Corporate Bond Price Reversals

Alexey Ivashchenko*

Abstract

I demonstrate that U.S. corporate bond dealers mitigate adverse selection risk by passing potentially informed transactions to institutional investors that become liquidity providers to informed traders. I obtain these results in a theoretically-motivated empirical setup that contrasts corporate bond price reversals in bonds with different information asymmetry, trading volume, and dealers' capital commitment. I find strong price reversals that become less pronounced following high-trading-volume days. The effect is the strongest when dealers' end-of-day inventory does not change and when information motives for trading are the most acute: in bonds with the highest information asymmetry and before issuers' earnings announcements. The results suggest that private information reveals itself in prices on high-volume days when dealers do not accept overnight inventory risk.

JEL classification: G12, G14.

Keywords: corporate bonds, trading volume, reversal, informed trading, dealer inventory

*VU Amsterdam. I thank Michael Rockinger (advisor), Edith Hotchkiss (discussant), Quan Wen (discussant), Vincent Bogousslavsky, Pierre Collin-Dufresne, Robert Czech, Jakub Hajda, Robert Kosowski, Maureen O'Hara, Norman Schürhoff, and Johan Walden for their helpful suggestions. I also thank seminar participants at HEC Montreal, Frankfurt School of Finance & Management, WU Vienna, Nova SBE, VU Amsterdam, the Bank of England, the University of Lausanne, the University of Geneva, the University of St. Gallen, and of the Microstructure Exchange seminar series for insightful comments and feedback. All remaining errors are my own.
Mail: Alexey.Ivashchenko@vu.nl; Web: ivasche.com

I. Introduction

Trading with a better-informed counterparty is a risky business. Liquidity providers in securities markets may lose money on transactions with informed traders and ask for remuneration to bear such adverse selection risk. But are all liquidity providers equal in their (in)ability to avoid trading with better-informed investors? I consider the case of the US over-the-counter (OTC) corporate bond market and two distinct liquidity providers: broker-dealers and institutional investors. My empirical analysis suggests that the latter are more likely to be adversely selected than the former.

The paper claims that bond dealers avoid trading with informed investors. When approached by a client who wants to trade, a corporate bond dealer chooses whether to provide liquidity herself (i.e., to trade using the dealer's inventory capacity) or to find another investor who wants to trade in the opposite direction and let the investor provide liquidity.[1] It turns out that bond prices are more likely to move against liquidity providers *after* the trade when end investors rather than dealers supply liquidity. This finding is stronger for bonds with higher information asymmetry (for instance, bonds with smaller outstanding amounts, fewer investors, and fewer intermediating dealers) and when information motives for trading are more acute (prior to issuers' earnings announcements). I obtain these results by contrasting corporate bond price reversals, measured as the first autocorrelation of returns, following days with different trading volumes and changes in dealers' inventory.

What is the link between price reversals and trading motives? Pure liquidity trading (non-informational trading) generates only a temporary price pressure that subsides, and the price reverts toward a pre-trade level. Trading driven by private information generates a price pressure that is more likely to persist. Now, assume that the information-driven trading volume increases with the strength of the informed investor's signal. Then the price reversal should be less pronounced (the return should be more persistent) following a high-

[1] The dealer executes both trades, but such pre-arranged transactions close fast, and bonds do not stay on the dealer's books for longer than several minutes.

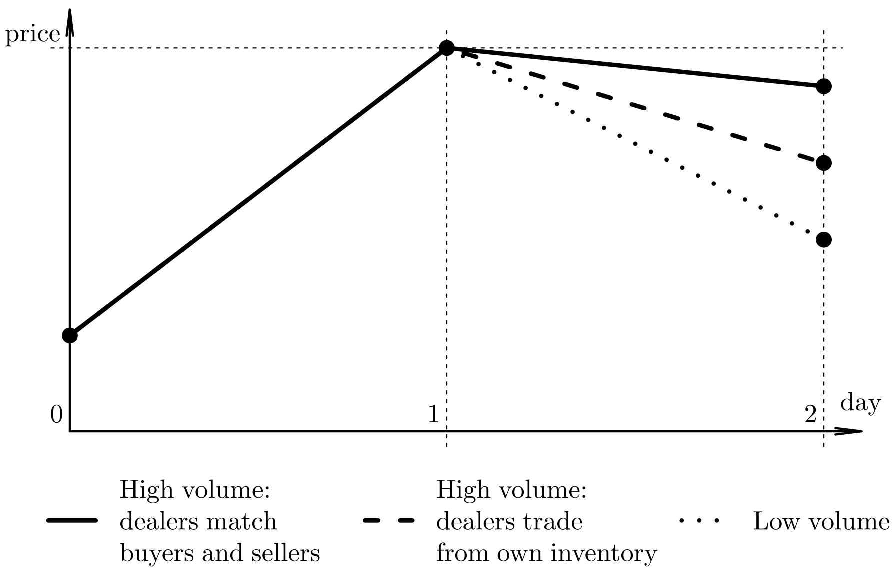

Figure 1. Stylized price reversal paths for a high-information-asymmetry bond. On day 1, the trading volume is either low or high. The solid line shows a reversal path on a high-volume day when dealers' end-of-day inventory (in this particular bond) does not change, and dealers buy from some investors as much as they sell to other investors. The dashed line refers to when trading volume on day 1 is high, and dealers trade a lot from their inventory. The 'Low volume' dotted line represents the average reversal path. For comparison purposes, I assume that the price change on day 1 is the same in all three cases.

trading-volume day if the information is indeed driving the trade.[2] I find that changes in U.S. corporate bond prices are the most persistent following high-volume days when dealers buy from some investors as much as they sell to other investors, and dealers' end-of-day inventory does not change. Figure 1 shows the stylized reversal paths I obtain for a typical high-information-asymmetry bond. Reversals are, on average, strong, but price changes become more persistent as trading volume increases, especially if dealers only match buyers and sellers and do not accept overnight inventory risk. The effect becomes stronger as information motives for trading intensify, both in the cross-section of corporate bonds and in the time series. Following high-volume days with unchanged dealers' inventory, the bonds with higher information asymmetry exhibit more persistent price changes than low-

[2]I assume that new public information affects prices without inducing abnormally high trading volumes.

asymmetry bonds, especially before issuers' earnings announcements. When dealers trade using their inventory capacity, there is no such effect.

Formally, my empirical analysis proceeds in two steps. In the first step, I use TRACE data from 2005–2018 aggregated to the daily frequency to estimate the following volume-return relationship for individual corporate bonds:

$$R_{t+1} = \beta_0 + \underbrace{\left(\beta_1 + \beta_2 \cdot \text{Client-to-client volume}_t + \beta_3 \cdot \text{Client-to-dealer volume}_t\right)}_{\text{Return autocorrelation}} R_t + \epsilon_{t+1}, \quad (1)$$

Above, R_{t+1} stands for total corporate bond return on day $t+1$. 'Client-to-client' (CtC) volume is the volume of investors' purchases from dealers matched by investors' sales to dealers within business day t; it does not add to dealers' aggregate end-of-day inventory in this bond. The difference between investors' purchases and sales is the change in dealers' inventory on day t: it stays on dealers' books until day $t+1$. I call the absolute value of the change in dealers' inventory a 'Client-to-dealer' (CtD) volume. I scale CtC and CtD volumes for individual bonds so that, in (1), β_1 measures the reversal on an average-volume day, while β_2 and β_3 capture how the reversal changes following high-volume days with different dealers' capital commitment. The volume-return relationship (1) stems from a theoretical model where risk-averse investors trade corporate bonds with each other for either liquidity or informational reasons, and inventory fluctuates independently of the arrival of the news.[3] In the second step, I run a cross-sectional regression of estimated volume-return coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ on information asymmetry proxies that are either individual issue and issuer characteristics or compound information asymmetry indicators.

I find that $\hat{\beta}_1$ is negative. Bond prices tend to revert following average-volume days. For a typical high-asymmetry bond, $\hat{\beta}_1$ stands at around -0.4; if the price increases by 100 b.p. on an average-volume day, it falls by 40 b.p. the next day. For the same high-asymmetry bond, $\hat{\beta}_2$ is positive. For every additional standard deviation of the CtC volume, return autocorrelation increases (reversal reduces) by 0.08. $\hat{\beta}_3$ is about two times smaller

[3] I discuss the theoretical model in Appendix A. It is an extension of the model of Llorente, Michaely, Saar, and Wang (2002).

than $\hat{\beta}_2$ for the high-asymmetry bond. The results suggest that bond price changes are the most persistent when trading volumes are high, but dealers are reluctant to trade from their inventory capacity. Remarkably, I find that $\hat{\beta}_1$ and $\hat{\beta}_3$ decrease while $\hat{\beta}_2$ increases as information asymmetry grows in the cross-section of bonds. It means that of the two types of corporate bond liquidity providers, dealers and investors, the latter are more likely to see bond prices moving against them following a large trade. My results become stronger close to (but before) the issuers' earnings announcements. Then, the price of a typical high-asymmetry bond does not revert following a large C-to-C volume day, but it still reverts almost as strongly as the average -0.4 following a large C-to-D volume day. These findings suggest that information-motivated trading in corporate bonds does exist. It most likely occurs on high-volume days when dealers only match buyers and sellers and do not accept additional inventory risk. Related, I find a wedge in systematic reversal returns between portfolios of high- and low-asymmetry bonds, which is not subsumed by trading costs.

My paper contributes to several streams of literature. The paper investigates the impact of adverse selection on corporate bond liquidity provision and, with this regard, introduces novel empirical evidence to the literature on liquidity provision in OTC markets. Adrian, Boyarchenko, and Shachar (2017), Bessembinder, Jacobsen, Maxwell, and Venkataraman (2018), Choi and Huh (2022), Dick-Nielsen and Rossi (2018), and Berndt and Zhu (2019) study how the capital commitment of U.S. corporate bond dealers has changed over time. This literature has documented that liquidity provision has been shifting from dealer banks, subject to stricter regulatory requirements, to less constrained bond investors, which has implications for corporate bond illiquidity, trading costs, and market quality. Goldstein and Hotchkiss (2020) further discuss how dealers' capital commitment varies in the cross-section of bonds and finds that dealers tend to avoid holding inventory in riskier and less actively traded bonds. Dealers still intermediate trading in such bonds but act as pure brokers: they do not hold bonds on their books for more than a couple of minutes. My paper extends the findings of Goldstein and Hotchkiss (2020) with the analysis of *post-trade* bond price patterns. Between large CtC (investors provide liquidity) and CtD (dealers provide liquidity) volume days, corporate bond returns are more persistent after the former, especially in riskier and less actively traded bonds and ahead of major corporate information events. In other words, dealers mitigate adverse selection risk by passing it to non-dealer liquidity providers. I remain agnostic about the mechanisms behind dealers' ability to identify potentially informed transactions beforehand, but the non-anonymity of OTC trading seems to be the most likely explanation.

By identifying price and volume patterns that are consistent with the footprint of private information, my paper contributes to the debate on the presence of information-driven trading in the corporate bond market. Asquith, Au, Covert, and Pathak (2013) analyze the relationship between bond short interest and returns and find no evidence of information-based trading either in investment-grade or in high-yield bonds. Hendershott, Kozhan, and Raman (2019) use similar data on loaned bonds and conclude that information-driven trading is present in high-yield bonds but not in the investment-grade universe. In my paper, high-information-asymmetry bonds are not necessarily high-yield ones. My sample consists primarily of investment-grade bonds, yet information asymmetry proxies vary significantly in the sample. Therefore, I find evidence of information-based trading in investment-grade bonds. Han and Zhou (2014) discuss a positive relationship between microstructure-based information asymmetry measures and bond yield spreads and also argue that information motives are present in corporate bond trading. My paper further emphasizes the circumstances in which information likely stands behind bond price changes: when trading volumes are abnormally high and non-dealers provide liquidity.

Broader, my paper contributes to the discussion of intermediation frictions in OTC markets and the pricing implications of such frictions. Duffie, Gârleanu, and Pedersen (2005) present a theoretical framework where OTC market frictions drive asset illiquidity; Friewald and Nagler (2019) provide supporting empirical evidence from the corporate bond market. I demonstrate that conditional illiquidity, measured as the first autocorrelation of corporate bond returns, is a function of the underlying issue- and issuer-specific information asymmetry and dealers' capital commitment. The result has implications for the pricing of the cross-section of corporate bonds. Chordia, Goyal, Nozawa, Subrahmanyam, and Tong (2017) and Bai, Bali, and Wen (2019) show that a one-month lagged return is the strong return predictor in the cross-section of bonds. However, reversal portfolios have zero or negative Sharpe ratios after trading cost adjustment (Chordia et al. 2017). I obtain the same result for reversal portfolios constructed on low-information-asymmetry bonds. Yet, I show that reversal portfolios of high-asymmetry bonds survive trading cost adjustment.

Methodologically, my analysis of volume-return coefficients follows the tradition of Campbell, Grossman, and Wang (1993). In related work, Llorente et al. (2002) investigate the volume-return coefficients of U.S. stocks. I extend and adapt their motivating theoretical model and an empirical setup to the OTC corporate bond market.

The paper is organized as follows. Section II talks about the bond sample and the steps I take to estimate a volume-return relationship for individual bonds (the theoretical model that underlies my econometric strategy is presented in Appendix A). Section III presents estimated volume-return coefficients, and Section IV investigates its determinants, particularly information asymmetry proxies, in a cross-section of bonds. Section V explores how volume-return coefficients behave over time, around earnings announcements, and among bonds issued by the same firm. Section VI runs multiple robustness checks. Section VII discusses the implications of my results for reversal investment strategies. Section VIII concludes.

II. Data and measurements

A. Data sources

I construct the dataset of corporate bond prices and volumes from Enhanced TRACE tick-by-tick data. The sample is restricted to USD-denominated, fixed-coupon, not asset-backed, non-convertible corporate bonds. I apply the filters of Dick-Nielsen (2014) to clean the TRACE data. I calculate daily corporate bond prices as volume-weighted transaction prices within a given day. Bond characteristics are from Mergent FISD, issuer characteristics – from CRSP and IBES, bond holdings of mutual funds – from CRSP Mutual Funds, and the data on intermediating dealers – from the academic version of the TRACE dataset. I talk in more detail about the sample in Appendix B.

B. Sample filtering and 'active periods'

I estimate the dynamic volume-return relationship for each bond separately, which requires long enough time-series of returns and volumes for every bond. In a baseline specification of the volume-return relationship (1), I estimate four coefficients in an OLS regression. To avoid over-fitting, I require at least 60 daily observations per bond. However, corporate bonds experience waves of trading activity, as documented in Ivashchenko and Neklyudov (2018). The intervals between trading days with non-zero trading volume might be quite long. Asking for at least 60 consecutive business days is too restrictive, there are very few bonds that satisfy this criterion. Instead, I ask for more than 60 daily observations where every two successive observations are at most three business days apart.[4]

For some bonds, more than one sequence of trading days satisfies the criterion above. I call every such sequence an 'active period' and retain all active periods in the sample. I remove all days in between the active periods from the sample. Estimation of the volume-return relationship is carried out per bond per active period.

Also, I remove from the sample all active periods when a bond was either upgraded from high-yield (HY) to investment-grade (IG) territory or downgraded in the opposite direction. Bao, O'Hara, and Zhou (2018) analyze the corporate bond market liquidity around downgrades and find abnormal price and volume patterns associated with insurance companies selling bonds due to regulatory constraints. To ensure that downgrade anomalies do not drive my results, I remove all such periods from my sample. I also remove bonds with less

[4]Here I follow the methodology of Bao, Pan, and Wang (2011) who study the illiquidity of corporate bonds on the daily data and allow consecutive observations to be several days apart.

	Mean	Median	S.D.	Min	5th	25th	75th	95th	Max	N.Obs.
Issue size, mln $	992	750	802	9	150	500	1250	2500	15000	4569890
Maturity, years	8.09	5.50	7.68	1.08	1.50	3.25	9.00	27.08	87.17	4569890
Coupon rate, %	4.94	5.01	1.85	0.45	1.95	3.50	6.12	7.90	15.00	4569890
Rating	7.49	7.00	3.36	1.00	3.00	5.00	9.00	14.00	21.00	4569890
Age, years	4.06	3.08	3.79	0.08	0.33	1.42	5.50	11.25	28.92	4569890
Total bond return, %	0.02	0.02	1.13	-20.65	-1.30	-0.25	0.29	1.32	22.26	4569890
CtC volume, % of size	0.55	0.02	1.98	0.00	0.00	0.00	0.17	2.85	15.69	4569890
$-\Delta$Inventory, % of size	0.02	0.02	3.50	-19.40	-4.42	-0.19	0.36	4.21	18.48	4569890
CtD volume, % of size	1.50	0.28	3.17	0.00	0.01	0.06	1.24	7.85	19.40	4569890
Realized bond bid-ask, %	1.07	0.64	1.29	0.00	0.09	0.30	1.39	3.39	19.66	2795169
No. mutual fund owners	47.3	40.0	42.2	0.0	0.0	15.0	67.0	129.0	402.0	4569890
No. dealers	37.2	33.0	17.7	1.0	17.0	25.0	45.0	70.0	289.0	4569890
Issuer equity value, bln $	88.3	49.3	108.4	0.0	2.6	16.3	134.9	258.1	1103.5	4202245
Stock bid-ask, %	0.05	0.03	0.10	0.00	0.01	0.02	0.05	0.16	1.97	4202239
Days to earnings announcement	44.4	44.0	26.3	0.0	4.0	22.0	67.0	85.0	92.0	3449469

Table I. Summary statistics of the bond-day panel when only active trading periods are retained. The sample period is from Jan 4, 2005, to Dec 31, 2018. Bond transactions are from Enhanced TRACE (data on bond dealers – from Academic TRACE), bond characteristics – from Mergent FISD, stock prices and issuer characteristics – from CRSP and IBES, and fund holdings – from CRSP Mutual Funds. For every bond, I retain only long sequences (at least 60 trading days) of close daily observations (every two consecutive trading days are at most three business days apart). Besides, I exclude from the sample active periods that contain a crossing of the investment-grade/high-yield rating threshold. I keep only bonds with more than one year to maturity in the sample. The issue size is the outstanding notional amount. Rating is on a conventional numerical scale from 1 (AAA) to 21 (C). The total bond return consists of the change in the clean price and the accrued interest. The CtC (client-to-client) trading volume is the minimum between total client purchases and total client sales per bond per day; it is always positive. $-\Delta$Inventory is the difference between client purchases and client sales; it can be positive (dealers' inventory decreases) or negative (dealers' inventory increases) depending on which of the two is greater. Its absolute value is the CtD (client-to-dealer) trading volume (the absolute value of the change in aggregate broker-dealer inventory in a given bond). All trading volumes are expressed in percentages of the outstanding notional amount. The realized bond bid-ask spread is the difference between volume-weighted average client buy and sell prices, expressed as a percentage of the daily average price. It is computed only for the days with at least three trades, hence some missing observations. 'No. mutual fund owners' is the number of individual funds that hold the bond (according to the most recent fund holdings report) as of the bond trading date. 'No. dealers' is the number of unique dealers that intermediated trades in the bond in a given month. 'Stock bid-ask' is the difference between the closing bid and ask stock prices of the issuer, in % of the closing mid-price (all from CRSP). Quarterly earnings announcement days are from IBES and are only used in Section V.A (there, bond trading days more than 92 days before the next earnings announcement are excluded from consideration). For further details about the sample, see Appendix B.

than one year to maturity from the sample. Such bonds are excluded from major bond market indices, which also drives substantial institutional rebalancing and creates abnormal price patterns that are not the primary focus of this study.

Table I presents summary statistics of the bond-day panel where only active periods are retained in the sample. My filtered sample includes around 4.6 million bond-day observations that cover almost 16 thousand distinct active periods between 2005 and 2018 and 7 thousand different bonds issued by more than 1 thousand firms. An average bond in the sample is an investment-grade bond with an outstanding notional amount of around 1 billion USD, 8 years to maturity, and a 5% coupon rate. Its average daily total return is 2 b.p., and the realized bid-ask spread is about 1%. The bond is held by 47 mutual funds and is traded by 37 unique dealers. This paper's sample of active trading periods constitutes about 20% of the entire TRACE corporate bond records. The excluded bonds are less liquid, riskier, and have smaller outstanding amounts than the sampled bonds.

C. *Volume measures*

To measure the CtC trading volume, I first compute total daily client purchases from dealers and client sales to dealers; call it V_{it}^{buy} and V_{it}^{sell} respectively for bond i on day t. The minimum of the two is my measure of the CtC trading volume:

$$\text{CtC volume}_{it} = V_{it}^{(c)} = \min\left\{V_{it}^{\text{buy}}, V_{it}^{\text{sell}}\right\}.$$

It represents a trading volume that has no impact on aggregate dealers' inventory in bond i at the end of the trading day t as compared to day $t-1$; it is non-negative by construction. The difference between client purchases and client sales is a negative change in dealers' inventory:

$$-\text{Change in inventory}_{it} = V_{it}^{(s)} = V_{it}^{\text{buy}} - V_{it}^{\text{sell}}.$$

$V_{it}^{(s)}$ can be either positive or negative. Positive values represent net purchases by clients from dealers and correspond to a decrease in total broker-dealers' inventory in bond i on day

t. Conversely, negative values of $V^{(s)}$ are increases in dealers' inventory. In equation (1), I consider the absolute value of $V_{it}^{(s)}$, which I call the CtD trading volume:[5]

$$\text{CtD volume}_{it} = |-\text{Change in inventory}_{it}| = \left|V_{it}^{(s)}\right|.$$

Table I shows that the CtD volume is, on average, several times higher than the CtC volume.[6]

Table C1 in Appendix demonstrates that there is a positive statistical relationship between the CtD and the CtC trading volumes (active trading by investors coincides with big changes in dealers' inventories), but the corresponding correlation coefficient is relatively small (less than 0.2). Also, for about two-thirds of bond-active periods, we can not reject the hypothesis that $\text{Corr}\left(V_t^{(c)}, V_t^{(s)}\right) = 0$, i.e., bond inventory is equally likely to fall or to increase on high CtC volume days. The persistence of both the CtC and the CtD trading volumes is relatively small, too: the average time series autocorrelation of both measures of trading volume is less than 0.1.

D. *Proxies for information asymmetry*

In empirical tests, I use individual issue- and issuer-specific variables and the principal components of different groups of individual variables to proxy for the extent of information asymmetry between bond investors. Some variables are bond-level proxies:

- realized bond bid-ask spread;
- bond outstanding notional amount;
- the number of mutual funds that hold the bond;
- the number of dealers that intermediate trades in the bond.

Other variables are issuer-level information asymmetry proxies:

[5] Such an imposed symmetry of return autocorrelation conditional on increases and decreases in dealers' inventory is a simplification. However, it does not undermine an alleged dependence of a respective volume-return coefficient β_3 on information asymmetry. An investigation of the asymmetries in conditional price reversals is beyond the scope of this paper.

[6] Notice that a traditional measure of daily trading volume (excluding inter-dealer transactions), $V^{\text{buy}} + V^{\text{sell}}$, is equal to CtC volume $+ 2 \cdot$ CtD volume. I do not take into account inter-dealer trades when I construct my measures of the trading volume.

- issuer market capitalization;
- stock bid-ask spread.

The last two proxies are calculated only for traded companies. I assume that informed trading is more likely in bonds with wider (stock or bond) bid-ask spreads, fewer mutual fund holders and intermediating dealers, lower outstanding amounts, and that are issued by smaller firms. Below I justify in more detail the use of these variables as proxies for information asymmetry.

The number of mutual funds that own the bond is related to the number of buy-side analysts scrutinizing bond valuations and the credit quality of the issuer. As in equity literature, I assume that analyst coverage is negatively related to information asymmetry between investors. Similarly, **the number of broker-dealers** intermediating trades in the bond is positively related to sell-side analyst coverage and, hence, negatively related to information asymmetry. The number of active broker-dealers also measures competition among them in a given bond. The lack of competition likely affects an average-volume day reversal, β_1 in equation (1), similarly to high information asymmetry: prices of bonds traded in a less competitive market should revert more on average. However, there is no straightforward explanation for why prices for low-dealer-competition bonds should revert *less* following high-volume days (the positive relationship between β_2 in equation (1) and information asymmetry) unless low competition among dealers is due to high information asymmetry in the first place.

Issuer and issue sizes are typical proxies for trade informativeness in the literature. Both are related to a broader investor base and, again, more in-depth analyst coverage, which supposedly leads to a higher number of investors who are ready to arbitrage out bond misvaluations. As Table C3 in Appendix shows, issue and issuer sizes are indeed positively correlated with the numbers of intermediating dealers and mutual funds that own the bond.

Stock and bond bid-ask spreads are also classic measures of information asymmetry. In Glosten and Milgrom (1985), the bid-ask spread is positively related to the extent of informed trading. A dealer wants to be compensated ex-ante for the risk of being adversely selected and charges wider spreads to trade riskier securities. There is, however, a confounding effect of bid-ask spreads on conditional price reversals. The mere existence of bid-ask spreads implies price reversals as in Roll (1984), i.e., the 'bid-ask bounce' effect. It implies stronger reversals for bonds with wider spreads (even when, ex-post, there turns out to be only liquidity trading). Hence, the impact of the bid-ask bounce on the average-day return autocorrelation, β_1 in equation (1), is similar to the expected effect of information asymmetry. The impact of the bid-ask bounce on β_2 and β_3 in equation (1) is unclear because it depends on whether the effect becomes stronger or weaker with higher trading volumes. To avoid these concerns, compound information asymmetry indicators constructed further in this section utilize sets of proxies both with and without bond bid-ask spreads.

The set of information proxies considered above is not exhaustive. In unreported results, I extended it further with both bond-level (bond return volatility, yield spread) and issuer-level (availability of a single-name CDS contract on the issuer, equity analyst disagreement, stock return volatility) characteristics to find no change in key quantitative and qualitative results of the paper. Rather than extending the list of individual proxies (all of which are imperfect measures of information asymmetry), I now attempt to blend already discussed bond and stock characteristics in a single compound information asymmetry index.

E. Compound information asymmetry indicators (indices)

A compound cross-sectional information asymmetry characteristic serves two purposes in this paper. First, it limits the confounding impact of non-information components in individual bond and issuer characteristics on volume-return coefficients (the second stage of empirical analysis covered in Section IV). Second, it simplifies the presentation of results. I test for the impact of information asymmetry on volume-return coefficients and, eventually, conditional bond return autocorrelation. It is easier to interpret such a test when information asymmetry is a single metric in the cross-section of bonds.

I construct information asymmetry indices by extracting, in the cross-section of bonds, the first principal components from groups of bond-average values of individual asymmetry proxies discussed above.[7] In the cross-section, each individual proxy is standardized (demeaned and divided by a cross-sectional standard deviation) before the extraction of the principal components. The indices and respective groups are:

- PC_{all}: stock and bond bid-ask spreads, (negative) issuer and issue sizes, (negative) numbers of mutual fund holders and intermediating dealers.
- PC_{bond}: same as PC_{all}, but issuer-level characteristics (stock bid-ask, issuer size) excluded.
- $PC_{\text{bond-ex-ba}}$: same as PC_{bond}, but bond bid-ask spread excluded.

Issuer and issue size and the numbers of fund holders and intermediating dealers are taken with a negative sign to facilitate the interpretation of extracted principal components. The first principal component loads positively on all (scaled) individual characteristics in all three considered sets. For instance, PC_{bond} increases with the average bond realized bid-ask spread and decreases with issue size, number of mutual funds, and dealers. Table C2 in the Appendix presents the loadings of principal components on individual characteristics. Issuer-level characteristics have the lowest loadings, but they are still substantial. For instance, PC_{all} has the (lowest) loading of 0.25 on a standardized stock bid-ask spread and the (highest) loading of 0.55 on a (negative) standardized issue size. The issue size has the highest loading across all indices. These first principal components explain between 42% (PC_{all}) and 70% ($PC_{\text{bond-ex-ba}}$) of variance, which is substantial. Intuitively interpretable loadings and a high portion of explained variance highlight the validity of constructed indicators as compound cross-sectional information asymmetry proxies.

[7] In the baseline specification, individual proxies are averaged for each bond in the same active trading periods in which volume-return coefficients are estimated. To address a possible confounding effect of the measurement error in the second stage of my empirical analysis, I run multiple robustness checks in Section VI. In particular, I show that the main results of the paper hold when the principal components are extracted from the *initial* (observed) values of individual information asymmetry proxies (the values at the beginning of the first active trading period for each bond).

III. Volume-return relationship

I estimate equation (1) separately for every bond and every active period rescaling trading volumes such that β_1 measures the first return autocorrelation on average-volume trading days:

$$R_{t+1} = \beta_0 + \beta_1 R_t + \beta_2 R_t \tilde{V}_t^{(c)} + \beta_3 R_t \tilde{V}_t^{(s)} + \epsilon_{t+1}. \tag{2}$$

Above, R_{t+1} is the total bond return between t and $t+1$, $\tilde{V}_t^{(c)}$ is the CtC trading volume on day t, standardized[8] for every active period separately, and $\tilde{V}_t^{(s)}$ is the CtD trading volume (the absolute value of inventory change) on day t, also standardized.

On the days when both the CtC and the CtD trading volumes are at the average level for a given bond in a considered active period, the first return autocorrelation is β_1. On the days when the CtC volume is one standard deviation above the mean ($\tilde{V}_t^{(c)} = 0$) and the change in inventory is at the average level ($\tilde{V}_t^{(s)} = 0$), the first return autocorrelation is $\beta_1 + \beta_2$. Conversely, when only the CtD volume is 1 standard deviation above the average, return autocorrelation equals to $\beta_1 + \beta_3$. Negative values of β_1 would mean that prices revert following average-volume days. Positive values of β_2 and β_3 would mean that prices tend to revert less following high-volume days. In this short section, I present and discuss the estimated volume-return coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$, and in the next section, I investigate the relationship between the coefficients and information asymmetry proxies, which is the main focus of this study.

Table II gives a snapshot of $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ estimated for each bond in every active period. I truncate estimated volume-return coefficients in the sample of active trading periods at the 1% and the 99% levels to limit the impact of extreme estimates on the second-stage regression. The average bond-active period has the first return autocorrelation of approximately -0.33. If the price drops today by 100 b.p. and both trading volumes are at the average level, the price will tend to increase by 33 b.p. tomorrow. One-third of the initial

[8] De-meaned and divided by the sample standard deviation so that $\tilde{V}_t^{(c)}$ has a zero mean and a unit variance for each bond and each active period.

	Mean	Med.	No.>0	No.<0	No.>0*	No.<0*	No. Obs.
$\hat{\beta}_1$	-0.3346	-0.3488	177	15697	5	14273	15874
$\hat{\beta}_2$	0.0687	0.0588	11141	4733	2650	517	15874
$\hat{\beta}_3$	0.0531	0.0519	10774	5100	3203	772	15874

Table II. Summary statistics of the estimated volume-return coefficients of equation (2). Each estimated coefficient is per bond per active period. There are at most fifteen active periods per bond. Returns are total returns between t and $t+1$. Trading volumes are de-meaned and standardized per bond per active period. Mean and Med. are, respectively, sample average and sample median. 'No. $>$ ($<$) 0' is the number of positive (negative) coefficients. 'No. $>$ ($<$) 0*' is the number of positive (negative) coefficients significant at a 10% confidence level. The number of observations is the number of bond-active periods. The sample of estimated volume-return coefficients is truncated at the 1% and the 99% levels.

price decrease reverts the next day. The average $\hat{\beta}_2$ of 0.07 suggests that following high CtC volume days, prices tend to revert less. In a previous example, if the initial 100 b.p. price decrease was accompanied by one standard deviation above-average CtC trading volume, then the next day reversal would be close to one-fourth rather than one-third. The average $\hat{\beta}_3$ of around 0.05 suggests that prices revert following high CtD volume days either. The difference between the average $\hat{\beta}_2$ and $\hat{\beta}_3$ is not statistically significant.

At this stage, we can not infer much from estimated volume-return coefficients. The signs and the magnitudes of the coefficients look reasonable. Strongly negative $\hat{\beta}_1$ is a reflection of the high illiquidity of the corporate bond market. The values of $\hat{\beta}_2$ and $\hat{\beta}_3$ are close; hence, both types of trading volume interact statistically similarly with reversals. Positive $\hat{\beta}_2$ and $\hat{\beta}_3$ can be consistent with the presence of informed trading, but can also reflect correlated trading volumes, or the interaction of the bid-ask bounce or bond riskiness with the trading volume. In the next section, I investigate explanatory factors of the cross-sections of volume-return coefficients with a particular focus on the impact of information asymmetry.

IV. Determinants of volume-return coefficients

A. *Empirical design*

In the introduction, I put forward an intuition on how volume-return coefficients β_1, β_2, and β_3 in equation (2) should vary with information asymmetry. In particular, I suggest that more information asymmetry implies lower β_1 (stronger reversals on average), higher β_2 (weaker reversals following high-volume days when dealers' inventory does not change), and no particular effect on β_3 (no difference in reversals between high- and low-information-asymmetry bonds following days when dealers' inventory is the only trading capacity). One gets the same relationship between volume-return coefficients and information asymmetry in a theoretical model a-là Llorente et al. (2002) extended with noisy changes in the secondary market supply (dealers' inventory) that are independent of the arrival of private news. I present such a model formally in Appendix A. In this section, I am testing the predictions of the model empirically in the cross-section of bonds.

The estimates $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ obtained in the previous section are per bond and per active period. There is more than one active period for every other bond in the sample, but there are at most fourteen active periods per bond. I take bond averages to obtain the cross-section of coefficients, and in the rest of this section, I fit explanatory linear models to this cross-section (time-varying volume-return coefficients will be discussed in Section V.C). Call $\hat{\beta}_{n,i}(k)$ a column-vector of estimates ($n = 1$, 2, or 3) for individual bonds $i \in \{1, \ldots, N\}$ with credit ratings $k \in \{1, \ldots, 21\}$.[9] As the baseline, I fit the following model for each n (i.e., the cross-section of each volume-return coefficient) separately:

$$\hat{\beta}_{n,i}(k) = c_n \cdot \text{Info asymmetry proxy (-ies)}_{n,i} + \text{Rating } k \text{ FE}_n + \epsilon_{n,i}, \tag{3}$$

[9] In this section, k is the prevalent rating in the active trading period in which $\hat{\beta}_n$ is estimated. In Section VI, I demonstrate that the results hold when k is the credit rating at the beginning of the active period.

where, for every n, $\epsilon_{n,i}$ is distributed as a zero-mean Normal. Rating fixed effects (FE) are control variables in the baseline specification (3).[10] If my intuition about the dependence of volume-return coefficients on information asymmetry proxies is correct, I should find $\hat{c}_1 < 0$, $\hat{c}_2 > 0$, and $\hat{c}_3$ at least smaller than $\hat{c}_2$ or, better, insignificantly different from zero or significantly negative.

	Mean	Median	S.D.	Min	5th	25th	75th	95th	Max	N.Obs.
$\hat{\beta}_1$	-0.32	-0.34	0.12	-0.63	-0.48	-0.40	-0.25	-0.10	0.07	7206
$\hat{\beta}_2$	0.06	0.06	0.12	-0.52	-0.12	0.00	0.12	0.27	0.85	7206
$\hat{\beta}_3$	0.05	0.05	0.10	-0.38	-0.12	-0.01	0.11	0.22	0.51	7206
Credit rating	7.78	8.00	3.32	1.00	3.00	6.00	9.00	14.00	21.00	7206
Bond bid-ask, %	1.13	0.78	0.99	0.07	0.23	0.45	1.50	3.16	14.81	7206
No. mutual fund owners	40.6	34.4	38.2	0.0	0.0	8.5	59.2	114.9	381.8	7206
Issue size, bln $	0.77	0.58	0.69	0.01	0.05	0.35	1.00	2.00	9.00	7206
No. dealers	32.1	28.7	13.1	7.7	17.4	23.3	37.5	58.0	120.7	7206
Issuer size, bln $	75.4	37.3	105.9	0.0	2.3	12.4	102.6	235.8	930.8	6671
Stock bid-ask, %	0.05	0.03	0.07	0.01	0.01	0.02	0.06	0.16	1.33	6671
PC_{all}	0.00	0.23	1.60	-12.50	-3.08	-0.70	0.98	2.13	8.16	6671
PC_{bond}	0.00	0.21	1.52	-13.85	-2.84	-0.66	0.93	2.10	5.91	7206
$\text{PC}_{\text{bond-ex-ba}}$	0.00	0.36	1.45	-14.44	-2.83	-0.51	0.94	1.57	2.27	7206

Table III. Summary statistics of the cross-section of volume-return coefficients and their predictors. The sample contains bond averages computed across all active periods in case there is more than one for a given bond. PC_{all}, PC_{bond}, and $\text{PC}_{\text{bond-ex-ba}}$ are the first principal components of (standardized) information asymmetry proxies (issuer and issue sizes, as well as numbers of dealers and mutual funds, are taken with a negative sign, so that higher covariate readings are associated with more information asymmetry). PC_{all} is extracted from the set of all six information asymmetry proxies. PC_{bond} is the first principal component of four bond-specific information asymmetry proxies (i.e., issuer size and stock bid-ask do not contribute to PC_{bond}). $\text{PC}_{\text{bond-ex-ba}}$ further excludes realized bond bid-ask from the list of factors.

Table III presents summary statistics of the cross-section of estimated volume-return coefficients with their potential explanatory factors. There are about 7 thousand individual bonds issued by 1 thousand firms in the cross-section. More than 90% of these bonds are issued by public firms. There is substantial variation in both the left-hand side and the right-hand side variables of regression (3) as Table III shows. Table C3 in Appendix further presents cross-sectional correlations of information asymmetry proxies.

[10]The model in Apppendix A shows that a tested cross-sectional relationship between volume-return coefficients and information asymmetry holds when bond riskiness remains constant. Rating fixed effects in this second-stage model control for bond riskiness. The results are quantitatively similar when credit ratings are replaced in the second-stage regression with realized bond return volatility (unreported).

B. Main results

Table IV presents estimated regressions (3) of volume-return coefficients on individual information asymmetry proxies. Issuer and issue sizes, as well as the numbers of mutual fund owners and intermediating dealers, are taken with a negative sign so that higher values of all right-hand side variables are associated with higher information asymmetry. Table IVa contains the results for $\hat{\beta}_1$. Observe that all information asymmetry proxies have a significantly negative impact on $\hat{\beta}_1$ if included in the regression separately. In a joint model 7, bond-specific information asymmetry proxies maintain significantly negative loadings. In a joint model 8 for public issuing firms, only the issuer's stock bid-ask spread flips the sign to positive. These results suggest that, on average, price reversals become more pronounced ($\hat{\beta}_1$ becomes more negative) for higher information asymmetry bonds: the bonds with fewer fund owners and intermediating dealers, lower issue and issuer size, and higher bid-ask spread.

Table IVb presents the results for $\hat{\beta}_2$. Recall that higher β_2 means less strong reversals following days when investors trade a lot essentially with each other and dealers do not hold any additional inventory by the end of the trading day. I expect $\hat{\beta}_2$ to be increasing in information asymmetry: reversals must be less strong for high asymmetry bonds when informed trading is most likely, i.e., after high CtC volume days. Observe first in Table IVb that, in line with the expectation, all bond-specific information asymmetry proxies enter the models for $\hat{\beta}_2$ significantly positively when included separately (models 1 to 4). The loadings on issuer-specific proxies (issuer size and stock bid-ask) are insignificant (models 5 and 6). In joint models 7 and 8, only the stock bid-ask spread turns significantly negative. Otherwise, the results in Table IVb suggest that higher-asymmetry bonds exhibit less strong price reversals following high CtC volume days.

Table IVc presents the regressions for $\hat{\beta}_3$. The interpretation of β_3 is analogous to β_2, but now we are talking about the reversals following days when changes in dealers' inventory are the only source of the trading volume. Higher β_3 means that prices tend to revert less following high CtD volume days. Unlike for β_2, I do not expect to find any particular dependence of β_3 on information asymmetry because dealers would rather pass high-asymmetry bonds to other investors and would not hold excess inventory in bonds with less transparent valuations.

Table IVc shows that there is indeed no clear-cut dependence of $\hat{\beta}_3$ on information asymmetry. For instance, bond bid-ask spread, the (negative) number of mutual fund bond owners, and the (negative) issue size have significantly negative loadings in models 1–3 (opposite to what we found for $\hat{\beta}_2$), while the (negative) number of dealers has a significantly positive loading (as for $\hat{\beta}_2$). In joint models 7 and 8 as well, there are both positive and negative loadings on the variables of interest.

Individual right-hand side variables in Table IV are noisy measures of information asymmetry. The interpretation of the resulting effect on volume-return coefficients is ambiguous when all individual proxies are included in the regressions jointly, as in models 7 and 8. To better summarize the relationship between information asymmetry and volume-returns coefficients, I regress $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ on compound information asymmetry indicators (indices). These are the first principal components extracted from different sets of information asymmetry proxies in the cross-section of bonds. The regression models are as in (3).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bond bid-ask	−0.035***						−0.009***	−0.015***
	(0.002)						(0.002)	(0.002)
–No. funds		−0.062***					−0.033***	−0.036***
		(0.002)					(0.003)	(0.003)
–Issue size			−0.063***				−0.030***	−0.026***
			(0.002)				(0.003)	(0.004)
–No. dealers				−0.039***			−0.009**	−0.008**
				(0.002)			(0.003)	(0.004)
–Issuer size					−0.037***			−0.009***
					(0.006)			(0.002)
Stock bid-ask						−0.009**		0.009***
						(0.004)		(0.002)
Rating FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	7,206	7,206	7,206	7,206	6,671	6,671	7,206	6,671
R^2	0.111	0.300	0.299	0.139	0.080	0.035	0.345	0.369

Note: *p<0.1; **p<0.05; ***p<0.01

(a) Models for $\hat{\beta}_1$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bond bid-ask	0.009***						0.006***	0.004
	(0.002)						(0.001)	(0.002)
–No. funds		0.011***					0.002	0.002
		(0.002)					(0.002)	(0.002)
–Issue size			0.013***				0.005**	0.005**
			(0.002)				(0.002)	(0.002)
–No. dealers				0.012***			0.008***	0.008***
				(0.001)			(0.002)	(0.002)
–Issuer size					0.001			−0.005
					(0.002)			(0.004)
Stock bid-ask						−0.002		−0.004**
						(0.002)		(0.002)
Rating FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	7,206	7,206	7,206	7,206	6,671	6,671	7,206	6,671
R^2	0.009	0.013	0.015	0.014	0.004	0.005	0.019	0.017

Note: *p<0.1; **p<0.05; ***p<0.01

(b) Models for $\hat{\beta}_2$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bond bid-ask	−0.026***						−0.029***	−0.025***
	(0.002)						(0.002)	(0.002)
–No. funds		−0.005**					0.010***	0.008**
		(0.002)					(0.002)	(0.003)
–Issue size			−0.004**				−0.007***	−0.008**
			(0.002)				(0.002)	(0.003)
–No. dealers				0.007***			0.008***	0.008***
				(0.002)			(0.002)	(0.002)
–Issuer size					0.010***			0.011***
					(0.002)			(0.002)
Stock bid-ask						−0.008**		−0.0003
						(0.004)		(0.004)
Rating FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	7,206	7,206	7,206	7,206	6,671	6,671	7,206	6,671
R^2	0.071	0.018	0.017	0.020	0.026	0.025	0.079	0.073

Note: *p<0.1; **p<0.05; ***p<0.01

(c) Models for $\hat{\beta}_3$

Table IV. Cross-sectional regressions of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ on individual information asymmetry proxies. Each model uses a fixed-effects estimator with rating-clustered standard errors. The regressors are standardized in the cross-section (have zero mean and unit variance).

	$\hat{\beta}_1$			$\hat{\beta}_2$			$\hat{\beta}_3$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PC_{all}	-0.043^{***}			0.007^{***}			-0.002^{*}		
	(0.001)			(0.002)			(0.001)		
PC_{bond}		-0.043^{***}			0.010^{***}			-0.005^{***}	
		(0.001)			(0.001)			(0.001)	
$PC_{bond\text{-}ex\text{-}ba}$			-0.044^{***}			0.010^{***}			-0.001
			(0.001)			(0.001)			(0.001)
Rating FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	6,671	7,206	7,206	6,671	7,206	7,206	6,671	7,206	7,206
R^2	0.342	0.337	0.322	0.012	0.018	0.017	0.022	0.020	0.016
Note:	*p<0.1; **p<0.05; ***p<0.01								

Table V. Cross-sectional regressions of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ on information asymmetry indices. Models (1)-(3) are for $\hat{\beta}_1$, models (4)-(6) – for $\hat{\beta}_2$, and models (7)-(9) – for $\hat{\beta}_3$. Each model uses a fixed-effect estimator with rating-clustered standard errors.

Table V presents the estimates. Models 1 to 3 regress $\hat{\beta}_1$ on three information asymmetry indices PC_{all}, PC_{bond}, and $PC_{bond\text{-}ex\text{-}ba}$. In all three regressions, the coefficients of interest are close to -0.04 and are significant. Models 4-6 in Table V confirm that $\hat{\beta}_2$ increases with information asymmetry. Compared with low-asymmetry bonds, high-asymmetry ones exhibit less pronounced price reversals following high CtC volume days. The size of the effect is comparable across different information asymmetry indices. Finally, models 7-9 in Table V suggest that the relationship between $\hat{\beta}_3$ and information asymmetry is either negative or absent.

The fact that the relationship between $\hat{\beta}_2$ and information asymmetry is positive and the one between $\hat{\beta}_3$ and information asymmetry is negative corroborates the hypothesis that the information content of bond prices on high CtC volume days differs from the one on high CtD volume days. Tables IV and V show that high-information-asymmetry bonds experience, on average, stronger price reversals than low-asymmetry bonds. However, following high CtC trading volume days, this 'gap' in reversals closes; such a thing does not happen following days with high CtD trading volume. Figure 2 plots the difference in reversals between high and low asymmetry bonds using the relationships between $\hat{\beta}_n$ and PC_{bond} from Table V. The results look similar when I use PC_{all} and $PC_{bond\text{-}ex\text{-}ba}$ (unreported).

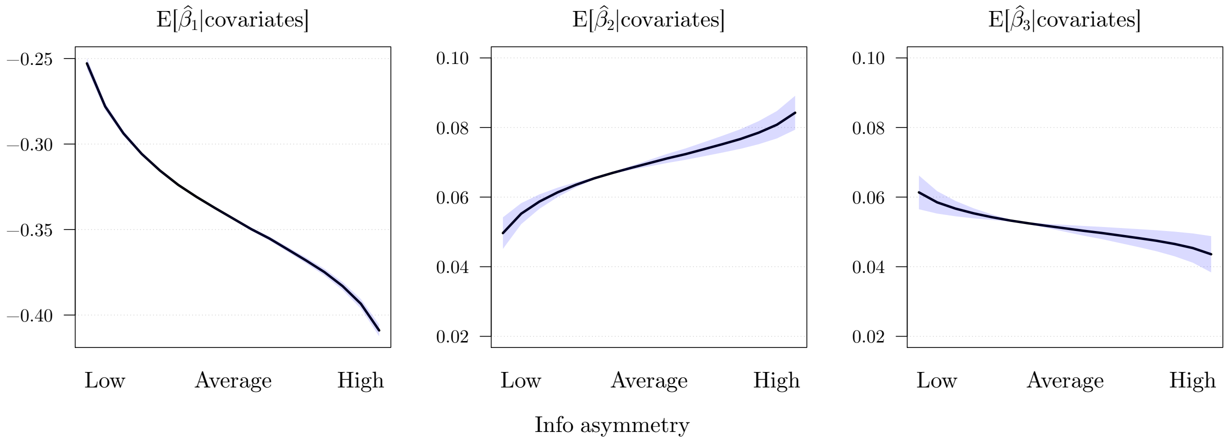

Figure 2. Point estimates and confidence intervals for the expected values of volume-return coefficients for a BBB-rated bond. The calculations are based on models with PC_{bond} from Table V. On the x-axes, from left to right, are the percentiles of PC_{bond}, from the 10^{th} ('Low' information asymmetry) to the 90^{th} ('High' information asymmetry). The credit rating remains fixed at the 'BBB' level. Solid lines are points estimates, and shaded areas around them are 95% confidence bands.

The left panel of Figure 2 shows the average values of $\hat{\beta}_1$ across percentiles of PC_{bond}. They are decreasing monotonically from -0.25 for the bonds with little information asymmetry (10^{th} percentile of PC_{bond}) to almost -0.4 for the bonds with high asymmetry (90^{th} percentile of PC_{bond}). The predicted reversal for high-asymmetry bonds is much stronger than for low-asymmetry bonds following average-volume days. The middle panel in Figure 2 shows an additional impact of high CtC volumes on next-day reversals. The average values of $\hat{\beta}_2$ are monotonically increasing from 0.05 for low-asymmetry to 0.08 for high-asymmetry bonds. It means that every additional standard deviation of the CtC volume reduces the difference in next-day reversals between high- and low-asymmetry bonds by about 0.03.[11] Finally, the right panel in Figure 2 demonstrates that predicted $\hat{\beta}_3$ is less sensitive to the degree of information asymmetry than $\hat{\beta}_2$ and decreases from 0.06 to 0.04 as information asymmetry grows. Figure A1 in Appendix A shows similar-shaped functions for the dependence of $\hat{\beta}_n$ on information asymmetry in a theoretical model in which informed trading *never* goes to

[11] Figure C1 in Appendix C shows that, following a day with the CtC trading volume 2 standard deviations above the average, the price reversal of high-asymmetry bonds is less strong than that of low-asymmetry bonds.

dealers. The evidence presented in this section suggests that information-driven trading in corporate bonds is much more likely when investors essentially trade with each other rather than when they trade with dealers.

V. Announcement, issuer, and time effects in volume-return coefficients

I have established the relationship between corporate bond price reversals, trading volume, and bond information asymmetry in the cross-section of bonds. Now, I extend the empirical evidence along several dimensions. In this section, I modify my baseline methodology to study how volume-return coefficients vary across time, within the issuer, and around corporate announcements. The results presented here confirm the main qualitative finding of the paper: corporate bond dealers avoid potentially informed client trade flow.

A. *Pre-announcement effects*

In Section IV, I contrasted price reversals following high CtC and CtD volume days among bonds with different information asymmetry. Information motives for trading are not constant over time. Information-driven trading is likely more intense around earnings announcements (Dechow, Sloan, and Zha 2014). Then, one should find the least strong corporate bond price reversals right before earnings announcements and in the bonds with the least transparent valuations. I test for such an effect with the following modification of my baseline methodology. I modify the volume-return relationship (2) to separate days close to quarterly earnings announcements from all other trading days:

$$R_{t+1} = \beta_0 + \beta_1 R_t + \beta_2 R_t \tilde{V}_t^{(c)} + \beta_3 R_t \tilde{V}_t^{(s)} + \beta_4 R_t \tilde{V}_t^{(c)} \mathbb{1}_t^{\text{EA}} + \beta_5 R_t \tilde{V}_t^{(s)} \mathbb{1}_t^{\text{EA}} + \epsilon_{t+1}. \quad (4)$$

In equation (4), $\mathbb{1}_t^{\text{EA}}$ is a dummy variable that takes the value of 1 if, from day t, there is at least one and at most ten days *to* the following quarterly earnings announcement for a given bond issuer (otherwise, the dummy is zero). It changes the interpretation of volume-return coefficients. Here, $\beta_1 + \beta_2$ is the average reversal following a far-from-announcement trading day t with the CtC volume one standard deviation above average for that bond and that active period. For a close-to-announcement trading day t with the same CtC volume, the average reversal is $\beta_1 + \beta_2 + \beta_4$. Similarly, following a CtD volume one standard deviation above average, the value β_5 measures the difference in average reversals between days close to and distant from earnings announcements.

I estimate equation (4) for the same subset of individual bonds issued by public firms and the same active trading periods as in Section IV. As before, the distributions of estimated volume-return coefficients (including $\hat{\beta}_4$ and $\hat{\beta}_5$ here) across bonds and active trading periods are truncated at the 1$^{\text{st}}$ and 99$^{\text{th}}$ percentiles to limit the impact of extreme observations on the second-stage results. Table C4 in Appendix C summarizes the cross-section for the second-stage analysis of this section. There is little difference to the cross-section in Section IV. The cross-sectional averages of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ are at -0.31, 0.06, and 0.05, respectively (virtually unchanged from the baseline analysis). The average values of $\hat{\beta}_4$ and $\hat{\beta}_5$ are, respectively, 0.03 and 0.02. For the second stage, I use the same regression model (3) as before.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
PC_{bond}	-0.046^{***}	0.007^{***}	-0.001	0.018^{**}	-0.011^{***}
	(0.001)	(0.001)	(0.001)	(0.006)	(0.004)
Rating FE	YES	YES	YES	YES	YES
Observations	5,051	5,051	5,051	5,051	5,051
R^2	0.353	0.011	0.016	0.005	0.006
Note:	*p<0.1; **p<0.05; ***p<0.01				

Table VI. Cross-sectional regressions of extended volume-returns coefficients on information asymmetry indices. Volume-return coefficients $\hat{\beta}_1$–$\hat{\beta}_5$ are estimated as in (4). The cross-section of volume-return coefficients and predictors is summarized in Table C4 in Appendix C. PC_{bond} is the first principal component extracted from the cross-section of standardized bond-specific information asymmetry proxies (number of fund owners, intermediating dealers, and the issue size are taken with a negative sign). Higher values of PC_{bond} are associated with higher information asymmetry. Each model uses a fixed-effect estimator with rating-clustered standard errors.

Table VI presents the regressions of volume-return coefficients $\hat{\beta}_1$–$\hat{\beta}_5$ on the information asymmetry index $\mathrm{PC}_{\mathrm{bond}}$ with credit rating fixed effects. For the clarity of exposition, I omit the results for other information asymmetry indices – they are quantitatively similar. For $\hat{\beta}_1$, I find almost the same negative loading on $\mathrm{PC}_{\mathrm{bond}}$ as in the baseline case in Table V. The higher the information asymmetry, the stronger the average bond price reversal is. The regressions for $\hat{\beta}_2$ and $\hat{\beta}_4$ tell a more nuanced story about bond price reversals following high CtC volume days that are far from ($\hat{\beta}_2$) and close to ($\hat{\beta}_4$) quarterly issuer earnings announcements. I find that the loadings on $\mathrm{PC}_{\mathrm{bond}}$ for both $\hat{\beta}_2$ and $\hat{\beta}_4$ are significantly positive, but the latter is almost two and a half times higher than the former. It means that the information content of CtC trades is the highest close to earnings announcements, as we expected. Also, notice that the respective estimate in Table V is 0.01, which is indeed between 0.007 and 0.018 (info asymmetry loading for $\hat{\beta}_2$ and $\hat{\beta}_4$, respectively) in Table VI.

Similar to the results for $\hat{\beta}_2$ and $\hat{\beta}_4$, there is a stark difference between volume-return coefficients that capture 'extra' bond return persistence after high CtD volume days that are far from ($\hat{\beta}_3$) or close to ($\hat{\beta}_5$) earnings announcements. Table VI claims that $\hat{\beta}_3$ is unrelated to $\mathrm{PC}_{\mathrm{bond}}$: there is no evidence of information-driven client-to-dealer trading even far from earnings announcements. The same applies to close-to-announcements days as $\hat{\beta}_5$ is negatively related to $\mathrm{PC}_{\mathrm{bond}}$ in Table VI. If there was information-driven trading in client-to-dealer transactions just before earnings announcements, prices would be more persistent following high CtD volume days, and more so in bonds with higher information asymmetry. We find the opposite: when dealers do trade from their inventory, prices are more likely to move in favor of bond dealers in high-asymmetry bonds relative to low-asymmetry bonds, even in pre-announcement periods. It confirms that bond dealers are capable of identifying and avoiding information-driven trade flow.

B. Within-issuer effects

There are firms that have many issued and not-yet-matured bonds. These bonds may differ in coupon rates, maturity, embedded options, and other characteristics. I investigate how volume-return coefficients differ across bonds of the same issuer. In Table VII, I present the estimates of a modification of model (3) only for firms with more than fifteen outstanding bonds. On top of credit rating fixed effects, I include issuer fixed effects in the regression models. Thus, Table VII shows within-firm dependence of volume-return coefficients on information asymmetry.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
PC_{bond}	-0.038^{***}	0.013^{***}	-0.006^{***}
	(0.001)	(0.002)	(0.002)
Rating, Issuer FE	YES	YES	YES
Issuer-clustered SE	YES	YES	YES
Observations	3,516	3,516	3,516
R^2	0.438	0.120	0.123
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table VII. Issuers with many bonds outstanding: cross-sectional regressions of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ on information asymmetry. The cross-section of bonds is restricted to issuers with at least fifteen outstanding bonds. Each model uses a fixed effects estimator with rating and issuer fixed effects. Standard errors are issuer-clustered.

I find that the signs of the impact of information asymmetry on $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ hold for the bonds of the same issuer. In the cross-section of bonds, $\hat{\beta}_1$ and $\hat{\beta}_3$ decrease in information asymmetry while $\hat{\beta}_2$ does increase, even controlling for issuer-specific effects. The loadings on PC_{bond} in Table VII are similar in size to those in Table V (without issuer fixed effects).

It suggests that private information some investors might possess is not only issuer-level (which is most likely private news about the credit quality of the issuer) but also bond-level. The bond-level information can be, for instance, private knowledge about liquidity trades of other investors, which yields a better estimate of price pressures and subsequent price reversals.[12] It can also be private knowledge about the exercise probability of embedded

[12] I remain agnostic about a mechanism through which some investors may learn valuable information about price pressures. Barbon, Di Maggio, Franzoni, and Landier (2019) suggest that there is information leakage from brokers to clients in the equity market.

options. Most bonds in my sample are callable; issuers have a right to redeem them at pre-specified dates before maturity. An early call changes the duration of a bond and, therefore, its risk profile. Superior knowledge about the likelihood of an early call gives an advantage in predicting bond returns prior to call announcements.

C. Time-varying volume-return coefficients

The evidence presented in the paper so far referred to an entire sample of 2005–2018. Volume-return coefficients, the subject of this study, may not be persistent over time, and the results may differ across time subsamples. To account for time variability in volume-returns coefficients, I re-estimate equation (2) for individual bonds within each calendar year-quarter. There are around 63 trading days per quarter. Defining an active period here, as in previous sections, as a sequence of at least 60 trading days close to each other would be too restrictive: the resulting sample would only include bonds that are effectively traded every day, which is not a representative subset of bonds. Instead, in this section, I define an active period as a sequence of at least 40 trading days (days with non-zero trading volume) within a calendar quarter where every two consecutive trading days are at most three business days apart. Then there is a unique active trading period (if any) per bond per calendar quarter. Therefore, by re-estimating equation (2) for individual bonds within each quarter, I obtain bond i – quarter t panels of volume-return coefficients $\hat{\beta}_{1,i,t}$, $\hat{\beta}_{2,i,t}$, and $\hat{\beta}_{3,i,t}$. I then explain the panels of volume-return coefficients with the fixed effects models that are analogous to (3) up to the inclusion of year-quarter fixed effects. Table C5 in Appendix C presents summary statistics for the second-stage year-quarter sample. In key bond and issuer characteristics, it does not differ much from the baseline sample. Table VIII presents the second-stage estimates.

In Table VIIIa, I fit rating-year-quarter fixed-effects models to the panels of volume-return coefficients. The loadings on PC_{bond} have the same signs as in the cross-sectional estimation ($\hat{\beta}_1$ and $\hat{\beta}_3$ decrease with information asymmetry while $\hat{\beta}_2$ increases). The point

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
PC_{bond}	-0.049^{***}	0.008^{***}	-0.006^{**}
	(0.002)	(0.001)	(0.003)
Rating FE	YES	YES	YES
Time FE	YES	YES	YES
Observations	78,314	78,314	78,314
R^2	0.177	0.006	0.014
Note:	*p<0.1; **p<0.05; ***p<0.01		

(a) Full sample (Jan 2005 – Dec 2018)

	$\hat{\beta}_1$		$\hat{\beta}_2$		$\hat{\beta}_3$	
	Pre-GFC	Post-GFC	Pre-GFC	Post-GFC	Pre-GFC	Post-GFC
PC_{bond}	-0.063^{***}	-0.049^{***}	0.010^{**}	0.008^{***}	-0.016^{**}	-0.004^{*}
	(0.004)	(0.003)	(0.004)	(0.001)	(0.007)	(0.002)
Rating FE	YES	YES	YES	YES	YES	YES
Observations	12,262	59,593	12,262	59,593	12,262	59,593
R^2	0.145	0.188	0.009	0.005	0.028	0.011
Note:	*p<0.1; **p<0.05; ***p<0.01					

(b) Pre- (Jan 2005 – Jun 2008) and post-GFC (Jan 2010 – Dec 2018) samples

Table VIII. Regressions of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ on information asymmetry indices in bond-quarter panel. In part (a), Models (1)-(3) are for $\hat{\beta}_1$, model (4)-(6) – for $\hat{\beta}_2$, models (7)-(9) – for $\hat{\beta}_3$. Each model in both (a) and (b) uses a fixed-effect estimator with rating-quarter double-clustered standard errors.

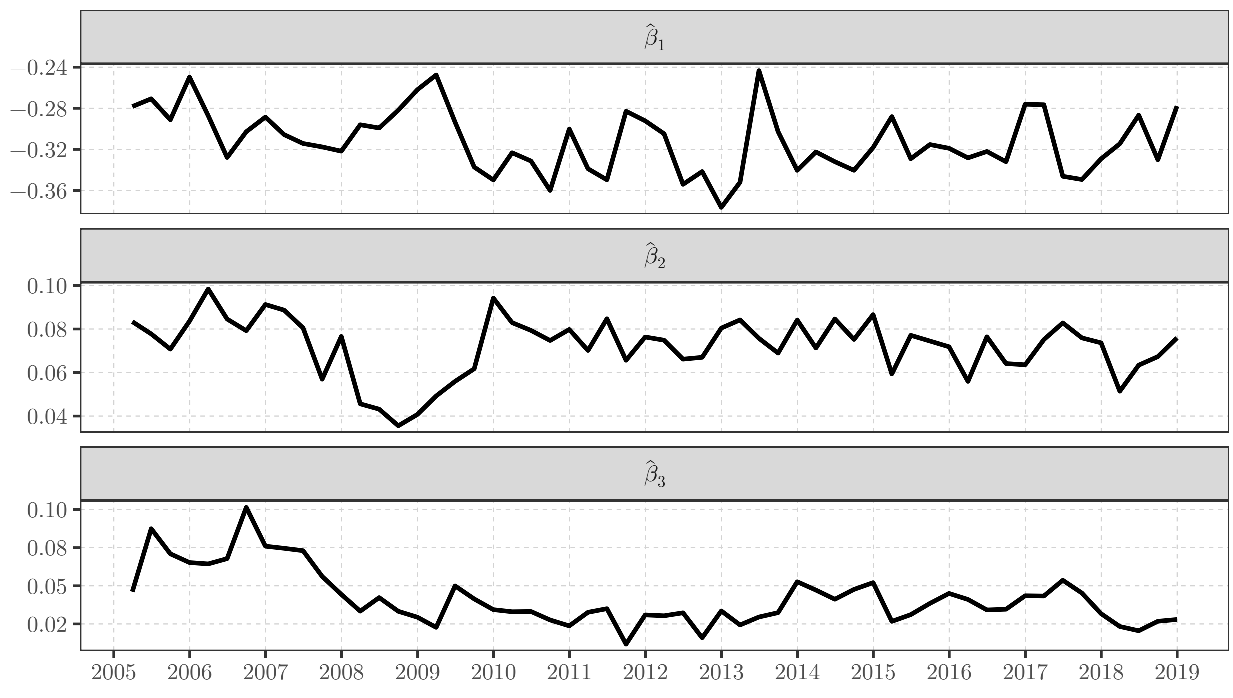

Figure 3. Time fixed effects for volume-return coefficients extracted from the models of Table VIIIa.

estimates are also close to the previously obtained values. Figure 3 presents time fixed effects extracted from the models of Table VIIIa. It turns out that $\hat{\beta}_1$ and $\hat{\beta}_2$ are stable over time (the time series of both variables do not contain a unit root according to conventional tests) while the level of $\hat{\beta}_3$ drops considerably around the GFC, from 0.08 to 0.03 on average. This result is in line with the evidence of a reduced risk-bearing capacity of corporate bond dealers (most of which are regulated banks) post-GFC (see, for instance, Adrian et al. 2017). Pre-GFC, bond dealers were more willing to accept the risk of being adversely selected and bond prices were more likely to move against dealers following large CtD trades than post-GFC. This result does not imply that the GFC and follow-up changes in regulation affected dealers' ability to identify the informed trade flow. Table VIIIb estimates the second-stage regression for pre- and post-GFC periods separately and finds that the dependence of $\hat{\beta}_2$ on information asymmetry does not differ much in these two subsets. If the dealers' ability to identify likely-informed trades pre-execution is due to the non-anonymity of OTC corporate bond trading (which I think is the case, but the formal investigation is beyond the scope of this paper), the result should not come as a surprise since the non-anonymity of trading has not been affected by the GFC. The non-positivity of the loading on PC_{bond} in the regressions for $\hat{\beta}_3$ in Table VIIIb, both pre- and post-GFC, also supports this result.

VI. Robustness

I run several robustness tests for the paper's main empirical result, which is a positive and non-positive dependence of $\hat{\beta}_2$ and $\hat{\beta}_3$ on information asymmetry, respectively. I report the results of these robustness tests in Table IX with PC_{bond} as the information asymmetry index. The independent modifications to the baseline methodology (as represented by the results in Table V) that I consider are the following.

- *Log-volumes in the first-stage model.* Here, I apply a $\log(x + \text{small constant})$ transformation to trading volumes before standardizing them per bond per active period.

Therefore, $\tilde{V}^{(c)}$ and $\tilde{V}^{(s)}$ in equation (2) become standardized log-volumes. It attenuates the impact of the largest trades on $\hat{\beta}_2$ and $\hat{\beta}_3$.

- *The simple average of the volume-weighted buy and sell prices instead of the VWAP.* On high CtD volume days, there is an imbalance between the volume of client purchases and sales of bonds. The VWAP is closer to an (unobserved) bid or ask price on such days. Therefore, returns computed from the VWAPs contain the effect of the bid-ask bounce, which may interfere with the identification of the relationship between information asymmetry and volume-return coefficients, as discussed in Llorente et al. (2002). Executable bid-ask spreads are unavailable for the majority of corporate bonds in the sample (thanks to the OTC market structure); therefore, the mid-price is not available either. To make the time series of individual bond prices less exposed to the bid-ask bounce, I take a simple average of the volume-weighted buy and sell prices rather than VWAPs (whenever the CtD volume is above zero).
- *Exclusion of retail-sized trades.* Small (retail-sized) trades in corporate bonds are priced unfavorably (Edwards, Harris, and Piwowar 2007). Thus, the reversal in bond prices may be due to the prevalence of retail-sized transactions on some trading days. To control for such an effect, I remove all trades smaller than $10k from consideration. The sample gets smaller and concentrates on more liquid bonds here because an active trading period now needs to consist of trading days with non-zero non-retail volume.
- *Market return in the first-stage model.* Here, I add the market return as a linear term to the right-hand side of equation (2). The market here is a size-weighted basket of all corporate bonds in the sample. The inclusion of market returns in the first stage corrects for a possible omitted-variable bias in $\hat{\beta}_i$.
- *Trading volumes in the first-stage model.* I add $\tilde{V}^{(c)}$ and $\tilde{V}^{(s)}$ as linear terms on the right-hand side of equation (2) while keeping the interaction variables of return and volume in the model too. As with the inclusion of market returns, it potentially corrects an omitted-variable bias.

- *Information asymmetry indices extracted from initial bond characteristics.* The averaging of bond characteristics across active periods (if there is more than one) for individual bonds introduces some measurement error to the right-hand side of equation (3). To limit its impact on the second-stage estimates, I use observed initial values (i.e., values at the beginning of the first active trading period) of information asymmetry proxies rather than time-series averages for individual bonds.
- *Weighted second-stage regression.* In the first-stage regression, volume-return coefficients are estimated with different precision for individual bonds. I assign higher weights to more precise estimates to limit the impact of high-variance estimates of $\hat{\beta}_i$ on the second-stage results. The weights are the inverse variance of the first-stage estimates.
- *Control for volume persistence in the second-stage model.* Assume that an investor executes a large buy trade over two business days.[13] On each day, her trades have a price impact, and returns tend to persist (or revert less). So, correlated volumes would generate a relationship between volumes and future returns similar to one of the asymmetric information and returns. I control for this alternative explanation by including the first autocorrelations of $\tilde{V}_t^{(c)}$ and $\tilde{V}_t^{(s)}$ (averages for individual bonds) in the second-stage model.

Table IX presents the results of these robustness tests. Each coefficient in the table is the estimated loading on PC_{bond} in the fixed-effects models for a respective $\hat{\beta}_i$. The first column demonstrates that $\hat{\beta}_1$ remains significantly negative across all alternative specifications, and the effect size does not change much (except for the inclusion of market returns in the first stage). The same applies to $\hat{\beta}_2$. Here, the effect size varies more across specifications (0.008–0.016; the baseline estimate is 0.010), but the loading on PC_{bond} remains significantly positive. The third column of Table IX shows that $\hat{\beta}_3$ either depends negatively on informa-

[13]This hypothesis is questionable since one gets better execution prices trading higher volumes on the corporate bond market, as shown in Edwards et al. (2007). Related, the average autocorrelation of $\tilde{V}_t^{(c)}$ is relatively low in the data (see Table C1 in Appendix).

tion asymmetry or there is no significant link between the two. Similar robustness tests with an alternative information asymmetry index $PC_{\text{bond-ex-ba}}$ in Table C6 in Appendix present a qualitatively similar picture. These results support the main findings of the paper.

Additionally, I re-estimate the baseline empirical model of the paper (Table V with PC_{bond}) independently in different sample splits. Table C7 in Appendix demonstrates that the results hold both in the investment-grade (IG) and the high-yield (HY) subsamples. They are somewhat stronger for HY bonds, as one would expect, as they are likely more information-sensitive than the IG ones. Table C8 in Appendix C shows that the results hold both for bonds issued by industrial companies and financial firms and are, thus, not due to a particular industry effect.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
A. Different inputs in the 1st stage			
Log-volumes	-0.043***	0.008***	-0.006***
	(0.001)	(0.001)	(0.002)
Avg. of VW buy and sell prices	-0.048***	0.012***	-0.003***
	(0.001)	(0.002)	(0.001)
B. Different models in the 1st stage			
Retail trades excluded	-0.044***	0.008***	0.000
	(0.001)	(0.001)	(0.001)
Market return added	-0.029***	0.011***	-0.005***
	(0.001)	(0.001)	(0.001)
Volumes added linearly	-0.041***	0.016**	-0.001
	(0.001)	(0.006)	(0.002)
C. Different 2nd stage			
PCs extracted from initial obs.	-0.042***	0.009***	-0.004**
	(0.002)	(0.002)	(0.002)
Weighted observations	-0.044***	0.006***	-0.001
	(0.001)	(0.001)	(0.001)
Vlm. correlation controls	-0.042***	0.008***	-0.006***
	(0.001)	(0.001)	(0.002)

Table IX. Regressions of volume-return coefficients on information asymmetry: robustness tests. Each line in the table presents loadings on the information asymmetry index PC_{bond} in fixed-effects models for the cross-section of volume-return coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$. Fixed effects are bonds' credit ratings. Standard errors (in parentheses) are also rating-clustered. 'Log-volumes' line is for the case when standardized and truncated trading volumes are replaced with standardized log-volumes in the first-stage regression. 'Avg. of VW buy and sell prices' is the case when returns are computed for the simple average of the volume-weighted buy and sell prices rather than for volume-weighted average prices. 'Retail trades excluded' removes trades below $10k USD notional from consideration. 'Market return added' extends the first-stage model with market returns. 'Volumes added linearly' extends the first-stage model with CtC and CtD volumes as linear right-hand side terms. 'PCs extracted from initial obs.' changes the inputs for the information asymmetry index PC_{bond}. Instead of bond averages, initial values (i.e., values at the beginning of the first active period) of information asymmetry proxies are used to construct the index. 'Weighted observations' is the weighted second-stage fixed-effects estimator. Observations (in the cross-section of bonds) are weighted with the inverse of the variance of $\hat{\beta}_i$ obtained at the first stage. 'Vlm. correlation controls' adds average individual bond time-series correlation of the CtC and the CtD trading volumes as control variables to the second-stage regression. Table C6 in Appendix presents an analogous table with $PC_{bond\text{-}ex\text{-}ba}$ as the information asymmetry index.

VII. Implications for investment strategies

Corporate bond price reversals depend on the extent of information asymmetry in a given bond, as my empirical analysis shows. What does it imply for the design of the short-term corporate bond reversal strategy? In this section, I show that the reversal strategy earns more if information asymmetry is taken into account in portfolio formation.

I start by constructing reversal portfolios as in Bai et al. (2019). At every rebalancing date (which is monthly), the bonds are double sorted on the previous month's credit rating and return. In Bai et al. (2019), each sorting is into quintiles, but since my sample is smaller I sort into rating terciles and return quintiles, a total of 15 bins. I only consider the long part of the reversal portfolio: this is a simple average of size-weighted returns in the top reversal quintile (lowest past returns) across three rating terciles.[14] The rebalancing is at the end of each month. I consider an unfiltered bond-month sample, i.e., I do not restrict the sample to active periods and do not remove the crossing of the IG/HY threshold (I would introduce a look-ahead bias if I did so). I do require the bonds to have, as of the sorting date, an outstanding amount of at least 200 mln USD and a 12-month backward-looking average of the realized bid-ask spread of at most 100 b.p. The latter helps to bring down the transaction cost of the reversal strategy, which is usually very high due to high portfolio turnover. I use the 12-month average of the realized bid-ask spread to account for transaction costs.

In addition to a long-reversal portfolio, I consider its two sub-portfolios separately. The first sub-portfolio contains the bonds with a below-median number of mutual fund bondholders six months prior to the sorting date. This sub-portfolio contains bonds with supposedly more information asymmetry. The second sub-portfolio contains the bonds with an above-median number of mutual fund bondholders (less information asymmetry). The results of the previous section suggest that in-sample and following average-volume periods, the reversals are stronger for bonds with more information asymmetry. So, one might expect the reversal

[14] I do not consider a short leg here. In the sample I work with, shorting top-performing corporate bonds was not profitable. See Ivashchenko and Kosowski (2022) for more details.

portfolio with more information asymmetry to outperform the reversal portfolio with less information asymmetry out-of-sample.

	Cum trading costs				Net trading costs			
	Mean	S.D.	SR	IR	Mean	S.D.	SR	IR
Long reversal (LR)	7.37	6.06	1.15	1.78	0.94	5.95	0.12	0.02
LR: many funds	6.69	7.03	0.90	1.16	0.02	6.87	-0.04	-0.20
LR: few funds	8.58	5.99	1.36	1.99	2.22	5.90	0.34	0.38
Market	1.65	3.49	0.29		0.88	3.47	0.06	

Table X. Performance statistics of the long leg of the reversal strategy for corporate bonds with monthly rebalancing. Mean is a sample average of monthly returns, in % per annum. S.D. is the standard deviation of monthly returns, in % per annum. SR is the Sharpe ratio relative to the 3-month Treasury Bill. IR is the information ratio relative to the market. The sample is from Jan 2006 to Dec 2018. For portfolio construction, I apply the following filters to the sample: a) the previous month's outstanding amount is greater than 200 mln USD, b) the previous month's backward-looking 12-month moving average of the realized bid-ask spread is below 100 b.p. Reversal portfolios are obtained from the double-sorting of bonds on the previous month credit rating (three terciles) and total return (five quintiles). For each of the 15 bins, the average bond return weighted by the previous month's outstanding amount is computed. The long-reversal (LR) return is a simple average return across three rating terciles for the top reversal (lowest past returns) quintile. 'LR: few funds' is the reversal portfolio with a below-median number of fund owners. 'LR: many funds' is the reversal portfolio with an above-median number of fund owners. The market return is the value-weighted return of the bonds in the sample. Trading costs are assumed to be half of the 12-month average of the realized bid-ask spread (average bid-ask spread in Table I).

Table X presents performance measures of three reversal portfolios in comparison to the market portfolio. Between 2005 and 2018, average long-reversal portfolio returns unadjusted for trading costs were around 7.4% per year. The sub-portfolio with many fund owners earned around 6.7% while the portfolio with few fund owners earned around 8.6%, which is considerably more than the market portfolio. The volatility of the sub-portfolio with few fund owners was also lower, which translates into the superior risk-adjusted performance of the reversal strategy for bonds with more information asymmetry. Once I account for trading costs, the performance of reversal portfolios becomes considerably lower because of high portfolio turnover. However, the sub-portfolio with few fund owners still earns almost 2.2% per year after trading cost adjustment, which is 2.5 times more than the corporate bond market. The information ratio of the reversal portfolio with few fund owners amounts to almost 0.4 (annualized) relative to the corporate bond market. The return on the reversal portfolio with many fund owners is effectively zero after trading cost adjustment.

The evidence presented in this section demonstrates that conditioning on ex-ante information asymmetry considerably affects the performance of reversal strategies in practice. Reversals tend to be stronger for bonds with more information asymmetry, and long-reversal portfolios with less mutual fund ownership, for instance, can outperform the corporate bond market after adjustment for trading costs. Given these findings, one can further investigate different information asymmetry signals and potentially improve the performance of a corporate bond reversal strategy.

VIII. Conclusion

In this paper, I estimate a dynamic volume-return relationship for individual bonds and explore the determinants of estimated volume-return coefficients in a cross-section of bonds. A particular focus of my analysis is on the impact of information asymmetry on volume-return coefficients.

The hypotheses that I test arise from a stylized theoretical trading model in which dealers never trade with an informed counterparty, while investors occasionally do. The model suggests that bonds with high information asymmetry have stronger price reversals than bonds with low information asymmetry, but less so following high-volume days when dealers' inventory *does not change*, and investors are essentially trading with each other. Conversely, following days with substantial changes in dealers' inventory, the difference in reversals between high- low-asymmetry bonds remains.

I find strong empirical support for these hypotheses in the data. Bonds with high information asymmetry exhibit stronger price reversals than low-asymmetry bonds, but less so following days when trading volumes are high, but dealers' inventory does not change at the end of the day (clients purchases equal client sales). If one considers, instead, the reversals following days when trading volume is high, but it is due to substantial changes in dealers' inventory, then the difference in reversals between bonds with high and low information asymmetry remains at the average-volume day level. High-asymmetry bonds, in my analysis, are the bonds that are owned by few mutual funds and intermediated by few dealers, have smaller outstanding amounts and bid-ask spreads, and are issued by smaller firms. The 'extra' persistence in high-asymmetry bond returns is the highest right before corporate earnings announcements. The effect is present both pre- and post-GFC, and in different subsamples of the corporate bond universe (including within bonds issued by the same firm).

These results are consistent with the assumption that trading volume in high-asymmetry bonds is more likely to come from investors who possess private information. Since dealers typically know their clients well and might be able to detect informed investors, they let other investors provide liquidity for such trades. Overall, my results suggest that there might be informed trading in corporate bonds, but when it happens, dealers are not providing liquidity and are not adversely selected.

My findings have implications for the design of investment strategies exploiting corporate bond reversals. In particular, I show that long-reversal portfolios of high-asymmetry bonds outperform long-reversal portfolios of low-asymmetry bonds both before and after adjustment for trading costs. Hence, illiquidity does not fully explain reversal returns. Moreover, reversal portfolios of high-asymmetry bonds outperform the corporate bond market after trading cost adjustment. An investor considering an implementation of a bond reversal strategy might profit from additionally sorting bonds on information asymmetry proxies.

My results also have implications for the regulation of OTC markets trading. Dealers in a non-anonymous OTC market likely avoid informed trade flows. It suggests that the dealers in this market are in a fundamentally different economic position than in a central limit-order book anonymous market. In the latter, dealers balance sheets have to absorb informed flows. Trading regulations in these two markets may not fully reflect such a discrepancy.

REFERENCES

Adrian, Tobias, Nina Boyarchenko, and Or Shachar, 2017, Dealer balance sheets and bond liquidity provision, *Journal of Monetary Economics* 89, 92 – 109.

Asquith, Paul, Andrea S. Au, Thomas Covert, and Parag A. Pathak, 2013, The market for borrowing corporate bonds, *Journal of Financial Economics* 107, 155 – 182.

Bai, Jennie, Turan G. Bali, and Quan Wen, 2019, Common risk factors in the cross-section of corporate bond returns, *Journal of Financial Economics* 131, 619 – 642.

Bao, Jack, Maureen O'Hara, and Xing (Alex) Zhou, 2018, The Volcker Rule and corporate bond market making in times of stress, *Journal of Financial Economics* 130, 95 – 113.

Bao, Jack, Jun Pan, and Jiang Wang, 2011, The illiquidity of corporate bonds, *The Journal of Finance* 66, 911–946.

Barbon, Andrea, Marco Di Maggio, Francesco A. Franzoni, and Augustin Landier, 2019, Brokers and order flow leakage: Evidence from fire sales, *The Journal of Finance* 74, 2707–2749.

Berndt, Antje, and Yichao Zhu, 2019, Dealer inventory, short interest and price efficiency in the corporate bond market, Working paper.

Bessembinder, Hendrik, Stacey Jacobsen, William Maxwell, and Kumar Venkataraman, 2018, Capital commitment and illiquidity in corporate bonds, *The Journal of Finance* 73, 1615–1661.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905–939.

Campbell, John Y., and Robert J. Shiller, 1988, Stock prices, earnings, and expected dividends, *The Journal of Finance* 43, 661–676.

Choi, Jaewon, and Yeso Huh, 2022, Customer liquidity provision: Implications for corporate bond transaction costs, *Management Science (Forthcoming)* .

Chordia, Tarun, Amit Goyal, Yoshio Nozawa, Avanidhar Subrahmanyam, and Qing Tong, 2017, Are capital market anomalies common to equity and corporate bond markets? An empirical investigation, *Journal of Financial and Quantitative Analysis* 52, 1301–1342.

Dechow, Patricia M., Richard G. Sloan, and Jenny Zha, 2014, Stock prices and earnings: A history of research, *Annual Review of Financial Economics* 6, 343–363.

Dick-Nielsen, Jens, 2014, How to clean Enhanced TRACE data, Working paper.

Dick-Nielsen, Jens, and Marco Rossi, 2018, The cost of immediacy for corporate bonds, *The Review of Financial Studies* 32, 1–41.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen, 2005, Over-the-counter markets, *Econometrica* 73, 1815–1847.

Edwards, Amy K., Lawrence E. Harris, and Michael S. Piwowar, 2007, Corporate bond market transaction costs and transparency, *The Journal of Finance* 62, 1421–1451.

Friewald, Nils, and Florian Nagler, 2019, Over-the-counter market frictions and yield spread changes, *The Journal of Finance* 74, 3217–3257.

Glosten, Lawrence R., and Paul R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71 – 100.

Goldstein, Michael A., and Edith S. Hotchkiss, 2020, Providing liquidity in an illiquid market: Dealer behavior in us corporate bonds, *Journal of Financial Economics* .

Han, Song, and Xing Zhou, 2014, Informed bond trading, corporate yield spreads, and corporate default prediction, *Management Science* 60, 675–694.

Hanson, Samuel G, Robin Greenwood, and Gordon Y Liao, 2018, Asset price dynamics in partially segmented markets, *The Review of Financial Studies* 31, 3307–3343.

Hendershott, Terrence, Roman Kozhan, and Vikas Raman, 2019, Short selling and price discovery in corporate bonds, *Journal of Financial and Quantitative Analysis* 139.

Ivashchenko, Alexey, and Robert Kosowski, 2022, Transaction costs and capacity of systematic corporate bond strategies, Working Paper.

Ivashchenko, Alexey, and Artem Neklyudov, 2018, (In)frequently traded corporate bonds, Working paper.

Llorente, Guillermo, Roni Michaely, Gideon Saar, and Jiang Wang, 2002, Dynamic volume-return relation of individual stocks, *The Review of Financial Studies* 15, 1005–1047.

Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *The Journal of Finance* 39, 1127–1139.

Wang, Jiang, 1994, A model of competitive stock trading volume, *Journal of Political Economy* 102, 127–68.

Appendix A. The model

In this Appendix, I present a model of competitive bond trading volume that builds on the same premises as my empirical analysis above: investors trading bonds with each other are occasionally adversely selected, while dealers avoid information-driven trade flow. The model justifies equation (2) and yields predictions about the dependence of volume-return coefficients on information asymmetry that closely match the empirical results I have discussed above. One can view the model of this section as the formal presentation of the intuition behind volume-return relationships I analyze in the paper.

The model is a modification of Llorente et al. (2002) which is a simplified version of Wang (1994) in its turn. In these models, two types of investors, informed and uninformed ones, are trading with each other for liquidity reasons and on private information. My model differs from Llorente et al. (2002) in two ways: I tailor the arithmetic of returns to defaultable bonds and I introduce noisy bond supply.

Changing a dividend-paying stock for a perpetual coupon-paying defaultable bond within the model requires approximations to keep the analysis tractable. In Llorente et al. (2002), private information is information about dividends, which is an additive component of dollar returns. In my model, private information relates to default risk, which is not an additive term in returns calculation. To make returns linear in a default loss and simplify the learning problem for uninformed traders, I consider a log-linear approximation of defaultable bond returns as in Hanson, Greenwood, and Liao (2018). Given that daily bond returns in my sample are small numbers (see Table I) with 5th and 95th percentiles close to 1.3% in the absolute value, the log-linear approximation of returns should not undermine the relevance of theoretical results for my empirical analysis.

I introduce noisy bond supply to the model to generate the additional trading volume that is not due to liquidity or informational signals the agents receive. In the model, I *assume* that supply changes that proxy for changes in dealers' bond inventory are independent of the arrival of private news. Table C1 suggests that such an assumption is not at odds with the data; the correlation between client-to-client and client-to-dealer daily volume measures in my sample is low. In the model, supply changes are publicly observed, unlike private liquidity signals. Under these assumptions, I can derive the dynamic volume-return relationship similar to (2) and provide additional implications for my empirical analysis compared to the baseline model of Llorente et al. (2002).

The economy

The discrete-time economy has two traded securities: a riskless bond in unlimited supply at a constant interest rate that is set to 0 for simplicity and a risky perpetual bond that pays a coupon C every period. Hanson et al. (2018) demonstrate that Campbell and Shiller (1988) decomposition applied to such a bond yields the log return r_t of the following form:

$$r_{t+1} \approx \kappa + c(1-\theta) + \theta p_{t+1} - p_t - d_{t+1}, \tag{A1}$$

where $p_t \equiv \log P_t$ is the log ex-coupon price of the bond, θ and κ are deterministic functions of the log-coupon $c \equiv \log C$, and d_{t+1} is the log default loss at time $t+1$.[15]

I assume that the log default loss consists of two additive components:

$$d_{t+1} = f_t + g_t.$$

f_t is publicly known at time t while g_t is a private time t information of a subset of investors. At time $t+1$, the value of d_{t+1} becomes publicly observed.

The risky bond is traded in a competitive bond market with noisy supply s_t, which is public knowledge. The market is populated with two classes of investors, $i = 1, 2$, with relative population weights ω and $1-\omega$. The investors are identical within each class, and each investor's initial endowment of the risky bond is set to 0 for simplicity. Type 1 investors are informed; they observe g_t. Type 2 investors do not observe g_t but learn it from the bond price using the Bayes rule. In addition, Type 1 investors have a random exposure z_t to

[15]For the derivation see Appendix D.

some non-traded asset that generates a log return of n_{t+1} in the subsequent period.[16] Type 2 investors do not know the exposure of type 1 investors to the non-traded risk. Overall, the information set of the informed investors at time t is $\{d, p, n, f, s, g, z\}_{0,\ldots,t}$ while the information set of the uninformed investors is $\{d, p, n, f, s\}_{0,\ldots,t}$.

I assume that n_t, g_t, and z_t are time-independent zero-mean normally distributed random variables with variances σ_n^2, σ_g^2, σ_z^2 respectively. I further assume that f_t is also time-independent and normally distributed with the mean $m_f = \kappa + c(1-\theta)$ and the variance σ_f^2.[17] All of n_t, g_t, z_t, and f_t are contemporaneously uncorrelated except for n_t and f_t that have a time-invariant negative covariance, which means that default losses are low when non-traded asset returns are high. This implies a constant positive covariance between r_t and n_t that equals σ_{rn}. Finally, the supply of the risky bond follows an AR(1) process

$$s_{t+1} = \delta s_t + \epsilon_{t+1}, \tag{A2}$$

where $|\delta| < 1$ and ϵ_t is normally distributed with zero mean and variance σ_s^2; it is independent over time and is independent from n_t, g_t, z_t, and f_t.

The investors of both types $i = 1, 2$ maximize the next period conditional expected utility $\mathbb{E}_t\left[-e^{-W_{t+1}^{(i)}}\right]$ derived from the next period wealth $W_{t+1}^{(i)}$ by choosing the demand $X_t^{(i)}$ for the risky bond.[18] To keep the model tractable, I need to take the log-linear approximation of the wealth dynamics, which under the assumptions of the model is

$$\begin{aligned} W_{t+1}^{(1)} &\approx W_t^{(1)} + X_t^{(1)} r_{t+1} + z_t(1 + n_{t+1}), \\ W_{t+1}^{(2)} &\approx W_t^{(2)} + X_t^{(2)} r_{t+1}. \end{aligned}$$

The model setup is different from Llorente et al. (2002) in two ways. First, I work with log returns approximated in (A1) around $\bar{p} \equiv 0$ and linearized wealth dynamics instead of

[16] Here I follow Llorente et al. (2002) assuming for simplicity that only one type of investors has income from a non-traded asset. It is enough to generate price reversals due to liquidity trading.

[17] The mean of f_t is chosen such that the long-term mean of the log bond price is 0 and the contributions of coupons and public news about future defaults to returns cancel one another on average.

[18] As in Llorente et al. (2002), the risk aversion is set to 1 since it only enters the expressions for investors‘ demands as the multiple of the variances of all exogenous shocks. Hence, one can implement higher or lower risk aversion in the model by proportionally scaling variances of all shocks up or down.

dollar returns and non-linearized wealth dynamics. Second, more importantly, I assume a noisy supply (A2) instead of a constant zero supply. Noisy supply allows me to decompose the trading volume in the model into two components: the first one is related to trading between informed and uninformed investors, and exogenous changes in asset supply drive the second one. Empirical counterparts of these two components are, respectively, the volume of corporate bonds purchased by clients matched by client sales in a given period and net changes in broker-dealer inventory.

Model equilibrium

I solve for the rational expectations equilibrium of the model assuming a linear pricing function for the log bond price. Define the log price adjusted for the publicly known credit loss component as $\tilde{p}_t \equiv p_t + (f_t - m_f)$ and assume it is linear with respect to g_t, z_t, and s_t:

$$\tilde{p}_t = -a(g_t + bz_t + es_t). \tag{A3}$$

Observe that the steady-state level of log bond price is 0 as in the linear approximation of log return (A1).

Given the pricing function (A3), the equation for returns (A1) re-writes as:[19]

$$r_{t+1} = -\theta\left(f_{t+1} - m_f\right) + \theta\tilde{p}_{t+1} - \tilde{p}_t - g_t. \tag{A4}$$

The expression for conditional expected returns follows from (A4):

$$\mathbb{E}_t^{(i)}\left[r_{t+1}\right] = -\tilde{p}_t - \mathbb{E}_t^{(i)}\left[g_t\right] - ae\theta\delta s_t.$$

The informed investors know g_t, hence $\mathbb{E}_t^{(1)}\left[g_t\right] = g_t$. The uninformed investors observe $\tilde{p}_t$ and s_t and estimate $\mathbb{E}_t^{(2)}\left[g_t|\tilde{p}_t, s_t\right]$. I show in Appendix D that

$$\mathbb{E}_t^{(2)}\left[g_t|\tilde{p}_t, s_t\right] = \gamma(g_t + bz_t), \tag{A5}$$

where $\gamma = \frac{\sigma_g^2}{\sigma_g^2 + b^2\sigma_z^2} > 0$. One can further show that conditional return variances for two types of investors are constant over time.

[19] In what follows, I replace an approximate equality in (A1) with the exact one.

With conditional expected return linear in g_t, z_t, and s_t and conditional return variance constant for both types of investors, the demand for risky bonds, $X_t^{(1)}$ and $X_t^{(2)}$, is also linear in g_t, z_t, and s_t[20]. The market for risky bonds clears:

$$\omega X_t^{(1)}(g_t, z_t, s_t) + (1-\omega) X_t^{(2)}(g_t, z_t, s_t) = s_t,$$

which must hold for any values of g_t, z_t, and s_t, implying a system of three non-linear equations for yet undetermined coefficients a, b, and e. One can show that if the parameters of the model are such that the system has real-valued solutions then it must be that a, b, and e are all positive, moreover, $\omega + \gamma - \omega\gamma < a < 1$ and $b = \sigma_{rn}$. I demonstrate in Appendix D that under mild restrictions on the parameters (that boil down to σ_s^2 being not 'too big') the model always has real-valued solutions, of which a unique triple of $\{a^*, b^*, e^*\}$ has economically reasonable values.

Trading volume in the model

Consider the aggregate difference in risky bond holdings in the economy at time t

$$\omega \Delta X_t^{(1)} + (1-\omega)\Delta X_t^{(2)} = \Delta s_t.$$

Using the equilibrium conditions one can decompose it as

$$\omega \Delta X_t^{(1)} + (1-\omega)\Delta X_t^{(2)} = \underbrace{V_{c,t}^{(1)}(\Delta g_t, \Delta z_t) + V_{c,t}^{(2)}(\Delta g_t, \Delta z_t)}_{=0} + \underbrace{V_{s,t}^{(1)}(\Delta s_t) + V_{s,t}^{(2)}(\Delta s_t)}_{=\Delta s_t},$$

where

$$\left|V_{c,t}^{(1)}(\Delta g_t, \Delta z_t)\right| = \left|V_{c,t}^{(2)}(\Delta g_t, \Delta z_t)\right| = \left|\alpha\left(\Delta g_t + \sigma_{rn}\Delta z_t\right)\right|, \tag{A6}$$

and $\alpha = {}^{\omega(a-1)}/_{\sigma_r^2}$. Here, $V_c^{(1)}$ and $V_c^{(2)}$ represent the volume of trading *between* informed and uninformed investors. This trading volume is due to changes in a private signal about credit loss Δg (information-driven trading) and the position in a non-traded asset Δz (liquidity-driven trading). $V_c^{(1)}$ and $V_c^{(2)}$ always have opposite signs but are equal in absolute value. For the convenience of notation, I will denote this trading volume $v_{c,t} = \left|\alpha\left(\Delta g_t + \sigma_{rn}\Delta z_t\right)\right| \geqslant 0$.

[20]See Appendix D.

An econometrician observing bond trading records in the TRACE database can compute what the client buy volume matched by the client sell volume was at the time t.[21] It is an empirical proxy for $v_{c,t}$.

Two other components, $V_s^{(1)}$ and $V_s^{(2)}$, represent trading due to changing bond supply. One can show that in equilibrium, these two components are always of the same sign, and they represent the proportion in which two types of agents absorb additional bond supply Δs. By construction, a change in bond supply is the buy volume that was not matched by the sell volume of the opposite sign. Its absolute value is equal to the absolute value of a change in aggregate dealers' inventory. The latter is an empirical counterpart of $v_{s,t} \equiv |\Delta s_t|$. What the model assumes is that $v_{c,t}$ and Δs_t are independent since the latter is uncorrelated with Δg and Δz that drive the former. Table C3 has demonstrated that this assumption largely holds in the data. The key takeaway of this paragraph is that I assume that an econometrician knows $v_{c,t}$ and $v_{s,t}$, and these two quantities are defined within the model as stated above.

Volume-return relationship and information asymmetry

Assume an econometrician observes the time-series of bond returns r_t and two types of volume, $v_{c,t}$ and $v_{s,t}$, as discussed above. Then the conditional expectation of future returns, given current returns and volume, can be approximated as

$$\mathbb{E}_t \left[r_{t+1} | r_t, v_{c,t}, v_{s,t} \right] \approx \left(\beta_1 + \beta_2 v_{c,t}^2 + \beta_3 v_{s,t}^2 \right) r_t, \tag{A7}$$

the derivation is presented in Appendix D. This volume-return relationship is a theoretical counterpart of equation (2) estimated in the empirical part of the paper. Unlike equation (2), equation (A7) contains squared volumes. In the data, squared volumes are extremely right-skewed, hence from an econometric standpoint, it is reasonable to estimate the volume-return relationship as in (2) with volume entering the equation without a square (Llorente

[21] All records in TRACE represent trading *between* a broker-dealer and a client and can be of two types only: a purchase by a client from a dealer or a sale to a dealer.

et al. 2002 follow the same approach). It does not change an economic interpretation of volume-return coefficients.[22]

Now, I would like to discuss how coefficients β_1, β_2, and β_3 change in the model as the extent of informed trading changes. In the benchmark model Llorente et al. (2002), both β_1 and β_2 are negative, but β_1 is decreasing and β_2 is increasing with the extent of information asymmetry proxied by σ_g^2. β_1 measures the first return autocorrelation, and negative β_1 decreasing with σ_g^2 means that for two equally risky bonds returns will *revert more* for the one with more information asymmetry. β_2 measures the impact of volume on the first autocorrelation, and negative β_2 increasing with σ_g^2 means that for two equally risky bonds, returns will *revert less following high-volume days* for the one with more information asymmetry. These theoretical results find empirical support in the U.S. stock market, as Llorente et al. (2002) shows.

Unlike in the benchmark model, I can not make a general statement about the signs of volume-return coefficients and their dependence on σ_g^2; I need to solve the model numerically first. In Figure A1, I present the relationships between information asymmetry σ_g^2 and β coefficients for the model calibrated to an average bond in TRACE. The bond has a coupon rate of 5%, high persistence of a supply shock $\delta = 0.95$, and a daily standard deviation of returns of 1%.[23] The latter stays fixed in all numerical solutions; this is an additional constraint I impose on the solutions of the model.[24] Figure A1 represents the cross-section of bonds with the same unconditional risk but different contributions of public, private, and liquidity shocks to return variance.

[22]Since an econometrician knows the sign of inventory changes, she could write an analog of equation (A7) conditioning additionally on this piece of knowledge. It would change the form of the equation slightly, and the loadings on two types of volume would become incomparable. An important part of my empirical analysis consists of a direct comparison of coefficients β_2 and β_3, and for that, I need to condition in (A7) on the absolute value of inventory changes.

[23]In Figure A1, I set $\delta = 0.95$ which roughly corresponds to $\text{Corr}(\Delta s_t, \Delta s_{t-1}) = -0.03$ because in the model $\text{Corr}(\Delta s_t, \Delta s_{t-1}) = -\frac{1}{2}(1-\delta)$. In the model, δ measures the persistence of supply, which is roughly the persistence of inventory. $\delta = 0.95$ implies the half-life of broker-dealer inventory of about 13 days. Further (unreported) estimations show, in line with the results of Dick-Nielsen and Rossi (2018), that dealers revert deviations from their target inventory faster post-crisis.

[24]Llorente et al. (2002) impose the same restriction on the total unconditional variance of returns.

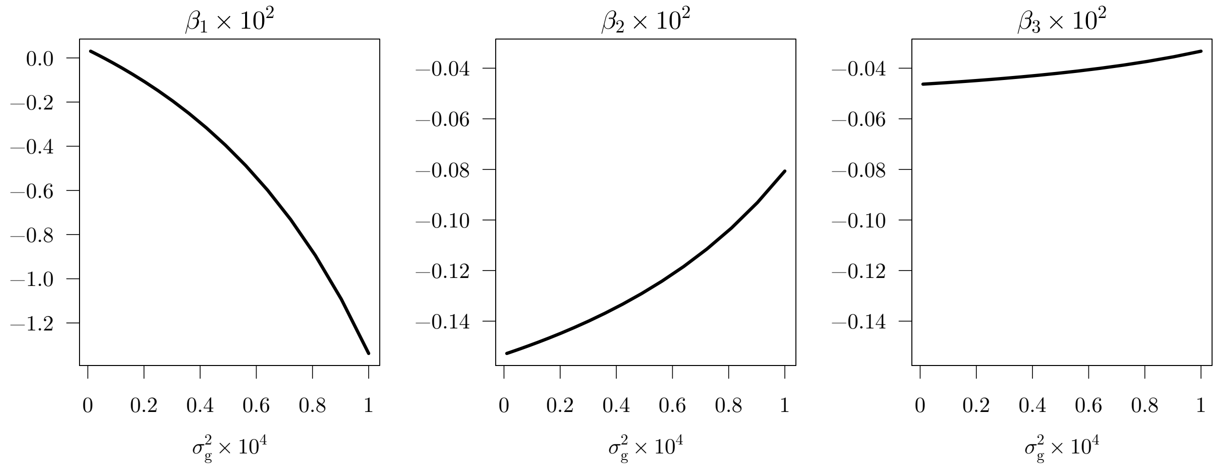

Figure A1. Dependence of β_1, β_2, and β_3 on information asymmetry σ_g^2 holding total return variance fixed. Each point on the curves is a numerical solution of the model. I obtain the relationships between σ_g^2 and β coefficients by varying σ_g from 0 to 1% holding an unconditional standard deviation of returns at 1%, which is a daily standard deviation of bond returns in the TRACE data. I choose the following parameters of the model to match a median bond in the sample: coupon rate $C = 5\%$, the persistence of a supply shock $\delta = 0.95$. The fraction of informed investors is $\omega = 0.05$, the correlation between traded and non-traded asset returns is $\sigma_{rn} = 0.3$, the variance of the supply shock is $\sigma_s^2 = 0.1$. I first solve the model for a very small value of σ_g, 5 b.p. here. Then, I hold the equilibrium value of a fixed in all subsequent solutions for $\sigma_g > 5$ b.p; I allow e to change. Thus, the comparative statics plotted here is a collection of solutions of the system of equations of three variables (σ_z^2, σ_f^2, and e): two model equilibrium equations plus an additional restriction on the total return variance.

The left and central panels in Figure A1 deliver the same message as the benchmark model. With more informed trading, returns tend to revert more, but less so following days when investors trade a lot with each other. On the left panel, which presents reversals following no-volume days, there is no reversal when σ_g is zero, and returns are due to public news that is fully priced within the same period. As σ_g increases, no-volume reversals intensify due to a greater impact of uninformed investors' errors in estimating g_t on returns.[25] On the central panel, the reversal following high-volume days is the strongest when σ_g is zero

[25]Here is the intuition for this result. With no volume, time t returns are not driven by liquidity shocks since Δz_t and Δg_t must be zero. Assume $z_{t-1} > 0$ and informed investors are net sellers of bonds. From (A4) and (A5) one finds that r_t is negative when $\frac{a}{\gamma}\mathbb{E}^{(2)}_{t-1}[g_{t-1}] < g_{t-1}$ other things being equal, i.e., when actual losses in default are higher than previously expected by uninformed investors. But that means that in $t-1$ informed investors' demand for bonds was lower than required by their hedging needs; so it is in t since the volume is zero. Hence, time t price is low and time $t+1$ expected return is high. Higher information asymmetry amplifies this effect.

because the entire trading volume between informed and uninformed investors represents, in this case, liquidity trading. Liquidity trading has a price impact but does not reveal any new information about the asset payoff; hence, the price reverts in the next period. As σ_g increases, it's more and more likely that some part of the between-investors trading volume comes from Δg and conveys the information about future returns; hence the reversal tends to decrease (β_2 tends to increase). The right panel in Figure A1 shows that β_3 that measures an additional component of reversals following days when inventory changes a lot is relatively insensitive to σ_g. It does not look surprising given that Δs in the model is uncorrelated with other motives for trading. One would expect β_3 to be flat with respect to σ_g in such case; a slightly upward sloping line on the right panel of Figure A1 is due to equilibrium e (price impact of inventory-changing trades) changing with σ_g.

The shape of the lines in Figure A1 closely matches the shape of their empirical counterparts presented in Figure 2. In the model, as it is in the data, β_1 decreases, and β_2 increases with information asymmetry, while β_3 is insensitive to information asymmetry. It gives additional support for the premises of the model: client-to-client trading volume may be due to private information, but client-to-dealer trading volume is likely driven by liquidity needs only.

As in Llorente et al. (2002), the limitation of my extended model is that β_2 stays negative for all reasonable model calibrations and does not turn positive (same applies to β_3, which is not part of the benchmark model). In reality, as Section III has shown, β_2 is positive for most corporate bonds. It does not undermine the main idea suggested by the model and tested in the empirical part of the paper. As the extent of informed trading increases, returns following high-volume days are less likely to revert, especially when dealers are not trading from their inventory capacity.

Appendix B. Data and sample

Sample selection

I apply some filters to the TRACE database *after* cleaning it as in Dick-Nielsen (2014). Here are the criteria I use to select the bonds in the sample:

- The bond is nominated in USD;
- It is a fixed coupon (including zero-coupon), non-asset backed, non-convertible, non-enhanced bond;
- Not privately issued and not issued under Rule 144A;
- Of one of the following types according to the Mergent FISD classification: CMTN (US Corporate MTN), CDEB (US Corporate Debentures), CMTZ (US Corporate MTN Zero), CZ (US Corporate Zero), RNT (Retail Note), USBN (US Corporate Bank Note), PS (Preferred Security), UCID (US Corporate Insured Debenture);
- Bond price is $\geqslant 5$ and $\leqslant 1000$ (for a face value of 100) at least once in the bond's lifetime.

Four additional criteria must be jointly satisfied to keep a trade record in the sample:

- The trade is executed between Jan 1, 2005, and Dec 31, 2018;
- Executed at eligible times (time stamps of the trades are between 00:00:00 and 23:59:59; there is a small number of trades in TRACE with misreported times that do not fall into this range, I remove them from the sample);
- Executed on or after the dated date of the bond (the date when the interest starts to accrue) and before the maturity date.

Agency transactions with commissions are retained in the sample.

Winsorization

To ensure that my results are not driven by extreme observations, I winsorize some variables. In particular, in the original bond-day panel (before active periods are determined) I winsorize:

- CtC trading volume at 99%;
- CtD trading volume at 1% and 99%;
- Credit spread at 99.9%;
- Bid-ask spread at 99.9%;
- Total daily returns at 0.1% and 99.9%.

Further, I truncate the estimates of $\hat{\beta}_n$ at 1% and 99% in the largest sample (before averaging across active periods).

Appendix C. Additional Tables and Charts

	Mean	Med.	No.>0	No.<0	No.>0*	No.<0*	No. Obs.
$\text{Corr}(V_t^{(c)}, \lvert V_t^{(s)} \rvert)$	0.188	0.188	13949	1925	9813	41	15874
$\text{Corr}(V_t^{(c)}, V_t^{(s)})$	-0.052	-0.043	5784	10090	1516	4805	15874
$\text{Corr}(V_t^{(c)}, V_{t-1}^{(c)})$	0.056	0.020	8997	6877	4361	21	15874
$\text{Corr}(\lvert V_t^{(s)} \rvert, \lvert V_{t-1}^{(s)} \rvert)$	0.095	0.089	12187	3687	6296	43	15874

Table C1. Correlation coefficients between different measures of the trading volume. $V^{(c)}$ is the CtC trading volume, $V^{(s)}$ is the change in dealers' inventory, and $|V^{(s)}|$ is the CtD trading volume. Each correlation coefficient is estimated per bond per active period. 'Mean' and 'Med.' are sample average and median values. 'No. $>$ ($<$) 0' is the number of positive (negative) correlation coefficients. 'No. $>$ ($<$) 0*' is the number of positive (negative) coefficients significant at a 10% confidence level. The number of observations is the number of bond-active periods.

	PC_{all}	PC_{bond}	$PC_{\text{bond-ex-ba}}$
Bond bid-ask	0.37	0.38	
–No. mutual fund owners	0.49	0.55	0.56
–No. dealers	0.39	0.44	0.53
–Issue size	0.55	0.60	0.64
–Issuer size	0.32		
Stock bid-ask	0.25		
Share of explained variance, %	0.42	0.57	0.70

Table C2. Loadings of principal components on standardized bond and issuer characteristics. Rows are individual characteristics, each is de-meaned and scaled by the cross-sectional standard deviation. The last line is the share of total variance explained by the first principal component for each group of individual variables.

	Bond bid-ask	No. funds	Issue size	No. dealers	Issuer size	Stock bid-ask	PC_{all}	PC_{bond}
No. funds	-0.44***							
Issue size	-0.39***	0.67***						
No. dealers	-0.04***	0.33***	0.61***					
Issuer size	-0.17***	0.14***	0.37***	0.30***				
Stock bid-ask	0.44***	-0.23***	-0.15***	0.08***	-0.19***			
PC_{all}	0.59***	-0.78***	-0.88***	-0.62***	-0.52***	0.40***		
PC_{bond}	0.57***	-0.84***	-0.91***	-0.66***	-0.33***	0.22***	0.96***	
$PC_{bond\text{-}ex\text{-}ba}$	0.35***	-0.81***	-0.92***	-0.76***	-0.33***	0.13***	0.92***	0.97***

Table C3. Cross-sectional correlation of information asymmetry indicators. The total number of bonds in the sample is 7206. *, **, and *** stand for 10%, 5%, and 1% significance respectively.

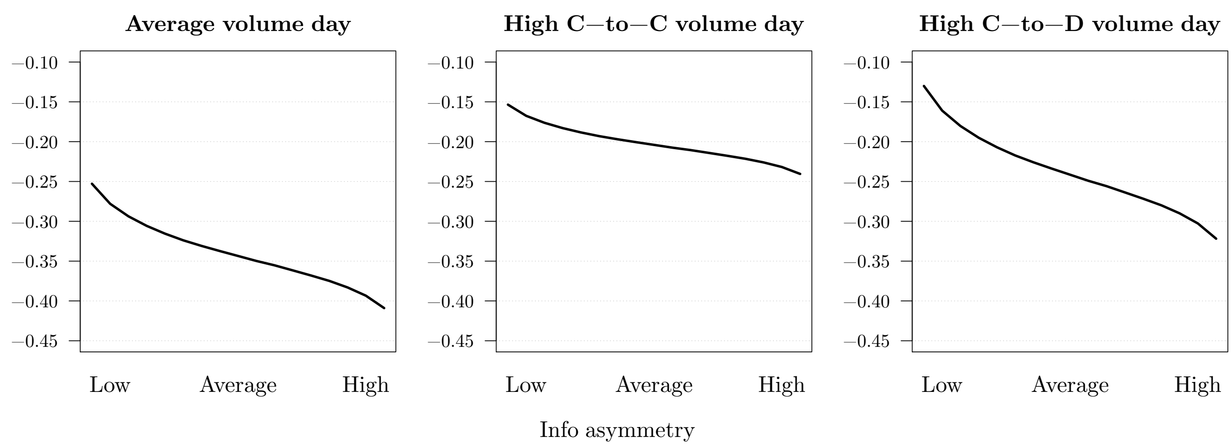

Figure C1. Point estimates for high-volume day reversals for a BBB-rated bond. The calculations are based on models with PC_{bond} from Table V. On the x-axes from left to right are the percentiles of PC_{bond}, from the 20^{th} ('Low' information asymmetry) to the 80^{th} ('High' information asymmetry). Credit rating remains fixed at the 'BBB' level. High CtC volume day is the day with CtC volume 2 standard deviations above the average (and average CtD volume); its reversal is $\mathbb{E}[\hat{\beta}_1|\text{covariates}] + 2\mathbb{E}[\hat{\beta}_2|\text{covariates}]$. High CtD volume day is the day with CtD volume 2 standard deviations above the average (and average CtC volume); its reversal is $\mathbb{E}[\hat{\beta}_1|\text{covariates}] + 2\mathbb{E}[\hat{\beta}_3|\text{covariates}]$. The reversal on the average volume day is simply $\mathbb{E}[\hat{\beta}_1|\text{covariates}]$.

	Mean	Median	S.D.	Min	5th	25th	75th	95th	Max	N.Obs.
$\hat{\beta}_1$	-0.31	-0.32	0.12	-0.62	-0.48	-0.39	-0.23	-0.09	0.07	5051
$\hat{\beta}_2$	0.06	0.05	0.13	-0.63	-0.14	-0.01	0.11	0.27	1.01	5051
$\hat{\beta}_3$	0.05	0.05	0.10	-0.38	-0.12	-0.00	0.11	0.22	0.51	5051
$\hat{\beta}_4$	0.03	0.02	0.76	-4.85	-1.06	-0.24	0.27	1.19	4.72	5051
$\hat{\beta}_5$	0.02	0.00	0.44	-2.86	-0.64	-0.16	0.19	0.67	2.80	5051
Credit rating	7.57	7.00	3.00	1.00	3.00	6.00	9.00	13.00	21.00	5051
Bond bid-ask, %	0.98	0.70	0.83	0.07	0.22	0.42	1.26	2.76	10.06	5051
No. mutual fund owners	46.6	40.0	38.9	0.0	0.0	18.2	65.5	121.8	272.3	5051
Issue size, bln $	0.87	0.68	0.70	0.01	0.20	0.45	1.00	2.25	8.34	5051
No. dealers	33.5	30.3	13.2	8.2	18.4	24.4	39.6	59.6	120.8	5051
Issuer size, bln $	79.2	41.7	110.1	0.2	2.9	15.2	109.1	231.6	920.3	5051
Stock bid-ask, %	0.04	0.03	0.05	0.01	0.01	0.02	0.06	0.13	0.61	5051
PC_{all}	0.00	0.20	1.62	-10.65	-3.12	-0.74	1.00	2.22	8.47	5051
PC_{bond}	0.00	0.25	1.51	-11.53	-2.95	-0.65	0.95	2.03	4.74	5051
$PC_{bond\text{-}ex\text{-}ba}$	0.00	0.36	1.45	-11.93	-2.99	-0.55	0.96	1.61	2.41	5051

Table C4. Summary statistics of the cross-section of volume-return coefficients and their predictors (accounting for proximity to earnings announcements). The table is analogous to Table III but summarizes the second-stage sample in the extension of the baseline model presented in Section V.A.

	Mean	Median	S.D.	Min	5th	25th	75th	95th	Max	N.Obs.
$\hat{\beta}_1$	-0.31	-0.32	0.17	-0.70	-0.57	-0.43	-0.20	0.01	0.21	78314
$\hat{\beta}_2$	0.07	0.06	0.24	-0.82	-0.30	-0.06	0.19	0.48	1.19	78314
$\hat{\beta}_3$	0.05	0.04	0.19	-0.58	-0.27	-0.07	0.16	0.38	0.71	78314
Credit rating	7.61	7.00	3.26	1.00	3.00	5.00	10.00	14.00	21.00	78314
Bond bid-ask, %	1.22	0.81	1.20	0.04	0.23	0.46	1.58	3.43	18.61	78314
No. mutual fund owners	46.4	39.3	42.0	0.0	0.0	13.0	66.7	128.5	386.3	78314
Issue size, bln $	0.96	0.75	0.79	0.01	0.14	0.50	1.25	2.50	15.00	78314
No. dealers	38.1	33.8	17.6	3.7	18.5	25.9	45.9	71.5	232.4	78314
Issuer size, bln $	82.2	43.7	104.9	0.0	2.4	13.8	123.3	250.2	1010.6	71520
Stock bid-ask, %	0.05	0.03	0.08	0.00	0.01	0.02	0.06	0.17	1.97	71520
PC_{all}	0.00	0.17	1.50	-14.84	-2.76	-0.70	0.91	2.04	10.88	71520
PC_{bond}	0.00	0.23	1.43	-16.61	-2.72	-0.64	0.91	1.91	6.31	78314
$PC_{bond\text{-}ex\text{-}ba}$	0.00	0.33	1.38	-16.90	-2.77	-0.57	0.94	1.59	2.33	78314

Table C5. Summary statistics of the year-quarter panel of volume-return coefficients and their predictors. The table is analogous to Table III but summarizes the second-stage sample in the extension of the baseline model presented in Section V.C.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
A. Different inputs in the 1st stage			
Log-volumes	-0.043***	0.008***	-0.002
	(0.001)	(0.001)	(0.001)
Avg. of VW buy and sell prices	-0.050***	0.012***	-0.001
	(0.001)	(0.002)	(0.001)
B. Different models in the 1st stage			
Retail trades excluded	-0.042***	0.008***	0.003***
	(0.002)	(0.001)	(0.001)
Market return added	-0.029***	0.010***	-0.001
	(0.002)	(0.001)	(0.001)
Volumes added linearly	-0.042***	0.016**	0.003**
	(0.001)	(0.007)	(0.001)
C. Different 2nd stage			
PCs extracted from initial obs.	-0.043***	0.010***	-0.000
	(0.001)	(0.002)	(0.001)
Weighted observations	-0.043***	0.006***	0.001*
	(0.001)	(0.001)	(0.001)
Vlm. correlation controls	-0.043***	0.009***	-0.001
	(0.001)	(0.001)	(0.001)

Table C6. Regressions of volume-return coefficients on information asymmetry: robustness tests. Each line in the table presents loadings on the information asymmetry index $PC_{\text{bond-ex-ba}}$ in fixed-effects models for the cross-section of volume-return coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$. Otherwise, the table is analogous to Table IX.

	$\hat{\beta}_1$		$\hat{\beta}_2$		$\hat{\beta}_3$	
	IG	HY	IG	HY	IG	HY
PC_{bond}	−0.042***	−0.050***	0.010***	0.010***	−0.003**	−0.020***
	(0.001)	(0.005)	(0.001)	(0.003)	(0.001)	(0.003)
Rating FE	YES	YES	YES	YES	YES	YES
Observations	6,105	1,101	6,105	1,101	6,105	1,101
R^2	0.346	0.280	0.017	0.017	0.011	0.064

Note: *p<0.1; **p<0.05; ***p<0.01

Table C7. Investment-grade and high-yield subsamples: cross-sectional regressions of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ on information asymmetry. Each models uses a fixed-effects estimator (credit rating fixed effects) with rating-clustered standard errors.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
		Industrial			Financial			Utility	
PC_{bond}	-0.048^{***}	0.003^{***}	0.001	-0.039^{***}	0.015^{***}	-0.008^{***}	-0.061^{***}	-0.005	-0.012
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.005)	(0.008)	(0.007)
Rating FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	3,784	3,784	3,784	2,889	2,889	2,889	492	492	492
R^2	0.355	0.010	0.032	0.345	0.045	0.032	0.334	0.025	0.056

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table C8. Broad industry subsamples: cross-sectional regressions of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ on information asymmetry. Each models uses a fixed-effects estimator (credit rating fixed effects) with rating-clustered standard errors. Industries to which bond issuers belong to are from Mergent FISD.

Appendix D. Additional Aspects of the Model

Log-linear approximation of returns

Consider a homogeneous portfolio of perpetual defaultable bonds with invoice price P_t and coupon rate C. Its next period return R_{t+1} is:

$$1 + R_{t+1} = \frac{(1 - D_{t+1})(P_{t+1} + C)}{P_t},$$

where $D_{t+1} = h_{t+1}L_{t+1}$, and h_{t+1} represents a default rate and $L_{t+1} \in [0, 1]$ represents loss given default for bonds in the portfolio at time $t + 1$.[26] Define $r_t \equiv \log(1 + R_t)$, $p_t \equiv \log(P_t)$, $c \equiv \log(C)$, and $-d_t \equiv \log(1 - D_t)$. Then

$$\begin{aligned} r_{t+1} &= -d_{t+1} - p_t + \log(P_{t+1} + C) \\ &= -d_{t+1} - p_t + p_{t+1} + \log\left(1 + \frac{C}{P_{t+1}}\right) \\ &= -d_{t+1} - p_t + p_{t+1} + \log\left(1 + e^{c - p_{t+1}}\right) \end{aligned}$$

Notice that the first-order Taylor expansion of $\log\left(1 + e^{c-x}\right)$ around $c - \bar{x}$ yields:

$$\log\left(1 + e^{c-x}\right) \approx \log\left(1 + e^{c-\bar{x}}\right) + \frac{e^{c-\bar{x}}}{1 + e^{c-\bar{x}}}\left((c - x) - (c - \bar{x})\right).$$

Then the expression for returns becomes:

$$\begin{aligned} r_{t+1} &= -d_{t+1} - p_t + p_{t+1} + \underbrace{\log\left(1 + e^{c-\bar{p}_{t+1}}\right) + \frac{e^{c-\bar{p}}}{1 + e^{c-\bar{p}}}(c - p_{t+1}) - \frac{e^{c-\bar{p}}}{1 + e^{c-\bar{p}}}(c - \bar{p})}_{\text{Call } \theta = \frac{1}{1+e^{c-\bar{p}}} \Rightarrow \frac{e^{c-\bar{p}}}{1+e^{c-\bar{p}}} = 1-\theta} \\ &= -d_{t+1} - p_t + p_{t+1} - \log\theta + (1 - \theta)(c - p_{t+1}) - (1 - \theta)(c - \bar{p}) \\ &= \theta p_{t+1} - p_t - d_{t+1} + (1 - \theta)c + \underbrace{\left(-\log\theta - (1 - \theta)\log\left(\theta^{-1} - 1\right)\right)}_{\equiv \kappa}, \end{aligned}$$

which is equation (A1). I set $\bar{p} = 0$ (the steady-state bond price is par), then $\theta = \frac{1}{1+C}$.

[26] With probability $1 - h_{t+1}$ the bond pays $P_{t+1} + C$ and with probability h_{t+1} it pays $(1 - L_{t+1})(P_{t+1} + C)$.

Learning by uninformed investors

The uninformed investor is a Bayesian agent learning about g_t and z_t at time t by observing $\tilde{p}_t$ and s_t. Recall that

$$\tilde{p}_t = -a(g_t + bz_t + es_t).$$

Hence, the agent knows $g_t + bz_t$ and an estimate of g_t immediately gives an estimate of z_t. The conditional distribution of $\tilde{p}_t$ given g_t and s_t is

$$\tilde{p}_t | g_t, s_t \sim N\left(-a(g_t + es_t), a^2 b^2 \sigma_z^2\right).$$

The unconditional distribution of g_t is $N(0, \sigma_g^2)$. Bayes theorem implies that $g_t | \tilde{p}_t, s_t$ is also Normal with a PDF $f_{g|\tilde{p},s}$:

$$f_{g|\tilde{p},s} \propto \exp \underbrace{\left(-\frac{(\tilde{p}_t + a(g_t + es_t))^2}{2a^2 b^2 \sigma_z^2} - \frac{g_t^2}{2\sigma_g^2}\right)}_{\equiv -\frac{1}{2}K}.$$

Expanding the square and collecting terms, one gets:

$$K = \frac{g_t^2 - 2g_t\left[-\frac{a\sigma_g^2 \tilde{p}_t + a^2 \sigma_g^2 e s_t}{a^2\left(\sigma_g^2 + b^2 \sigma_z^2\right)}\right] + \Lambda(\tilde{p}_t, s_t)}{\frac{b^2 \sigma_z^2 \sigma_g^2}{\sigma_g^2 + b^2 \sigma_z^2}},$$

where $\Lambda(\tilde{p}_t, s_t)$ does not depend on g_t. Plug in the expression for the pricing function $\tilde{p}_t = -a(g_t + bz_t + es_t)$ to get:

$$\mathbb{E}_t^{(2)}\left[g_t | \tilde{p}_t, s_t\right] = \underbrace{\frac{\sigma_g^2}{\sigma_g^2 + b^2 \sigma_z^2}}_{\equiv \gamma}(g_t + bz_t),$$

$$\mathbb{V}_t^{(2)}\left[g_t | \tilde{p}_t, s_t\right] = (1-\gamma)\sigma_g^2.$$

Optimal demands

The informed investor is solving the following problem:

$$\max_{X_t^{(1)}} \mathbb{E}_t\left[e^{-\left(W_t^{(1)} + X_t^{(1)} r_{t+1} + Z_t(1+n_{t+1})\right)}\right],$$

where the distributions of r_{t+1} and n_{t+1} given the informed investor's information set at time t are both Normal with means $\mathbb{E}_t^{(1)}[r_{t+1}]$ and 0, and variances $\mathbb{V}_t^{(1)}[r_{t+1}]$ and σ_n^2 correspondingly. The covariance between r_{t+1} and n_{t+1} is time-invariant and equals σ_{rn} by assumption. The solution of the informed investor's optimization problem is

$$X_t^{(1)} = \frac{\mathbb{E}_t^{(1)}[r_{t+1}] - \sigma_{rn} Z_t}{\mathbb{V}_t^{(1)}[r_{t+1}]}.$$

The optimization problem for the uninformed investor (who does not own the non-traded asset by assumption) is the same up to Z_t component in the wealth dynamic and yields

$$X_t^{(2)} = \frac{\mathbb{E}_t^{(2)}[r_{t+1}]}{\mathbb{V}_t^{(2)}[r_{t+1}]}.$$

Conditional variances $\mathbb{V}_t^{(1)}[r_{t+1}]$ and $\mathbb{V}_t^{(2)}[r_{t+1}]$ are constant:

$$\mathbb{V}_t^{(1)}[r_{t+1}] = \theta^2(\sigma_f^2 + \sigma_{\tilde{p}}^2),$$
$$\mathbb{V}_t^{(2)}[r_{t+1}] = \theta^2(\sigma_f^2 + \sigma_{\tilde{p}}^2) + (1-\gamma)\sigma_g^2,$$

Now, call $\sigma_r^2 \equiv \theta^2(\sigma_f^2 + \sigma_{\tilde{p}}^2)$ and plug in the expressions for conditional expected returns and variances into the expressions for optimal demand to get:

$$X_t^{(1)} = \frac{a-1}{\sigma_r^2} g_t + \frac{b(a-1)}{\sigma_r^2} z_t + \frac{ae(1-\theta\delta)}{\sigma_r^2} s_t,$$
$$X_t^{(2)} = \frac{a-\gamma}{\sigma_r^2 + (1-\gamma)\sigma_g^2} g_t + \frac{b(a-\gamma)}{\sigma_r^2 + (1-\gamma)\sigma_g^2} z_t + \frac{ae(1-\theta\delta)}{\sigma_r^2 + (1-\gamma)\sigma_g^2} s_t.$$

Existence of the equilibrium

The equilibrium conditions imply the following system of three non-linear equations in a, b, and e:

$$\frac{\omega(a-1)}{\sigma_r^2} + \frac{(1-\omega)(a-\gamma)}{\sigma_r^2 + (1-\gamma)\sigma_g^2} = 0,$$
$$\frac{\omega(ab-\sigma_{rn})}{\sigma_r^2} + \frac{(1-\omega)(a-\gamma)b}{\sigma_r^2 + (1-\gamma)\sigma_g^2} = 0,$$
$$\frac{\omega ae(1-\theta\delta)}{\sigma_r^2} + \frac{(1-\omega)ae(1-\theta\delta)}{\sigma_r^2 + (1-\gamma)\sigma_g^2} = 1.$$

The second equation immediately implies that $b = \sigma_{rn}$ is the only possible solution for b. The system of two remaining equations for a and e can be re-written as

$$0 = \phi_1(a,e) \equiv (a-\bar{a})(\sigma_r^2 + \omega(1-\gamma)\sigma_g^2) - (1-\bar{a})\omega(1-\gamma)\sigma_g^2,$$
$$0 = \phi_2(a,e) \equiv ae(1-\theta\delta)\omega(1-\gamma) - \sigma_r^2(a-\gamma),$$

where $\bar{a} = \omega + \gamma - \omega\gamma > \gamma > 0$. Observe from the first equation that $\phi_1(\bar{a},e) < 0$ and $\phi_1(1,e) > 0$. Hence, if the solution a^* exists, it must be that $a^* \in (\bar{a},1)$. Then, take the derivative of the first equation with respect to a treating e as a function of a:

$$\frac{d}{da}\left[\phi_1(a,e(a))\right] = \sigma_r^2 + \omega(1-\gamma)\sigma_g^2 + (a-\bar{a})(\sigma_g^2 + b^2\sigma_z^2 + \sigma_s^2 e^2 + \sigma_s^2 ae\frac{d}{da}[e(a)]),$$

which is positive for $a \in (\bar{a},1)$ if $e^*(a)$ that solves the second equation $0 = \phi_2(a,e)$ grows in a. In this case we would have a unique positive solution $a^* \in (\bar{a},1)$. Now, I am going to establish the conditions under which this is indeed the case.

The second equation can be re-written as a quadratic equation with respect to e:

$$0 = \phi_2(a,e) = \left(a^2(a-\gamma)\theta^2\sigma_s^2\right)e^2 - \left(a(1-\theta\delta)\omega(1-\gamma)\right)e + (a-\gamma)\theta^2(\sigma_f^2 + a^2(\sigma_g^2 + b^2\sigma_z^2)).$$

Since $a^* > \bar{a} > \gamma$, it must be that $\phi_2(a,0) > 0$, and if the solution e^* exists it must be that $e^* > 0$. Two candidate solutions of the quadratic equation can be written as:

$$\begin{aligned}
e^*(a) &= v(a) \pm v(a)k(a) \text{ where} \\
v(a) &\equiv \underbrace{\frac{(1-\theta\delta)(1-\gamma)\omega}{2\theta^2\sigma_s^2}}_{\equiv 1/B}\frac{1}{a(a-\gamma)}, \\
k(a) &\equiv \sqrt{1 - B^2\psi(a)}, \\
\psi(a) &\equiv (a-\gamma)^2\left(\frac{\sigma_f^2}{\sigma_s^2} + \frac{\sigma_g^2 + b^2\sigma_z^2}{\sigma_s^2}a^2\right);
\end{aligned}$$

and for $a \in (\bar{a},1)$ $v > 0, v' < 0, 0 < k < 1, k' < 0, \psi > 0, \psi' > 0$. For the solutions to exist it must be that $\psi < B^{-2}$ for $a \in (\bar{a},1)$. Observe that

$$\begin{aligned}
\psi &= (a-\gamma)^2\left(\frac{\sigma_f^2}{\sigma_s^2} + \frac{\sigma_f^2}{\sigma_s^2}a^2\right) < (1-\gamma)^2\left(\frac{\sigma_f^2}{\sigma_s^2} + \frac{\sigma_f^2}{\sigma_s^2}a^2\right) \text{ and} \\
B^{-2} &= \frac{(1-\theta\delta)^2(1-\gamma)^2\omega^2}{4\theta^4\sigma_s^4}.
\end{aligned}$$

So, it is suffice to impose the following restriction on model parameters:

$$\frac{(1-\theta\delta)^2\omega^2}{4\theta^4}\frac{1}{\sigma_s^2\left(\sigma_f^2+\sigma_g^2+b^2\sigma_z^2\right)} > 1,$$

to guarantee that the discriminant is non-negative and the quadratic equation for e has solutions. The condition is easy to obey since the shocks in the left-hand side denominator are small numbers. From now on I assume that the condition is satisfied.

Of the two roots of the quadratic equation for e, I am going to focus on the smaller one, $e^*(a) = v(a) - v(a)k(a)$. First, it is the root that guarantees that $e^*(a)$ grows with a when $a \in (\bar{a}, 1)$ as I am about to prove. Second, for reasonable parameters values $v(a)$ is a fairly large number (in a numerical example in Section A it is around 60) and a positive root $v(a) + v(a)k(a)$ does not make much economic sense.

The smaller root $e^*(a) = v(a) - v(a)k(a)$ grows with $a \in (\bar{a}, 1)$ if $\frac{d}{da}\left[e^*(a)\right] > 0$, i.e.:

$$\begin{aligned}
v' - v'k - vk' > 0 &\Leftrightarrow \\
v'(1-k) > vk' &\Leftrightarrow \\
\frac{v'}{v} > \frac{k'}{1-k} &\Leftrightarrow \\
\frac{v'}{v} > \frac{k'(1+k)}{1-k^2} &\Leftrightarrow \\
\frac{v'}{v} > \frac{-\frac{1}{2k}B^2\psi'(1+k)}{B^2\psi} &\Leftrightarrow \\
\frac{v'}{v} > -\frac{1}{2}\frac{\psi'\left(1+\frac{1}{k}\right)}{\psi} &\Leftrightarrow \\
-\frac{2a-\gamma}{a(a-\gamma)} > -\frac{1}{2}\frac{\psi'\left(1+\frac{1}{k}\right)}{\psi} &\Leftrightarrow \\
-\frac{2a-\gamma}{a(a-\gamma)} > -\frac{\frac{\sigma_f^2}{\sigma_s^2}(a-\gamma) + \frac{\sigma_g^2+b^2\sigma_z^2}{\sigma_s^2}a(a-\gamma)(2a-\gamma)}{(a-\gamma)^2\left(\frac{\sigma_f^2}{\sigma_s^2} + \frac{\sigma_g^2+b^2\sigma_z^2}{\sigma_s^2}a^2\right)}\left(1+\frac{1}{k}\right) &\Leftrightarrow \\
2 - \frac{\gamma}{a} < \frac{\frac{\sigma_f^2}{\sigma_s^2} + \frac{\sigma_g^2+b^2\sigma_z^2}{\sigma_s^2}a(2a-\gamma)}{\frac{\sigma_f^2}{\sigma_s^2} + \frac{\sigma_g^2+b^2\sigma_z^2}{\sigma_s^2}a^2}\left(1+\frac{1}{k}\right) &\text{ and observe that}
\end{aligned}$$

$$2-\frac{\gamma}{a}<2<1+\frac{1}{k}<\frac{\frac{\sigma_f^2}{\sigma_s^2}+\frac{\sigma_g^2+b^2\sigma_z^2}{\sigma_s^2}a(2a-\gamma)}{\frac{\sigma_f^2}{\sigma_s^2}+\frac{\sigma_g^2+b^2\sigma_z^2}{\sigma_s^2}a^2}\left(1+\frac{1}{k}\right),$$

which is indeed true.

To sum up, under the condition

$$\frac{(1-\theta\delta)^2\omega^2}{4\theta^4}\frac{1}{\sigma_s^2\left(\sigma_f^2+\sigma_g^2+b^2\sigma_z^2\right)}>1$$

the equation $0=\phi_2(a,e)$ always has a root $e^*(a)>0$ that grows with $a\in(\bar{a},1)$, and it leads to the unique solution $a^*\in(\bar{a},1)$ of $0=\phi_1(a,e^*(a))$.

Derivation of the volume-return relationship

Plug in the expression for the pricing function $\tilde{p}_t=-a(g_t+bz_t+es_t)$ into (A4) to get

$$r_t=-\theta(f_t-m_f)-a\theta g_t-a\theta bz_t-a\theta es_t+(a-1)g_{t-1}+abz_{t-1}+aes_{t-1}.$$

Assume an econometrician also observes $v_{c,t}=|\alpha(\Delta g_t+\sigma_{rn}\Delta z_t)|$ and $v_{s,t}=s_t-s_{t-1}$. Now, the goal is to compute $\mathbb{E}_t\left[r_{t+1}|r_t,v_{c,t},v_{s,t}\right]$.

Call, for the sake of convenience of notations, $x\equiv r_{t+1}$, $y\equiv r_t$, $v\equiv\alpha(\Delta g_t+\sigma_{rn}\Delta z_t)$, and $u\equiv v_{s,t}$. The unconditional distribution of (x,y,v,u) is Gaussian:

$$(x,y,v,u)'\sim\mathcal{N}\left(0,\begin{bmatrix}\Sigma_{11}&\Sigma_{12}\\\Sigma_{12}'&\Sigma_{22}\end{bmatrix}\right),$$

where $\Sigma_{11}=\sigma_{xx}$, $\Sigma_{12}\equiv[\sigma_{xy}\quad\sigma_{xv}\quad\sigma_{xu}]$ and

$$\Sigma_{22}\equiv\begin{bmatrix}\sigma_{yy}&\sigma_{yv}&\sigma_{yu}\\\sigma_{yv}&\sigma_{vv}&0\\\sigma_{yu}&0&\sigma_{uu}\end{bmatrix}.$$

The projection theorem for multivariate Normal distributions implies:

$$\mathbb{E}\left[x|y,v,u\right]=\beta_{xy}y+\beta_{xv}v+\beta_{xu}u,$$

where $(\beta_{xy}\quad\beta_{xv}\quad\beta_{xu})=\Sigma_{12}\Sigma_{22}^{-1}$.

Now consider $\mathbb{E}[x|y,|v|,u]$. First, apply the law of iterated expectations:

$$\begin{aligned}\mathbb{E}[x|y,|v|,u] &= \mathbb{E}[\mathbb{E}[x|y,v,u]\,|y,|v|,u] \\ &= \mathbb{E}[\beta_{xy}y+\beta_{xv}v+\beta_{xu}u|y,|v|,u] \\ &= \beta_{xy}y+\beta_{xv}\mathbb{E}[v|y,|v|,u]+\beta_{xu}u.\end{aligned}$$

Notice that $\mathbb{E}[v|y,|v|,u]=\mathbb{E}[v|y,|v|]$ since $\sigma_{vu}=0$. Now, use the fact that for any random variable Q with a PDF $f_Q(q)$:

$$\mathbb{E}[Q||q|]=|q|\frac{f_Q(|q|)-f_Q(-|q|)}{f_Q(|q|)+f_Q(-|q|)}.$$

In this case, it implies:

$$\mathbb{E}[v|y,|v|]=|v|\frac{f_{v|y}(|v|)-f_{v|y}(-|v|)}{f_{v|y}(|v|)+f_{v|y}(-|v|)},$$

where

$$v|y\sim\mathcal{N}\left(\frac{\sigma_{yv}}{\sigma_y y}y,\sigma_{vv}-\frac{\sigma_{yv}^2}{\sigma_{yy}}\right).$$

After straightforward algebra, one finds that

$$\mathbb{E}[v|y,|v|]=|v|\frac{e^{\rho|v|y}-e^{-\rho|v|y}}{e^{\rho|v|y}+e^{-\rho|v|y}}\approx\rho_{yv}|v|^2y$$

for small values of v, where $\rho_{yv}=\frac{\sigma_{yv}}{\sigma_{vv}\sigma_{yy}-\sigma_{yv}^2}$.

Assembling altogether:

$$\mathbb{E}[x|y,|v|,u]\approx\left(\beta_{xy}+\rho\beta_{xv}|v|^2\right)y+\beta_{xu}u.$$

Since v and u are assumed independent, an additional conditioning on $|u|$ in the expectation sign is straightforward:

$$\mathbb{E}[x|y,|v|,|u|]\approx\left(\beta_{xy}+\rho_{yv}\beta_{xv}|v|^2+\rho_{yu}\beta_{xu}|u|^2\right)y,$$

which is the analogue of (A7). Above, $\rho_{yu}=\frac{\sigma_{yu}}{\sigma_{uu}\sigma_{yy}-\sigma_{yu}^2}$. To compute the coefficients in this relationship given model parameters one needs to compute the covariance matrix Σ. Direct calculations yield:

$$\sigma_{xx} = \theta^2\sigma_f^2 + \left((a\theta)^2 + (a-1)^2\right)\sigma_g^2 + (ab)^2(\theta^2+1)\sigma_z^2 + \frac{(ae)^2(\theta^2+1-2\theta\delta)}{1-\delta^2}\sigma_s^2;$$

$$\sigma_{xy} = (1-a)a\theta\sigma_g^2 - (ab)^2\theta\sigma_z^2 + \frac{(ae)^2(\theta\delta(1-\delta)+\delta-\theta)}{1-\delta^2}\sigma_s^2;$$

$$\sigma_{xv} = \alpha(a(\sigma_g^2 + b^2\sigma_z^2) - \sigma_g^2);$$

$$\sigma_{xu} = \frac{ae(1-\theta\delta)}{1+\delta}\sigma_s^2;$$

$$\sigma_{yy} = \sigma_{xx};$$

$$\sigma_{yv} = \alpha(1-a(1+\theta))\sigma_g^2 - \alpha ab^2(1+\theta)\sigma_z^2;$$

$$\sigma_{yu} = -\frac{ae(1+\theta)}{1+\delta}\sigma_s^2;$$

$$\sigma_{vv} = 2\alpha^2(\sigma_g^2 + b^2\sigma_z^2);$$

$$\sigma_{uu} = \frac{2}{1+\delta}\sigma_s^2.$$