Corporate Bond Price Reversals

Alexey Ivashchenko*

Abstract

I demonstrate empirically that 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 by contrasting price reversals following days with abnormal trading volume across bonds with different information asymmetry. When traders are informed, the part of the price reversal that arises after high-volume days should increase with bond information asymmetry. When traders are uninformed, there should be no such effect. Following high-volume days when investors provide liquidity, the reversal patterns are consistent with the former case. When dealers trade from their inventory, I observe the latter. The results suggest that the informational content of bond prices is higher 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 (doctoral 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 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. The question is whether all liquidity providers are 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.

This paper claims that bond dealers avoid trading with informed investors. Consider a dealer who is approached by an investor willing to sell a bond. The dealer must decide whether to provide liquidity for the transaction herself or to let another investor supply liquidity. In the first case, the dealer buys the bond and holds it as inventory for an ex-ante unknown period. In the second case, the dealer finds another investor willing to buy the bond. The dealer then uses her balance sheet to transfer the bond from the seller to the buyer at prearranged terms, but the bond only stays on the dealer's book for a few minutes. Are bond prices equally likely to reveal private information in these two cases?

I argue that the information content of prices is higher when investors rather than dealers supply liquidity. Assume that there are only two trading motives: private information and liquidity needs. Observed trading volume is the mixture of volumes generated by each of the trading motives. How can we infer the prevalent trading motive from the history of transaction prices and volumes?[1] I use a theoretically grounded empirical methodology that links the prevalent trading motive to the cross-sectional dependence between bond information asymmetry and a particular component of bond price reversal specific to days with abnormally high trading volume. I call such component the '*volume-reversal offset*'

[1]Historical high-quality pre-trade price benchmarks (executable dealer quotes) are unavailable for US corporate bonds, which renders classic microstructure methods to answer the question, like Hasbrouck (1991), inapplicable.

and measure it, for individual bonds, as the difference in price reversals between high- and low-volume days. I argue that, in the cross-section of bonds, the volume-reversal offset must increase with bond information asymmetry if trading is occasionally information-driven. If trading is never due to private information, the volume-reversal offset should not depend on bond information asymmetry. I find that the latter is indeed the case for trading volumes generated by dealer liquidity provision, while client liquidity provision is strongly associated with the former case.

More formally, I first estimate, for individual corporate bonds, a linear dependence between daily return autocorrelation (a measure of bond price reversal) and trading volumes generated by client and dealer liquidity provision:

$$\text{Return autocorrelation}_{t,t+1} = \beta_1 + \beta_2 \times \text{CtC volume}_t + \beta_3 \times \text{CtD volume}_t, \tag{1}$$

where the CtC (client-to-client) and the CtD (client-to-dealer) volumes correspond, respectively, to client and dealer liquidity provision on day t. The volumes are scaled such that β_1 is the price reversal following a trading day with average CtC and CtD volumes. β_2 and β_3 are the volume-reversal offsets. They capture the difference in reversals between high- and low-volume days separately for two types of liquidity providers.

The cross-sections of estimated $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ contain information about the underlying trading motives. In particular, volume-reversal offsets $\hat{\beta}_2$ and $\hat{\beta}_3$ must increase with information asymmetry if transaction prices reveal private information.[2] If trading is only liquidity-driven, then the respective volume-reversal offset should be unrelated to information asymmetry. Hence, my main empirical tests are the regressions of $\hat{\beta}_2$ and $\hat{\beta}_3$ on information asymmetry in the cross-section of US corporate bonds. I consider multiple bond- and issuer-specific information asymmetry proxies (e.g., issue, issuer size, and bid-ask spreads;

[2] The key assumption needed to generate such result in a stylized model of trading a-là Llorente, Michaely, Saar, and Wang (2002) is that the volume of information-driven trading must increase in the strength of the private information signal. I present such a model in Appendix A, while Section II develops the underlying intuition.

institutional ownership; the number of unique dealers) as well as compound information asymmetry indicators that deliver qualitatively similar results.

I find that average $\hat{\beta}_2$ and $\hat{\beta}_3$ are positive. Both indeed represent reversal 'offsets' because the average $\hat{\beta}_1$ stands at around $-1/3$: a third of the original percentage price change reverts the next trading day. Following trading days with above-average volume, the reversal is less strong. The magnitude is such that one-standard-deviation above-average volume on day t takes the price reversal between days t and $t+1$ closer to $-1/4$ (the average offsetting effect is slightly more pronounced for the CtC than for the CtD volume). Most importantly, $\hat{\beta}_2$ increases in the cross-section of bonds (from 0.05 for the lowest info-asymmetry decile to almost 0.09 for the highest decile). That is, the CtC-volume offset is the strongest for those bonds in which information-driven trading is the most likely. In contrast, $\hat{\beta}_3$ decreases with information asymmetry (from 0.06 for the lowest info-asymmetry decile to 0.04 for the highest decile). I argue that such a pattern is expected when informed volume loads onto client liquidity supply while uninformed volumes go to dealers. Economic mechanisms beyond informational trading that link trading volume and reversal, such as search and bargaining functions, the relationship 'size discount', and the bid-ask bounce, would not explain such results. I remain agnostic about the mechanism behind dealers' ability to filter out potentially informed transactions, but the non-anonymity of OTC trading seems to be the most likely explanation.

In an additional test supporting the main result of the paper, I evaluate an extended version of the reversal model (1) that separates days prior to scheduled corporate announcements from the rest of the trading days. If CtC volumes are more likely to be informed, then the relationship between information asymmetry and the CtC volume-reversal offset must be the strongest before earnings announcements when informational motives for trading are the most acute. I indeed find that the pre-announcement CtC volume-reversal offset grows twice faster with information asymmetry than the non-announcement CtC offset. On the contrary, CtD volumes appear even less informed prior to earnings announcements than in non-announcement periods. I also demonstrate that the main result of the paper is robust to alternative econometric specifications and that it holds in various bond subsamples. In addition, I derive implications for investment strategies exploiting corporate bond price reversals.

My paper contributes to several streams of literature. The paper introduces novel empirical evidence about the link between adverse selection and dealer liquidity provision in the US corporate bond market. 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 recent changes in dealers' capital commitment. 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. My paper extends the findings of Goldstein and Hotchkiss (2020) with the analysis of *post-trade* bond price patterns. I find that the realisation of adverse selection risk is more likely following large trades that dealers did not actively intermediate.

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.

More broadly, this paper contributes to the discussion of intermediation frictions in OTC markets and the pricing implications of such frictions. For example, Duffie, Gârleanu, and Pedersen (2005) presented a theoretical framework where OTC market frictions drive asset illiquidity, while Friewald and Nagler (2019) provided 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 commitments. The results in this work have implications for the pricing of the cross-section of corporate bonds. For instance, Chordia, Goyal, Nozawa, Subrahmanyam, and Tong (2017) and Bai, Bali, and Wen (2019) show that a one-month lagged return is a strong return predictor in the cross-section of bonds. However, reversal portfolios have zero or negative Sharpe ratios after adjusting for trading costs (Chordia et al. 2017). I obtain the same result for reversal portfolios constructed on low-information-asymmetry bonds. However, I show that reversal portfolios of high-asymmetry bonds survive the trading cost adjustment.

Methodologically, my analysis of the volume-return offset follows the tradition of Campbell, Grossman, and Wang (1993). In related work, Llorente et al. (2002) investigate the volume-reversal offset 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 organised as follows. Section II develops an intuition that links reversals, trading volume, and information asymmetry and maps it into the empirical strategy of the paper. Section III discusses the bond sample and the estimation of the volume-reversal offset, while the estimates are discussed in Section IV. Section V investigates the key cross-sectional relationship between volume-reversal offsets and information asymmetry. Section VI explores how the results change over time, around earnings announcements, and among bonds issued by the same firm. Section VII runs multiple robustness checks. Section VIII discusses the implications of my results for reversal investment strategies, and Section IX concludes.

II. Hypothesis development and empirical framework

Consider a securities market in which investors trade either on private information or for liquidity (i.e., non-informational) reasons, and the market price never fully reveals the reason for trading at the time of the trade. If a liquidity provider, *pre-trade*, suspects an information-driven transaction, she asks for an additional price concession that adds to a non-informational cost of liquidity provision. Such a price concession represents the remuneration for bearing an adverse selection risk. *Post-trade*, private information becomes public, and the liquidity provider learns whether the original trade was information- or liquidity-driven. That is, the adverse selection risk is or is not realised. In the latter case, the security's fundamental value is unaffected, and the price reverts. The price does not revert as much in the former case because the fundamental value has changed. Hence, the magnitude of the post-trade price reversal contains information about the original trading motive. The question remains whether one can extract this information from the history of trading outcomes in a cross-section of securities that differ in the likelihood of information-driven trading.

To answer this question, consider first the case where trading is occasionally driven by private information. For a given security, assume that the likelihood of having at least *some* information-driven trading is higher on a high-trading-volume day than on a low-volume day.[3] One should then expect the price reversal to be less strong following a high-volume day than following a low-volume day because, in the former case, the original price change is partly due to the change in the security's fundamental value. I will call such a difference

[3] Campbell et al. (1993) and Llorente et al. (2002), for instance, formalize this assumption by making information-driven security demand increasing in the strength of the private information signal.

between price reversals on high- and low-volume days the *volume-reversal offset*, which is not constant in the cross-section of securities. The higher the ex-ante likelihood of information-driven trading, the stronger the volume-reversal offset. A positive relationship between the volume-reversal offset and information asymmetry is the footprint of information-driven trading in the cross-section of individual securities because similarly high volumes are more likely to convey at least some private information in high-asymmetry than low-asymmetry bonds.

Consider a case where trading is only liquidity-driven. For a given security, the difference in post-trade reversals between high- and low-volume days is not due to the revelation of private information. Therefore, when there is no information-driven trading, one should not expect a particular relationship between the volume-reversal offset and information asymmetry in the cross-section of securities. A formal presentation of the above intuition in a stylised trading model built upon Llorente et al. (2002) appears in Appendix A. The model implies that occasionally-informed trading yields a strong relationship between the volume-reversal offset and the underlying information asymmetry in the cross-section of securities. Such a relationship breaks for purely non-informational trading.

I exploit this prediction to contrast the prevalence of information-driven trading in two distinct types of corporate bond transactions. The first is liquidity provision by dealers. In such a case, dealers buy bonds into inventory and hold them at least until the next trading day or sell bonds that have been in the inventory at least since the previous trading day. The second type is liquidity provision by bond investors. Here, dealers connect buying and selling investors by holding the respective inventory only intra-day. I call trading volumes attributed to these two liquidity provision regimes the client-to-dealer (CtD) and the client-to-client (CtC) volume, respectively. I hypothesise that these two types of trading volumes exhibit different informational content: the CtC volume is more likely to contain information about the fundamental value of the bond compared to the CtD volume, which is predominantly liquidity driven. I test the hypothesis by contrasting the cross-sectional dependence of the volume-return offset on bond information asymmetry between CtC and CtD volumes.

The empirical analysis proceeds in two steps. In the first step, I estimate the relationship between trading volume and subsequent price reversal for individual corporate bonds:

$$R_{t+1} = \beta_0 + \underbrace{\left(\beta_1 + \beta_2 \text{CtC volume}_t + \beta_3 \text{CtD volume}_t\right)}_{\text{Conditional return autocorrelation}} R_t + \epsilon_{t+1}. \tag{2}$$

In equation (2) above, R_t and R_{t+1} represent total corporate bond returns on trading days t and $t+1$, respectively. CtC and CtD volumes measure dealer and client liquidity provision. Equation (2) estimates the price reversal (conditional return autocorrelation) between trading days t and $t+1$ as a linear function of the CtC and the CtD trading volume on day t. I standardise CtC and CtD volumes for individual bonds so that $\hat{\beta}_1$ is the average-volume-day reversal. Following a trading day t with one-standard-deviation above-average CtC volume (and keeping the CtD volume at the average level), the reversal is $\hat{\beta}_1 + \hat{\beta}_2$. Hence, $\hat{\beta}_2$ is a measure of the *CtC volume-reversal offset*, that is, the difference in reversal between days with different CtC trading volumes. Likewise, $\hat{\beta}_3$ is the *CtD volume-reversal offset*.

In the second step of my empirical analysis, I estimate the dependence between first-step estimates $\hat{\beta}_n$ and proxies of information asymmetry in the cross-section of individual bonds. The impact of information asymmetry on volume-reversal offsets $\hat{\beta}_2$ and $\hat{\beta}_3$ is of primary interest. If private information is more likely to drive CtC rather than CtD volumes, then $\hat{\beta}_2$ should exhibit a stronger cross-sectional relationship with information asymmetry than $\hat{\beta}_3$. This is a key empirical test of the paper. A $\hat{\beta}_2$ growing with information asymmetry in the cross-section of bonds, unlike $\hat{\beta}_3$, would identify CtC volumes as more information-rich than CtD volumes.

Economic mechanisms beyond information-driven trading might affect the relationship between trading volume and price reversal, hence, the above identification. The question remains whether CtC volumes are uninformed in practice and if the search and bargaining cost of liquidity provision specific to OTC markets (Duffie et al., 2005) drives the empirical result of the paper. As a hypothetical, a liquidity provider contemplates taking a corporate bond position. Search and bargaining costs would make it expensive to offset such a position in the future, even more if the original position is sizable and the bond is infrequently traded. Hence, the cost of the liquidity provision should increase in trade size and bond trading infrequency (which correlates with typical measures of information asymmetry). The post-trade price reversal should then also become stronger for high-asymmetry bonds and following high-volume days. This prediction is the opposite of the private information channel prediction. Hence, search and bargaining costs are unlikely to generate a positive relationship between information asymmetry and CtC volume-reversal offsets.

Moreover, the relationship between non-anonymous dealers and investors in OTC markets affects trading costs. In particular, clients generating a lot of trading volume and, hence, revenue for liquidity providers tend to receive tighter bid-offers than clients trading infrequently and in smaller amounts. Such 'size discounts' in OTC trading have been widely discussed in the literature.[4] Size discounts translate into positive volume-reversal offsets in (2). Hence, positive values of $\hat{\beta}_2$ and $\hat{\beta}_3$ (which are positive for U.S. corporate bonds, on average) might be an implication of the relationship motives in liquidity provision even in the absence of information-driven trading. However, there is no obvious or documented link between client-dealer relationship concerns and the dependence of the volume-reversal offset on information asymmetry in the cross-section of bonds. It is unlikely that liquidity suppliers would give the largest size discounts precisely in the riskiest of bonds.

A random arrival of buying and selling investors trading at given bid and offer prices also generates a price reversal even when trading is purely liquidity-driven. This phenomenon is known as the bid-ask bounce. A bond with wider bid-ask spreads would exhibit stronger

[4] For instance, Green, Hollifield, and Schürhoff (2007) and Edwards, Harris, and Piwowar (2007) quantify size discounts in municipal and corporate bond markets, respectively. Recently, Pintér, Wang, and Zou (2022) demonstrated that the size discount disappears in the UK government bond transaction records after controlling for both dealer and client identities.

price reversals due to the bid-ask bounce. In my empirical setup, this translates into a more negative $\hat{\beta}_1$ for higher-asymmetry bonds (as reported herein in the sample of TRACE bonds). However, the literature has not documented a mechanism linking the bid-ask bounce to the cross-section of volume-reversal offsets, and there is no obvious candidate for such a mechanism. In addition, my second-stage cross-sectional results on the dependence of $\hat{\beta}_2$ and $\hat{\beta}_3$ on information asymmetry hold qualitatively even for information asymmetry proxies that are the least correlated with bond bid-offers.

III. 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.[5] I talk in more detail about the sample in Appendix B.

[5] The Academic Corporate Bond TRACE Data set has been obtained directly from FINRA under the standard data agreement. It contains masked identifying information about corporate bond dealers. Such information is missing in the Enhanced TRACE dataset available through WRDS. In this paper, the information on masked dealer identities only feeds into the calculation of bond-specific information asymmetry proxies in Section III.D.

B. Sample filtering and 'active periods'

I estimate equation (2) for each bond separately, which requires a long enough time series of returns and volumes for every bond. To avoid over-fitting, I require at least 60 daily observations per bond. These are the days with at least one dealer-to-client transaction reported in TRACE; zero-trading days are removed from the sample. 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; very few bonds satisfy this criterion. Instead, I ask for more than 60 daily observations where every two successive observations are at most three business days apart.[6]

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) 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 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

[6]Here, I follow the approach 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	4570902
Maturity, years	8.09	5.50	7.68	1.08	1.50	3.25	9.00	27.08	87.17	4570902
Coupon rate, %	4.94	5.01	1.85	0.45	1.95	3.50	6.12	7.90	15.00	4570902
Rating	7.49	7.00	3.36	1.00	3.00	5.00	9.00	14.00	21.00	4570902
Age, years	4.06	3.08	3.79	0.08	0.33	1.42	5.50	11.25	28.92	4570902
Total bond return, %	0.02	0.02	1.13	-20.64	-1.30	-0.25	0.29	1.32	22.26	4570902
CtC volume, % of size	0.55	0.02	1.98	0.00	0.00	0.00	0.17	2.85	15.70	4570902
$-\Delta$Inventory, % of size	0.02	0.02	3.51	-19.40	-4.42	-0.19	0.36	4.21	18.48	4570902
CtD volume, % of size	1.50	0.28	3.17	0.00	0.01	0.06	1.24	7.85	19.40	4570902
Realized bond bid-ask, %	1.07	0.64	1.29	0.00	0.09	0.30	1.39	3.39	19.66	2795698
No. mutual fund owners	47.3	40.0	42.2	0.0	0.0	15.0	67.0	129.0	402.0	4570902
No. dealers	37.2	33.0	17.7	1.0	17.0	25.0	45.0	70.0	289.0	4570902
Issuer equity value, bln $	88.3	49.3	108.4	0.0	2.6	16.3	135.0	258.1	1103.5	4203170
Stock bid-ask, %	0.05	0.03	0.10	0.00	0.01	0.02	0.05	0.16	1.99	4203164
Days to earnings announcement	44.4	44.0	26.3	0.0	4.0	22.0	67.0	85.0	92.0	3450335

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 non-negative by construction. $-\Delta$Inventory is the difference between client purchases and client sales; it can be positive (aggregate dealers' inventory in the bond decreases) or negative (aggregate dealers' inventory in the bond 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 VI.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.

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 bln USD and a 5% coupon rate, observed 4 years since issuance and 8 years prior to maturity. 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*

Each transaction record in TRACE is a report by a bond dealer about an individual bond transaction. The dealer indicates whether a trading counterparty is a client or another bond dealer. To measure two types of aggregate trading volume per bond per day, I use only TRACE transactions between dealers and clients. A dealer also reports whether she was a buyer or a seller in each such transaction. 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:

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

CtC volume$_{it}$ 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 $t-1$. CtC volume$_{it}$ is zero on the days when either V_{it}^{buy} or V_{it}^{sell} is zero; otherwise it is greater then zero.[7]

[7]When I estimate equation (2), I further standardize CtC and CtD volumes separately for each bond and each active trading period when I estimate equation (2). Hence, a zero-CtC-volume observation translates into $(-1)\times$ the average CtC volume scaled by the standard deviation of the CtC volume for that bond and that active trading period.

The difference between client purchases and client sales is a negative change in dealers' inventory:

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

$V_{it}^{(s)}$ 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 (2), I consider the absolute value of $V_{it}^{(s)}$, which I call the CtD trading volume:[8]

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

Table I shows that the CtD volume is, on average, several times higher than the CtC volume. Notice that a traditional measure of daily trading volume (excluding inter-dealer transactions), $V^{\text{buy}} + V^{\text{sell}}$, is equal to CtD volume $+ 2 \cdot$ CtC volume.

CtC volume$_{it}$ and CtD volume$_{it}$, as defined above, treat bonds dealers in their entirety. Assume that on the day t, clients sold \$10 mln worth of bond i to dealers and purchased \$8 mln from dealers. Assume further that these trading volumes represent, respectively, 1% and 0.8% of the total outstanding amount in bond i. Then, CtC volume$_{it} = 0.8\%$, and CtD volume$_{it} = 0.2\%$. The latter number shows that dealers' inventory in bond i, aggregated across all dealers, has increased by 0.2% of the outstanding amount on day t. It can be that \$10 mln and \$8 mln were sold to and bought from different dealers. I do not construct and analyze individual dealer inventory in this paper. My focus is on the use of aggregate dealers' balance sheet space as opposed to client liquidity provision and how both relate to the underlying trading motives. On a practical level, it means that I evaluate CtC volume$_{it}$ and CtD volume$_{it}$ with a standard academic version of the TRACE dataset; individual dealer identities are not used in this calculation.

[8] 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 volume-reversal offsets on information asymmetry. An investigation of the asymmetries in conditional price reversals is beyond the scope of this paper.

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 $\mathrm{Corr}\left(V_t^{(c)}, V_t^{(s)}\right) = 0$, 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 variables to proxy for the extent of information asymmetry in the cross-section of bonds. 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:

- 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 (2), 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 bonds with lower dealer competition should exhibit higher volume-reversal offsets 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 a confounding non-informational 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 (2), is similar to the expected effect of information asymmetry. The impact of the bid-ask bounce on β_2 and β_3 in equation (2) is unclear because it depends on whether the effect becomes stronger or weaker with higher trading volumes. Compounding information asymmetry indicators constructed later in this section utilize sets of proxies with and without bond bid-ask spreads to address these concerns. Also, a) Section V.B discusses the effect of individual proxies on β_i and demonstrates that it extends beyond what realized bid-ask spreads capture, and b) Section VII establishes the robustness of results when bond returns are calculated using a simple average of volume-weighted client buy and sell prices, which attenuates the effect of the bid-ask bounce at least on the days when both buy and sell client trades happen.

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 $\hat{\beta}_n$ in the second-stage cross-sectional regressions. Second, it simplifies the presentation of results. I test for the impact of information asymmetry on volume-reversal offsets. It is easier to interpret such a test when information asymmetry is a scalar 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.[9] 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.
- $\text{PC}_{\text{bond-ex-ba}}$: 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% ($\text{PC}_{\text{bond-ex-ba}}$) 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.

[9] 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 VII. In particular, I show that the main results of the paper hold when the principal components are extracted from the *initial* values of individual information asymmetry proxies (the values at the beginning of the first active trading period for each bond).

IV. Volume-return relationship

I estimate equation (2) separately for every bond and every active period rescaling trading volumes such that β_1 measures the first return autocorrelation on the average-volume days:

$$R_{t+1} = \beta_0 + \beta_1 R_t + \beta_2 R_t \tilde{V}_t^{(c)} + \beta_3 R_t \tilde{V}_t^{(s)} + \epsilon_{t+1}. \tag{3}$$

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[10] 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 one 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 increases by 33 b.p., on average, the next trading day. The average

[10] 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.
$\hat{\beta}_1$	-0.3345	-0.3489	179	15702	6	14277	15881
$\hat{\beta}_2$	0.0687	0.0587	11146	4735	2651	518	15881
$\hat{\beta}_3$	0.0531	0.0519	10780	5101	3206	773	15881

Table II. Summary statistics of the estimated volume-return coefficients of equation (3). 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.

$\hat{\beta}_2$ 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 $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$. Strongly negative $\hat{\beta}_1$ is a reflection of the high illiquidity of the corporate bond market, be it due to informational or non-informational frictions. The values of $\hat{\beta}_2$ and $\hat{\beta}_3$ are close; hence, both types of trading volume contribute similarly to price reversals. Positive $\hat{\beta}_2$ and $\hat{\beta}_3$ can be consistent with the presence of informed trading but can also reflect relationship size discounts (Pintér et al., 2022). In the next section, I investigate explanatory factors of the cross-sections of $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_3$ with a particular focus on the impact of information asymmetry.

V. Determinants of volume-return coefficients

A. *Empirical design*

In Section II, I put forward an intuition on how volume-return coefficients β_1, β_2, and β_3 in equation (3) should vary with information asymmetry. I suggest that more information asymmetry implies lower β_1 (stronger reversals on average), higher β_2 (a stronger CtC volume-reversal offset), and no particular effect on β_3 (a weaker link between information asymmetry and the CtD offset as compared to the CtC offset). 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 VI.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 $k \in \{1, \ldots, 21\}$.[11] As the baseline, I fit the following model for each n (i.e., the cross-section of each volume-return coefficient) separately:

$$\hat{\beta}_{n,i}(k) = c_n \cdot \text{Info asymmetry proxy (-ies)}_{n,i} + \text{Rating } k \text{ FE}_n + \epsilon_{n,i}, \tag{4}$$

[11] In this section, k is the prevalent rating in the active trading period in which $\hat{\beta}_n$ is estimated. In Section VII, 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 (4).[12] 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 $\hat{c}_3$ either smaller than $\hat{c}_2$ or insignificantly different from zero, or negative.

Table III presents summary statistics of the cross-section of estimated volume-return coefficients with their potential explanatory factors. There are about seven thousand individual bonds issued by one 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 (4) as Table III shows. Table C3 in Appendix further presents cross-sectional correlations of information asymmetry proxies.

	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	7212
$\hat{\beta}_2$	0.06	0.06	0.12	-0.52	-0.12	0.00	0.12	0.27	0.85	7212
$\hat{\beta}_3$	0.05	0.05	0.10	-0.38	-0.12	-0.01	0.11	0.22	0.51	7212
Credit rating	7.78	8.00	3.33	1.00	3.00	6.00	9.00	14.00	21.00	7212
Bond bid-ask, %	1.13	0.78	0.99	0.07	0.22	0.45	1.50	3.16	14.81	7212
No. mutual fund owners	40.6	34.3	38.2	0.0	0.0	8.5	59.2	114.9	381.8	7212
Issue size, bln $	0.77	0.58	0.69	0.01	0.05	0.35	1.00	2.00	9.00	7212
No. dealers	32.1	28.7	13.0	7.7	17.4	23.3	37.4	58.0	120.7	7212
Issuer size, bln $	75.4	37.3	105.8	0.0	2.2	12.4	102.6	235.8	930.8	6676
Stock bid-ask, %	0.05	0.03	0.07	0.01	0.01	0.02	0.06	0.16	1.33	6676
PC_{all}	0.00	0.23	1.60	-12.50	-3.07	-0.70	0.98	2.13	8.16	6676
PC_{bond}	0.00	0.21	1.52	-13.85	-2.84	-0.66	0.93	2.10	5.91	7212
$\text{PC}_{\text{bond-ex-ba}}$	0.00	0.36	1.45	-14.44	-2.83	-0.51	0.94	1.57	2.27	7212

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. $\text{PC}_{\text{bond-ex-ba}}$ further excludes realized bond bid-ask from the list of factors.

[12] 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 (4) 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 issuers, 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. However, the results on $\hat{\beta}_1$ are also consistent with explanations beyond the private information channel.

Table IVb presents the results for $\hat{\beta}_2$. Recall that higher β_2 means a stronger volume-reversal offset 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. 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 stronger CtC volume-reversal offsets. Observe also, in models 7 and 8, that the effects of the issue size and the number of intermediating dealers on $\hat{\beta}_2$ extend beyond the effect of the realized bond bid-ask, which would likely not be the case if the variation in the cross-section of $\hat{\beta}_2$ was solely due to the bid-ask bounce.

	(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	7,212	7,212	7,212	7,212	6,676	6,676	7,212	6,676
R^2	0.112	0.301	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	7,212	7,212	7,212	7,212	6,676	6,676	7,212	6,676
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	7,212	7,212	7,212	7,212	6,676	6,676	7,212	6,676
R^2	0.071	0.018	0.017	0.020	0.026	0.025	0.080	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).

Table IVc presents the regressions for $\hat{\beta}_3$. The interpretation of β_3 is analogous to β_2, but now the focus is on the CtD volume-reversal offset. 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.

	$\hat{\beta}_1$			$\hat{\beta}_2$			$\hat{\beta}_3$		
	(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	6,676	7,212	7,212	6,676	7,212	7,212	6,676	7,212	7,212
R^2	0.342	0.337	0.322	0.012	0.018	0.017	0.022	0.021	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.

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 (4).

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 have larger CtC volume-reversal offsets. 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.

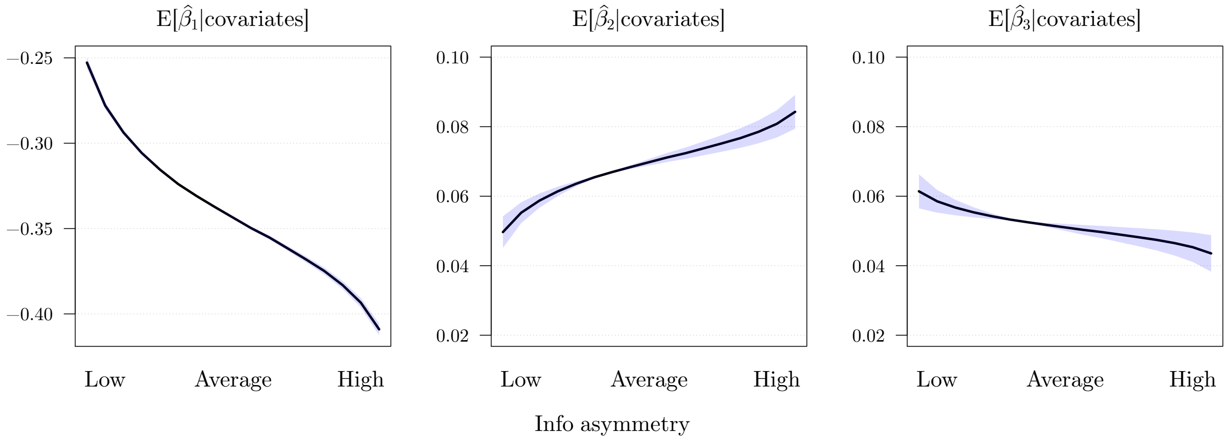

Figure 1. Point estimates and confidence intervals for the expected values of volume-return coefficients for a BBB-rated bond. The calculations are based on models with PC_{bond} from Table V. On the x-axes, from left to right, are the percentiles of PC_{bond}, from the 10^{th} ('Low' information asymmetry) 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.

Figure 1 plots the relationships between $\hat{\beta}_n$ and PC_{bond} from Table V. The left panel of Figure 1 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 middle panel in Figure 1 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 almost 0.09 for high-asymmetry bonds. Finally, the right panel in Figure 1 demonstrates that the 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.[13] 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.

VI. 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 an informed trade flow better than non-dealer liquidity providers.

A. Pre-announcement effects

In Section V, 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 a stronger dependence between the CtC volume-reversal offset and information asymmetry right before earnings announcements. I test for such an effect with the following modification of my baseline methodology. I modify the volume-return relationship (3) to separate days close to quarterly

[13]Figure A1 in Appendix A shows similar-shaped functions for the dependence of $\hat{\beta}_n$ on information asymmetry in a theoretical model that distinguishes occasionally-informed and never-informed trading volumes.

earnings announcements from all other trading days:

$$R_{t+1} = \beta_0 + \beta_1 R_t + \beta_2 R_t \tilde{V}_t^{(c)} + \beta_3 R_t \tilde{V}_t^{(s)} + \beta_4 R_t \tilde{V}_t^{(c)} \mathbb{1}_t^{\text{EA}} + \beta_5 R_t \tilde{V}_t^{(s)} \mathbb{1}_t^{\text{EA}} + \epsilon_{t+1}. \tag{5}$$

In equation (5), $\mathbb{1}_t^{\text{EA}}$ 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 (5) for the same subset of individual bonds issued by public firms and the same active trading periods as in Section V. 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 1st and 99th 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 V. 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 (4) as before.

Table VI presents the regressions of volume-return coefficients $\hat{\beta}_1$–$\hat{\beta}_5$ on the information asymmetry index PC_{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 PC_{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 CtC volume-reversal offsets for the days that are far from ($\hat{\beta}_2$) and close to ($\hat{\beta}_4$) quarterly issuer earnings announcements. I find

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$
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	5,054	5,054	5,054	5,054	5,054
R^2	0.354	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 (5). 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.

that the loadings on PC_{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 I 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 the CtD volume-reversal offset far from ($\hat{\beta}_3$) or close to ($\hat{\beta}_5$) earnings announcements. Table VI shows that $\hat{\beta}_3$ is unrelated to PC_{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 PC_{bond} in Table VI. If there was information-driven trading in client-to-dealer transactions just before earnings announcements, $\hat{\beta}_5$ would increase in information asymmetry. I find the opposite. It confirms that bond dealers can identify and avoid information-driven trade flows.

B. *Within-issuer effects*

There are firms that have many outstanding bonds at each point in time. 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 (4) 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 $\hat{\beta}_2$ increases. The loadings on PC_{bond} in Table VII are similar in size to those in Table V (without issuer fixed effects).

This result 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. It can also be private knowledge about the exercise probability of embedded 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 (3) 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 trade almost 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 (3) for individual bonds within each quarter, I obtain bond i – year-quarter q panels of volume-return coefficients $\hat{\beta}_{1,i,q}$, $\hat{\beta}_{2,i,q}$, and $\hat{\beta}_{3,i,q}$. I then explain the panels of volume-return coefficients with the fixed effects models that are analogous to (4) 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 $\mathrm{PC}_{\mathrm{bond}}$ 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 estimates are also close to the previously obtained values. Figure 2 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

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
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	78,332	78,332	78,332
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	12,263	59,610	12,263	59,610	12,263	59,610
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.

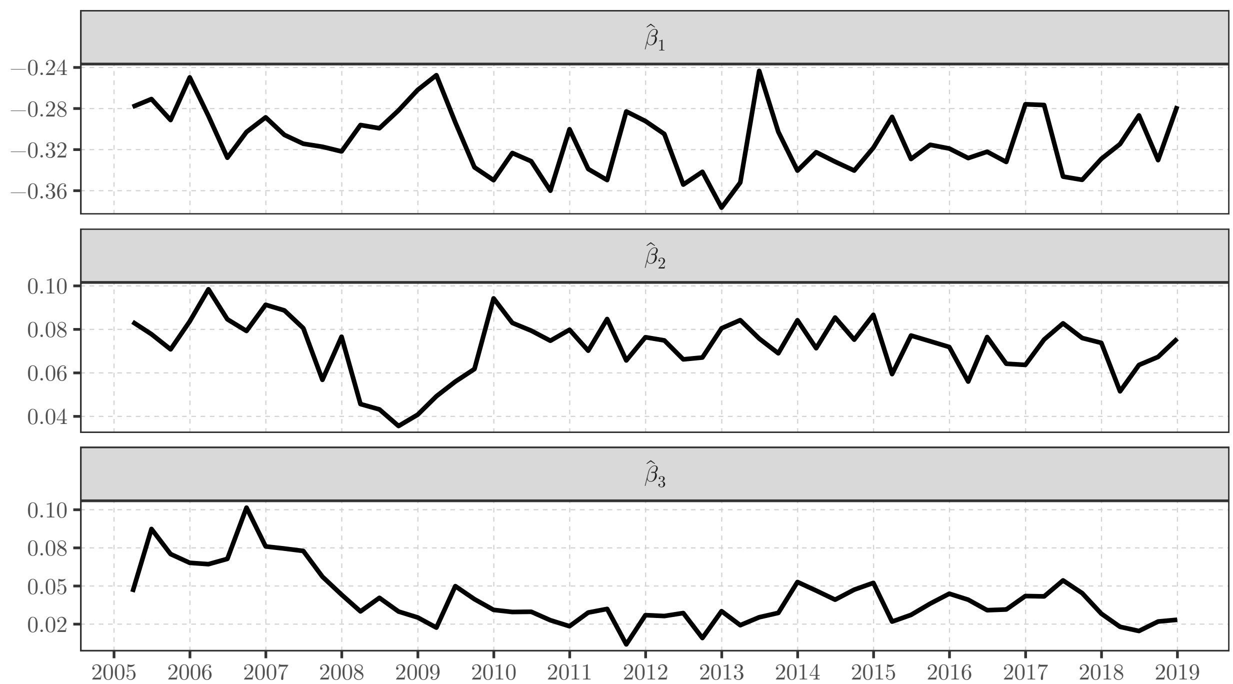

Figure 2. Time fixed effects for volume-return coefficients extracted from the models of Table VIIIa.

(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. 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.

VII. Robustness

I run several robustness tests for the paper's main empirical result, which is the dependence of $\hat{\beta}_2$ and $\hat{\beta}_3$ on information asymmetry. I report the results of these robustness tests in Table IX with PC_{bond} as the information asymmetry index. I consider the following modifications to the baseline methodology.

- *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 (3) 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. 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 et al. 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 either \$10k or \$100k notional from consideration. The cut-off for retail-sized trades is somewhat ambiguous, and both aforementioned values have been used in the academic literature. The industry practice is to use \$10k as the cut-off. In private conversations, several largest institutional bond investors have confirmed that their rebalancing trades are very often below \$100k notional (but almost all are above \$10k), especially in the case of 'portfolio trading' (as opposed to transacting individual bonds). With either cut-off, the sample gets smaller and concentrates on more liquid bonds here because an active trading period now consists only of trading days with non-zero non-retail volume.
- *Exclusion of retail notes.* 7.5% of the sample (see Table C8 in Appendix C) is retail notes. These instruments are similar to corporate bonds, but at issuance, they can be purchased directly from an issuing corporation and in smaller sizes than for a typical corporate debenture. In the secondary market, retail notes are not necessarily traded by retail investors only. Yet, it is likely that the fraction of retail investors in this subset of bonds is greater than in other instruments in the sample. To account for a possible bias that retail notes might introduce, I remove them from the sample and re-run the baseline model on the sample that does not contain retail notes.
- *Market return in the first-stage model.* Here, I add the market return as a linear term to the right-hand side of equation (3):

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

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 $\hat{\beta}_n$.

- *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 (3):

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

  As with the inclusion of market returns, it potentially corrects for 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 (4). 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 $\hat{\beta}_n$ 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 order over two business days.[14] 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

[14] 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).

including the first autocorrelations of $\tilde{V}_t^{(c)}$ and $\tilde{V}_t^{(s)}$ (averages for individual bonds) in the second-stage model.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$
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	-0.047***	0.052***	-0.001
	(0.004)	(0.013)	(0.002)
Trades less than \$10k excluded	-0.046***	0.032***	0.007***
	(0.002)	(0.007)	(0.001)
Trades less than \$100k excluded	-0.034***	0.009**	0.005***
	(0.002)	(0.004)	(0.001)
Retail notes excluded	-0.045***	0.004***	0.002
	(0.001)	(0.001)	(0.001)
B. Different models in the 1st stage			
Market return added	-0.029***	0.010***	-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. 'Trades less than … excluded' remove trades below an indicated threshold (in notional amount) from consideration. 'Retail notes excluded' removes retail notes from the sample. '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 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 (0.006–0.052; 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 information asymmetry or there is no significant link between the two, except for the cases when retail-sized trades are removed. In both such cases, despite the sample zooming into the most liquid and actively traded bonds, $\hat{\beta}_2$ is several times more sensitive to the information asymmetry index than $\hat{\beta}_3$. This is also in line with the explanation that CtC volume is more likely to be informed, as suggested by the model in Appendix A, in particular, Figure A1. Overall, these robustness tests 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 C6 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 C7 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.

VIII. 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.[15] The rebalancing is at the end of each month. Here, I consider a raw 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 portfolio with more information asymmetry to outperform the reversal portfolio with less information asymmetry out-of-sample.

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

[15] 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.

	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	1.17	0.02	6.87	-0.04	-0.20
LR: few funds	8.57	5.99	1.36	1.98	2.21	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).

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.

IX. Conclusion

In this paper, I estimate individual corporate bond return autocorrelation as a linear function of the trading volume and explore the determinants of the estimated relationship in the cross-section of TRACE bonds. My analysis focuses on the impact of information asymmetry on the volume-reversal offset: the difference in bond price reversals between high- and low-volume days.

In the cross-section of bonds, the volume-reversal offset must increase with the underlying bond information asymmetry when trading is occasionally driven by private information. When trades are uninformed, there should be no such dependence. I use this prediction to identify the informational content of trading volumes attributed to either dealer or client liquidity provision.

I find that bonds with higher information asymmetry exhibit more substantial volume-reversal offsets when dealers' inventory does not change, client purchases equal client sales, and bond investors are, in fact, liquidity providers. I find the opposite when dealers supply liquidity and trading volumes mirror dealers' bond inventory changes. In this case, the volume-reversal offset either decreases in or does not depend on information asymmetry. The result gets stronger when informational motives for trading are more acute, for instance, before earnings announcements.

These results suggest that the informational content of bond prices is higher when investors, rather than intermediating dealers, supply liquidity to the corporate bond market. Since OTC dealers typically know their clients well and might detect informed traders, the dealers let other investors supply liquidity for informed trades. As a result, dealers and non-dealer liquidity providers have a different exposure to an adverse selection risk in the US corporate bond market. I also derive implications for the design of investment strategies exploiting corporate bond price reversals.

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.

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* 135, 16–40.

Green, Richard C., Burton Hollifield, and Norman Schürhoff, 2007, Financial intermediation and the costs of trading in an opaque market, *The Review of Financial Studies* 20, 275–314.

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.

Hasbrouck, Joel, 1991, Measuring the information content of stock trades, *The Journal of Finance* 46, 179–207.

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.

Pintér, Gábor, Chaojun Wang, and Junyuan Zou, 2022, Size discount and size penalty: trading costs in bond markets, Working Paper 970, Bank of England.

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 justifies equation (3) and yields predictions about the dependence of volume-return coefficients on information asymmetry conditional on the informativeness of the trading volume.

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. Noisy bond supply generates an additional trading volume that is not due to liquidity or informational signals the agents receive. In the model, I *assume* that supply changes are independent of the arrival of private news. Under this assumption, I can contrast the dependence of price reversal on the trading volume between occasionally-informed and uninformed order flows.

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, Greenwood, and Liao (2018) demonstrate that the 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 $t+1$.[16]

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

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

[16]For the derivation see Appendix D.

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 some non-traded asset that generates a log return of n_{t+1} in the subsequent period.[17] 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.[18] 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.

---

[17] 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.

[18] 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.

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.[19] 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

$$W_{t+1}^{(1)} \approx W_t^{(1)} + X_t^{(1)} r_{t+1} + z_t(1 + n_{t+1}),$$

$$W_{t+1}^{(2)} \approx W_t^{(2)} + X_t^{(2)} r_{t+1}.$$

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 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. The second is driven by exogenous changes in asset supply. The 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).

[19] 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.

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

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

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.

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, 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

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

[20]See Appendix D.

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 $v_{c,t} = |\alpha\left(\Delta g_t + \sigma_{rn}\Delta z_t\right)| \geqslant 0$. 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. 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 in Appendix C demonstrates that this assumption largely holds in the data.

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 (3) estimated in the empirical part of the paper. Unlike equation (3), 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 (3) with volume entering the equation without a square (Llorente et al. 2002 follow the same approach). It does not change an economic interpretation of volume-return coefficients.[21]

In the model, coefficients β_1, β_2, and β_3 change as the extent of informed trading changes. I solve the model numerically to study the dependence of the volume-return coefficients on σ_g^2. In Figure A1, I present the relationship between information asymmetry σ_g^2 and volume-return 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%.[22] The latter stays fixed in all numerical solutions; this is an additional constraint I impose on the solutions of the model.[23] 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.

---

[21]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.

[22]In Figure A1, I set $\delta = 0.95$ which roughly corresponds to $\mathrm{Corr}(\Delta s_t, \Delta s_{t-1}) = -0.03$ because in the model $\mathrm{Corr}(\Delta s_t, \Delta s_{t-1}) = -\frac{1}{2}(1-\delta)$. 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.

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

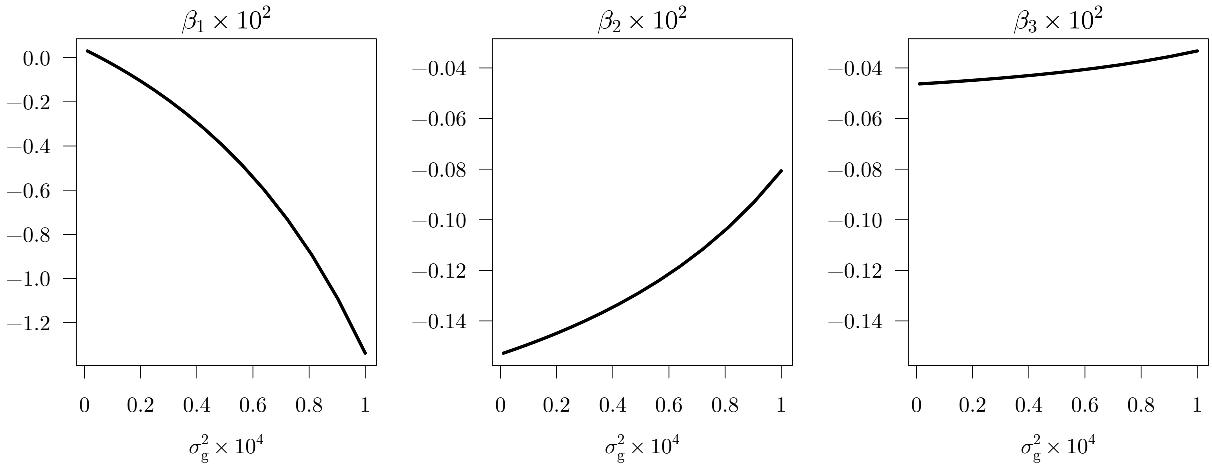

Figure A1. Dependence of β_1, β_2, and β_3 on information asymmetry σ_g^2 holding total return variance fixed. Each point on the curves is a numerical solution of the model. I obtain the relationships between σ_g^2 and β coefficients by varying σ_g from 0 to 1% holding an unconditional standard deviation 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 of Llorente et al. (2002). With more informed trading, returns tend to revert more, but the volume-reversal offset increases with the extent of occasionally-informed trading as suggested by the dependence of β_2 on σ_g^2. 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 are 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.[24] On

[24]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

the central panel, the reversal following high-volume days is the strongest when σ_g is zero 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 1. 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 main result of the paper: client-to-client trading volume may be due to private information, but client-to-dealer trading volume is primarily driven by liquidity needs.

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 IV 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, the volume-reversal offset must also increase.

---

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.

Appendix B. Data and sample

Sample selection

I apply 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 (except for 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);[25]
- 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.

[25] The distribution of types of debt securities in the final sample, after active trading periods are singled out, is in Table C8 in Appendix C.

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 99%;
- Realized 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.
Corr($V_t^{(c)}, \lvert V_t^{(s)}\rvert$)	0.188	0.188	13956	1925	9820	41	15881
Corr($V_t^{(c)}, V_t^{(s)}$)	-0.052	-0.043	5786	10095	1517	4806	15881
Corr($V_t^{(c)}, V_{t-1}^{(c)}$)	0.056	0.020	9004	6877	4364	21	15881
Corr($\lvert V_t^{(s)}\rvert, \lvert V_{t-1}^{(s)}\rvert$)	0.095	0.089	12193	3688	6300	43	15881

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}	$\text{PC}_{\text{bond-ex-ba}}$
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	0.33		
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***	0.38***	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.

	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	5054
$\hat{\beta}_2$	0.06	0.05	0.13	-0.63	-0.14	-0.01	0.11	0.27	1.01	5054
$\hat{\beta}_3$	0.05	0.05	0.10	-0.38	-0.12	-0.00	0.11	0.22	0.51	5054
$\hat{\beta}_4$	0.03	0.02	0.76	-4.85	-1.06	-0.24	0.27	1.19	4.72	5054
$\hat{\beta}_5$	0.02	0.00	0.44	-2.87	-0.64	-0.16	0.19	0.67	2.80	5054
Credit rating	7.57	7.00	3.01	1.00	3.00	6.00	9.00	13.00	21.00	5054
Bond bid-ask, %	0.98	0.70	0.83	0.07	0.22	0.42	1.26	2.76	10.06	5054
No. mutual fund owners	46.7	40.0	39.2	0.0	0.0	18.2	65.5	121.9	360.5	5054
Issue size, bln $	0.87	0.68	0.71	0.01	0.20	0.45	1.00	2.25	9.00	5054
No. dealers	33.5	30.3	13.3	8.2	18.4	24.4	39.6	59.7	120.8	5054
Issuer size, bln $	79.2	41.8	110.1	0.2	2.9	15.2	109.1	231.6	920.3	5054
Stock bid-ask, %	0.04	0.03	0.05	0.01	0.01	0.02	0.06	0.13	0.61	5054
PC_{all}	-0.00	0.20	1.62	-11.94	-3.11	-0.73	1.00	2.21	8.43	5054
PC_{bond}	-0.00	0.25	1.51	-13.40	-2.94	-0.64	0.95	2.02	4.71	5054
$PC_{bond\text{-}ex\text{-}ba}$	-0.00	0.36	1.46	-13.82	-2.97	-0.55	0.96	1.60	2.39	5054

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 VI.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	78332
$\hat{\beta}_2$	0.07	0.06	0.24	-0.82	-0.30	-0.06	0.19	0.48	1.19	78332
$\hat{\beta}_3$	0.05	0.04	0.19	-0.58	-0.27	-0.07	0.16	0.38	0.71	78332
Credit rating	7.61	7.00	3.26	1.00	3.00	5.00	10.00	14.00	21.00	78332
Bond bid-ask, %	1.22	0.81	1.20	0.04	0.23	0.46	1.58	3.42	18.61	78332
No. mutual fund owners	46.4	39.3	42.0	0.0	0.0	13.0	66.7	128.5	386.3	78332
Issue size, bln $	0.96	0.75	0.79	0.01	0.14	0.50	1.25	2.50	15.00	78332
No. dealers	38.1	33.8	17.6	3.7	18.5	25.9	45.9	71.5	232.4	78332
Issuer size, bln $	82.2	43.7	104.9	0.0	2.4	13.8	123.4	250.2	1010.6	71536
Stock bid-ask, %	0.05	0.03	0.08	0.00	0.01	0.02	0.06	0.17	1.98	71536
PC_{all}	0.00	0.17	1.50	-14.84	-2.76	-0.70	0.91	2.04	10.90	71536
PC_{bond}	0.00	0.23	1.43	-16.61	-2.72	-0.64	0.91	1.91	6.31	78332
$PC_{bond\text{-}ex\text{-}ba}$	0.00	0.33	1.38	-16.90	-2.77	-0.57	0.94	1.59	2.33	78332

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 VI.C.

	$\hat{\beta}_1$		$\hat{\beta}_2$		$\hat{\beta}_3$	
	IG	HY	IG	HY	IG	HY
PC_{bond}	-0.042^{***}	-0.051^{***}	0.010^{***}	0.010^{***}	-0.003^{**}	-0.020^{***}
	(0.001)	(0.004)	(0.001)	(0.003)	(0.001)	(0.003)
Rating FE	YES	YES	YES	YES	YES	YES
Observations	6,108	1,104	6,108	1,104	6,108	1,104
R^2	0.346	0.282	0.017	0.018	0.011	0.065
Note:	*p<0.1; **p<0.05; ***p<0.01					

Table C6. 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	3,788	3,788	3,788	2,891	2,891	2,891	492	492	492
R^2	0.355	0.010	0.033	0.345	0.045	0.032	0.334	0.025	0.056
Note:	*p<0.1; **p<0.05; ***p<0.01								

Table C7. 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.

	% of sample
US Corporate Debentures	78.1
US Corporate Medium-Term Notes	12.7
US Retail Notes	7.5
US Corporate Bank Note	1.7

Table C8. Sample composition by type of debt instrument, in % of the total number of individual bonds in the sample. This corresponds to the sample in Table I, i.e., after active trading periods are singled out.

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 $f_{g|\tilde{p},s}$:

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

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:

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

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

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

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

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

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

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

In this case, it implies:

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

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

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

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

Assembling altogether:

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

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

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

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:

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

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

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

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

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

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

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

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

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