and value would have higher residual volatility in their stock returns. However, as the authors note, residual volatility, in fact, has two components: the information asymmetry component and a market innovation component. To the extent we ignore the market innovation component, the residual volatility overestimates the information asymmetry component. In the present context, we are examining a market where liquidity effects may very well dominate both information asymmetry and market innovations.

VI. Conclusions

In this paper, we examine the relationship between quoted bid-ask spreads and volatility for the ten most actively traded corporate bonds on the NYSE’s Automated Bond System

[14] These results are available from the authors upon request. The power to detect a relationship between volatility and spread also declines, but the basic pattern observed in the results presented here on the full sample are maintained.

(ABS). We find a negative and statistically significant relationship between volatility and observed spreads for six out of the ten bonds in our sample. The other four bonds reveal no statistically significant relationship, with two of them negative and two of them positive.

The type of market we observe probably drives our results. The ABS is a retail- and odd lot-driven electronic exchange. Bid-ask spreads appear to be dominated by liquidity effects that are not captured in the classical decomposition provided by microstructure theory. Our results point to the need to develop a decomposition of bid-ask spreads that encompass liquidity effects. Even very liquid markets have periods where individual securities require long time intervals for transactions to take place. The thinness of the ABS market allows us to see that lack of liquidity may lead to a relationship between volatility and spread that is counter to the one that existing decompositions would suggest.

There continues to be no evidence of a positive relationship between volatility and spread.

References:

Alexander, G.J., A.K. Edwards and M.G. Ferri, 1999. "Trading Volume and Liquidity in Nasdaq's High Yield Bond Market", Manuscript, Security and Exchange Commission, Office of Economic Analysis.

Bollerslev, T. and M. Melvin, 1994 " Bid-ask spreads and volatility on the foreign exchange market, Journal of International Economics, 36, 355-72

Dueweke, D., Hyland, M. and F. Siesel, "Measuring the New York Stock Exchange's Share of Corporate Trading Volume." Extra Credit, The Journal of High Yield Research, September/October 1992, 9-16. Merrill Lynch Global Securities Research & Economics Group.

Engle, R. and J. Russell, 1997, "Forecasting the frequency of changes in quoted foreign exchange prices with the ACD model", Journal of Empirical Finance, 4, 187-212

Engle, R. and J. Russell, 1998, "ACD: a new model for irregularly space transaction data", Econometrica, Vol. 66, 5, 1127-62

Glosten, L, 1987," Components of the bid-ask spread and the statistical properties if transaction prices" Journal of Finance, 42, 1293-1307

Glosten, L. and L. Harris, 1988, " Estimating the components of bid-ask spread" Journal of Financial Economics, 21, 123-42

Hausman, J., A. Lo and C. MacKinlay, 1992 "An ordered probit analysis of transaction stock prices", Journal of Financial Economics, 31, 319-79

Hong, G. and A. Warga, 1999, "An empirical study of bond market transactions", Forthcoming, Financial Analysts Journal.

Hotchkiss, E. and T. Ronen, 1999. "Informational links between bond and stock markets: An intradaily analysis". Manuscript, Boston College.

Huang, R. and H. Stoll, 1997 "The components of bid-ask spread: A general approach" Review of Financial Studies, Vol.10, 4, 995-1034

Jones, C., G. Kaul, and M. Lipson, 1994, "Transactions, Volume and Volatility" Review of Financial Studies", Vol. 7, 4, 631-51

Krishnaswamy, S. and Subramaniam, V. 1999, "Information asymmetry, valuation, and the corporate spin-off decision", Journal of Financial Economics, 53, 73-112

Krishnaswamy, S., Spindt, P. and Subramaniam, V. 1999, "Information asymmetry, monitoring, and the placement structure of corporate debt", Journal of Financial Economics, 51, 403-434

Lin, J., G. Sanger and G. Booth, 1995, “Trade size and components of bid-ask spread” Review of Financial Studies, Vol8, 4, 1153-83

Roll. R, 1984, “ A simple effective measure of the effective bid-ask spread in an efficient market” Journal of Finance, 4, 1127-39

Table I

Sample quotes on the ABS screen. The sample bond is issued by Penn Traffic Company. The bid and ask prices with the respective quantity of bonds bid are reported.

Ticker Symbol	Date	Time	Bid Price	# of bonds	Ask Price	# of bonds
PNF 05	961001	102834	67.625	50	68	30
PNF 05	961001	102850	67.625	50	68	20
PNF 05	961001	103510	67.625	50	67.75	16
PNF 05	961001	104142	67.625	40	67.75	16
PNF 05	961001	111056	67.625	40	67.75	41
PNF 05	961001	111129	67.625	20	67.75	41
PNF 05	961001	111146	67.625	5	67.75	41
PNF 05	961001	111151	67	67	67.625	15
PNF 05	961001	111213	67	67	67.5	10
PNF 05	961001	111228	67.125	20	67.5	10
PNF 05	961001	111229	67.125	35	67.5	10
PNF 05	961001	111305	67.25	5	67.5	10
PNF 05	961001	112228	67.5	20	67.625	15
PNF 05	961001	112303	67.5	4	67.625	15
PNF 05	961001	121915	67.5	4	67.75	45
PNF 05	961001	134212	67.5	4	67.75	95
PNF 05	961001	135033	67.5	4	67.75	105
PNF 05	961001	140218	67.375	25	67.5	6
PNF 05	961001	140240	67.5	14	67.75	105
PNF 05	961001	141735	67.5	14	67.75	80
PNF 05	961001	141746	67.5	14	67.75	60
PNF 05	961001	141755	67.5	14	67.75	10
PNF 05	961001	141804	67.5	14	68	20
PNF 05	961001	143520	67.5	34	68	20
PNF 05	961001	144308	67.5	34	68	10

Table II

The top ten bonds along with the respective number of trades[15], quotes and trade matched quotes are listed below.

	Ticker Symbol	CUSIP #	Quotes	Trades
1	PNF 05	707832AD	10597	6035
2	SME 04	817587AC	6727	4586
3	CLAR 02	180476AA	6610	3466
4	BBY 00	086516AB	5756	3095
5	PCS 03	704378AD	4563	2226
6	STO 01	861589AK	3892	2932
7	HDS 03	431692AA	2706	1260
8	WHX 03	963142AH	2344	1343
9	AGY 04	040228AE	2339	1102
10	SME 01	817587AD	2255	955

Table III

Bond Descriptors

	Company Name	CUSIP #	Coupon	Maturity	Rating
1	Penn Traffic Co.	707832AD	9 5/8	4/15/2005	Caa2
2	Service Merchandise Inc.	817587AC	9	12/15/2004	B2
3	Claridge Hotel	180476AA	11 ¾	2/1/2002	Ca
4	Best Buy Co.	086516AB	8 5/8	10/1/2000	B2
5	Payless Cashways Inc.	704378AD	9 1/8	4/15/2003	B3
6	Stone Container Corp.	861589AK	9 7/8	2/1/2001	B1
7	Hills Stores Co.	431692AA	10 ¼	9/30/2003	Unrated
8	Wheeling-Pittsburgh Corp.	963142AH	9 3/8	11/15/2003	B1
9	Argosy Gaming Co.	040228AE	13 ¼	6/1/2004	B1
10	Service Merchandise Inc.	817587AD	8 3/8	1/15/2001	Ba

[15] We have excluded all quotes with negative spreads. There are 2-4 such quotes for every bond.

Table IV

Frequency table for spreads for the top ten bonds. The first three columns indicate the frequencies and percentage of frequencies for the indicated spread, for each bond. The last two columns indicate the number of ordered groups along with the respective spreads used for ordered probit model.

Spreads	# of quotes	% of quotes	Groups	Spreads
		PNF 05		
0.125	554	16.1563138	1	0.125-0.25
0.25	408	11.8985127	2	0.375-0.5
0.375	379	11.0527851	3	0.625-0.875
0.5	457	13.3275007	4	>0.875
0.625	277	8.0781569		
0.75	297	8.66141732		
0.875	198	5.77427822		
1	196	5.71595217		
(1.0 - 2.0]	516	15.048119		
(2.0 - 3.0]	94	2.741324		
(3.0 - 4.0]	23	0.67074949		
(4.0 - 5.0]	11	0.32079323		
>5	19	0.5540974		
		SME 04		
0.125	1871	27.8132897	1	0.125
0.25	1733	25.7618552	2	0.25
0.375	1173	17.4371934	3	0.375-0.5
0.5	797	11.8477776	4	>0.5
0.625	374	5.55968485		
0.75	262	3.89475249		
0.875	145	2.15549279		
1	114	1.6946633		
(1.0 - 2.0]	231	3.433923		
(2.0 - 3.0]	20	0.29730935		
(3.0 - 4.0]	2	0.02973094		
(4.0 - 5.0]	2	0.02973094		
>5	3	0.044596403		
		CLAR 02		
0.125	829	12.5397065	1	0.125-0.25
0.25	1049	15.8674936	2	0.375-0.5
0.375	874	13.2203903	3	0.625-1
0.5	764	11.5564967	4	>1
0.625	590	8.92451974		
0.75	558	8.44047799		
0.875	454	6.86734231		
1	384	5.80850098		
(1.0 - 2.0]	940	14.2187264		
(2.0 - 3.0]	122	1.84540917		
(3.0 - 4.0]	21	0.3176524		

(4.0 - 5.0]	15	0.22689457		
>5	11	0.16638935		
		BBY 00		
0.125	968	16.8172342	1	0.125-0.25
0.25	1208	20.9867964	2	0.375-0.5
0.375	958	16.6435024	3	0.625-2
0.5	835	14.5066018	4	>2
0.625	468	8.13064628		
0.75	378	6.56706046		
0.875	253	4.39541348		
1	167	2.90132036		
(1.0 - 2.0]	459	7.9742877		
(2.0 - 3.0]	41	0.71230021		
(3.0 - 4.0]	9	0.15635858		
(4.0 - 5.0]	5	0.08686588		
>5	7	0.12161223		
		PCS 03		
0.125	676	14.8148148	1	0.125-0.25
0.25	763	16.7214552	2	0.375-0.5
0.375	777	17.0282709	3	0.625-1
0.5	715	15.6695157	4	>1
0.625	470	10.3002411		
0.75	349	7.64847688		
0.875	216	4.73372781		
1	172	3.76944992		
(1.0 - 2.0]	396	8.67850099		
(2.0 - 3.0]	23	0.50405435		
(3.0 - 4.0]	1	0.02191541		
(4.0 - 5.0]	3	0.06574622		
>5	2	0.04383081		
		STO 01		
0.125	2205	56.6546763	1	0.125
0.25	1179	30.2929085	2	0.25
0.375	287	7.37410072	3	>0.25
0.5	100	2.56937307		
0.625	53	1.36176773		
0.75	24	0.61664954		
0.875	14	0.35971223		
1	14	0.35971223		
(1.0 - 2.0]	15	0.38540596		
(2.0 - 3.0]	1	0.02569373		
(3.0 - 4.0]	0	0		
(4.0 - 5.0]	0	0		
>5	0	0		
		HDS 03		
0.125	177	6.55070318	1	0.125-0.375
0.25	276	10.2146558	2	0.5-0.75

0.375	238	8.80829016	3	0.875-2.0
0.5	385	14.2487047	4	>2.0
0.625	174	6.43967432		
0.75	186	6.88378979		
0.875	156	5.77350111		
1	180	6.66173205		
(1.0 - 2.0]	620	22.945966		
(2.0 - 3.0]	180	6.66173205		
(3.0 - 4.0]	77	2.84974093		
(4.0 - 5.0]	15	0.55514434		
>5	38	1.40636566		
		WHX 03		
0.125	596	25.4266212	1	0.125
0.25	588	25.0853242	2	0.25
0.375	405	17.278157	3	0.375-0.5
0.5	323	13.7798635	4	>0.5
0.625	155	6.61262799		
0.75	95	4.05290102		
0.875	62	2.64505119		
1	34	1.45051195		
(1.0 - 2.0]	79	3.37030717		
(2.0 - 3.0]	7	0.29863481		
(3.0 - 4.0]	0	0		
(4.0 - 5.0]	0	0		
>5	0	0		
		AGY 04		
0.125	163	6.96879008	1	0.125-0.375
0.25	232	9.9187687	2	0.5-0.75
0.375	231	9.87601539	3	0.875-1.0
0.5	264	11.2868747	4	>1
0.625	174	7.43907653		
0.75	214	9.14920906		
0.875	197	8.42240274		
1	165	7.05429671		
(1.0 - 2.0]	506	21.6331766		
(2.0 - 3.0]	131	5.60068405		
(3.0 - 4.0]	34	1.45361265		
(4.0 - 5.0]	3	0.12825994		
>5	25	1.06883283		
		SME 01		
0.125	136	6.03104213	1	0.125-0.375
0.25	285	12.6385809	2	0.5-0.75
0.375	235	10.421286	3	0.875-1.0
0.5	277	12.2838137	4	>1
0.625	161	7.13968958		
0.75	200	8.8691796		
0.875	154	6.82926829		
1	147	6.51884701		

(1.0 - 2.0]	504	22.3503326
(2.0 - 3.0]	117	5.18847007
(3.0 - 4.0]	28	1.24168514
(4.0 - 5.0]	6	0.26607539
>5	5	0.22172949

Table V
This table presents the maximum likelihood estimates for the WACD(1,1) model for the top ten bonds

	Quotes				
	log likelihood	g	w	a	b
PNF 05	-29987	0.64483	47.009	0.14523	0.80739
T stat		*83.750*	*5.1600*	*8.1113*	*34.851*
SME 04	-15759	0.69519	11.727	0.055598	0.93408
T stat		*58.316*	*2.3952*	*5.3680*	*77.341*
CLAR 02	-10998	0.61828	24.393	0.26247	0.7447
T stat		*50.117*	*3.4527*	*6.8995*	*24.241*
BBY 00	-14900	0.63751	29.701	0.10322	0.87042
T stat		*57.212*	*2.9593*	*5.1660*	*35.567*
PCS 03	-12778	0.61356	38.149	0.057043	0.90592
T stat		*52.511*	*1.9078*	*2.6247*	*23.944*
STO 01	-6577.6	0.67456	42.272	0.13257	0.83210
T stat		*36.926*	*2.1364*	*3.8959*	*(23.433)*
HDS 03	-6176.5	0.59636	114.83	0.25161	0.65924
T stat		*36.791*	*1.7348*	*3.0297*	*5.3000*
WHX 03	-5344.4	0.62930	64.372	0.031918	0.90773
T stat		*32.599*	*0.55568*	*0.90782*	*6.5023*
AGY 04	-6362.9	0.61899	113.74	0.12118	0.76654
T stat		*36.915*	*2.3012*	*3.0701*	*10.407*
SME 01	-6811.0	0.58249	85.784	0.086505	0.83147
T stat		*38.212*	*1.8954*	*2.2608*	*11.863*

Notes: The likelihood function of the WACD(1,1) model is given by

$$L(q) = \sum_{i=1}^{N(T)} \ln\left(\frac{g}{x_i}\right) + g\ln\left(\frac{\Gamma(1+\frac{1}{g})\, x_i}{y_i}\right) - \left(\frac{\Gamma(1+\frac{1}{g})\, x_i}{y_i}\right)^{g}$$

where

$y_i = w + a x_i + b y_{i-1}$, for i > 1

$y_i = \frac{w}{(1-b)}$ for i = 1

$q = (g, w, a, b,),\ a, b \geq 0, w > 0,\ a + b < 1$

Γ is a Gamma distribution

(4)

The log likelihood function in (4) is maximized iteratively with respect to parameters in θ^1.

- The log likelihood values are reported using the optimal parameter estimates.

Table VI
Maximum likelihood estimates for the ordered probit model

		Quotes		
		d_0	d_1	d_2
PNF 05		-2.9734812	-0.3279581	-0.7891552
	Std. Error	*0.336274*	*0.064345*	*0.025239*
	p value	*0.0001*	*0.0001*	*0.0001*
SME 04		2.2318643	-0.164741	-0.6526056
	std. Error	*0.660119*	*0.124688*	*0.036027*
	p value	*0.0007*	*0.1864*	*0.0001*
CLAR 02		-2.3285669	-0.2164398	-0.5077492
	std. Error	*0.336413*	*0.064521*	*0.036208*
	p value	*0.0001*	*0.0008*	*0.0001*
BBY 00		-2.5105426	-0.269538	-0.7891176
	std. Error	*0.577093*	*0.109402*	*0.037253*
	p value	*0.0001*	*0.0137*	*0.0001*
PCS 03		-3.1117839	-0.3995745	-0.4982693
	std. Error	*1.004133*	*0.189267*	*0.037753*
	p value	*0.0019*	*0.0348*	*0.0001*
STO 01		-2.3363578	-0.2436609	-0.6870954
	std. Error	*0.927175*	*0.175453*	*0.080013*
	p value	*0.0117*	*0.1649*	*0.0001*
HDS 03		-0.8685193	0.01372671	-0.6683769
	std. Error	*0.803229*	*0.15132*	*0.048008*
	p value	*0.2796*	*0.9277*	*0.0001*
WHX 03		0.19543474	0.1941718	-0.0434465
	std. Error	*3.901435*	*0.728514*	*0.058579*
	p value	*0.96*	*0.7898*	*0.4583*
AGY 04		-3.9196685	-0.6744519	0.79172219
	std. Error	*1.366998*	*0.256642*	*0.052428*
	p value	*0.0041*	*0.0086*	*0.0001*
SME 01		-2.7411886	-0.4446156	0.33550559
	std. Error	*1.429023*	*0.268166*	*0.045817*
	p value	*0.0551*	*0.0973*	*0.0001*

Notes:

- The estimated ordered probit model is of the form

$$k_t^* = d_0 + d_1 h_{t-1} + d_2 k_{t-1} + e_t$$

$$e_t \sim N(0,1)$$