Internet Appendix to "Price Discovery in Illiquid Markets: Do Financial Asset Prices Rise Faster Than They Fall?"*

RICHARD C. GREEN

DAN LI

and

NORMAN SCHURHOFF

*Citation format: Green, Richard C., Dan Li, and Norman Schürhoff, Internet Appendix to "Price Discovery in Illiquid Markets: Do Financial Asset Prices Rise Faster Than They Fall?," *Journal of Finance,* DOI 10.1111/j.1540-6261.2010.01590.x. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

Part A. Dealer Intermediation between Retail and Institutional Investors and Price Asymmetries.

This section elaborates on how the timing and structure of dealer intermediation between large investors on the bid side of the market, and smaller investors on the ask side produces asymmetries in prices and dealer spreads.

Dealers in the municipal market operate differently than intermediaries in the equities market or the Treasury market, where dealers are routinely holding inventories of specific securities, or are net short almost as often as they are net long. In those markets, trade is high-frequency on the time-series dimension. Dealers can have some confidence that if they sell a specific security, they can later repurchase the same security to replenish inventory or cover a short position.

In the market we study, this is not the case. There are very few pairs of transactions where the dealer sells bonds to a customer from inventory and later purchases bonds for customers to replenish inventory. Trade in specific bonds is sufficiently infrequent that the dealer cannot with confidence anticipate being able to find the bond in question. Dealers are never short because of this.

Since we cannot identify individual dealers in our data, we cannot verify directly in the data that dealers are never, or rarely, short. It is obvious, however, from simply looking at the raw data that purchases from customers almost invariably precede sales to customers in time. We provide an example of the raw data in Table I.

We can also provide indirect evidence that matched pairs of transactions are rarely sales by dealers from inventory, with subsequent purchases covering or replenishing the sale. Table II below shows all the transactions in our database we could identify as "round-trip" transactions: a purchase from a customer followed by a sale of the same bond in the same par amount. There are over three and a half million of these. It also shows all of the pairs that reverse the ordering. These are sales to customers followed by a purchase of the same bond in the same par amount. There are about one million of these. The table sorts all these pairs by the time between the two sides of the trade. Note that most of the "reverse round trip" pairs are time stamped within less than one minute of each other (15.94%), or are more than five days apart (48.06%). This strongly suggests that the trade was either pre-arranged, and the two legs are essentially simultaneous, or that the two sides of the pairs are not, in fact, going through the same dealer. Notice that the distribution of durations for the round trip pairs are very different. It peaks at 1 day.

Table I
History of Trades in One Bond.

Trade Date	Time	Price	Size	Type
4/19/2002	8:45	103.75	10K	Purchase from Customer
4/22/2002	10:33	106.28	10K	Sale to Customer
4/24/2002	9:53	103.25	25K	Purchase from Customer
4/25/2002	14:25	106.23	25K	Sale to Customer
6/11/2002	14:14	103.5	15K	Purchase from Customer
6/12/2002	9:02	105.98	15K	Sale to Customer
8/12/2002	15:59	102.38	10K	Inter-dealer
8/12/2002	15:59	99.75	10K	Purchase from Customer
8/12/2002	16:32	103.88	10K	Inter-dealer
8/13/2002	6:07	104.88	10K	Sale to Customer
8/15/2002	9:12	103	20K	Purchase from Customer
8/16/2002	10:00	105.64	20K	Sale to Customer
8/19/2002	11:23	104	10K	Purchase from Customer
8/20/2002	9:43	105.62	10K	Sale to Customer
8/26/2002	11:05	102.75	25K	Purchase from Customer
8/27/2002	9:02	105.5	25K	Sale to Customer
10/18/2002	17:21	98.43	25K	Purchase from Customer
10/18/2002	17:21	100.18	25K	Inter-dealer
10/21/2002	12:00	102.18	25K	Inter-dealer
11/5/2002	8:27	103.97	25K	Sale to Customer
11/11/2002	9:11	101.75	25K	Purchase from Customer
11/12/2002	15:08	104.55	25K	Sale to Customer
11/15/2002	12:03	101.5	25K	Purchase from Customer
11/18/2002	10:19	103.5	25K	Sale to Customer
11/20/2002	12:44	102.75	20K	Purchase from Customer
11/20/2002	15:48	99.98	10K	Purchase from Customer
11/20/2002	15:48	101.36	10K	Inter-dealer
11/21/2002	11:00	102.86	10K	Inter-dealer
11/21/2002	11:16	103.76	10K	Sale to Customer
11/22/2002	8:42	104.4	20K	Sale to Customer
1/14/2003	7:52	102.25	30K	Purchase from Customer
1/15/2003	9:04	103.71	30K	Sale to Customer
1/15/2003	12:28	101.25	10K	Purchase from Customer
1/16/2003	8:54	103.74	10K	Sale to Customer
2/3/2003	16:59	101.39	20K	Inter-dealer
2/4/2003	0:00	102.39	20K	Inter-dealer
2/4/2003	9:24	103.28	20K	Sale to Customer
2/4/2003	9:45	101.39	20K	Purchase from Customer
4/30/2003	12:09	102.1	60K	Purchase from Customer
4/30/2003	13:04	102.2	60K	Sale to Customer
5/5/2003	15:16	102	10K	Purchase from Customer

Table II
Distribution of Durations for Round-Trips and Reverse Round-Trips.

	Round-trip			Reverse Roundtrip		
Time gap	N	pct	cpct	N	pct	cpct
<1 min	154,794	4.32%	4.32%	163,087	15.94%	15.94%
<2 min	122,997	3.44%	7.76%	35,615	3.48%	19.42%
<5 min	81,378	2.27%	10.03%	21,248	2.08%	21.50%
<15 min	137,920	3.85%	13.88%	30,068	2.94%	24.44%
<1 hour	331,875	9.27%	23.15%	60,816	5.94%	30.38%
<2 hours	215,585	6.02%	29.18%	39,878	3.90%	34.28%
<5 hours	217,971	6.09%	35.26%	50,629	4.95%	39.22%
>5 hours	34,240	0.96%	36.22%	17,727	1.73%	40.96%
1 day	960,002	26.81%	63.04%	44,116	4.31%	45.27%
2 days	210,101	5.87%	68.90%	18,588	1.82%	47.09%
3 days	242,931	6.79%	75.69%	17,759	1.74%	48.82%
4 days	169,814	4.74%	80.43%	16,100	1.57%	50.39%
5 days	128,877	3.60%	84.03%	15,768	1.54%	51.94%
>5 days	571,610	15.97%	100.00%	491,792	48.06%	100.00%
Total	3,580,095	100.00%		1,023,191	100.00%	

We now illustrate through simulation that the asymmetric behaviors we document for prices and dealer spreads are mutually consistent if pricing on the bid side of the market is more likely to reflect transactions with informed, institutional investors. This is clearly the case. The aggregate par value of purchases from customers is roughly equal to the value of sales to customers in our sample of seasoned bonds, but the number of sales is twice as large. On average, therefore, purchases are twice the size of sales. Still, there are many purchases by dealers at par values that suggest the intermediaries are dealing with retail investors. If dealers have a timing option that allows them to delay recognition of movements in intrinsic value for retail investors, but those investors are more predominant when dealers are selling than when they are buying, then we will see half spreads on both sides of the market rising when prices move in a direction that favors the dealer, while our measure of the price, the mid-point of the implicit inside spread, will rise faster than it falls.

The results below expand on the simple numerical example in Section III of the paper and Table II. We simulate the intrinsic value of an hypothetical bond using the Vasicek process. Using our data, we calibrate arrival rates for trades of various sizes on both sides of the market. Dealers then earn profits from two sources. They mechanically earn a proportional half-spread that varies with transaction size. This is just compensation for transaction services they provide. They also have a timing option. They can, with a certain probability, succeed in buying or selling at the lagged intrinsic value. Otherwise they trade at the current intrinsic value. The probability they succeed can depend on whether the trade is a buy or a sale, and on the size of the trade. (In the numerical example, the probability is 1 for all retail trades and zero for all institutional trades.) We then use these arrival rates and price determination rules to simulate transaction histories, prices, and inventories, subject to a constraint that the dealer's inventory never goes negative. (If a buyer arrives and the dealer has no bonds, no trade occurs.) Using this artificial sample of trades, we can calculate midpoints and half-spreads as we do in the paper, and run tests on the data that are analogous to those in Table VI and Table VIII in the paper. Generating the basic asymmetries is not difficult. Matching the magnitudes of our coefficients requires a substantial time lag (see the row of the table in bold face), suggesting that retail investors on the ask side of the market are quite uninformed about prices.

Below we outline the procedure we use to simulate intrinsic values, midpoints, and half spreads.

Step 1 - we compiled the subsample of bond-days used in the tests that require v_{it} for two consecutive days. Pooling all these bond-days, we can estimate the probability a buyer and a seller arrive and trade a specific quantity within a day. If the arrivals are Poisson, this gives us parameters for the arrival intensities, $\lambda_s(q)$ and $\lambda_b(q)$, $q = 10K$, $15K$, $25K$, $50K$, $100K$, $250K$, $1,000K$.

Step 2 - Use the Vasicek model to simulate bond prices for a 20-yr maturity bond. This gives us a simulated series of daily intrinsic values p_t^* or yield spreads, y_t^*.

Step 3 - Assume dealer profits on trades are due to two components. They earn a proportional half-spread that varies with transaction size $\phi(q)$. We use values of 100 basis points for $q = 10K - 15K$, 75 b.p. for $q = 25K - 50K$, 10 b.p. for $q = 100K$, and 2 b.p. for $q = 250K - 1,000K$.

Dealers also have a timing option that allows them to charge a buyer $\max\{p_t^*, p_{t-l_b}^*\}$ with probability $\pi_b(q)$, and p_t^* with probability $1-\pi_b(q)$. The index $l_b \geq 1$ denotes the information lag of an uninformed buyer. We vary l_b between 1 to 10 days. Similarly, a dealer can charge a seller $\min\{p_t^*, p_{t-l_s}^*\}$ with probability $\pi_s(q)$, and p_t^* with probability $1 - \pi_s(q)$. The index $l_s \geq 1$ denotes the information lag of an uninformed seller. We start with $\pi_b(q) = \pi_s(q) = 1$ for trades of less than 100K, and $\pi_b(q) = \pi_s(q) = 0$ for $q \geq 100K$.

Step 4 - On each day, use the arrival rates from Step 1 and the price determination rules in Step 2 to simulate transaction prices, subject to the constraint that dealer inventory does not go negative. If a buyer arrives with quantity q, and the dealer inventory is bigger than q, the buyer pays the price given by the dealer's timing option plus the fixed half spread. If a seller arrives, the dealer always buys, and the seller receives a price given by the timing option less $\phi(q)$.

Step 5 - For each simulated day with sufficient transactions, we compute the midpoint as the average of the minimum customer buy price and the maximum customer sale price.

Compute the average customer buy price a_t and average customer sale price b_t. Then we compute the ask- and bid-side half spreads, $(a_t - v_t)/v_t$ and $(b_t - v_t)/v_t$.

Step 6 - We now have an artificial sample on which to run two tests that are analogous to the central results in our paper in Tables 6 and 8. Analogous to Table 6, we run the regression:

$$sp_t = \beta^+ (\Delta v_t)^+ + \beta^- (\Delta v_t)^- + \epsilon_t.$$

Corresponding to Table 8, we run

$$\Delta y_t = \delta^+ (y_t^* - y_{t-1})^+ + \delta^- (y_t^* - y_{t-1})^- + \varepsilon_t.$$

Predicted results - We should obtain asymmetries in the coefficients like we observe in the tables. If we manipulate the arrival rates and timing option so the arrival rates are symmetric for institutionally sized trades on the ask and bid side, and so that $\pi(q) = 0$ for trades of all size. The asymmetries in the coefficients should then disappear.

For the subsample of bond-days used in the tests that require v_{it} for two consecutive days, we estimated the arrival intensities, $\lambda_s(q)$ and $\lambda_b(q)$, for $q = 10K$, $15K$, $25K$, $50K$, $100K$, $250K$, $1,000K$ by pooling all bond-days. The corresponding probabilities for ask trades of size 10, 15, 25, 50, 100, 250, 1,000 are 16.49, 4.40, 12.12, 9.89, 5.95, 3.48, and 4.09 percent. The corresponding probabilities for bid trades of size 10, 15, 25, 50, 100, 250, 1,000 are 10.83, 2.94, 8.29, 7.89, 5.30, 3.44, and 4.88 percent. That means, 48.86 (35.25) percent of all trades in the estimation sample are dealer sales to (purchases from) retail customers (defined by trade of 50 or fewer bonds), and 7.57 (8.32) percent of all trades in the estimation sample are dealer sales to (purchases from) institutional clients (defined by trade of 100 or more bonds). Finally, we calibrated the arrival intensities by scaling the probabilities to match the average number of transactions per day given trading takes place, which is 3.3 trades per day.

Table III reports the results. The first row replicates the data moments found in Tables 6 and 8. The second row shows that when all traders are informed ($\pi_s(q) = \pi_b(q) = 0$ for all q), the coefficient estimates are symmetric, as we would expect. In the last column we compute a measure of fit as the squared distance between the empirical moments and the simulated model moments. A smaller number means better fit. In the first simulation panel we vary the fraction of uninformed buyers from zero to 100 percent. Across rows, the coefficient β^- captures the widening ask half-spread in falling markets and becomes closer to the data. However, the coefficients for the bid half-spread and for rising markets cannot be captured by the presence of uninformed buyers only. In the next two panels, we vary the fraction of uninformed sellers from zero to 100 percent. The coefficient β^+ for the widening bid half-spread in rising markets can now be captured but the other parameters are still mismatched. An informational disadvantage for retail investors of only one day, hence, can be refuted.

Next we increase the information lag of uninformed retail buyers and, respectively, sellers from one to two, five, and ten days. The simulations are now able to capture the price elasticities observed in the data. To match the data one needs a sizeable fraction of uninformed investors buying bonds at inflated prices and selling bonds at deflated prices. The informational disadvantage of retail investors is striking. At the specification that best matches the data, the row in bold face, the information lag is between five to ten trading days.

The simulations also reveal that when information is symmetrically distributed among retail buyers and sellers, the simulated data does not generate sufficient asymmetry in the coefficients. The reason is that, contrary to our illustrative example, there is not enough asymmetry in transaction sizes at the bid compared to the ask when calibrated to the data. In unreported results we show that it is easy to generate the observed price asymmetry when transaction sizes are sufficiently asymmetric.

In summary, the simulations show that the data is broadly consistent with a time delay in the information of retail investors. The portion of uninformed investors is particularly large at the ask. The information lag is substantial and reaches up to two weeks. This is consistent with our findings in the early part of the paper (Figure 1).

Part B. Supplementary and Robustness Tests

Table III

INFORMATION LAG & PROBABILITY OF INFORMATION IN SIMULATED DATA.

The table reports parameter estimates from simulated data for the elasticity of the bid-ask spread and for the speed of adjustment in the yield spread to changing market conditions. The simulated prices and yields are for a sample of 10,000 bonds with twenty years to maturity simulated for one year. Across panels, we vary the probability of uninformed retail traders at the ask and, respectively, the bid, (π_b, π_s), and the information lag in days of retail traders at the ask and, respectively, bid, (l_b, l_s), as specified in the first column. The coefficients are from the regressions

$$sp_t = \beta^+(\Delta v_t)^+ + \beta^-(\Delta v_t)^- + \epsilon_t,$$

$$\Delta y_t = \delta^+(y_t^* - y_{t-1})^+ + \delta^-(y_t^* - y_{t-1})^- + \varepsilon_t.$$

In the last column, we compute a measure of the distance between the empirical moments and the simulated model moments by summing the squared deviations. The specification with the minimum distance is highlighted.

Parameters				Bid-Ask Spread		Ask Spread		Bid Spread		Yield Spread Adj. Speed		Moment Distance
π_b	π_s	l_b	l_s	β^+	β^-	β^+	β^-	β^+	β^-	δ^+	δ^-	
Data moments				10	-30	3	-40	10	-5	0.92	0.74	–
0	0	1	1	-2	-2	-6	-6	4	5	1.00	1.01	3,100
.5	0	1	1	-11	-9	-10	-18	-1	9	1.00	1.01	2,609
1	0	1	1	-22	-12	-14	-24	-7	13	1.01	1.01	3,263
0	.5	1	1	1	8	-10	-2	11	10	1.00	1.01	4,178
.5	.5	1	1	-7	-1	-14	-14	7	13	1.00	1.01	3,257
1	.5	1	1	-17	-4	-17	-21	0	16	1.00	1.01	3,478
0	1	1	1	1	19	-15	5	15	14	1.00	1.01	5,960
.5	1	1	1	-6	9	-18	-7	11	17	1.00	1.01	4,565
1	1	1	1	-15	5	-20	-15	6	19	1.01	1.01	4,370
1	0	2	1	-14	-7	-11	-22	-3	15	1.08	0.89	2,625
1	.5	2	1	-11	0	-15	-19	3	19	1.01	0.90	3,029
1	1	2	1	-12	7	-19	-15	7	22	0.96	0.92	3,966
1	0	5	1	14	-13	3	-28	11	15	1.23	0.73	1,776
1	.5	5	1	15	-7	0	-25	15	18	1.12	0.74	1,744
1	1	5	1	14	-1	-4	-22	18	21	1.05	0.75	2,127
1	0	10	1	46	-25	20	-38	26	13	1.36	0.64	4,154
1	.5	10	1	46	-20	17	-36	29	15	1.21	0.65	3,365
1	1	10	1	46	-16	14	-33	31	17	1.12	0.66	3,108
1	0	5	2	14	-13	3	-28	11	15	1.23	0.73	1,776
1	.5	5	2	10	-7	-3	-26	12	18	1.05	0.74	1,469
1	1	5	2	5	-2	-9	-22	14	20	0.91	0.75	1,953
1	0	10	2	46	-25	20	-38	26	13	1.36	0.64	4,154
1	.5	10	2	40	-20	15	-36	26	16	1.13	0.65	2,371
1	1	10	2	36	-15	9	-33	27	17	0.97	0.66	1,834
1	0	5	5	14	-13	3	-28	11	15	1.23	0.73	1,776
1	.5	5	5	6	-17	-4	-30	11	13	0.97	0.72	670
1	1	5	5	-1	-13	-15	-25	14	12	0.76	0.73	1,485
1	0	5	10	14	-13	3	-28	11	15	1.23	0.73	1,776
1	**.5**	**5**	**10**	**11**	**-40**	**-3**	**-39**	**14**	**-1**	**0.93**	**0.72**	**179**
1	1	5	10	7	-41	-15	-35	21	-6	0.69	0.75	1,130
1	0	10	10	46	-25	20	-38	26	13	1.36	0.64	4,154
1	.5	10	10	29	-39	8	-43	21	4	0.93	0.63	779
1	1	10	10	16	-35	-7	-36	23	1	0.68	0.64	1,070

Table IV
Granger Causality test for yields of all maturity.

The body of the table reports p-values for Granger tests of causality between municipal yield indices and treasuries of matched maturity. The first row compare the Bond Buyer 40 yield to call to the 10 year Treasury rate. The remaining rows consider Lehman Brothers indicative yields to treasuries of comparable maturity.

Maturity	P-value	Trsy Causes Muni	P-value	Muni Causes Trsy
BBI vs. 10Y Treasury Note	0.001	TRUE	0.459	FALSE
3M	0.000	TRUE	0.006	TRUE
6M	0.000	TRUE	0.840	FALSE
1Y	0.000	TRUE	0.175	FALSE
2Y	0.000	TRUE	0.960	FALSE
3Y	0.000	TRUE	0.930	FALSE
5Y	0.000	TRUE	0.719	FALSE
7Y	0.000	TRUE	0.830	FALSE
10Y	0.000	TRUE	0.646	FALSE
20Y	0.000	TRUE	0.256	FALSE

Table V
Effects of Macroeconomic News on Yield Spreads.

The table documents the effect of macro announcements on municipal bond yields and spreads between municipal bonds and treasuries. The yield log-spread is the natural logarithm of the ratio of the maturity-matched Treasury rate and the midpoint yield on the muni bond. Yields are measured in basis points. The explanatory variables capturing the effect of macro news are the standardized surprise component in the macro announcement. We compute the standardized surprise, as in **?**, as the actual value minus the consensus forecast divided by their standard deviation across all observations. Additional control variables are omitted from the table, including macro announcement dummies that equal one if there is an announcement of the corresponding item on the given day and zero otherwise, and various bond controls. These control variables are the bond and issuer characteristics described in Appendix **??**, order flow variables (bond-level trading volume and aggregate order imbalances over the past ninety sessions), dummies for the calendar year and the US state of issuance, and control variables for the par sizes and the daily changes in the size of the trades used to measure muni yields. The estimation results are from a cross-sectional regression on the announcement day, and standard errors are adjusted for heteroskedasticity. The sample is restricted to investment-grade bonds. *, ** and *** indicate that the corresponding p-values are less than 0.10, 0.05, and 0.01, respectively.

	Δ Yield Log-Spread	Δ Yield	Δ Treasury Yield
Positive Macro Surprise:			
Advance Retail Sales	0.41***	0.07	1.92***
Capacity Utilization	0.12**	−0.27	0.74***
Nonfarm Payrolls	1.23***	1.56***	7.58***
Consumer Price Index	0.14***	0.42**	0.97***
GDP Annualized	0.36***	0.39***	2.28***
Industrial Production	−0.02	−0.34	−0.46***
Producer Price Index	−0.42***	−0.04	−2.53***
Consumer Confidence	0.09***	0.08	0.72***
Jobless Rate	0.20***	−0.22	0.60***
FOMC Rate Decision	−0.03	0.16	−0.01
Negative Macro Surprise:			
Advance Retail Sales	0.25***	0.38**	1.95***
Capacity Utilization	0.44***	0.12	1.97***
Nonfarm Payrolls	0.56***	0.41***	2.88***
Consumer Price Index	0.44***	0.20	2.29***
GDP Annualized	−0.03	0.16	−0.01
Industrial Production	−0.35***	0.41**	−1.36***
Producer Price Index	0.15***	0.18	1.27***
Consumer Confidence	−0.02	0.13	−0.03
Jobless Rate	−0.01	−0.15	−0.30***
FOMC Rate Decision	−0.01	0.10	0.08***
Observations	226,065	228,019	228,019
R^2	0.13	0.12	0.14

Part C. Panel Estimator of Asymmetric and Sluggish Response

Denote by y_{it} the realization of the dependent variable at time t for bond i, and let its latent equilibrium value be y^*_{it}. The variable y_{it} can be the bid-ask spread or the yield spread of municipals over treasuries. Let x_{it} and z_{it} be observable characteristics of bond i at time t. The observed sample is (x_{it}, y_{it}, z_{it}) for bond $i = 1, ..., N$ and date $t = 1, ..., T$.

The adjustment of y_{it} towards its new equilibrium value can be expressed as follows:

$$\Delta y_{it} = f(y^*_{it} - y_{it-1}; z_{it}), \tag{C1}$$

where $f : \mathbb{R} \to \mathbb{R}$ is a monotonic invertible function with $f(0) = 0$, $f'(\epsilon) > 0$ for $\epsilon \in \mathbb{R}$. The model (C1) allows the speed of adjustment to depend on the endogenous deviation from equilibrium, $y^*_{it} - y_{it-1}$, and on exogenous determinants z_{it} through their effects on the functional form of f. Let δ be the set of parameters that determine f.

Assume

$$y^*_{it} = \beta' x_{it} + \alpha_i + \varepsilon_{it}, \tag{C2}$$

where α_i is a bond-specific unobserved effect and $\varepsilon_{it} = y^*_{it} - E(y^*_{it}|x_{it}, \alpha_i)$ is the residual. We now make a standard random effects assumption. Assume that, as in Ahn and Thomas (2006), the random vector $(y_{i1}, \alpha_i, (\varepsilon_{it})_{t=2,...,T})' \in \mathbb{R}^{T+1}$ exhibits a Gaussian distribution:

$$\begin{pmatrix} y_{i1} \\ \alpha_i \\ (\varepsilon_{it})'_{t=2,...,T} \end{pmatrix} = \mathcal{N}\left[\begin{pmatrix} m \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_y & \sigma_{y\alpha} & 0 \\ \sigma_{y\alpha} & \sigma^2_\alpha & 0 \\ 0 & 0 & \Sigma_{it} \end{pmatrix}\right]. \tag{C3}$$

Introduce the parameters $\gamma_0 = \frac{-\sigma_{y\alpha}}{\sigma_y^2} m$ and $\gamma_1 = \frac{\sigma_{y\alpha}}{\sigma_y^2} - 1$, and let σ be the set of parameters that determine Σ_{it}, $\forall i, t$. After inverting and first-differencing, equation (C1) yields a system of equations with autocorrelated error $\Delta\varepsilon_{it}$:

$$\begin{aligned} &f^{-1}(\Delta y_{it}) - f^{-1}(\Delta y_{it-1}) + \Delta y_{it-1} = \beta' \Delta x_{it} + \Delta\varepsilon_{it}, \quad &&\text{for } t = 3, ..., T, \\ &f^{-1}(\Delta y_{it}) - \gamma_1 y_{it-1} = \gamma_0 + \beta' x_{it} + u_i, &&\text{for } t = 2. \end{aligned} \tag{C4}$$

In (C4), the error term $u_i = \alpha_i + \varepsilon_{i2} - E(\alpha_i + \varepsilon_{i2}|y_{i1})$, with $E(\alpha_i + \varepsilon_{i2}|y_{i1}) = \gamma_0 + (1+\gamma_1)y_{i1}$, has zero mean and variance equal to $\sigma_{iu}^2 = \sigma_{i2}^2 + s_u^2$ with $s_u^2 = \sigma_\alpha^2 - \frac{\sigma_{y\alpha}^2}{\sigma_y^2}$.

Under the random-effects assumption (C3), the error distribution $\mathbf{f}$ of $(u_i, \Delta\varepsilon_i')'$ in (C4) is jointly Gaussian. Let $\mu_{it} = E(\Delta\varepsilon_{it}|\Delta\varepsilon_{it-1}, ..., \Delta\varepsilon_{i3}, u_i)$ and $s_{it}^2 = V(\Delta\varepsilon_{it}|\Delta\varepsilon_{it-1}, ..., \Delta\varepsilon_{i3}, u_i)$ be the conditional mean and variance of the error given past observations. The properties of the normal distribution yield the following recursive structure: $\mu_{i3} = \frac{-\sigma_{i2}^2}{\sigma_{iu}^2} u_i$, $s_{i3}^2 = \sigma_{i3}^2 + \sigma_{i2}^2 - \frac{(\sigma_{i2}^2)^2}{\sigma_{iu}^2}$, and for $t = 4, ..., T$,

$$\mu_{it} = \frac{-\sigma_{it-1}^2}{s_{it-1}^2}(\Delta\varepsilon_{it-1} - \mu_{it-1}), \qquad s_{it}^2 = \sigma_{it}^2 + \sigma_{it-1}^2 - \frac{(\sigma_{it-1}^2)^2}{s_{it-1}^2}. \tag{C5}$$

We can now derive the log-likelihood of the parameter vector $\theta = (\beta, \delta, \gamma, \sigma, s_u)$ given the data $(y_{i2}, ..., y_{iT})'$ for bond $i = 1, ..., N$ conditional on the initial observation y_{i1}, as y_{i1} is uncorrelated with the error vector $(u_i, \Delta\varepsilon_i')'$:

$$\ln \mathcal{L}(\theta|y) = \sum_{i=1}^{N} \ln \mathcal{L}(\theta|y_i), \tag{C6}$$

where

$$\begin{aligned}
\ln \mathcal{L}(\theta|y_i) &= \ln \mathbf{f}(u_i|y_{i1}) + \sum_{t=3}^{T} \ln \mathbf{f}(\Delta\varepsilon_{it}|\Delta\varepsilon_{it-1}, ..., \Delta\varepsilon_{i3}, u_i, y_{i1}) + \sum_{t=2}^{T} \ln \left| \frac{\partial f^{-1}(\Delta y_{it})}{\partial \Delta y_{it}} \right| \\
&= -(T-1)\ln\sqrt{2\pi} - \ln \sigma_{iu} - \frac{1}{2}\frac{(u_i)^2}{\sigma_{iu}^2} \\
&\quad + \sum_{t=3}^{T} \left(-\ln s_{it} - \frac{1}{2}\frac{(\Delta\varepsilon_{it} - \mu_{it})^2}{s_{it}^2} \right) + \sum_{t=2}^{T} \ln \left| \frac{\partial f^{-1}(\Delta y_{it})}{\partial \Delta y_{it}} \right|. \qquad \text{(C7)}
\end{aligned}$$

Standard error estimates A reasonable concern is whether the observations are independent in expression (C6). Many of the observations took place on the same days, and many issuers have several bond issues outstanding. In addition, the errors may be heteroskedastic and autocorrelated for given bonds even after controlling for unobserved bond-specific effects. In the presence of clustered errors, our MLE estimates are still consistent but standard errors can be understated, overstating significance and leading to incorrect inference in finite sample.

The common correction is to compute cluster-robust standard errors that generalize the White (1980) heteroskedasticity-consistent variance estimator (see Petersen (2009)). This permits both heteroskedasticity and error correlation of unknown form within clusters. In our setting, the natural adjustment to the variance of the estimator is to allow for two-way non-nested clustering. Following Cameron, Gelbach and Miller (2007), we cluster errors across bonds for particular calendar days (capturing time effects) and across bond-days for particular bond issuers (capturing municipality effects). Bond issue effects (that are not already captured by the random effects α_i) are nested within issuer clusters and need not be considered separately.

Denote the score vector by $g(\theta) = D \ln \mathcal{L}(\theta|y)$ and the Hessian by $H(\theta) = D^2 \ln \mathcal{L}(\theta|y)$.

Under standard assumptions, the estimated coefficient vector $\widehat{\theta}$ is asymptotically normal with covariance matrix $Var[\widehat{\theta}] = \{-E[H(\theta)]\}^{-1} Var[g(\theta)] \{-E[H(\theta)]\}^{-1}$. The cluster-robust estimator for the variance of the coefficient estimates equals

$$\widehat{Var}[\widehat{\theta}] = \{-H(\widehat{\theta})\}^{-1} \widehat{Var}[g(\widehat{\theta})] \{-H(\widehat{\theta})\}^{-1}, \tag{C8}$$

where $\{-H(\widehat{\theta})\}^{-1}$ is the standard covariance estimate and $\widehat{Var}[g(\widehat{\theta})] = \widehat{V}_{10} + \widehat{V}_{01} - \widehat{V}_{11}$ where

$$\widehat{V}_r = \sum_m \sum_n g_m(\widehat{\theta}) g_n(\widehat{\theta})' I_r(m,n) \tag{C9}$$

and $I_r(m,n)$ is an indicator that takes the value one if observations m and n share all clusters referenced by r, and zero otherwise.

REFERENCES

Ahn, Seung C. and Gareth M. Thomas, 2006, Likelihood Based Inference for Dynamic Panel Data Models, Working paper, Arizona State University.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, 2007, Robust Inference with Multi-way Clustering, Working paper, University of California - Davis.

Petersen, Mitchell A., 2009, Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches, *Review of Financial Studies* 22, 435-480.

White, Halbert, 1980, A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity, *Econometrica* 48, 817-838.