Bid-Ask Spread, Volatility, and Volume in the Corporate Bond Market

MADHU KALIMIPALLI AND ARTHUR WARGA

MADHU KALIMIPALLI
is an assistant professor of finance at the School of Business and Economics at Wilfrid Laurier University in Waterloo, Ontario.
mkalimip@wlu.ca

ARTHUR WARGA
is Dean and Judge James A. Elkins Endowed Chair of Banking and Finance at the University of Houston's Bauer College of Business in Texas.
warga@uh.edu

This article provides the first microstructure analysis of the time series relationship between bid-ask spreads and volatility and volume in the market for corporate bonds. The market we examine is the New York Stock Exchange, which maintains the Automated Bond System (ABS). ABS is a fully automated electronic trading and information system with completely transparent schedules of bid and ask prices. 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. In general, trading on ABS is relatively thin, with typically retail-sized trades (under 100 bonds).

The National Association of Security Dealers (NASD) is in the process of introducing an experimental hourly reporting system for 500 large-issue corporate bonds through its Bond Market Transparency Initiative. It will still be some time before transaction or quotation streams are available for a broad sample of corporate bonds that will permit the kind of detailed examination of price dynamics we carry out.

Despite the size of the ABS market, the most actively traded issues not only rival the dealer market in terms of both frequency of trade and dollar volume of trade, but also in some cases even dominate the dealer market.[1] This means we can get an early start on validating procedures for measuring volatility in bond markets, and also establish some benchmark results to eventually permit a comparison with the dealer market.

We examine the time series relationships for the ten most actively traded bonds in the ABS market. The ABS market, unlike more liquid markets, is characterized by long non-trading intervals and time-clustering of quotes. For markets like ABS (and perhaps bond markets in general), estimation of generalized autoregressive conditional heteroscedasticity (GARCH)-based conditional volatilities would potentially entail a severe loss of information. This is because GARCH estimates are constructed by first transforming data to regularly spaced intervals, which could mask features of the data. Therefore we base our volatility measure on the time duration between trades (Engle and Russell [1998]).

We also explore the relationship between trading volume and bid-ask spreads to ensure that our volatility measure is not capturing the independent effect of trading volume. Finally, to check the robustness of our results, we examine GARCH-based volatilities.

For the ten most frequently traded bonds on the ABS, we find results that support the theory. Latent volatility and observed spreads are largely positively related. Volume is significantly negatively related to spreads. Our results are robust to both duration- and GARCH-based volatilities, although the duration-based approach uses more information and admits more observations that are apparently lost in data aggregation for GARCH models.

I. LITERATURE

Other researchers have examined the links between volatility and volume and the bid-ask spreads in more liquid markets. 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. In a two-step process, they obtain GARCH estimates of the underlying volatility of the exchange rate process to use as inputs in an ordered probit model. Bollerslev and Melvin find a positive relationship between latent volatility and observed spreads on the deutschemark-dollar exchange market.

Using a similar framework, Gwilym, Clare, and Thomas [1998] find a positive relationship between bid-ask spreads for stock index options traded on the London International Financial Futures Exchange market and the volatility of the underlying stock market index. Bessembinder [1994] also finds spreads in the foreign exchange market to widen using GARCH-based return volatilities and proxies for liquidity costs. McInish and Wood [1992] report that NYSE equity spreads widen (narrow) with underlying volatility (trading volume and trade size) over time.

The market microstructure theory suggests that the bid-ask spread can be decomposed into a transitory component representing operating and inventory carrying costs and a permanent component representing adverse selection costs. Higher volatility of underlying asset returns implies two things for market makers. First, they face a higher inventory risk on account of portfolio imbalances resulting from uncertainty in the market; and second, they face a higher possibility of trading with informed traders. As a result, bid-ask spreads are wider.

Unlike volatility and spreads, the relationship between volume and spreads in a time series context is an open empirical question, as Lee, Mucklow, and Ready [1993] suggest. On the one hand, higher trading volume could imply a higher adverse selection component to a market maker, thereby leading to a wider spread or less market quote depth. On the other hand, as Harris and Raviv [1993] suggest, if there is a lack of consensus among market participants, and limit orders are arriving on both sides of the bid-ask spread, then spreads may narrow and quote depth might increase.

II. THE AUTOMATED BOND SYSTEM (ABS)

The NYSE maintains a fully automated electronic trading and information system for bonds known as the Automated Bond System. 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, although at times they could trade for their own accounts. 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. The "nine-bond rule" obligates member brokers to direct trades of nine or fewer bonds to the ABS market. With few exceptions, corporate bonds traded on the ABS have par values of $1,000.

Most institutional trading in corporate bonds occurs in the dealer market. This means that trades initiated by uninformed (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. Even for the ten most actively traded issues on ABS that we examine, retail-sized trades dominate even if the total volume of trade is similar to the dealer market.

Transaction and quote data on the ABS market are irregularly spaced. The average time interval between two quotes for the most actively traded bond in our sample is about nine minutes, while for the tenth-most actively traded bond it is about 36 minutes. The distribution of the time intervals between quotes (also referred to as durations) for a given bond is skewed to its right; i.e., long durations are more likely than short durations.[2] We also observe that long (short) quote intervals for a given bond are followed by long (short) quote intervals, a phenomenon referred to as duration-clustering.

As shown in Hong and Warga [2000], pricing on the ABS is similar to 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 record transactions only by the day they occur.[3]

III. MODELS FOR IRREGULARLY SPACED DATA

Autoregressive Conditional Duration Model

We measure volatility as an inverse function of the time interval between two successive quotations. When the time interval between two successive quotes is short, this implies there is new information flowing into the market rapidly, implying a high volatility. Conversely, when the time interval between two successive quotes is long, there is no new information being released to the market. This in turn implies a low volatility. Our volatility measure, therefore, is related to the occurrence of quotes.

Higher volatility by this measure could imply higher trading activity as described by the mixture of distributions hypothesis (MDH).[4] Using volume (as measured by average of bid and ask quote depth), we check to see if our volatility measure still has explanatory power.

Quotes in the ABS market do not arrive at equal intervals. This raises the question of how to 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]. Choice of too short an interval, however, might lead to many intervals with no new information and even some form of heteroscedasticity in the data. Choice of too long an interval might smooth out features of the data and lead to a loss of information. These problems are certain to be more severe in the ABS market than in markets where aggregating data over a fixed interval of time is successful (e.g., foreign exchange or equities).

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 durations, where "durations" refer to the time interval between quote 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 intensity conditional on past activity. The ACD model corrects for duration-clustering in the data.[5] Duration-clustering describes conditions where long (short) durations are followed by long (short) durations.

We describe the ACD model and our implementation of it more fully in the appendix.

Steps 1-4 describe the filters used to clean the data before estimating the ACD model:

1. Eliminate first quotes on a given day to remove any extra information that has accumulated since the last market close.
2. Define a new price process as an average of bid-ask prices. The averaging filters out bid-ask bounce in the data.
3. Next, filter out any noise in the data by "thinning" the price process: keeping a quote only if its average price differs from the average price of the previous quote by at least 1/8. This eliminates quotes that are repeated, and quotes where one side (bid or ask) has changed by the minimum tick of 1/8 and the other side remains unchanged. The motivation is to eliminate possibly 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*.
4. Deseasonalize the thinned price durations for any possible seasonal (time-of-the-day and day-of-the- week) patterns. We use three dummy variables to account for time-of-the-day effects and five dummy variables to account for day-of-the-week effects.[6] In all, we have 15 dummy variables (three time-of-the day dummies times five 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 a).

Using output from Step 4, we estimate the ACD model and obtain the latent conditional duration of the quotes in the next step:

5. Choose an ACD(1, 1) model with Weibull distribution [WACD(1, 1)] as the benchmark model for the data. Based on the estimate for γ, which is significantly < 1 in Equation (A-3) of the appendix, we choose the Weibull distribution for the error term ε_i. We choose the ACD(1, 1) structure on conditional duration ψ, as it is a fairly general specification and a good starting point for any specification search (Engle and Russell [1997]). Fit the WACD(1, 1) model for the thinned price process and obtain the conditional expected price durations (ψ_i).

Using the conditional duration from Step 5, we construct latent volatility in the next step:

EXHIBIT 1
Sample Quotes on the ABS Screen

Ticker Symbol	Date	Time	Bid Price	# of Bonds	Ask Price	# of Bonds
PNF 05	961001	102834	67.625	50	68.000	30
PNF 05	961001	102850	67.625	50	68.000	20
PNF 05	961001	103510	67.625	50	67.750	16
PNF 05	961001	104142	67.625	40	67.750	16
PNF 05	961001	111056	67.625	40	67.750	41
PNF 05	961001	111129	67.625	20	67.750	41
PNF 05	961001	111146	67.625	5	67.750	41
PNF 05	961001	111151	67.000	67	67.625	15
PNF 05	961001	111213	67.000	67	67.500	10
PNF 05	961001	111228	67.125	20	67.500	10
PNF 05	961001	111229	67.125	35	67.500	10
PNF 05	961001	111305	67.250	5	67.500	10
PNF 05	961001	112228	67.500	20	67.625	15
PNF 05	961001	112303	67.500	4	67.625	15
PNF 05	961001	121915	67.500	4	67.750	45
PNF 05	961001	134212	67.500	4	67.750	95
PNF 05	961001	135033	67.500	4	67.750	105
PNF 05	961001	140218	67.375	25	67.500	6
PNF 05	961001	140240	67.500	14	67.750	105
PNF 05	961001	141735	67.500	14	67.750	80
PNF 05	961001	141746	67.500	14	67.750	60
PNF 05	961001	141755	67.500	14	67.750	10
PNF 05	961001	141804	67.500	14	68.000	20
PNF 05	961001	143520	67.500	34	68.000	20
PNF 05	961001	144308	67.500	34	68.000	10

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

6. The $\psi_I\{I = 1, 2, \ldots, N\}$ is 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

$$\sigma_i^2 = \frac{c^2}{\psi_i} \tag{1}$$

Equation (1) gives us implied conditional volatilities, which along with corresponding bid-ask spreads and volume are used as inputs in the ordered probit model.[7]

Ordered Probit Model

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

Our ordered probit model can be written as:

$$\begin{aligned} k_t^* &= \delta_0 + \delta_1 h_t + \delta_2 k_{t-1} + \delta_3 Vol_t + \varepsilon_t \\ \varepsilon_t &\sim N(0,\ 1) \end{aligned} \tag{2}$$

(with variables expressed in log form), where the conditional volatility of the bond prices h_t (estimated as in Step 6), lagged value of observed spread (k_{t-1}), and volume (measured as logarithm of the average of bid and ask quantities) jointly explain the conditional mean and variance of k^*_t, the unobserved spread.

We assume a normal distribution for the error term in Equation (2). The unobserved spread can be partitioned into a set of m ordered disjoint intervals. A size of $m = 4$ seems to be appropriate for all bonds with the exception of one (STO 01), where we set $m = 3$.

IV. DATA

The bid-ask quote data on corporate bonds cover a period of approximately five months extending from October 1, 1996, through February 21, 1997. The data include bonds traded on the NYSE from 9:30 AM to 4:00 PM every trading day. We present a sample of quotes for one bond (issued by Penn Traffic Co.) in Exhibit 1.

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 sorts are presented in Exhibit 2. A description of each bond appears in Exhibit 3.

All the bonds examined are non-investment-grade bonds (below Baa).[8] Clearly, a data filter based on trade frequency limits our ability to generalize our results to all traded bonds, but in this case that is unavoidable.

Exhibit 4 reports trade volume data for five of the bonds in our sample that are also tracked by Nasdaq's Fixed Income Pricing System (FIPS). The sum of Nasdaq's

EXHIBIT 2
Top Ten Bonds

	Ticker Symbol	CUSIP #	Number of Quotes	Number of 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

Quotes with negative spreads excluded. There are 2-4 such quotes for every bond.

EXHIBIT 3
Bond Descriptors

	Company Name	Ticker Symbol	CUSIP #	Coupon	Maturity	Rating
1	Penn Traffic Co.	PNF 05	707832AD	9 ⅝	4/15/2005	Caa2
2	Service Merchandise Inc.	SME 04	817587AC	9-00	12/15/2004	B2
3	Claridge Hotel	CLAR 02	180476AA	11 ¼	2/1/2002	Ca
4	Best Buy Co.	BBY 00	086516AB	8 ⅝	10/1/2000	B2
5	Payless Cashways Inc.	PCS 03	704378AD	9 ⅛	4/15/2003	B3
6	Stone Container Corp.	STO 01	861589AK	9 ⅞	2/1/2001	B1
7	Hills Stores Co.	HDS 03	431692AA	10 ¼	9/30/2003	Unrated
8	Wheeling-Pittsburgh Corp.	WHX 03	963142AH	9 ⅜	11/15/2003	B1
9	Argosy Gaming Co.	AGY 04	040228AE	13 ¼	6/1/2004	B1
10	Service Merchandise Inc.	SME 01	817587AD	8 ⅜	1/15/2001	Ba

EXHIBIT 4
Trading Volumes 10/1/96-2/23/97

	SME 04	BBY 00	PCS 03	STO 01	WHX 03
ABS Par Volume	\$119,452,000	\$70,243,000	\$43,169,000	\$138,445,000	\$31,818,000
FIPS Par Volume	\$102,506,000	\$41,485,000	\$35,294,000	\$348,174,000	\$112,163,000
ABS/(FIPS + ABS)	53.82%	62.87%	55.02%	28.45%	22.10%

EXHIBIT 5
Ordered Probit Model Groups and Spreads

	Groups			
	1	2	3	4
PNF 05	0.125-0.25	0.375-0.5	0.625-0.875	>0.875
SME 04	0.125	0.25	0.375-0.5	>0.5
CLAR 02	0.125-0.25	0.375-0.5	0.625-1	>1
BBY 00	0.125-0.25	0.375-0.5	0.625-2	>2
PCS 03	0.125-0.25	0.375-0.5	0.625-1	>1
STO 01	0.125	0.25	>0.25	-
HDS 03	0.125-0.375	0.5-0.75	0.875-2	>2
WHX 03	0.125	0.25	0.375-0.5	>0.5
AGY 04	0.125-0.375	0.5-0.75	0.875-1	>1
SME 01	0.125-0.375	0.5-0.75	0.875-1	>1

Spreads are in increments of 1/8.

FIPS and the NYSE's ABS constitutes all the trading in a given bond.

As is apparent in the third line of the table reporting the ABS share of total par trade volume (dollar volume is similar), three of the five bonds represent at least half of all trading volume on ABS. It is noteworthy that the highest-volume ABS bond is not a FIPS issue, and hence it should be expected that ABS would represent a significant portion of the total market for that bond.[9]

For the ordered probit analysis, the observed spread has to be grouped into a finite number of ordered categories. Exhibit 5 presents the groups and the respective spread intervals we adopt for each bond. A group size of four seems to be appropriate for all bonds, with the exception of one bond (STO 01) where the size is three.

V. RESULTS

We first present results from ACD-based volatility estimates and then results from GARCH-based volatility estimates.

ACD-Based Results

Exhibit 6 reports results for the ACD model. The WACD(1, 1) estimates for all bonds are highly significant at the 5% level.[10] 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 1.00. This indicates that long durations are more likely than short durations, and hence the Weibull distribution is suitable.

Exhibit 7 reports results from the ordered probit model with ACD-based volatility, lagged spread, and volume as the explanatory variables. Our main interest is in the sign and the statistical significance of the parameter δ_1, which determines the relationship between bid-ask spreads and latent volatility. The partition coefficients for the ordered probit model (not reported) are all highly significant.

Examining the p-values for parameter δ_1 in the unobserved spread equation, we can distinguish *three groups* of bonds. *Four bonds* (PNF 05, CLAR 02, BBY 00, and PCS 03) have a positive and statistically significant relationship (all at the 5% level) between the bid-ask spread

EXHIBIT 6

Maximum-Likelihood Estimates for WACD(1, 1) Model

	Log Likelihood	Quotes γ	ω	α	β
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	−6578	0.67456	42.272	0.13257	0.83210
T stat		36.926	2.1364	3.8959	23.433
HDS 03	−6177	0.59636	114.83	0.25161	0.65924
T stat		36.791	1.7348	3.0297	5.3000
WHX 03	−5344	0.62930	64.372	0.031918	0.90773
T stat		32.599	0.55568	0.90782	6.5023
AGY 04	−6363	0.61899	113.74	0.12118	0.76654
T stat		36.915	2.3012	3.0701	10.407
SME 01	−6811	0.58249	85.784	0.086505	0.83147
T stat		38.212	1.8954	2.2608	11.863

The likelihood function of the WACD(1, 1) model is described in Equation (A-3). The log likelihood values are reported using the optimal parameter estimates.

and latent volatility of the bond price process. *Three bonds* (SME 04, STO 01, and HDS 03) have a positive but not statistically significant relation. The remaining *three bonds* have a negative relationship (significant for AGY 04, but insignificant for WHX 03 and SME 01).

In general, we see a positive relation between volatility and spreads that could indicate either a higher adverse selection component or a transitory component, or both. Since the ABS market is mainly retail trade-driven, we expect the positive relation to be the result of a higher transitory component (order processing and inventory costs).

The coefficient δ_2 for lagged spread is positive and statistically significant for all the bonds. The coefficient δ_2 in general is of very high magnitude (the only exception is WHX 03, which has a coefficient of 0.45). This implies that that a current-period positive shock to the bid-ask spread will lead to higher spreads in consecutive periods, with a high degree of persistence. High persistence implies slow decay, and hence a tendency for the spread to wander away from a stationary mean (implying low mean reversion).

EXHIBIT 7

Maximum-Likelihood Estimates for the Ordered Probit Model with ACD Volatility, Lagged Spread, and Volume as the Variables

		δ_0	δ_1	δ_2	δ_3	LR Ratio
PNF 05		3.41631	0.33850	0.78070	−0.10409	1282.9988
T = 3980	*p value*	0.0001	0.0001	0.0001	0.0001	
SME 04		3.05655	0.17408	0.62814	−0.20970	393.5475
T = 2041	*p value*	0.0004	0.154	0.0001	0.0001	
CLAR 02		2.97740	0.17105	0.85593	−0.21984	998.1749
T = 2401	*p value*	0.0001	0.0015	0.0001	0.0001	
BBY 00		2.94341	0.27882	0.77952	−0.10842	514.4126
T = 1957	*p value*	0.0001	0.0110	0.0001	0.004	
PCS 03		3.82288	0.48756	0.67851	−0.04583	347.5695
T = 1675	*p value*	0.0001	0.010	0.0001	0.193	
STO 01		1.59827	0.19932	0.72049	0.12027	86.5878
T = 856	*p value*	0.096	0.248	0.0001	0.017	
HDS 03		1.70325	0.05729	0.68825	−0.07554	261.6017
T = 819	*p value*	0.026	0.692	0.0001	0.063	
WHX 03		0.05618	−0.23107	0.44693	−0.02630	55.3344
T = 694	*p value*	0.989	0.754	0.0001	0.618	
AGY 04		−2.13380	−0.72304	0.76886	−0.17265	293.8758
T = 835	*p value*	0.086	0.002	0.0001	0.0001	
SME 01		−0.29137	−0.32971	0.86570	−0.10932	310.8092
T = 900	*p value*	0.845	0.232	0.0001	0.0680	

T gives sample size. LR ratio gives the likelihood ratio statistic assessing the overall fit (slope coefficients). LR χ^2 distribution table value at 5% significance with three degrees of freedom is 7.81.

Coefficients for the volume variable are negative in nine cases (five significant at the 5% level and two significant at the 10% level) and positive in one case (but insignificant). A strong negative relation between trading volume and bid-ask spreads is consistent with a low adverse selection component and a high transitory component. The likelihood ratio statistic for the overall fit of the model (LR ratio) is very high in all cases.

Overall, Exhibit 7 indicates a positive volatility-spread relation and a strong negative volume-spread relation. We also find that correlations between ACD volatility estimates and volume proxies are of very low order, which reinforces the conclusion that our volatility measure is not proxying for volume.[11]

EXHIBIT 8

Maximum-Likelihood Estimates for the Ordered Probit Model with GARCH Volatility, Lagged Spread, and Volume as the Variables

		δ_0	δ_1	δ_2	δ_3	LR Ratio
PNF 05		1.75934	0.06612	0.58896	−0.07804	134.8282
T = 1057	*p value*	0.0001	0.049	0.0001	0.059	
SME 04		2.11894	0.03586	0.79627	−0.03487	183.1173
T = 1082	*p value*	0.0001	0.582	0.0001	0.440	
CLAR 02		4.27258	0.33191	0.98047	−0.07124	345.4225
T = 944	*p value*	0.0001	0.0001	0.0001	0.179	
BBY 00		3.48554	0.24775	0.73913	−0.11019	187.7800
T = 949	*p value*	0.0001	0.011	0.0001	0.035	
PCS 03		4.19147	0.38103	0.71536	0.02063	154.7159
T = 930	*p value*	0.009	0.048	0.0001	0.660	
STO 01		2.99413	0.12680	1.29667	0.09298	188.3252
T = 943	*p value*	0.0001	0.099	0.0001	0.109	
HDS 03		1.65219	0.09066	0.85736	0.00450	218.3138
T = 687	*p value*	0.115	0.525	0.0001	0.913	
WHX 03		4.88515	0.22748	1.05775	−0.11154	270.0982
T = 784	*p value*	0.0001	0.076	0.0001	0.026	
AGY 04		1.60759	−0.00677	1.13732	−0.10308	287.2179
T = 585	*p value*	0.153	0.962	0.0001	0.040	
SME 01		4.36753	0.36313	0.94309	−0.04309	239.7630
T = 617	*p value*	0.0001	0.001	0.0001	0.497	

T gives sample size. LR ratio gives the likelihood ratio statistic assessing the overall fit (slope coefficients). LR χ^2 distribution table value at 5% significance with three degrees of freedom is 7.81.

GARCH-Based Results

To test the robustness of our results to an alternative volatility specification, we perform the volatility estimation using a more traditional GARCH model. We begin by constructing return series for each bond as follows:

1. Eliminate the first quotes on a given day. This is intended to remove any extra information in the first quote that has accumulated since the previous market close. Aggregate quotes into 30-minute intervals each day (there are 13 such intervals on a given day between 9:30 AM and 4:00 PM).
2. For a given 30-minute interval, choose the last quote of that interval (for example, if between 9:30 AM and 10:00 AM there are two quotes, one at 9:56 AM and the other at 9:58 AM, we choose the latter). Report the bid-ask spread, average of bid and ask prices (to proxy for the transaction price), and bid and ask quantities for that last quote. If there are no quotes in a given 30-minute interval, drop that interval from the data. For example, if there is only one time interval with no quotes on a given day, we will have 12 data points for spreads and prices on that day.[12]

3. From Step 2, we have time series data at 30-minute intervals for bid-ask spreads, prices, and trading volume during the sample period. Compute half-hour returns from the price series adjusting for accrued interest using the coupon rate.
4. Obtain residuals from an AR(1) process of the returns (this is to filter any possible autocorrelation). All the AR coefficients are highly significant. Then obtain the conditional volatilities using a GARCH(1, 1) model (all the GARCH coefficients are highly significant, and there is a very high persistence in volatility for six out of ten bonds).[13]
5. Spreads, GARCH-based conditional volatilities, and volume (measured as a logarithm of the average of bid and ask quantities) are used as inputs in the ordered probit model.

Exhibit 8 reports results from the ordered probit model. The sample sizes in general are much smaller in Exhibit 8 compared to Exhibit 7 (the only exceptions are STO 01 and WXH 03). This implies that GARCH models could result in the loss of potentially informative observations.

According to the p-values for coefficient δ_1 of the volatility variable, there again are three groups of bonds. Five bonds show a positive and statistically significant relationship (at the 5% level) between the bid-ask spread and latent volatility. Four bonds show a positive but not statistically significant relation. One bond has a negative and insignificant relation. In general, we see a somewhat stronger positive volatility-spread relation with the GARCH volatilities in Exhibit 8 than that found in Exhibit 7.

The lagged spread coefficients δ_2 in general are large and positive, and have very low p-values, comparable to results in Exhibit 7. Volume is negative for seven bonds (significant at the 5% level for four) and positive for the other three bonds (largely insignificant). In general, we see a much weaker negative volume-spread relation in Exhibit 8.

The likelihood ratio statistic over all estimates of the model is high in all cases. As with the ACD-based model, low correlations between volume and volatility lead us to conclude that our volatility measure is not capturing volume.[14]

Overall, the GARCH results in Exhibit 8 confirm the positive volatility-spread relation and the negative volume-spread relation shown in Exhibit 7, thus implying that results are robust to alternative volatility specifications. For almost all bonds, the significance levels across Exhibits 7 and 8 would appear to be driven by power resulting from whichever test uses more observations.

VI. CONCLUSIONS

For the ten most active corporate bonds on the NYSE's Automated Bond System, we find a largely positive and statistically significant relationship between volatility and observed spreads. There is a largely negative and statistically significant relationship between volume and observed spreads. The results appear to be robust to alternative volatility specifications. Our results provide some confidence that time-based volatility measures for thinly traded markets like corporate bonds are a feasible alternative to traditional GARCH-based volatilities. The GARCH approach causes elimination of great amounts of data and the possible loss of information.

Since the ABS market is mainly retail trade-driven, it shares many of the characteristics of a regional stock exchange. The observed positive volatility-spread relation and negative volume-spread relation is consistent with a strong transitory component and a weak adverse selection component. These results are in agreement with those reported by Lin, Sanger, and Booth [1995], who examine NYSE stock trades executed on regional exchanges.

The absence of evidence that there are significant adverse selection costs would not be expected in the dealer market for high-yield debt. The answer to this question will have to wait for the introduction of mandatory bond dealer reporting.

APPENDIX
ACD Model

According to the ACD(m, q) model, the duration x_i between the i-th and the $(i-1)$-th quote can be written as:

$$x_i = \psi_i \varepsilon_i$$
$$\psi_i = \omega + \sum_{j=1}^{m} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j}$$
$$\text{for } \alpha_j, \beta_j \geq 0,\ \omega > 0, \text{ for all } i,\ i = 1, \ldots, N \qquad \text{(A-1)}$$

where ψ_i is the conditional expected duration, given past information and the parameter set θ, where $\theta = \{\alpha_j, \beta_j$, parameters of the error distribution$\}$; that is:

$$\psi_i = E[x_i | x_{i-1}, \ldots, x_1;\ \theta] \qquad \text{(A-2)}$$

m and q refer to the orders of the lags used, and $\{\varepsilon_i\}$ is the error term, which is i.i.d. and follows a specified distribution (for details, see Engle and Russell [1997, 1998]).

The conditional intensity for the ACD model is a function of both conditional duration and error distribution and therefore can capture rich dynamics of the data. Our bond data suggest that the Weibull distribution fits the data better than an exponential distribution, implying that long durations are *more* likely to occur than short durations.

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

$$L(\theta) = \sum_{i=1}^{N(T)} \ln\left(\frac{\gamma}{x_i}\right) + \gamma \ln\left(\frac{\Gamma(1+\frac{1}{\gamma})\, x_i}{\psi_i}\right) - \left(\frac{\Gamma(1+\frac{1}{\gamma})\, x_i}{\psi_i}\right)^{\gamma} \quad \text{(A-3)}$$

where

$\psi_i = \omega + \alpha x_i + \beta\psi_{i-1}$, for $i > 1$;

$\psi_i = \dfrac{\omega}{(1-\beta)}$ for $i = 1$;

$\theta = (\gamma, \omega, \alpha, \beta)$, $\alpha, \beta \geq 0$, $\omega > 0$, $\alpha + \beta < 1$; and

Γ is a gamma distribution.

The log likelihood function in (A-3) is maximized iteratively with respect to parameters in θ and subject to the non-negativity and stationarity conditions stated above. Maximum likelihood estimation is done using the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) algorithm in GAUSS.

ENDNOTES

[1]As documented in Dueweke, Hyland, and Siesel [1992], the ABS accounts for a significant percentage of the number of transactions in the most active issues of bonds traded there. Indeed, as we show, half of our sample bonds are Fixed Income Pricing (FIPS)-listed issues, and the ABS trade in these bonds represents a significant portion of all trade.

[2]The term duration in this context should not 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.

[3]A small sample of high-yield bond transactions (67 bonds) beginning in 1994 is available on Nasdaq's Fixed Income Pricing System (FIPS). See Alexander, Edwards, and Ferri [2000] and Hotchkiss and Ronen [1999] for a description and analysis of this data. Only hourly reports are currently available for analysis.

[4]MDH postulates that volume and volatility are jointly determined by the number of information events. 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 (Jones, Kaul, and Lipson [1994]). Lamoureux and Lastrapes [1990] show that GARCH effects on returns disappear when volume is included in the variance equation.

[5]Analogously, the GARCH model corrects for volatility clustering in the data.

[6]The first dummy variable indicates the morning interval between 9:30 AM and 12:00 M; the second dummy variable indicates the interval from 12:00 M to 2:00 PM; and the third dummy variable indicates the interval from 2:00 PM to 4:00 PM on a weekday. 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.

[7]All the variables are used in log form. Recall $c = 1/8$ in our model.

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

[9]A simultaneous analysis of listed equity is beyond the scope of this article, although we should note that two of the companies represented here did not have listed equity during the period examined (Wheeling Pittsburgh and Claridge Hotels).

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

[11]The correlations for the ten bonds in Exhibit 7 are: 0.0076062249, –0.0085882183, 0.039659061, 0.0086629291, –0.024021605, 0.085762320, –0.046031102, 0.19420845, –0.0072828517, and –0.022074182.

[12]On average, on a given day, there was one interval without any quotes for the most liquid bond, and four for the least liquid bond.

[13]GARCH(1, 1) is used as a benchmark specification.

[14]The correlations for the ten bonds in Exhibit 8 are: 0.042107675, –0.054413137, –0.030664694, 0.020520565, –0.0091080019, –0.15116095, 0.069466626, -0.035707913, –0.059511209, and –0.18346703.

REFERENCES

Alexander, G.J., A.K. Edwards, and M.G. Ferri. "The Determinants of Trading Volume of High-Yield Corporate Bonds." *Journal of Financial Markets*, 3, No. 2 (May 2000), pp. 177-204.

Bessembinder. H. "Bid-Ask Spreads in the Interbank Foreign Exchange Markets." *Journal of Financial Economics*, 35 (1994), pp. 317-348.

Bollerslev, T., and M. Melvin. "Bid-Ask Spreads and Volatility on the Foreign Exchange Market." *Journal of International Economics*, 36 (1994), pp. 355-372

Dueweke, D., M. Hyland, and F. Siesel. "Measuring the New York Stock Exchange's Share of Corporate Trading Volume." *The Journal of High Yield Research*, September/October 1992, pp. 9-16.

Engle, R., and J. Russell. "ACD: a New Model for Irregularly Spaced Transaction Data." *Econometrica,* 66, 5 (1998), pp. 1127-1162.

——. "Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the ACD Model." *Journal of Empirical Finance,* 4 (1997), pp. 187-212.

Gwilym, O., A. Clare, and S. Thomas. "The Bid-Ask Spread on Stock Index Options: An Ordered Probit Analysis." *Journal of Futures Markets,* 18, No. 4 (1998), pp. 467-485.

Harris, M., and A. Raviv. "Differences of Opinion Make a Horse Race." *Review of Financial Studies,* 6, 3 (1993), pp. 473-506.

Hausman, J., A. Lo, and C. MacKinlay. "An Ordered Probit Analysis of Transaction Stock Prices." *Journal of Financial Economics,* 31 (1992), pp. 319-379.

Hong, G., and A. Warga. "An Empirical Study of Bond Market Transactions." *Financial Analysts Journal,* 56, No. 2 (March/April 2000), pp. 32-46.

Hotchkiss, E., and T. Ronen. "The Informational Efficiency of the Corporate Bond Market: An Intraday Analysis." Working paper, Boston College, 1999.

Jones, C., G. Kaul, and M. Lipson. "Transactions, Volume and Volatility." *Review of Financial Studies,* 7, 4 (1994), pp. 631-651.

Lamoureux, C., and W. Lastrapes. "Heteroscedasticity in Stock Return Data: Volume vs. GARCH Effect." *Journal of Finance,* Vol XIV (1990), pp. 221-229.

Lee, C., B. Mucklow, and M. Ready. "Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis." *Review of Financial Studies,* 6, 2 (1993), pp. 345-374.

Lin, J., G. Sanger, and G. Booth. "Trade Size and Components of Bid-Ask Spread." *Review of Financial Studies,* 8, 4 (1995), pp. 1153-1183.

McInish, T., and R. Wood. "An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks." *Journal of Finance,* XLVII, 2 (1992), pp. 753-764.