The Importance of Investor Heterogeneity: An Examination of the Corporate Bond Market

Jian Li* and Haiyue Yu†

Jan 2022

Abstract

We show secondary market frictions now have larger effects on credit spreads than in 2005. We provide a model connecting this to the rapid growing mutual fund shares in the corporate bond market. The model features investors with different trading needs, choosing among a risk-free asset and heterogeneous illiquid bonds. As the risk-free rate declines, short-term investors enter the bond market, reaching for yields. Although they provide liquidity, their greater trading needs amplify the sensitivity of credit yields to bid-ask spreads, leading to larger liquidity components. We test the model's predictions using U.S. investor holdings data and find consistent evidence.

We thank Zhiguo He, Ralph Koijen, Douglas Diamond, Lars Peter Hansen, and Anil Kashyap for their invaluable guidance. We are grateful for discussions with many people, including Eliot Abrams, Markus Brunnermeier, Wenxin Du, John Heaton, Moritz Lenel, Yueran Ma, Stefan Nagel, Lubos Pastor, Anna Pavlova, Jonanthan Payne, Batchimeg Sambalaibat, Wei Xiong and Motohiro Yogo. We thank seminar participants at National University of Singapore, University of Washington Foster, Tsinghua University Graduate School of People's Bank of China, Purdue University Krannert, UCSD Rady, Rice University Jones, Atlanta Fed, Indiana University Kelley, HKU, Yale SOM, Bank of Canada, University of Utah David Eccles, Texas A&M Mays, Harvard Business School, Chicago Fed, Columbia Business School, MIT Sloan, Wisconsin School of Business, Arizona State University Carey, Dallas Fed, UCLA Anderson, Richmond Fed, Wharton, Copenhagen Business School and NYU Stern for helpful comments.

*Columbia Business School. Email: jl5964@columbia.edu. Corresponding author.

†Department of Economics, Princeton University. Email: haiyuey@princeton.edu.

Conflict-of-interest disclosure statement

Jian Li
I have nothing to disclose.

Haiyue Yu
I have nothing to disclose.

1 Introduction

Corporate bond markets are a crucial financing channel for firms in the U.S. and worldwide. The corporate bond market disruption in March 2020 certainly suggests that the market has become more fragile compared to a decade ago. However, it has been difficult to find any measures indicating such change in the market. Preceding this event, bid-ask spreads and other standard measures implied that liquidity was weakly improving.[1] In contrast to the existing literature that focuses on measuring the level of liquidity, we take a different perspective and study the sensitivity of credit yields to liquidity and find that the sensitivity has increased fourfold over the past 15 years. Despite the trend in the level of liquidity, small differences in bid-ask spreads are now associated with much larger differences in credit yields. As a result, liquidity components—bid-ask spreads multiplied by the sensitivity—are now a significantly larger fraction of credit yields. A direct implication of this finding is that credit spreads, as well as firms' financing cost, are now more sensitive to secondary market frictions.

We explore to what extent the changes in the sensitivity coefficient can be explained by changes in investor composition over the past decade. The U.S. corporate bond market has indeed experienced large shifts in investor composition since the early 2000s. As shown in Figure 1, the share of mutual funds has grown from less than 8% during the early 2000s to more than 20% today.[2] Given the significant difference in trading patterns and the massive change in the market composition, it is important to understand how this trend affects the bond markets.

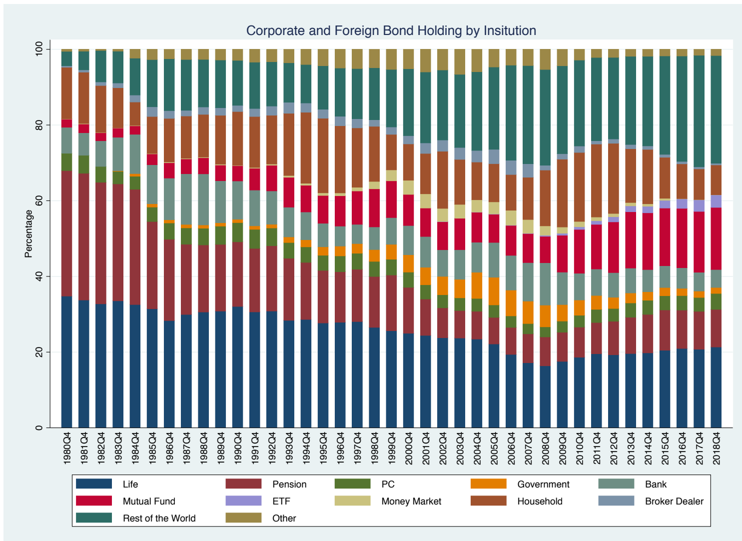
There are many institutional differences among corporate bond investors. We focus on one of the most prominent differences: trading frequencies. Mutual funds typically have much higher trading frequencies than insurance companies and pension funds. Insurance companies and pension funds are mainly viewed as long-term investors with low turnover rates.[3] We develop a model with heterogeneous investors. We show that as the risk-free

[1] Adrian, Fleming, Shachar and Vogt. "Has U.S. Corporate Bond Market Liquidity Deteriorated?" *Liberty Street Economics*, Oct 5, 2015; Adrian, Fleming, Vogt and Wojtowicz. "Corporate Bond Market Liquidity Redux: More Price-Based Evidence. *Liberty Street Economics*, Fed 9, 2016; Adrian, Fleming, Vogt and Wojtowicz. "Further Analysis of Corporate Bond Market Liquidity. *Liberty Street Economics*, Fed 10, 2016.

[2] Many papers have documented this increase in mutual funds, particularly in the post-crisis period (see e.g. Goldstein, Jiang, and Ng, 2017; Feroli, Kashyap, Schoenholtz, and Shin, 2014).

[3] We estimate the average quarterly turnover of mutual funds is 0.15-0.2, compared with the 0.05 turnover

Figure 1: Corporate and Foreign Bond Market Breakdowns



Source: Flow of Funds. This figure plots quarterly share of holding of corporate and foreign bonds by investor type from the flow of funds. The sample period is from 1990Q1 to 2018Q4. "Life" stands for life insurance companies and "PC" stands for property and casualty insurance companies.

rate declines, more short-term investors, such as mutual funds, reach for yields and enter the illiquid bond market. While more market participants increase liquidity and lower the bid-ask spreads, the short-term investors, because of their higher trading needs, amplify the effect of bid-ask spreads on credit yields. We then test the model's predictions in the cross section of bonds using detailed quarterly investor holdings data in the U.S. and aggregate international data, and find consistent evidence.

We establish our motivating fact by considering a cross-sectional regression of credit yields on bid-ask spreads, controlling for firm and bond characteristics. The sensitivity of credit yields to bid-ask spreads (the regression coefficient) has grown four times in the last 15 years. Before the 2008 global financial crisis, a one standard deviation in bid-ask spreads is associated with a 13 basis points (bps) difference in credit yields. However, in 2019, a one standard deviation in bid-ask spreads is associated with more than a 50 bps difference in credit yields. As a result, the median liquidity component, defined as the sensitivity coefficient multiplied by its bid-ask spreads, has grown from 5% of the credit spread at the beginning of 2005 to more than 25% of the credit spread at the end of 2019. This trend

rate of insurance companies.

holds for both investment-grade and high-yield bonds.

Motivated by the large turnover differences among investors, we build a model featuring a continuum of heterogeneous investors choosing among a continuum of illiquid corporate bonds, together with a risk-free liquid asset. Investors differ in their frequency of liquidity shocks, which generally come from the financial institutions' liability side. For example, insurance companies may need to make large payouts on existing policies; similarly, mutual funds may face sudden inflows and outflows.[4] Corporate bonds are heterogeneous in their maturities and default probabilities. This dimension of bond heterogeneity allows us to derive and test cross-sectional predictions.

Each period, a certain measure of new investors enter the economy as potential buyers of the bonds. Each bond has a separate sub-market with different liquidity conditions and prices. After observing the menu of sub-markets, new investors decide which sub-market to enter and engage in a search process to find a seller. Once a match is found, trade happens, and part of the trade surplus is taken away by a dealer in the background as the bid-ask spreads. We assume bid-ask spreads are increasing in the ratio of sellers to buyers in the market. The rationale is that when there are more sellers, the dealers have to hold the bonds on their balance sheet for longer periods of time. Because of either regulatory constraints or inventory costs, dealers would charge higher bid-ask spreads to compensate.[5] Bondholders experience liquidity shocks at different frequencies. Once a liquidity shock hits, the bondholder becomes a seller and starts searching for buyers. After selling the bond, the investor leaves the economy.

Our framework corresponds to a directed search setup. Relative to a random search, adopting this approach has two benefits. First, a directed search framework allows us to derive rich cross-sectional predictions, which can later be tested in the data. This approach is also more realistic, because, in practice, most investors know which types of bonds they want to buy or sell; then they contact dealers to find out about the supply or demand. The

[4]Although some of the mutual fund flows could be endogenous to bond market conditions, in normal times, the outflows are mostly a result of end investors' idiosyncratic liquidity needs. Here, we take a reduced-form approach in modeling how the liability sides affect their trading behavior.

[5]We do not model dealers explicitly in our baseline model. The bid-ask spreads capture their interaction with the investors. Intuitively, the bid-ask spreads are higher when the imbalance between the number of sellers and buyers is large. For the market and time period that we are studying, we believe it is more likely to be a buyer's market, in which case the bid-ask spread is increasing in the seller-buyer ratio. Appendix C micro-founds this assumption using dealer inventory cost and capacity constraints.

occurrence of actual trades may take time and could be uncertain; hence, we have search frictions within each sub-market.

We prove the equilibrium features a cutoff strategy: investors with liquidity shock frequencies below a certain threshold participate in the bond market, and others hold the liquid risk-free asset. Furthermore, short-term investors sort into sub-markets with short-term and risky bonds, whereas long-term investors sort into long-term and low-risk bonds. In addition to the sorting pattern, our model also predicts that the credit yields of bonds with more short-term investors are more sensitive to the bid-ask spreads. The correlation between the sensitivity coefficient and investor composition is partly a result of the heightened trading needs, but more importantly, it is because sub-markets with high bid-ask spreads are also sub-markets with high selling pressure. In the cross section, both reasons contribute to a positive cross-derivative of credit spreads with respect to bid-ask spreads and investor trading frequencies.

As the risk-free rate drops, the cutoff increases and more short-term investors participate in the illiquid bond market, reaching for yields. As more short-term investors enter, the seller-buyer ratio reduces, which lowers the bid-ask spreads in all sub-markets. However, because short-term investors have higher trading needs, the transaction costs are encountered more times before bonds mature, and the credit spreads become more sensitive to the bid-ask spreads.

The mechanism in our model is fundamentally different from that in Amihud and Mendelson (1986), where bid-ask spreads are exogenously fixed and investors choose which sub-market to join based on the bid-ask spreads. In our model, investors sort along bond characteristics such as maturity and default probability. Bid-ask spreads are in turn endogenously determined by the bond type and the type of investors participating in the sub-markets. Quantitatively, we argue that a model with exogenous bid-ask spreads cannot rationalize the large change in the sensitivity coefficient given the change in the average investor turnover rate. In a calibration exercise, we show our model, which endogenizes bid-ask spreads, can match the change in the liquidity component reasonably well.

We then test the model's predictions in the cross section of bonds using detailed quarterly investor holdings data in the U.S.[6] We measure investor turnover in any given quarter as

[6] We obtain investor information from eMaxx special reports and standard bond information from WRDS, Mergent FISD and FINRA's Trade Reporting and Compliance Engine (TRACE).

the lagged average of the net percentage change in its bond holdings over the last four quarters. We then take the value-weighted average and aggregate it to an "investor composition" measure at the bond level. Consistent with the model, yields of bonds held by short-term investors are more sensitive to differences in bid-ask spreads than bonds held by long-term investors. We sort bonds into five groups according to their average investor composition and regress credit spreads on bid-ask spreads group by group, controlling for bond and firm characteristics. We find that groups in which bonds have more short-term investors have higher coefficients in front of bid-ask spreads, and they also have larger median liquidity components as a fraction of credit spreads. This is perhaps the strongest evidence demonstrating the importance of investor heterogeneity. In a traditional homogeneous investor model, the realized difference in turnover is purely from noise. As a result, if we sort bonds based on investor turnover, we should find no difference in the sensitivity coefficients across bond groups. The positive correlation between the sensitivity coefficient and investor composition is a unique prediction of our heterogeneous investor model (relative to homogeneous investor models).

We also verify the empirical sorting patterns are consistent with the model's prediction. All time-series and cross-sectional results are robust to alternative measures of liquidity from the literature.[7] We use bid-ask spreads in the main paper because it is a direct empirical counterpart of what is in our model.

On the aggregate level, we find that countries that experienced larger declines in risk-free rates saw greater growth in mutual fund shares. Using data on fixed-income funds from Morningstar, we indeed find higher growth in mutual fund shares in bond markets is associated with a larger decline in risk-free rates over the same periods, controlling for macroeconomic conditions (GDP, inflation, and unemployment).

We conclude by calibrating the model to show it can account for the large changes in the U.S. data and by conducting two counterfactual experiments. We focus on the segment of bonds with maturity between 0.5 and 20 years. Model parameters are calibrated to match the size of the liquidity component in 2005, the average investor turnover rate, the sensitivity

[7] In particular, we try the measure from Dick-Nielsen, Feldhütter, and Lando (2012). They measure liquidity as a factor that loads evenly on various price measures (Roll's measure, Ahmed's measure, imputed round costs, and standard deviation of each) and on quantity measures (days with no trading for the firm and for the bond, bond turnover rate). Our bid-ask spread measure is highly correlated with this alternative liquidity measure. Details are explained in Appendix D.

coefficient of bid-ask spreads, and the variance of credit spreads and bid-ask spreads in 2005 and 2019. Given the change in the risk-free rate, our model can match the investor composition change, the growth in the liquidity component, and the size of the sensitivity coefficient reasonably well. Models in which bid-ask spreads are exogenous cannot match the size of the sensitivity coefficient quantitatively given the change in investor composition in the data.

In the first counterfactual exercise, we investigate the interactive effect of investor composition and recent dealer regulation changes. As part of the 2010 Dodd-Frank Wall Street Reform and Consumer Protection Act, the Volcker Rule intends to limit banks' risk-taking activities. However, an unintended consequence of the rule could be diminished bond market-making activities (Duffie, 2012) because it raises the balance sheet costs for dealers to hold bonds in inventory. The question we seek to answer is whether changes in investor composition amplify or alleviate the regulation impact. While short-term investors have higher trading needs, they are also liquidity providers in the market. On net, we find that the entry of short-term investors partially masks the impact of dealer regulation changes, consistent with the empirical literature finding mixed results on the impact of the Volcker Rule. In the second counterfactual exercise, we use the calibration to evaluate changes over time in the market's response to aggregate liquidity shocks. We exogenously raise the seller-buyer ratio in each sub-market by 1% and examine how credit yields respond in the scenario of 2005 and 2019. In both the cross section and over time, sub-markets with more short-term investors experience sharper increases in credit yields because short-term investors with higher trading needs amplify the impact of secondary market frictions on credit yields. Our results shed light on why the corporate bond market experienced large disruptions in March 2020.

The remainder of the paper is organized as follows. We discuss the related literature below. Section 2 presents the new fact we identify in detail. Section 3 describes the environment of the model. Section 4 analyzes the simple one-bond case and the heterogeneous-bond case. Section 5 tests the implications from the full model, and Section 6 conducts a simple calibration exercise. Section 7 concludes.

1.1 Related Literature

A growing literature on the corporate bond market has shown that liquidity can explain a significant part of the common component in yield variations (Bao, Pan, and Wang, 2011; Dick-Nielsen, Feldhütter, and Lando, 2012; Friewald, Jankowitsch, and Subrahmanyam, 2013; Friewald and Nagler, 2019; Goldstein, Hotchkiss, and Pedersen, 2019). In light of the concern from market participants and regulators, many papers have studied how the level of liquidity has changed since the global financial crisis. The evidence is mixed (Anderson and Stulz, 2017). Adrian, Boyarchenko, and Shachar (2017) show that bonds that are actively traded by constrained institutions are less liquid, and such effect is higher for high-yield bonds. Bretscher, Schmid, Sen, and Sharma (2020) adopt a demand system approach to study the evolution of liquidity over time and find a higher price impact in the post-crisis period. Cost of immediacy may also have increased since the regulatory changes (Dick-Nielsen and Rossi, 2018). On the other hand, Adrian, Fleming, Shachar, and Vogt (2017) show limited evidence for liquidity deterioration. In one of the counterfactual exercises, we analyze the interaction effect of changes in investor composition and changes in regulation.

We focus on the sensitivity of credit yields to bid-ask spreads and how investor composition affects that sensitivity. In a contemporaneous working paper, Wu (2020) documents the same aggregate trend and examines whether this trend is caused by post-crisis changes in dealer regulations. We instead provide a model that explains this trend through observed changes in investor composition.[8] Amihud and Mendelson (1986) is one of the first papers to analyze how investor preference affects the impact of transaction costs in the equity market. They take transaction costs as exogenous and allow investors with different investment horizons to sort into assets with different bid-ask spreads. Apart from the difference in asset class studies, we endogenize bid-ask spreads as a function of the investor types. Instead of sorting along transaction costs, investors sort along characteristics that are more exogenous to the secondary market (maturities or default probabilities). This endogeneity turns out to be important for quantitatively matching the sensitivity coefficient in the data.

In light of the bond market disruption in March 2020, recent papers have investigated the

[8] In our framework, dealer regulation changes do not directly contribute to the increase in the sensitivity coefficient. That said, dealer regulation changes can have an indirect effect through altering the investor composition.

changes in liquidity during the COVID-19 crisis (see e.g. Kargar, Liu, Lester, Weill, Lindsay, and Zúñiga, 2020; O'Hara and Zhou, 2021). Ma, Xiao, and Zeng (2021) show that fixed-income mutual funds are important contributors to the high selling pressure. Consistent with their findings, in our last counterfactual exercise, we find that short-term investors amplify aggregate shocks in the sub-markets in which they dominate.

An alternative way to measure the liquidity component is to follow Longstaff, Mithal, and Neis (2005) and use the credit default swap (CDS) basis (Beber, Brandt, and Kavajecz, 2009; Oehmke and Zawadowski, 2016; Schwarz, 2018). Choi and Shachar (2014) and Choi, Shachar, and Shin (2019) discuss the relationship between the CDS basis and liquidity provision by dealers before and during the financial crisis. A recent paper documents that the CDS basis has become more negative (Bai and Collin-Dufresne, 2019), consistent with our finding that the liquidity component has been growing over time. We do not measure liquidity this way for two reasons: first, not all firms have an actively traded CDS, and second, CDS's themselves are not perfectly liquid; their pricing is heavily influenced by the funding situation of dealers.

Recent papers have documented empirical clientele effects in certain aspects of the corporate bond market. Chen, Huang, Sun, Yao, and Yu (2020) find that among insurance company investors, the different preferences for liquidity affect the credit spread of the bonds. Musto, Nini, and Schwarz (2018) show that long-horizon investors take advantage of the liquidity premium. We conduct a similar analysis with a focus on short-term investors such as mutual funds. In addition, the sorting of investors into different bonds is endogenous in our case.

In terms of the theoretical framework, our model accounts for both investor and bond heterogeneity. Vayanos and Wang (2007) introduce a model with heterogeneous investors and identical assets. They show that there exists a (separating) clientele equilibrium where all short-horizon investors search for the same asset. Our framework allows us to analyze sorting. Hugonnier, Lester, and Weill (2019) allow for arbitrary heterogeneity in dealers' valuations to study the intermediation process in over-the-counter markets. We focus on heterogeneity in the trading needs of final investors. Furthermore, in contract to the papers that use random research techniques to study over-the-counter markets (Duffie, Gârleanu, and Pedersen, 2005; Lagos and Rocheteau, 2007; Lagos, Rocheteau, and Weill, 2011; Hugonnier, Lester, and Weill, 2019), we adopt a direct search framework in the full model to capture investors' behaviors more realistically (Eeckhout and Kircher, 2010).

Other papers have examined the issue of illiquidity from a structural perspective. For example, Ericsson and Renault (2006) link liquidity to bankruptcy renegotiation. He and Milbradt (2014) focus on the interdependence between liquidity and default components in credit yields, and a subsequent paper quantifies the relevant channels (Chen, Cui, He, and Milbradt, 2017). In an environment with search frictions, Feldhütter (2012) shows that the difference between prices paid by large and small traders identifies liquidity crises. Our paper focuses on the relationship between liquidity and investor composition, abstracting away from the interplay between liquidity and default risks.

2 New Fact

In this section, we establish the fact that the sensitivity of credit yields to bid-ask spreads has been increasing since 2005, leading to an increase in the liquidity component from 5% to over 25%. Although the level of bid-ask spreads has been weakly decreasing, it has been playing a much more important role in determining the credit spreads.

2.1 Data

For the corporate bond characteristics, we use data from Mergent FISD. Following the literature, we focus on US corporate debentures with a fixed coupon rate that are non-convertible, non-puttable and non-exchangeable. From Mergent, we get information on bond age, rating, offering yield, maturity, and other bond features. In terms of rating, we mainly use the rating from Standard and Poor's (S&P).[9] We then merge the bond information with the Center for Research in Security Prices (CRSP) for equity price and with Compustat for firm accounting information.

In terms of the transaction data, we use the enhanced Trade Reporting and Compliance Engine (TRACE) maintained by the Financial Industry Regulatory Authority (FINRA). We filter the TRACE data following Dick-Nielsen (2014). Moreover, we also use the price filter to clean the extreme price observations in TRACE. We obtain the reported yield from TRACE and then calculate the credit spread by subtracting the yield of the corresponding Treasury security. We follow Gürkaynak, Sack, and Wright (2007) in calculating the yield curve. For

[9]If the rating information from S&P is missing, we use Moody's, and lastly Fitch.

Table 1: Summary Statistics at Bond-Quarter Level – All Bonds

	Mean	Min	Max	Std
Bid-ask spread (WRDS)	.0053561	0	.0965653	.0050625
Credit spread	0.021	0.002	0.182	0.023
Quarterly transaction volume (million USD)	156.5	0.58	10280	275.4
The par value of debt (USD)	685274	6500	15000000	614380
Time to Maturity (30/360 Convention)	9.131	0.003	99.962	8.991
Age	4.044	0.008	60.518	3.695
Annual interest rate (%)	5.377	0.000	15.000	1.943
Observations	174739			

This table presents bond-quarter level summary statistics. Bond data is obtained from WRDS and Mergent FISD. We include only US Corporate debenture with fixed coupon rate, non-convertible, non-puttable, non-exchangeable. Sample period is from 2005Q2 to 2019Q2.

each bond, we use the average credit spread on the last trading day in each quarter as the dependent variable $CS_{i,t}$. To avoid extreme observations, the credit spreads are winsorized at 0.5% and 99.5%. We drop the bonds that have no trades in the last month of the quarter.

We use the quarterly bid-ask spreads calculated by WRDS directly for the analysis in this section. The WRDS bid-ask spread is calculated as the difference between the volume-weighted customer sell prices and buy prices, divided by the average of the two.

Our sample period is from 2005Q2 to 2019Q2, and the sample contains 1,824 firms and 15,567 bonds. Table 1 presents the summary statistics for our bond universe. The average bid-ask spread is 53 bps, with a standard deviation of around 50 bps. The average credit spread is 2.1%, with a standard deviation of around 2.3%.

2.2 Analysis and Results

We regress credit spreads on bid-ask spreads quarter by quarter, controlling for bond- and firm- level characteristics. Our baseline regression is shown in equation (2.1), where $CS_{i,t}$ denotes the credit spreads and $BA_{i,t}$ denotes the bid-ask spread for bond i in quarter t. The control vector $\mathbf{X_{i,t}}$ includes bond age, time to maturity, couple rate, offering amount, and ratings at the bond level, and leverage, size, profitability, equity volatility, and total asset value at the firm level. We also add industry fixed effects at the four digit level and control for the level and slope of the Treasury yield curve:

$$CS_{it} = \alpha_t + \beta_t BA_{it} + bond\ characteristics_{i,t} + \gamma_{\mathbf{t}}^{\top}\mathbf{X_{i,t}} + \epsilon_{i,t}. \tag{2.1}$$

The quarterly coefficients in front of the bid-ask spreads β_t are plotted in Figure 2a. The shaded region shows the 95% confidence interval for the estimates over time. The pattern shown in the introduction holds in the general bond universe under robust controls: the coefficient in front of bid-ask spreads has been increasing in the post-crisis period. More specifically, in 2005-2007, a 100 bps difference in bid-ask spreads is associated with a 20-30 bps difference in credit yields, whereas it is associated with a 1% difference in credit yields in 2019. The increase in the sensitivity coefficient is not caused by changes in the bid-ask spread variation—the same trend holds true if we instead plot the correlation coefficient between credit spreads and bid-ask spreads over time,[10] see Figure 2b. Although our coefficients are not causal estimates of how changes in bid-ask spreads affect credit yields, they are interesting moments that our model will try to match.

Several concerns could be raised regarding the robustness of our result. First of all, liquidity is difficult to measure. We show that our result is robust to using other liquidity measures proposed in the literature. The pattern is qualitatively similar if we use our own calculation of bid-ask spreads. The difference is that we exclude inter-dealer trades, agency trades, and retail trades.[11] We also find a similar result using an implied round-trip cost[12] and the illiquidity measure used in Dick-Nielsen, Feldhütter, and Lando (2012).

Another concern is that the change in relationship is driven by changes in default risk or the risk appetite of investors or both. To better control for default risk, we use a subset of bonds with credit default swaps and control for CDS spreads on the right-hand side. Our result is robust to this specification. Details are shown in Appendix B. Lastly, the

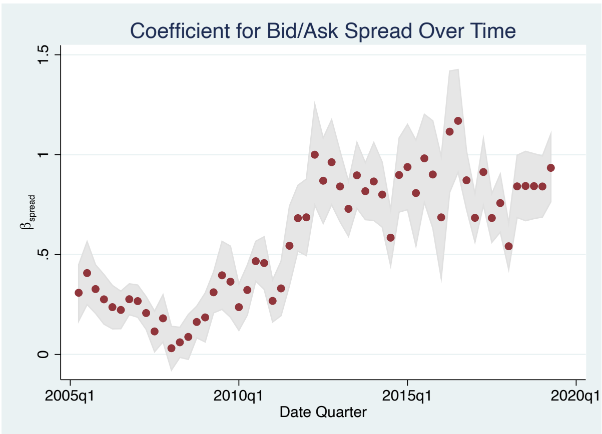
[10]To calculate the correlation coefficient, we regress credit spreads and bid-ask spreads on bond and firm characteristics respectively, quarter by quarter. We then calculate the correlation coefficient between the two residuals in each quarter.

[11]For each day, and for each bond, we first calculate the daily bid-ask spread: we take the weighted average buy price minus the weighted average sell price divided by the average of the two. We exclude inter-dealer trades, agency trades and trades with a size smaller than 100,000. We then average it over a quarter.

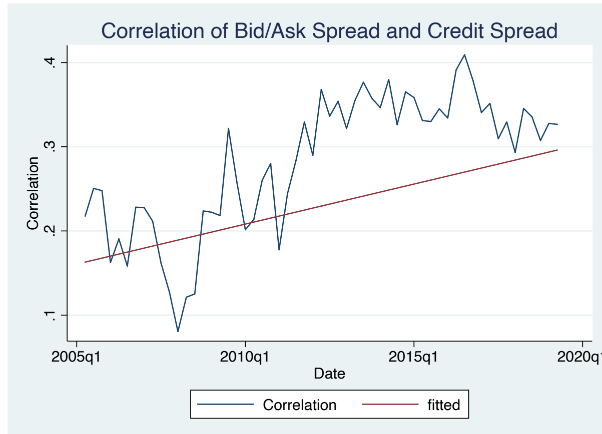
[12]We identify matched trades defined as the same bonds of the same size traded in the same day. We then take the weighted average buy price minus the weighted average sell price divided by the average of the two. We then take the average over a quarter.

Figure 2: Relationship between Credit Spreads and Bid-ask Spreads

(a) Regression Coefficient



(b) Correlation Coefficient



We regress credit spreads on bid-ask spreads quarter by quarter for all bonds, controlling for bond and firm characteristics. We then plot β_t over time in the left-hand-side panel. The shaded region indicates the 95% confidence interval. For the right-hand-side panel, we regress credit spreads and bid-ask spreads on bond and firm characteristics quarter by quarter for all bonds. We then calculate the correlation coefficient for the residuals and plot it over time. Data comes from WRDS Bond Return and Mergent FISD.

constitution of bonds may have changed over this time period. To partially address this concern, we also repeat our analysis including the bond fixed effect, and our results survive the additional test.

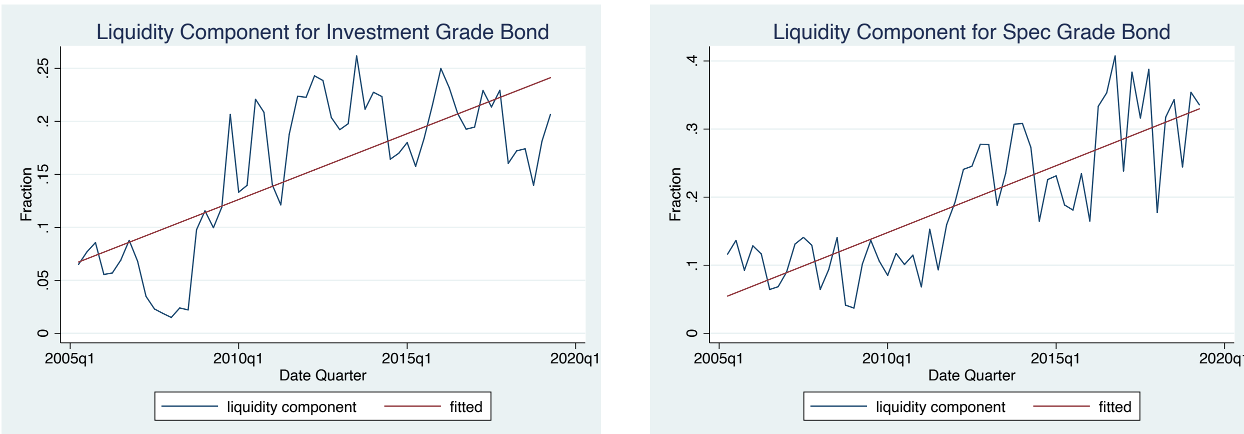
The finding on the coefficient naturally leads us to wonder how the liquidity component has changed over time. Following the literature, we calculate the liquidity component as

$$\text{liquidity component}_{i,t} = \frac{\beta_t \times BA_{it}}{CS_{it}}. \tag{2.2}$$

We plot the cross-sectional median over time, separately for investment-grade and high-yield bonds. The results are shown in Figure 3a and Figure 3b, respectively. The main contributing factor for the increase in the liquidity component is the growing β_t coefficient, given that the level of bid-ask spreads has weakly declined. In light of this new finding, the rest of the paper builds a model to explain this fact, linking it to the growth of mutual fund shares in the corporate bond market.

Figure 3: Liquidity Component as a Fraction of Credit Spreads

(a) Investment-grade Bonds (b) High-Yield Bonds



The figure plots the liquidity components of investment-grade bonds and high-yield bonds over time. For each quarter t, we first run 2.1 to get the coefficient of bid-ask spread β. Bond characteristic data is from Mergent FISD and WRDS Bond Returns, the price data is from TRACE. Next for each bond, define $liquidity\_component_{i,t} = \frac{\beta_t \times bid\_ask_{it}}{credit\_spread_{it}}$. The top figure plots the median liquidity component of investment-grade bonds over time and the bottom figure plots the corresponding trend for high-yield bonds. The red line is a fitted linear trend over time.

3 Model Environment

In this section, we describe the model environment in detail. Our model features heterogeneous investors who are endogenously matched with different assets (bonds). Investors differ in the frequency of liquidity shocks they receive, and bonds differ in their maturities or default probabilities (or both). Bonds are illiquid in the secondary market because of search frictions and transaction costs (bid-ask spreads). Once a bondholder receives a liquidity shock, she will try to sell the bond in the secondary market, which involves searching and also paying a bid-ask spread to the dealer intermediating the transaction. We use the model to analyze the endogenous investor composition of different bonds and how the pattern changes as the risk-free rate declines.

3.1 Setup

Time is continuous and indexed by $t \geq 0$. There are entrepreneurs and investors. Both types of agents are infinitely lived and risk neutral. Entrepreneurs have discount rate ρ_e and can start projects at $t = 0$. Each project requires one unit of investment initially and generates a perpetual cash flow of $x > 0$.

Corporate Bonds: Entrepreneurs borrow on the corporate bond market to fund their projects. There are N types of bonds, indexed by i, $i = \{1, 2...N\}$. Each bond has face value $F = 1$ with flow coupon payment r_i, which is endogenously determined in equilibrium.[13] Type i bond matures with probability δ_i, and upon maturity, the firm refinances by issuing bonds with the same characteristics. The default of bond i is modeled as a Poisson process, with intensity d_i. Upon default, investors can recover $0 \leq s_i < 1$ from bond i. We take bond characteristics δ_i, d_i, and s_i as exogenous, since our focus is on the secondary market.[14] We normalize the amount outstanding for all bonds to be 1. Throughout the analysis, the N bonds are ordered such that $\delta_1 + d_1 \leq \delta_2 + d_2 \leq ... \leq \delta_N + d_N$.

Investors: Each period, a measure m_I of heterogeneous investors enter the economy. When investors enter, they are patient with discount rate ρ. Investors can choose which corporate bonds to hold or just hold the risk-free asset. For simplicity, investors are restricted to holding at most one unit of asset at any time. Because investors are risk neutral, they choose to hold either one unit or zero units. Investor j who holds a corporate bond is subject to idiosyncratic liquidity shocks that arrive at Poisson rate θ_j. The liquidity shock frequency $\theta_j \in [\underline{\theta}, +\infty)$ is heterogeneous among investors, with cumulative distribution function $F(\cdot)$ and probability distribution function $f(\cdot)$. It is a permanent feature of the investor and is observed by everyone in the marketplace. Once hit by the liquidity shock, patient investors irreversibly become impatient and value the coupon payment Δ units less than before.[15] So the effective flow payment to an impatient bondholder is $r_i - \Delta$.[16] We assume that Δ is homogeneous for all bonds, but that assumption can easily be extended to allow for heterogeneity as well.

We view these liquidity shocks as mostly coming from the liability side of the investors

[13] We assume x is large enough such that $r_i < x$, so firms do not face liquidity issues themselves.

[14] Thinking through the transmission to the primary market is also an interesting exercise, but we leave it out of this paper.

[15] The model can be generalized to the case in which impatient investors become patient again with some probability. We can view some of the new patient investors here as existing impatient investors reverting back to being patient.

[16] This specification follows the over-the-counter literature led by the seminal work of Duffie, Gârleanu, and Pedersen (2005). Duffie, Gârleanu, and Pedersen (2007) show this is the limiting case of risk-averse investors as their risk-aversion coefficient approaches 0.

and may be exogenous to what is happening in the corporate bond market. In the case of insurance companies, they could come from policy payouts or portfolio adjustments to satisfy the regulatory requirements. For mutual funds, they could come from inflows, outflows, or index weight adjustments. However, for active mutual funds, their trades are often reactions to price adjustments in the market, as some of the inflow/outflows of certain passive mutual funds. Since our trading motive is exogenously specified, we cannot capture this behavior accurately, and we acknowledge this limitation of the model. Nonetheless, our key mechanism remains valid as long as differences exist in trading frequencies among investors, regardless of what the actual trading motives are.

Bond Markets: Bonds have a primary and a secondary market. In the *primary market*, bonds are placed to investors through a centralized auction in which all investors can participate and bid for the bonds. We introduce the primary market in order to endogenize the offering yield of the bonds.

Our focus is on the *secondary market*. In the secondary market, bondholders who become impatient (the sellers) search and trade with patient investors who do not have any bonds (the buyers). Since the shock is irreversible, once investors are hit by liquidity shocks, they become homogeneous except for the bonds they hold. So all sellers in the secondary market are heterogeneous only to the extent that they hold different types of bonds. We denote the measure of sellers holding bond i by $\alpha_{s,i}$. Buyers are heterogeneous in their expected liquidity shock frequencies, and we denote the measure of buyers with θ_j probability of liquidity shock as $\alpha_{b,j}$ or $\alpha_b(\theta_j)$. These measures are determined endogenously in equilibrium. In the searching process, the flow measure of matches is $m(\alpha_{s,i}, \alpha_{b,j})$, where $m_1 \geq 0$, $m_2 \geq 0$, $m(0, \cdot) = 0$, $m(\dot{;}0) = 0$, and is continuously differentiable in both arguments.

Each transaction includes a bid-ask spread, denoted as ξ, which captures the surplus taken by dealers during the transactions. Our model focuses on the trading needs of final investors and abstracts away from modeling the intermediate dealers explicitly. However, from the existing literature, we know the intermediary sector matters significantly for how well the bond market functions (He, Khorrami, and Song, 2020). Having the bid-ask spread ξ_i is our way of capturing the influence of that sector. We endogenize the bid-ask spread for each bond i as a function of the measures of sellers and buyers in that sub-market. The key assumption is laid out in Assumption 1 (i.e., the bid-ask spread is increasing in the seller-buyer ratio).

Assumption 1. *The bid-ask spread increases in the seller-buyer ratio:*

$$\xi = \xi_0 + \xi_1 g(\frac{\alpha_s}{\alpha_b}) \qquad g' > 0 \quad \xi_0, \xi_1 > 0.$$

The validity of this assumption warrants more discussion. At least two ways can be used to justify the bid-ask spreads increasing in the seller-buyer ratio. The first one is to assume that the dealer balance sheet capacity is fixed. In this case, when there is high selling pressure, the dealers with fixed capacity can extract more rent from intermediating trades via charging higher bid-ask spreads. The second one is to consider that dealers incur an inventory cost when holding the bonds on their balance sheets. If the selling pressure is large, more bonds have to be absorbed by the dealer sector. A large bond inventory on dealers' balance sheets requires the dealers to post a significant amount of capital against the bonds. It also increases the burden of funding these illiquid assets. Part of the bid-ask spreads is charged to cover those costs. In Appendix C, we derive a model with the dealer sector explicitly. There are two sets of sub-markets: one in which dealers search and trade with sellers and another one in which dealers search and trade with buyers. The two sets of sub-markets run parallel with each other. We rule out short sales and assume dealers incur a flow cost that is proportional to the inventory they carry. The gaps in the prices between the two sets of markets are the bid-ask spreads. We verify our main results go through in the extended model. The extended model is more complicated and difficult to derive intuition from; hence, we present the simpler model in the main text.

Earlier work that models dealers' optimal pricing problem has shown that the bid-ask spreads should be symmetric in response to excess supply as well as to excess demand. For example, in Amihud and Mendelson (1980), the bid-ask spread is symmetric around a positive optimal inventory level. However, recent work has shown that dealers' holding of the corporate bonds is a tiny fraction of the overall bonds outstanding (Dick-Nielsen and Rossi, 2018; He, Khorrami, and Song, 2020). This is particularly true over the past decade. We take the small inventory of dealers as suggestive evidence that dealers' optimal inventory level is close to zero.

Furthermore, recent empirical work has found significant bid-ask-spread spikes during periods of high selling pressure, both in the global financial crisis and in the ongoing COVID-19 crisis (Dick-Nielsen, Feldhütter, and Lando, 2012; Di Maggio, Kermani, and Song, 2017;

Kargar, Liu, Lester, Weill, Lindsay, and Zúñiga, 2020). On the other hand, none of the bid-ask spreads during good times have any drop that is nearly as close in magnitudes as the spikes in crisis periods. Empirically, the fluctuations of bid-ask spreads are not symmetric.[17]

The price mechanism in the bond markets is as follows. Sellers post prices at which they are willing to sell their bonds. Sellers cannot discriminate against buyers based on their types. Buyers, observing the combination of bond types, prices offered, and queue length, decide which sub-market in which to participate. Trades happen at the posted price when a buyer and a seller meet. As before, we assume there is a bid-ask spread (transaction cost) ξ between the price that the seller receives and the price that the buyer pays.

Risk-free Asset: As mentioned before, the investors can also hold the risk-free liquid asset. The trading of such risk-free asset is frictionless. In reality, this risk-free asset should be an outside option that mutual funds use as a liquidity buffer. In normal times, this asset could be Treasury bonds. However, during crises, even Treasury bonds could become illiquid, in which case we should think of this asset as cash or three-month Treasury bills. The return on this asset is what we call the risk-free rate and is denoted by r_f. Many papers have explored factors that affect the risk-free rates, such as international capital flows and demographic changes (Bernanke, Bertaut, DeMarco, and Kamin, 2011; Hamilton, Harris, Hatzius, and West, 2016; Gagnon, Johannsen, and Lopez-Salido, 2016; Rachel and Smith, 2017). Since these factors are outside of our model, we take the risk-free rate as exogenous. The value to the investors from holding this liquid asset is simply $V_f = \frac{r_f}{\rho}$.

3.2 Investors' Value Functions

Denote $V_{0,i}(\theta)$ as the value of a patient investor θ holding bond i, and $V_{s,i}$ as the value of an impatient investor holding bond i. We use $V_{b,i}(\theta)$ to denote the value that a patient investor θ attaches to searching in bond i's secondary market.

Only when an investor becomes impatient does she start to search for a buyer. When a seller meets a buyer, the surplus for the seller is $P_{s,i} - V_{s,i}$, where $P_{s,i}$ is the price that the

[17] In other models, when dealers' inventory is high, the levels of the bid price and ask price both increase by the same amount, such that the bid-ask spread does not change (Ho and Stoll, 1983). The key assumption is that dealers' risk aversion or per-unit inventory cost is constant in the level of inventory. If risk aversion is increasing in the amount of risky bond held, or if the total inventory cost is convex, the bid-ask spreads will be increasing in the level of inventory, consistent with Assumption 1.

seller receives. The surplus for the buyer is $V_{0,i} - P_{b,i} - V_{b,i}$, where $P_{b,i}$ is the price that the buyer pays. Lastly, the bid-ask spread is defined as $\xi_i \equiv P_{b,i} - P_{s,i}$.

The value function of a patient investor θ holding bond i is

$$\begin{aligned} V_{0,i}(\theta) =& E\Big[\int_0^{\min\{t_1,t_2,t_3\}} e^{-\rho\tau} r_i d\tau + \mathbf{1}_{t_1 \le \min\{t_2,t_3\}} e^{-\rho t_1} + \mathbf{1}_{t_3 \le \min\{t_2,t_1\}} e^{-\rho t_3} s_i \\ & + \mathbf{1}_{t_2 < \min\{t_1,t_3\}} e^{-\rho t_2} V_{s,i}\Big], \end{aligned}$$

where t_1 is the time when the bond matures, t_2 is the time when the agent becomes impatient, and t_3 is the time when the bond defaults.

The value of an impatient holder is

$$\begin{aligned} V_{s,i} =& E\Big[\int_0^{\min\{t_1,t_3,t_4\}} e^{-\rho\tau} r_i d\tau + \mathbf{1}_{t_1 \le \min\{t_3,t_4\}} e^{-\rho t_1} + \mathbf{1}_{t_3 \le \min\{t_1,t_4\}} e^{-\rho t_3} s_i \\ & + e^{-\rho t_4} \mathbf{1}_{t_4 < \min\{t_1,t_3\}} P_{s,i}(\theta)\Big], \end{aligned}$$

where t_4 is the time when the seller finds a buyer.

The value of a patient buyer in sub-market i is

$$V_{b,i} = E\big[e^{-\rho t_5}(V_{0,i}(\theta) - P_{b,i})\big],$$

where t_5 is the time when the buyer meets a seller.

Investors choose which bond market to participate in or choose to hold the risk-free asset. If they choose to participate in the bond market, they need to decide whether to participate in the primary market or secondary market. The value of participating in bond i's primary market is $V_{0,i} - F$, where F is the value that successful bidders pay the firm. Hence, the investors' problem is summarized as follows:

$$\max\{\quad \max_{1 \le i \le N}\{V_{0,i} - F, V_{b,i}\}, \quad V_f\}.$$

We study the steady state throughout the analysis. We focus on $N \to \infty$. This corresponds to a directed search setup with a continuum of heterogeneous bonds and a continuum of heterogeneous investors.

4 Analysis

In this section, we solve the equilibrium and verify that it is consistent with the time-series pattern we have identified. We then derive a set of cross-sectional predictions that can be tested further in the data. The model allows us to look at the exact empirical counterpart of the variables that we have constructed in the empirical section. We show endogenizing liquidity as a function of investor composition is important for rationalizing the empirical findings quantitatively.

The first result is the equilibrium sorting patterns between investor types and bond types: short-term investors are matched with short-term and high-default-probability bonds. In addition, sub-markets with short-term investors are associated with lower bid-ask spreads. Second, we show in the cross section, the credit yields of bonds held by short-term investors are more sensitive to bid-ask spreads. Lastly, when the risk-free rate decreases, more short-term investors enter the bond market, which simultaneously drives down the bid-ask spreads and drives up the sensitivity of credit spreads to bid-ask spreads.

4.1 Equilibrium

For each market, denote the seller-buyer ratio as $\lambda \in [0, \infty]$, which can be interpreted as the queue length. We assume the meeting intensity is only a function of λ. As before, denote $\mu_b(\lambda)$ as the probability that a buyer meets a seller. More sellers means a buyer is more likely to meet a seller, so μ_b is increasing. On the other hand, the probability that a seller meets a buyer μ_s is decreasing in the seller-buyer ratio. We adopt the most commonly used Cobb-Douglas matching function, which implies the meeting intensities take the following form:

$$\mu_b(\lambda) = \eta \lambda^{\gamma} \qquad \mu_s(\lambda) = \eta \lambda^{\gamma - 1},$$

where $\eta > 0$ and $\gamma \in (0, 1)$.

The seller's problem becomes

$$\rho V_{s,j} = r_j - \Delta + d_j(s_j - V_{s,j}) + \delta_j(1 - V_{s,j}) + \max_{\lambda, p_s} \mu_s(\lambda)(p_s - V_{s,j})$$

subject to the constraint that λ is consistent with the number of buyers choosing to participate in this market. The buyer's problem is

$$\rho V_b(\theta) = \max_{(p_b,\lambda,(\delta,d)_j)\in G} \mu_b(\lambda)(V_{0,j}(\theta) - V_b(\theta) - p_b),$$

where G is the support for all of the available combinations of $(p_b, \lambda, (\delta, d)_j)$. Furthermore,

$$\rho V_{0,j}(\theta) = r_j + d_j(s_j - V_{0,j}) + \delta_j(1 - V_{0,j}) + \theta(V_{s,j} - V_{0,j}).$$

The bid-ask spreads by definition equal the gaps between prices: $p_b - p_s = \xi$. Lastly, in equilibrium, each bond sub-market is matched with one type of investor, and the offering yield is determined by $V_{0,i} - F = V_{b,i}$ (i.e., buyers should be indifferent between purchasing from the secondary market and purchasing from the primary market).

Denote $U(\theta)$ as the maximum value that a buyer of type θ can get across all the sub-markets. Substituting in the prices with $U(\theta)$, we can rewrite the seller's problem as

$$\max_{\theta,\lambda} \quad \frac{\frac{\Delta}{\rho+\delta+d+\theta} - \xi(\lambda) - U(\theta)(1 + \frac{\rho}{\mu_b(\lambda)})}{\frac{1}{\mu_s(\lambda)} + \frac{1}{\rho+\delta+d+\theta}}.$$

Sellers take buyers' outside option $U(\theta)$ as given (i.e., they take other sellers' postings as given) and choose which types of buyers to attract, understanding the cost they need to pay to do so. The term $\frac{\Delta}{\rho+\delta+d+\theta} - \xi(\lambda)$ is the total surplus created from the transaction; $U(\theta)(1 + \frac{\rho}{\mu_b})$ is the part that buyer type θ gets in order for her to come to this sub-market. The remaining part is what the seller gets, adjusted for the expected turnover rate.

Note that in this expression, $\delta_j + d_j$ enters as a whole. This implies that $\tilde{\delta}_j = \delta_j + d_j$ is a sufficient statistic for the sorting pattern. This is because we assumed all investors are risk neutral and the cost of default is the same for all investors. The main intuition for the sorting result is that long-term investors have a comparative advantage in holding long-term and low-default-intensity bonds. For a short-term investor, the *relative* benefit of the bond maturing is higher, and the *relative* cost of the bond defaulting is lower. Hence, short-term investors are sorted into bonds with short maturity and high default intensity.[18]

[18]We can extend the model so that different investors have different levels of aversion to default losses. For example, it might be more expensive for insurance companies to hold risky bonds compared with mutual funds because of regulatory restrictions. Incorporating such features may break the linear substitutability

Lemma 1 shows that when the transaction cost is not too large, buyers and sellers do not wait, and they trade with the first counterparty they meet. The reason is that investors endogenously choose to enter the bond market, so the buyer's benefit V_0 is ensured to be high enough so that the trade surplus between the seller and any type of buyer is always positive. Lemma 1 also indicates limited heterogeneity among the endogenous group of buyers.

Lemma 1. *When ξ_0 and ξ_1 are small, investors do not wait and always trade with the first counterparty they meet:*

$$V_0 - V_s - V_b - \xi \geq 0.$$

Proof: See Appendix A.1.

Although we cannot solve the equilibrium in full closed form, we can still characterize the equilibrium features under special cases. We present our first main result in Proposition 1. The proposition states sufficient conditions for positive assortative matching (i.e., in equilibrium, short-term investors buy short-term and high-default-intensity bonds).

Proposition 1. *When $\gamma \to 1$, and ξ_0 and ξ_1 are small relative to Δ, the equilibrium features positive assortative matching. Formally,*

$$\theta'(\tilde{\delta}) > 0.$$

Short-term investors are matched with short-term and high-default-probability bonds.

Proof: See Appendix A.3.

The conditions in Proposition 1 are sufficient but not necessary. Numerically, we obtain positive assortative matching for a wide set of parameter values away from the sufficient conditions. The intuition is the following: Short-term bonds are valuable because holders are less likely to sell them in the secondary market with discounts. Short-term investors, who have a higher likelihood of having this selling need, value this benefit of short-term

between d_j and δ_j — summary statistics for sorting $\tilde{\delta}(\delta, d)$ may be a non-linear function in the maturity rate and default rate. But our forces described here still exist; under reasonable assumptions, the additional feature even reinforces our current sorting pattern. Empirically, it is hard to separate the channel described in our model from the regulatory requirement channel. The fact that our sorting patterns hold within the mutual funds indicates that our mechanism is at least partially at work.

bonds more than long-term investors. Hence, in equilibrium, short-term investors are paired with short-term debt. The case for default probability is the other side of the same coin. All investors dislike bonds with high default intensity. However, if a bond does not default, a short-term investor is likely to have to sell it on the secondary market at a discount. In this sense, the relative cost of a bond defaulting is smaller for a short-term investor than for a long-term investor. This match-value complementarity implies that short-term investors sort into bonds with high default probabilities.[19]

Positive assortative matching is not always true because search frictions push for the opposite direction – negative assortative matching. All buyers and sellers value trade immediacy, especially the long-term investors and sellers with long-term bonds. A long-term buyer prefers to join sub-markets with high seller-buyer ratios, but that would imply long waiting times for the sellers. A seller with short-term bonds is more likely to provide this trade security because her opportunity cost for waiting is low. The value for trade immediacy pushes for negative assortative matching. The condition provided in Proposition 1 guarantees the match-value complementarity dominates.

After establishing positive assortative matching, we can characterize the equilibrium in a pair of ODEs. Similar to Eeckhout and Kircher (2010), we are looking for a pair of functions $\lambda(\tilde{\delta})$ and $\theta(\tilde{\delta})$ that solves the sellers' problem and buyers' problem while respecting the initial distributions of investor types and bond types. The equilibrium $\lambda(\tilde{\delta})$ and $\theta(\tilde{\delta})$ can be written as the solution to a pair of boundary value problems. Lemma 2 fully characterizes the system of ODEs.

Lemma 2. *The solution to the problem can be characterized by the following system of ordinary differential equations:*

$$\begin{cases} \theta'(\tilde{\delta}) = \frac{\theta m_I f(\theta)\mu_b + \tilde{\delta}\lambda(\mu_s+\tilde{\delta})(\theta+\tilde{\delta})}{(\mu_s+\tilde{\delta})(\tilde{\delta}+\theta)m_I f(\theta)\lambda} \\ \lambda'(\tilde{\delta}) = -\frac{\rho\gamma\lambda U'}{\rho\gamma U - \xi''\eta\lambda^\gamma - \xi'\eta\gamma\lambda^{\gamma-1}} + \frac{\frac{1-\gamma}{\theta'}\left[\frac{\Delta}{\mu_s} - U(1+\frac{\rho}{\mu_b}) - \xi\right]}{[\rho\gamma U - \xi''\eta\lambda^\gamma - \xi'\eta\gamma\lambda^{\gamma-1}]\left[\frac{\rho+\delta+d+\theta}{\mu_s}+1\right]} \end{cases}$$

[19] Even though the maturity rate and default intensity have the same implication for sorting, they do not have the same impact on interest rates. Everything else equal, bonds with higher maturity rates have lower interest rates, whereas the opposite is true for the case of default probabilities.

with boundary conditions

$$\theta(\underline{\delta} + \underline{d}) = \underline{\theta} \qquad \theta(\overline{\delta} + \overline{d}) = \overline{\theta} \qquad V_b(\overline{\theta}) = V_f.$$

Proof: see Appendix A.2.

The decision to enter the corporate bond market features a cutoff strategy. Investors with frequencies of liquidity shocks less than $\overline{\theta}$ enter the corporate bond market, and investors with frequencies of liquidity shocks above that threshold hold risk-free assets.

Our second result concerns the correlation between investor composition and liquidity across different sub-markets. Corollary 1 shows that under certain conditions, the bid-ask spreads are lower in the sub-markets populated with short-term investors.

Corollary 1. *When positive assortative matching is true, we have*

$$\lambda'(\tilde{\delta}) < 0 \qquad \xi'(\tilde{\delta}) < 0.$$

That is, the bid-ask spreads are lower in markets with more short-term investors.

Proof: See Appendix A.4.

When a buyer chooses which sub-market to join, she trades off value and trade immediacy. The match surplus is higher for long-term bond (small $\tilde{\delta}$) sub-markets; both the seller's gain and the buyer's gain from trade are larger. The parameter γ governs the sensitivity of trading immediacy to the seller-buyer ratio. When γ is close to 1, the waiting time is very sensitive to the seller-buyer ratio, which increases the marginal cost of moving to the long-term bond sub-markets. Thus, the seller-buyer ratio in the long-term bond market is smaller, leading to higher bid-ask spreads.

The mechanism that generates this correlation between transaction costs and investor composition is fundamentally different from that in Amihud and Mendelson (1986). They assume bid-ask spreads are exogenous and investors choose sub-markets based on the pre-specified bid-ask spreads, whereas in our model, bid-ask spreads are an endogenous outcome given the types of investors in each sub-market. Instead of sorting along transaction costs, in our model, investors sort along features that are relatively exogenous to the secondary market, such as maturity rate and default intensity.

This distinction is important for our later analysis. When considering certain regulation changes that affect investor composition, the liquidity and transaction costs in *all* markets would endogenously change, whereas in Amihud and Mendelson (1986), the transaction costs are fixed. In Section 6, we show investor composition in the sub-markets has important interactive effects with regulation changes in the dealer sector.

Next, we move on to examine what happens when the risk-free rate declines. Proposition 2 shows that investor composition in all sub-markets becomes more short term; liquidity in all sub-markets also improves.

Proposition 2. *As the risk-free rate declines,*

- *More short-term investors will enter the illiquid bond market, i.e.* $\frac{d\bar{\theta}}{dr_f} < 0$.
- *Each bond (indexed by* $\tilde{\delta}$*) will be matched with a shorter-term investor; each sub-market will have smaller bid-ask spreads, i.e.* $\frac{d\theta(\tilde{\delta};\bar{\theta})}{d\bar{\theta}} > 0$ *and* $\frac{d\lambda(\tilde{\delta};\bar{\theta})}{d\bar{\theta}} < 0$

Proof: See Appendix A.5.

Investors' outside option decreases as the risk-free rate drops, which induces more short-term investors to participate in the bond market, reaching for higher yields. In this heterogeneous-bond setting, more short-term investors means the cutoff threshold $\overline{\theta}$ increases. As more short-term investors enter the bond market, the bond with the shortest maturity and highest default probability (highest $\tilde{\delta}$) gains more short-term investors. Eventually, it spills over into all the other bonds' sub-markets.

As more short-term investors enter, at any point in time, both the number of sellers and the number of buyers increase. However, the increase in the number of buyers is larger than that of sellers. The reason is that every new entrant needs to purchase a bond, whereas only bondholders who are hit by liquidity shocks need to sell. As a result, the entry of more investors relaxes the selling pressure on the secondary market and hence reduces the bid-ask spreads. On net, the increase in the number of buyers relative to sellers leads to higher liquidity.[20] All sub-markets now have a lower seller-buyer ratio and lower bid-ask spreads. This result captures the liquidity provision role of mutual fund entrants.

[20]Note that this result holds in the new steady state instead of just the transitioning stage where only the number of buyers increases.

4.2 Credit Spreads

To analyze the response of credit spreads to the entry of short-term investors, we first solve for the equilibrium expression for interest rates in Lemma 3.

Lemma 3. *In equilibrium, the interest rate for firm j is given by*

$$r_j = \underbrace{\rho}_{discount\ rate} + \underbrace{d_j(1-s_j)}_{default\ spread} + (\rho+\delta_j+d_j)U(\theta_j) + \frac{\mu_s(\lambda_j)\theta_j\left[\frac{\Delta}{\mu_s(\lambda_j)} + U(\theta_j)(1+\frac{\rho}{\mu_b(\theta_j)}) + \xi_j\right]}{\mu_s(\lambda_j)+\rho+\delta_j+d_j+\theta_j}.$$

Proof: See Appendix A.6.

The first part of the interest rate is to compensate for investors' discount rate; the second part is to compensate for investors' loss in defaults. The last two parts are a result of sorting and search frictions in secondary markets.

Exogenous liquidity shocks: Our main goal is to analyze the changes in the sensitivity of credit spreads to bid-ask spreads as investor composition changes. We first analyze exogenous movement in the bid-ask spreads. Proposition 3 shows that the effect of bid-ask spreads on credit spreads increases as more short-term investors enter the bond market.

Proposition 3. *As $\gamma \to 1$, the sensitivity of the credit yield to the bid-ask spread can be simplified to*

$$\frac{dr_j}{d\xi_0} = \frac{\theta\mu_s(\lambda_j)}{\theta\mu_s(\lambda_j)+\rho+\delta_j+d_j+\theta}. \tag{4.1}$$

In markets with shorter-term investors, credit yields are more sensitive to the bid-ask spreads,

$$\frac{d\left|\frac{dr_j}{d\xi_0}\right|}{d\theta}. > 0 \tag{4.2}$$

Proof: See Appendix A.7.

This effect works through two channels: the first is that when investors have higher frequencies of liquidity shocks and hence higher trading needs, each bond is traded more times before maturity. Thus, the underlying frictions are amplified more and show up more significantly in prices. This finding is similar to Amihud and Mendelson (1986). The second channel is unique to our framework, which incorporates search frictions. When short-term investors enter, they also serve as liquidity providers to the market. As a result, the waiting time decreases and trading frequencies increase. This channel also amplifies the effect of transaction costs on credit spreads.

Endogenous liquidity: In our model, bid-ask spreads are endogenous to the type of bonds and the type of investors holding the bonds. This is also likely to be the case in the data. To take this endogeneity issue seriously and capture the exact empirical counterpart in the model, we run the same regression using model-simulated data as we do in the actual data. More precisely, for each risk-free rate rf_t, we resolve the model to get the bid-ask spread $BA_{i,t}$ $(\xi_{i,t})$ and credit spread $CS_{i,t}$ for each bond i. We then regress the credit spread on maturity rate $\delta_{i,t}$, default intensity $d_{i,t}$, and bid-ask spread $BA_{i,t}$:

$$CS_{i,t} = \beta_0 + \beta_\xi BA_{i,t} + \beta_1 \delta_{i,t} + \beta_2 d_{i,t} + \epsilon_{i,t}.$$

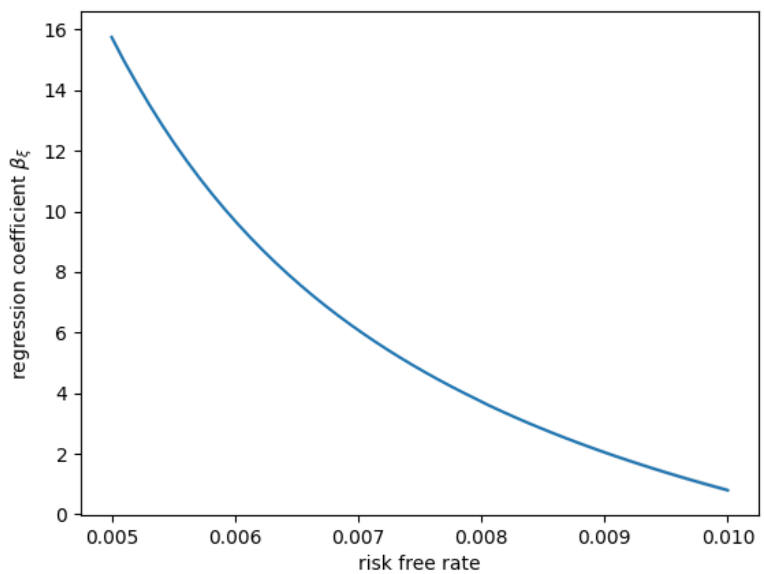
Then we get β_ξ for each given rf_t and compute the liquidity component in the same way as in Section 2. The goal is to see whether the model generates the same pattern as in the data.

Figure 4 shows how the sensitivity of credit spreads to bid-ask spreads changes with the risk-free rate. As the risk-free rate declines, more short-term investors enter the bond market, amplifying the sensitivity of credit yields to bid-ask spreads for all sub-markets. As a result, we see the coefficient rising as the risk-free rate drops. Correspondingly, even though bid-ask spreads are declining, the whole liquidity component is rising as more short-term investors enter the market. We are able to match qualitatively the aggregate trend that we identify in the data.

The pattern is similar when we use offering yield r instead of secondary market credit spreads. To unpack what is in the correlation between the bond interest rate and bid-ask spreads, we decompose the regression coefficient (with the offering yield as the left-hand-side variable) into two parts (after controlling for the maturity rate and default intensity):

$$\begin{aligned}\frac{r(\theta_j,\xi_j)-r(\theta_i,\xi_i)}{\xi_j-\xi_i} =& \frac{r(\theta_j,\xi_j)-r(\theta_j,\xi_i)+r(\theta_j,\xi_i)-r(\theta_i,\xi_i)}{\xi_j-\xi_i} \\ =& \underbrace{\frac{r(\theta_j,\xi_j)-r(\theta_j,\xi_i)}{\xi_j-\xi_i}}_{\phi_1(\theta_j)} + \underbrace{\frac{r(\theta_j,\xi_i)-r(\theta_i,\xi_i)}{\xi_j-\xi_i}}_{\phi_2}.\end{aligned}$$

Figure 4: Model Simulated Data– Over-time Regression Results



This figure illustrates comparative statics with respect to the risk free rate. The initial distribution of investor types is specified as a Pareto distribution, with the scale parameter equals to 0.0403 and the shape parameter equals to 0.1. For the rest of the parameters: $D = 1$, $m_I = 0.2$, $\Delta = 0.702$, $\rho = 0.009$, $\eta = 0.1135$, $\gamma = 0.9$, $\underline{\delta} + \underline{d} = 0.01$, $\overline{\delta} + \overline{d} = 0.25$. Lastly, the bid-ask spreads $\underset{\sim}{\xi}(\lambda) = 0.26 + 0.14\lambda^{1.84}$. The left hand side panel plots the matching between investor type θ and bond type $\tilde{\delta} = \delta + d$. High risk-free rate is specified as 0.01 and low-risk free rate is specified as 0.005. For each given risk-free rate, we solve the model and run the following regression: $r_i \sim constant + \beta_\xi \xi_i + bond\ characteristics$, bond characteristics include maturity and default probability. We then plot β_ξ for each risk-free rate. The shaded region represents 95% confidence interval.

The first part is the direct effect of bid-ask spread changes on credit yields:

$$\begin{aligned}\phi_1(\theta) &= \frac{\theta\rho\mu_b' - \frac{(\rho+\delta+d)\mu_b}{\rho+\mu_b}(\rho+\delta+d+\theta)(1-\gamma+\frac{\rho\mu_b'\lambda}{(\rho+\mu_b)\mu_b})}{(\rho+\delta+d+\theta)(\rho+\mu_b)(1-\gamma+\frac{\rho\mu_b'}{\rho+\mu_b}(\frac{1}{\rho+\delta+d+\theta}+\frac{\lambda}{\mu_b}))} \\ &< \frac{\theta}{1+\frac{\lambda}{\mu_b}(\rho+\delta+d+\theta)} < \theta.\end{aligned}$$

At the annual level, the average realized turnover increased from 0.32 to 0.37. A back-of-the-envelope calculation shows that θ is bounded above by 0.37, and hence the first direct impact is bounded in size and cannot match the magnitudes that we observe empirically (above 1 for investment-grade bonds and around 3 for high-yield bonds).

The second part ϕ_2 comes from the difference in investor composition across sub-markets and its impact on bid-ask spreads through endogenous seller-buyer ratios. Sub-markets with high bid-ask spreads are also sub-markets with high seller-buyer ratios, which also leads to higher yields. As a result, the second part is positive, $\phi_2 > 0$. Because of this endogenous relationship, the cross-sectional correlation between bid-ask spreads and interest rates is much stronger.

As the average investor turnover increases, the first part becomes larger (i.e., $\phi_1' > 0$). It is less obvious how ϕ_2 will change. The second part ϕ_2 is capturing the effect of the seller-buyer ratio on credit yields. As investors become more short term, the seller-buyer ratio and waiting time become more important, which leads to a higher value in ϕ_2. Although both ϕ_1 and ϕ_2 increase with the entry of more short-term investors, ϕ_2 accounts for the larger portion of the increase, whereas the change on the direct component ϕ_1 is quite limited. We show this in more detail through a calibration exercise in Section 6.

Remark 1. *While our model implies higher turnover rates for bonds over time, keeping everything else constant, it is no longer true when the amount outstanding for bonds is also increasing. The bond turnover rate is defined as*

$$turnover = \frac{transaction\ volume}{D}.$$

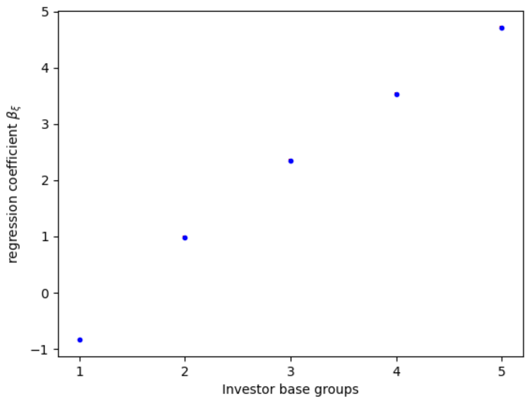
When D is increasing, bond turnover decreases for two reasons: (1) more investors acquire bonds from the primary market, so the transaction volume in the secondary market decreases (numerator is decreasing); (2) the total size of bonds is larger (denominator is increasing).[21] *Empirically, we do see higher transaction volume over time, but evidence on bonds' turnover rates has been mixed. Given that the total size of the bond market has been expanding massively over this period, the model's implication does not contradict the facts. Furthermore, if the amount outstanding increases, the model predicts a rising bid-ask spread, which is a counteracting force to the entry of short-term investors. This force could explain why we are not seeing a sharper decline in bid-ask spreads over time. However, the amount outstanding does not directly affect the sensitivity of credit yields to bid-ask spreads.*

Lastly, the same pattern holds true in the cross section as well: credit spreads are more sensitive to bid-ask spreads for bonds that have more short-term investors. We cannot prove

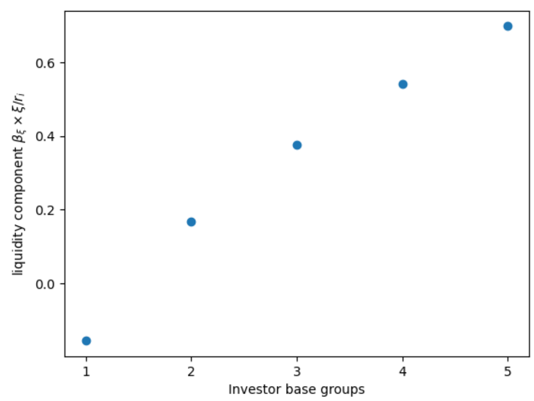
[21] A formal proof is available upon request.

Figure 5: Model Simulated Results

(a) Cross-sectional Regression



(b) Liquidity Component



This figure reports group regression results from model simulated data described in Section 4. The initial distribution of investor types is specified as a Pareto distribution, with the scale parameter equals to 0.0403 and the shape parameter equals to 0.1. For the rest of the parameters: $D = 1$, $m_I = 0.2$, $\Delta = 0.702$, $\rho = 0.009$, $\eta = 0.1135$, $\gamma = 0.9$, $\underline{\delta} + \underline{d} = 0.01$, $\overline{\delta} + \overline{d} = 0.25$. We divide the sub-markets into 5 groups, where group 1 contains bonds with the longest-term investors and group 5 contains bonds with the shortest-term investors. We run the following regression for each group: $credit\_spreads_i \sim \beta_0 + \beta_\xi \xi_i + \beta_1 \delta_i + \beta_2 d_i$ and plot β_ξ. The center point in the left-hand-side figure is the point estimate of the coefficient in front of ξ_i in each group, while the bands around them indicate the 1% confidence interval. We plot the median $\frac{\beta_i \xi_i}{credit\_spread_i}$ for each group in the right-hand-side.

this pattern for a general set of parameters in this model. Hence, we show this correlation numerically: in Figure 5a, we sort the bonds into five groups with an equal number of bonds in each group, indexed from 1 to 5. Group 1 bonds have the longest-term investors, and group 5 bonds have the shortest-term investors. We run the same regression group by group and plot the coefficient in Figure 5a. Bonds with more short-term investors have higher sensitivity coefficients. If we calculate the liquidity component by group, we also see higher liquidity components for groups of bonds with more short-term investors, as shown in Figure 5b. We show later that the empirical data yield similar patterns. This is an important prediction of the model with investor heterogeneity. If all investors are homogeneous, the realized difference in turnover would be purely due to noise. As a result, the sorting of bonds based on investor turnover would be equivalent to a random sort. We should observe no difference in the sensitivity coefficient of credit yields to bid-ask spreads. We show in Section 5 that we indeed observe the model-predicted relationship between the sensitivity coefficient and the average investor turnover, confirming the important role of investor heterogeneity.

5 Empirical Tests

In this section, we test the model predictions in the cross section using detailed investor holdings data for U.S. bonds and international fixed-income funds data. We first look at investor holdings and bond characteristics in the U.S. and then move on to the international evidence.

5.1 Bond-level Analysis

The model predicts that in equilibrium, short-term investors hold shorter-term risky bonds, and sub-markets endogenously have lower bid-ask spreads. Furthermore, in the sub-markets with more short-term investors, we should see a higher sensitivity of credit yields to bid-ask spreads. We use detailed investor holdings data from eMaxx to test the above predictions. We find sorting patterns between investor types and bond types consistent with the model prediction. The interaction term of investor composition and bid-ask spreads on credit yields is highly significant and positive, which implies short-term investors amplify the effect of transaction costs on credit yields. To further validate this, we sort bonds into groups according to their investor composition and find groups with a shorter-term investor composition have higher sensitivity coefficients, as our model predicts. Finally, we look at the issuance decision in the primary market and find that a shorter-term investor base in the previous four quarters is associated with shorter-term debt issued in this quarter. All of our results hold if we use an alternative measure of illiquidity, as in Dick-Nielsen, Feldhütter, and Lando (2012).

5.1.1 Data

As in Section 2, we use data from Mergent FISD and WRDS for bond characteristics. In addition to the filters applied in Section 2, we further restrict bonds whose eMaxx coverage exceeds 20% to ensure that we are capturing a significant portion of the bonds' investor base.

We use TRACE to calculate credit spreads and bid-ask spreads. Importantly, we drop the dealer-to-dealer trades because our model focuses on the transaction costs and liquidity conditions that end investors face. We also drop the retail trades (transactions with a face value less than 100,000 USD) because more and more retail trades happen on electronic trading platforms, where trades could happen directly between two customers.[22] Credit spreads are calculated as before.

To compute the bid-ask spread, we calculate the volume-weighted average customer buy price and customer sell price for each day and each bond. The bid-ask spread is defined as the difference between the two, normalized by the middle price. We then take the mean for the quarter as the bid-ask spread for that quarter.

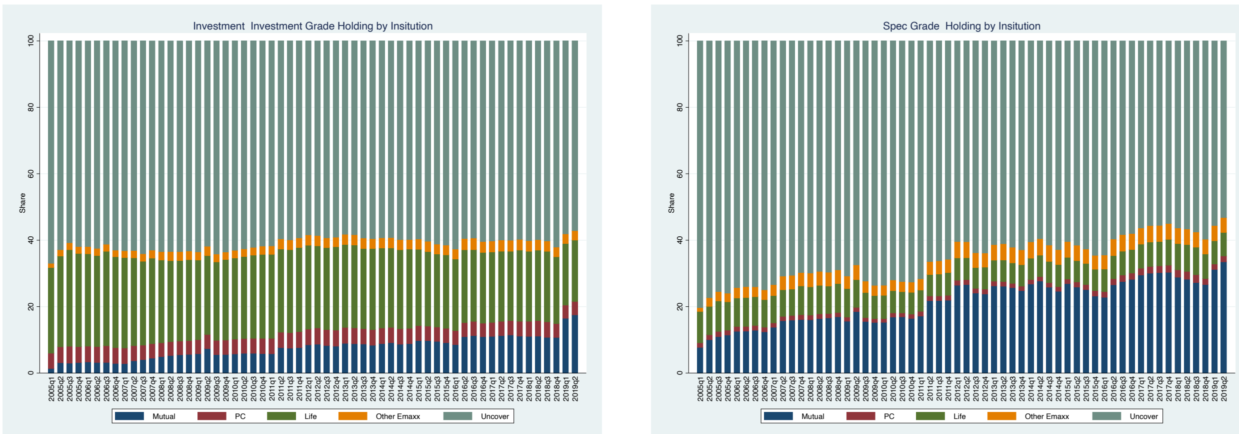
Since the measurement of bid-ask spreads may be noisy, we also construct a measure of liquidity following Dick-Nielsen, Feldhütter, and Lando (2012). This alternative measure of liquidity takes into account other price measures such as price impact, imputed round cost, and quantity measures such as the number of zero-trading days, trading volume and so on. We believe this liquidity measure likely has a smaller measurement error than empirically estimated bid-ask spreads. All of our results are robust to, if not stronger than, this alternative liquidity measure. The details are in Appendix D.

Our investor holdings data come from the Lipper eMaxx database of Thomson Reuters, which contains the corporate bond CUSIP-level holdings' data for insurance companies, mutual funds, pension funds, and other investors at the quarterly level. Figure 6a and Figure 6b plot the coverage of eMaxx for investment-grade and high-yield bonds. For our selected bond space, eMaxx covers 30%-40% of bonds' total outstanding amount consistently throughout the sample period. We do not find evidence of increasing coverage over time. Compared with the flow of funds data, the missing parties are mostly foreign accounts, banks, or both.

[22]Most electronic trading platforms (e.g., MarketAxess) offer both dealer-to-customer trades and all-to-all trading protocols, where customers trade directly with each other. The frictions on these centralized trading platforms are sometimes different from the search frictions that we describe in our model; hence, we exclude these trades from our empirical work. Focusing on large trades also partially addresses the concern that bid-ask spreads depend on the size of trades.

Figure 6: eMaxx Bond Coverage

(a) Investment-grade Bonds (b) High-yield Bonds



The figure plots quarterly share of holding of investment-grade and high-yield bonds covered by eMaxx from 2005Q2 to 2019Q2. The holding data is from eMaxx; the amount outstanding and rating data are from Mergent FISD. We classify the investors in eMaxx into life insurance company, P&C insurance company, mutual fund and others. The unclassified is calculated using the amount outstanding minus the total amount in eMaxx.

Consistent with the flow of funds aggregate data, the share of mutual funds has been increasing over time in our dataset, especially after 2008. Mutual funds have a larger share in high-yield bonds, whereas life insurance companies have a larger share in investment-grade bonds. Moreover, the growth of mutual fund shares is also more significant in the high-yield bonds than in the investment-grade bonds.

To measure investor composition for each bond, we first create a measure for individual investors and then aggregate it to the bond level. For individual investors, we use the percentage of net changes in bond holdings to approximate the frequencies of their liquidity shocks. The net transaction of fund j in quarter t is defined as

$$net\_transaction_{j,t} = \frac{\left| \sum_i holding_{i,j,t} - \sum_i holding_{i,j,t-1} \right|}{\sum_i holding_{i,j,t-1}}, \tag{5.1}$$

where $holding_{i,j,t}$ is the par value of bond i held by fund j in quarter t. A persistent higher net transaction indicates the fund experiences inflow or outflow (or both) more frequently, suggesting that the fund is subject to a higher frequency of liquidity shocks.

We then take the average $net\_transaction_{j,t}$ in the past four quarters to capture the persistent component of the fund's trading behavior,

$$NT_{j,t} = \frac{1}{4}\sum_{t'=1}^{4} net\_transaction_{j,t-t'}. \tag{5.2}$$

To measure investor composition for bond i, we take the value-weighted average net transaction of all its bondholders. For bond i in period t,

$$investor\_comp_{i,t} = \frac{\sum_j holding_{i,j,t} \times NT_{j,t}}{\sum_j holding_{i,j,t}}. \tag{5.3}$$

Higher $investor\_comp_{i,t}$ means more short-term investors are participating in the submarket for bond i.

Our final sample contains 1,739 firms and 14,645 bonds, covering 2005Q2 to 2019Q2. Table 2 shows the summary statistics of our bond-quarter observations. The average bid-ask spread is about 40 bps, and the standard deviation of the bid-ask spread is also 40 bps. The average credit spread in the sample is 2.2% with a standard deviation of 2.5%. It is worth pointing out that there is significant variation in bond maturity; hence, it is reasonable to consider investors sorting along this dimension. Lastly, in terms of ratings, about one-third of the bonds are rated AAA-A, and one-fourth of the bonds are high-yield bonds. The largest component is the BBB-rated bonds, accounting for nearly 40% of all bonds in our sample.

Furthermore, Table 3 shows the summary statistics of the portfolio information for mutual funds, insurance companies, and P&C insurance companies, which are the most important investor types in our sample. Life insurance companies tend to hold bonds with a longer time to maturity and higher ratings than mutual funds and P&C insurance firms. Mutual funds tend to hold higher shares of high-yield bonds. In terms of the average bid-ask spreads, the bonds held by insurance companies have higher bid-ask spreads, on average, than those held by mutual funds. We test this more formally later.

Table 2: Summary Statistics at Bond-Quarter Level

	Mean	Min	Max	Std
Investor comp	0.086	0.000	0.679	0.048
Bid-ask spread	0.004	0.000	0.073	0.004
Credit spread	0.022	0.002	0.182	0.025
Quarterly transaction volume (million USD)	150.1	0.2	1,058	269.1
Time to Maturity (30/360 convention)	9.468	0.252	100.153	8.997
Age	4.165	0.000	60.710	3.858
Annual interest rate (%)	5.522	0.000	15.0	1.917
Observations	184125			

This table presents bond-quarter level summary statistics. Bond data is obtained from WRDS and Mergent FISD. We include only US Corporate debenture with fixed coupon rate, non-convertible, non-putable, non-exchangeable. In addition, we only include bonds whose eMaxx coverage exceeds 20%.

5.1.2 Analysis and Results

To examine how investor composition affects credit spreads via liquidity, we set up the benchmark regression as follows:

$$CS_{i,t} = \alpha + \beta_1 BA_{i,t} + \beta_2 investor\_comp_{i,t} + \beta_3 BA_{i,t} \times investor\_comp_{i,t} + \gamma^\top \mathbf{X_{i,t}} + \epsilon_{i,t}. \tag{5.4}$$

The controls in $\mathbf{X_{i,t}}$ are the same as those in Section 2. We do not use CDS spreads in the main analysis because doing so would further restrict our sample to bonds that have CDS traded. Also, a single-name CDS itself is illiquid. We have verified that all of our results hold when subsetting to bonds with CDS and controlling for CDS spreads. Since we have panel data of yield spreads, the standard error is clustered at the quarter and firm level.

Table 4 reports the benchmark result. First, a higher bid-ask spread is associated with a higher credit spread: a one standard deviation increase in the bid-ask spread leads to an increase in the credit spread of around 38 bps. More importantly, once we add the interaction between investor type and bid-ask spreads into the regression, the sign of the bid-ask spreads switches, and the interaction term is highly positive and significant. This result indicates that when the bondholders are subject to more frequent liquidity shocks, transaction costs have a larger impact on credit spreads. As a robustness check, we include CDS spreads

Table 3: Summary Statistics by Investor Type

Investor type	Life insurance	Mutual funds	P&C
Net transaction (%)	0.093	0.203	0.132
Amount (Trillion)	1.88	1.71	0.36
Average bid-ask spreads (bps)	35.5	28.5	25.9
Average yield	2.61	4.33	3.11
Time-to-maturity	12.22	7.56	5.89
Coupon	4.77	5.01	3.99
Number of funds	1091	6498	2014
Fraction of AAA-A	0.432	0.263	0.462
Fraction of BBB	0.465	0.407	0.407
Fraction of high-yield	0.103	0.330	0.131

This table summarizes bond-quarter level statistics by life insurance companies, mutual funds and P&C insurance companies. Data source is WRDS, Mergent FISD and eMaxx. Sample period is from 2015Q1 to 2019Q2.

and only perform the test using the subset of bonds with CDS. The results are qualitatively similar. The details are shown in Appendix B.

We then sort the bonds by investor composition into five groups. Table 5 reports the summary statistics by investor quintile. Bonds in group 1 have mostly long-term investors (investors subject to less frequent liquidity shocks), whereas bonds in group 5 have mostly short-term investors. We then run the following regression group by group:

$$CS_{i,t} = \alpha + \beta_1 BA_{i,t} + \beta_2 investor\_comp_{i,t} + \gamma^{\top}\mathbf{X_{i,t}} + \epsilon_{i,t}. \tag{5.5}$$

Figure 7a plots the coefficient in front of bid-ask spreads (β_1) and the corresponding confidence intervals for each group. The coefficient is larger as we move to bond groups with more short-term investors. For example, a 100 bps increase in the bid-ask spread in group 1 indicates an increase in the credit spread of around 25 bps, whereas in group 5, the difference is around 1.5%, even after controlling for ratings.[23] This result means that bonds with more short-term investors are more affected by the transaction costs or illiquidity in the secondary market, in the sense that differences in bid-ask spreads are mapped into larger differences in

[23] This is not a result of differences in the standard deviations of bid-ask spreads; in fact, the standard deviations of bid-ask spreads for bonds in group 5 are lower than those in group 1.

Table 4: Credit Yields Regression on Bid Ask Spreads

	(1) Credit spread	(2) Credit spread	(3) Credit spread
Bid-ask spread	0.405***	0.395***	-0.387***
	(0.0824)	(0.0844)	(0.0889)
Investor comp		-0.0136*	-0.0556***
		(0.00646)	(0.00796)
Bid-ask spread × Investor comp			11.72***
			(1.217)
Time to maturity ($\times 10^{-2}$)	0.0150***	0.0143***	0.0146***
	(0.00176)	(0.00176)	(0.00170)
Coupon ($\times 10^{-2}$)	0.0863***	0.0860***	0.0797***
	(0.0158)	(0.0157)	(0.0157)
Offering amount ($\times 10^{-2}$)	4.25e-08	4.31e-08*	4.13e-08
	(2.18e-08)	(2.14e-08)	(2.08e-08)
N	153831	147242	147242
adj. R^2	0.756	0.760	0.767

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table presents the regression results for equation 5.4. Bid-Ask Spread is the measure of liquidity and "Investor Comp" is the measure of investor composition, higher value indicates the investor have higher frequency of liquidity shock, i.e. shorter term investors. We also control for bond age, firm leverage, fraction of long term debt, equity volatility and profitability, where equity volatility is the standard deviation of the equity return of the issuer in that quarter.

credit spreads.

Our results show that the cross derivative of credit spreads with respect to investor trading needs and bid-ask spreads is positive. This finding is direct evidence that investor heterogeneity plays an important role in pricing liquidity into credit spreads of bonds in the cross section.

Following this analysis, we compute the liquidity component for each group. The liquidity component is calculated as the sensitivity coefficient multiplied by the bid-ask spreads. The median liquidity component as a fraction of the credit spreads for each group is plotted

Table 5: Summary Statistics by Investor Composition Quantile

	(1)		(2)		(3)		(4)		(5)	
	mean	std	mean	std	mean	std	mean	std	mean	std
Investor comp	0.038	0.007	0.052	0.007	0.068	0.010	0.095	0.018	0.154	0.046
Bid-ask Spread	0.005	0.006	0.005	0.005	0.004	0.004	0.004	0.004	0.003	0.003
Alternative measure of illiquidity	0.153	0.839	0.008	0.702	-0.045	0.639	-0.099	0.593	-0.142	0.560
Credit spread	0.016	0.017	0.016	0.017	0.017	0.019	0.023	0.025	0.036	0.032
Quarterly transaction volume (million USD)	47.9	67.68	88.0	124.4	138.9	207.7	207.0	341.4	224.3	360.7
Time to maturity (30/360 Convention)	14.251	10.792	11.168	9.797	9.766	9.305	7.859	7.741	6.189	5.428
Age	6.925	4.774	5.104	3.965	4.006	3.538	3.186	3.127	2.659	2.669
Interest rate (%)	5.618	1.405	5.236	1.478	5.059	1.628	5.254	1.960	6.376	2.396
Fraction of AAA-A bonds	0.546	0.498	0.504	0.500	0.416	0.493	0.259	0.438	0.086	0.280
Fraction of BBB bonds	0.414	0.493	0.456	0.498	0.501	0.500	0.462	0.499	0.261	0.439
Fraction of speculative grade bonds	0.040	0.196	0.040	0.195	0.083	0.276	0.279	0.449	0.653	0.476
Observations	27645		35256		38531		40425		42268	

Bond-Quarter level summary statistics by investor composition quantile. For each quarter, we sort the bond into five groups based on their investor composition. Group 1 are bonds with the shortest-term investor, group 5 are bonds with longest-term investor. A-AAA is a dummy indicating the bond has rating A-AAA. Similar for BBB and Speculative.

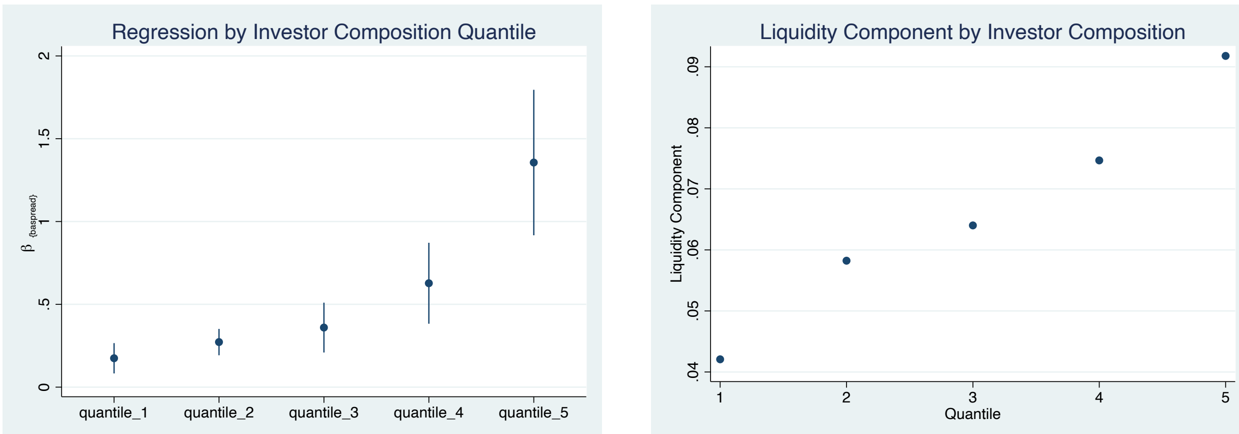
in Figure 7b. As expected, bonds with more short-term investors have a larger liquidity component. Using the offering yields, the results are qualitatively similar but less significant, partly because we have fewer observations for offering yields.[24]

We also verify that the sorting patterns between investor types and bond types match the model predictions. Consistent with the literature, we find that bonds held by short-term investors have a shorter time to maturity and a higher default probability. In Figure 8a, we sort bonds into 50 bins according to their investor composition and then plot the average time to maturity for each bin. The relationship is almost linear. To investigate the sorting pattern along default probabilities, we assign numeric values for bonds of different ratings: AAA bonds receive the lowest value, 1, and D-rated bonds receive the highest value, 22. We again sort bonds into 50 bins according to investor composition, and then plot the average numeric rating for each bin. Figure 8b shows the result. We find consistent evidence in regression analysis controlling for bid-ask spreads, as well as other bond and firm characteristics, as shown in Table 6. We acknowledge that other factors may be contributing to the sorting patterns outside of the model (for example, the regulatory constraint faced by

[24]The results for offering yields are more prominent when using the alternative measure of liquidity. We again sort the bonds into five groups based on their investor composition, repeating the exercise in the secondary market. The results are shown in Table 16 in Appendix D; the impact of expected illiquidity is higher when more short-term investors are trading the bond.

Figure 7: Regression by Investor Quantile

(a) Sensitivity of Credit Yields to Bid-Ask Spreads (b) Median Liquidity Component



We sort the bonds (in each quarter from 2005Q2 to 2019Q2) into five groups, by investor base turnover. Group 1 contains bonds whose investors have the lowest turnover rate and group 5 contains bonds whose investors have the highest turnover rate. In the left-hand-side, we plot the regression coefficients and 1% confidence interval of credit spread regressed on bid-ask spreads, controlling for bond and firm characteristics, group by group. The exact regression equation is in 5.5. For each bond, define $liquidity\_component_{i,t} = \frac{\beta \times bid\_ask_{it}}{credit\_spread_{it}}$, where β is the coefficient in front of bid-ask spreads. The right-hand-side figure plots the median liquidity component over credit spreads for each group.

insurance companies). We view the empirical findings as suggestive evidence supporting the model.

5.2 Cross-country Analysis

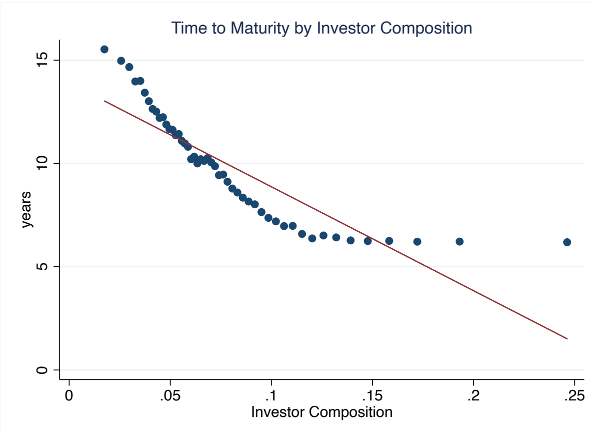
Our model predicts that the growth of mutual funds in illiquid asset holdings is due to the decline in the risk-free rate. To investigate the relationship between risk-free rates and mutual fund growth empirically, we conduct a cross-country analysis. We show that countries with a larger decline in risk-free rates have experienced higher growth in mutual fund shares in the bond markets.

5.2.1 Data

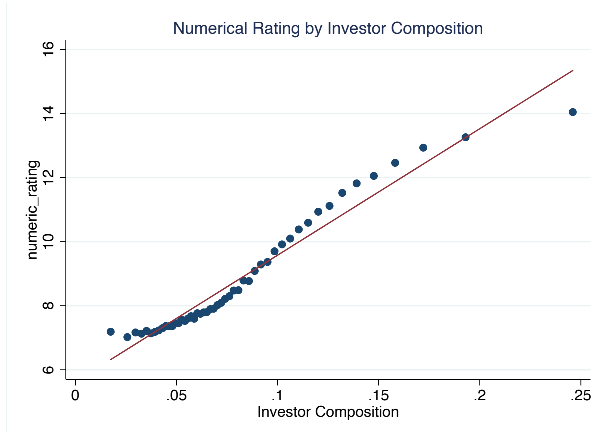
From Morningstar, we obtain monthly data on the size of the fixed-income mutual funds across 69 countries from January 2007 to August 2020. To match other controls we have, we

Figure 8: Correlation between Bond Characteristics and Investor Types

(a) Time to Maturity



(b) Bond Ratings



We sort bonds between 2005Q2 to 2019Q2 according to their investor composition, then plot the average time to maturity for each bin in the left-hand-side figure. In the right-hand-side figure, we plot the average numeric rating for each bin. The red lines are linear fitted lines. Bond information is obtained from WRDS and investor composition information is constructed from eMaxx.

average the monthly holdings to the quarter and annual level. We define the growth rate of short-term investor holdings as

$$flow_{i,t+1} = \frac{total\_asset_{i,t+1} - total\_asset_{i,t} - CA_{i,t+1}}{total\_asset_{i,t}},$$

where $total\_asset_{i,t}$ is country i's fixed-income funds' total asset value in period t, and $CA_{i,t+1}$ is the capital appreciation from period t to $t+1$. We subtract the net capital appreciation during the period to capture the net inflow more accurately.

The IMF's International Financial Statistics database has data on country interest rates and macroeconomic conditions: quarterly GDP growth, the unemployment rate, and inflation. For the risk-free rate, we use the country-specific government bond yields in the baseline analysis. This likely corresponds to the liquid asset that short-term investors would consider holding. Given the difference in capital flow patterns between developed and emerging markets, we include only OECD countries in the baseline case. In an extension, we include all countries available and use the "money market rate" or "saving rate" for countries where government bond yields are not available.

Table 6: Sorting of Investors

	(1) Time to maturity	(2) Bid-ask spread
Investor comp	-49.57***	-0.00952***
	(3.884)	(0.000836)
Age	-0.592***	0.000119***
	(0.0591)	(0.0000121)
Coupon	2.126***	-0.0358***
	(0.154)	(0.00471)
Offering amount	-0.000000431	-4.67e-10***
	(0.000000254)	(5.41e-11)
N	179965	151736
adj. R^2	0.200	0.275

Standard errors in parentheses: sym* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table presents correlation between investor composition and the maturity of bonds they hold, and the bid-ask spreads in each sub-market. The coefficients in the first column are from the regression: $ttm_{i,t} = \alpha + \beta_1 investor\_comp_{i,t} + \beta_2 BA_{i,t} + \gamma^\top \mathbf{X_{i,t}} + \epsilon_{i,t}$. The coefficients in the second column are from $BA_{i,t} = \alpha + \beta investor\_comp_{i,t} + \gamma^\top \mathbf{X_{i,t}} + \epsilon_{i,t}$. "Investor comp" is the measure of investor composition, higher value indicates the investor have higher frequency of liquidity shock, i.e. shorter term investors. We also control for bond age, rating, firm leverage, fraction of long term debt, equity volatility and profitability, where equity volatility is the standard deviation of the equity return of the issuer in that quarter.

5.2.2 Analysis and Results

We run the following regression to examine the relationship between the growth of short-term investors' holdings and the risk-free rate:

$$flow_{i,t} = \beta_0 + \beta_1 \Delta rf_{i,t} + \gamma^\top \mathbf{X_{i,t}} + \epsilon_{i,t}. \tag{5.6}$$

The controls $X_{i,t}$ here include macroeconomic conditions — GDP growth, unemployment rate, and inflation— as well as time and country fixed effects. Each period is either a quarter or a year.

The regression results in Table 7 show a strong cross-country relationship between risk-

Table 7: Relationship between Fund Growth and Risk-free Rates (OECD Countries)

	Dependent variable: fixed income fund growth					
	(1)	(2)	(3)	(4)	(5)	(6)
Change in risk-free rate	-0.046	-0.039*	-0.038*	-0.050**	-0.056***	-0.068***
	(0.029)	(0.022)	(0.021)	(0.020)	(0.020)	(0.021)
GDP growth		-0.011	-0.007		0.008	0.017
		(0.011)	(0.008)		(0.013)	(0.026)
Unemployment		0.002	-0.012		0.003	-0.041
		(0.003)	(0.009)		(0.007)	(0.039)
Inflation		0.043	0.044		0.031	0.038
		(0.033)	(0.030)		(0.050)	(0.101)
Country fixed effect	No	No	Yes	No	No	Yes
Year fixed effect	No	No	Yes	No	No	Yes
Frequency	Quarterly	Quarterly	Quarterly	Yearly	Yearly	Yearly
Observations	671	671	671	188	188	188
R^2	0.005	0.022	0.118	0.012	0.015	0.279
Adjusted R^2	0.004	0.016	0.069	0.006	-0.006	0.101
Residual Std. Error	0.374	0.372	0.362	0.648	0.652	0.616
F Statistic	2.540	1.261	29.544***	6.089**	4.297***	155.570***

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

Risk-free rate and macroeconomic data is obtained from IMF statistics. We get aggregate fund size data from Morningstar. Fixed income fund growth is defined as the growth in net flows over the total asset under management in the previous period. Sample period is from January 2007 to August 2020. We include only OECD countries whose fund size data is available.

free rates and fund growth. A larger decline in the risk-free rate is associated with a significantly higher fund growth rate in the fixed-income market.[25] This finding is consistent with the time-series pattern observed in the U.S. Moreover, one would expect the effect to be larger over a longer time horizon. That is indeed what we find: the effect of the risk-free rate on the fund growth rate at the yearly level is one and half times as large as that at the quarterly level.[26] All of our results still hold when we extend the data sample to all countries.

[25]The intuition that more short-term investors enter illiquid asset markets because of the low-risk-free-rate environment should extend to assets other than corporate bonds, which is indeed what we find. The above relationship also holds if we use the growth of all mutual funds, not just fixed-income funds.

[26]We are not looking at the growth of mutual fund shares over the full sample period because the sample periods are different for different countries.

6 Calibration

In this section, we conduct a simple calibration exercise and show that our model can account for the significant changes in the data. Using the calibrated model, we investigate how changes in investor composition interact with the recent regulation changes on dealers from the Volcker Rule. We find that the liquidity provision role of mutual funds has alleviated the impact of dealer regulations on the market. Lastly, we show that unexpected aggregate liquidity shocks have a larger impact on the market with more short-term investors today than before.

Our calibration is at the quarterly level. We focus on the bonds with maturity between 0.5 and 20 years; hence, we set $\underline{\delta} = 0.0125$ and $\overline{\delta} = 0.5$. Furthermore, we assume the ex ante investor types follow a Pareto distribution with scale parameter $\underline{\theta}$ and shape parameter α. We assume $\xi = \xi_0 + \xi_1 \lambda^{\nu}$. The secondary market yield is calculated as $CS_{i,t} = \frac{r_{i,t}+\delta_i(1-\bar{p}_{i,t})+d_i(s_i-\bar{p}_{i,t})}{\bar{p}_{i,t}} - rf_t$ where $\bar{p}_{i,t}$ is the mid-price.

We calibrate the parameters $(\Delta, \xi_0, \xi_1, \gamma, \eta, \underline{\theta}, \nu, \rho, \alpha)$ to match the size of the liquidity component in 2005, the average investor turnover rate, the sensitivity coefficient of bid-ask spreads, the variance of credit spreads, and the bid-ask spreads in 2005 and 2019. We target the liquidity component and sensitivity coefficients because they are the key moments of interest in the data. The average investor turnover is the key driver of all other changes; hence, we include it as well. In the data, we only observed realized turnover. We use the definition in Section 5 and construct the exact counterpart in the model, taking into account the waiting time. In our eMaxx sample, the average investor turnover rate increased from 0.08 to 0.95 from 2005 to 2019. Since the sensitivity coefficients are derived from a cross-sectional regression, it is influenced by the variances of credit spreads and bid-ask spreads. Thus, we include both as moment targets as well. We use the 10-year Treasury bond rate as our risk-free rate because it is the main liquid outside option that mutual funds hold. In 2005, the 10-year Treasury rate was around 4% annually, and in mid-2019, that rate decreased to around 2%. Table 8 shows the value of each moment that we are targeting, along with the model counterparts, and Table 9 shows the value of calibrated parameters. All the parameters are jointly determined.

Even though we have the same number of parameter values and moments, we cannot

Table 8: Target Moments and Model Fits

Moment	Empirical	Model simulated value
Ave. inv turnover in 2005	0.08	0.0881
Ave. inv turnover in 2019	0.95	0.0897
Coeff of BA spreads in 2005	0.25	0.251
Coeff of BA spreads in 2019	1.1	1.10
Liquidity comp in 2005	5%	4.75%
Std in BA spread in 2005	0.0032	0.00064
Std in BA spread in 2019	0.0025	0.00057
Std in credit spread in 2005	0.028	0.0204
Std in credit spread in 2019	0.015	0.0214

This table presents the targeted empirical moments and model produced moment values. The calibration is at the quarterly level. We focus on the bonds that are between 0.5-20 years of maturity, hence we set $\underline{\delta} = 0.0125$ and $\overline{\delta} = 0.5$. Furthermore, we assume the ex-ante investor types follow a Pareto distribution with scale parameter $\underline{\theta}$ and shape parameter α. We assume $\xi = \xi_0 + \xi_1 \lambda^{\nu}$. We feed in interest rate of 1% in 2005 and 0.5% in 2019, both at the quarterly level.

match all the moments exactly, because the model is quite stylized.[27] However, our model can match the large change in the sensitivity coefficient of credit yields to bid-ask spreads, given the change in the average turnover ratio. For this to be true, it is important to endogenize investor sorting behavior, the seller-buyer ratio, and the bid-ask spreads in each sub-market. To see why, consider a model similar to Amihud and Mendelson (1986), in which the bid-ask spreads are exogenous. The average sensitivity coefficient is given by

$$E[\frac{\theta}{1+\rho+\delta+d}]. \tag{6.1}$$

Given the value we used for calibration, it is bounded above by 0.12, much lower than what we observe empirically. It cannot generate the observed magnitudes of changes in the impact of bid-ask spreads. The reason our model can generate a much higher sensitivity coefficient is because bid-ask spreads are endogenous to the investor composition . As explained before, bid-ask spreads are higher in sub-markets with higher seller-buyer ratios, which also leads to higher interest rates in those sub-markets. Hence, the sensitivity coefficient in the model is

[27] For example, each bond sub-market is matched with only one type of investor. There is no diversification motive among investors. This simplification is necessarily to capture the sorting pattern in semi-closed form.

Table 9: Model Parameter Values

Parameter	Value
Liquidity discount Δ	0.502
Time discount ρ	0.764
Scale parameter of initial distribution $\underline{\theta}$	0.049
Shape parameter of initial distribution α	0.240
Constant in bid-ask spread ξ_0	0.174
Scale parameter in bid-ask spread ξ_1	0.0026
Power parameter in bid-ask spread ν	0.755
Scale parameter in matching function η	0.762
Power parameter in matching function γ	0.949

This table presents the calibrated model parameter values. The calibration is at the quarterly level. We focus on the bonds that are between 0.5-20 years of maturity, hence we set $\underline{\delta} = 0.0125$ and $\overline{\delta} = 0.5$. Furthermore, we assume the ex-ante investor types follow a Pareto distribution with scale parameter $\underline{\theta}$ and shape parameter α. We assume $\xi = \xi_0 + \xi_1 \lambda^{\nu}$. We feed in interest rate of 1% in 2005 and 0.5% in 2019, both at the quarterly level.

much larger and is able to match the empirical counterpart. This implies the waiting time is an important difference across the sub-markets and is priced into the credit spreads.

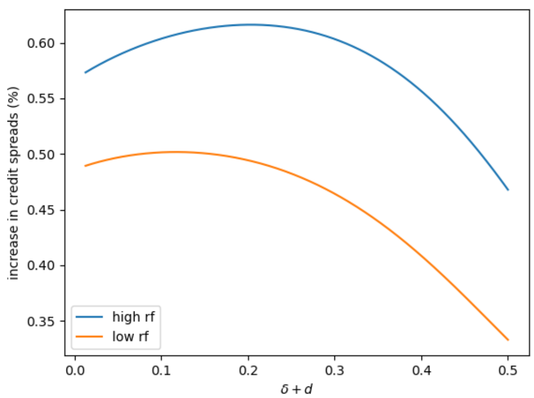
We calculate the liquidity component in 2019 given the calibrated model parameters. Note that this moment is not targeted in the calibration exercise. Our model predicts that the liquidity component should be 0.208 given the change in the risk-free rate and investor composition, which closely matches the $\sim$0.25 empirical counterpart.

Next, using the calibrated parameter values, we perform two exercises. We first look at what happens if ξ_1 increases by 1%. This experiment is motivated by the implementation of the Volcker Rule. It is widely perceived that the Volcker Rule has raised dealers' cost to hold inventories. To capture this, we increase the sensitivity of bid-ask spreads to seller-buyer ratio ξ, recalculate the equilibrium coupon rate, and allow investors to re-sort.[28] We perform the exercise under the market conditions of 2005 and 2019, respectively. The difference is the equilibrium investor composition in the bond market. Figure 9a illustrates the comparison of credit spread changes in response to a 1% increase in the sensitivity of bid-ask spreads to the seller-buyer ratio, in 2005 and 2019.

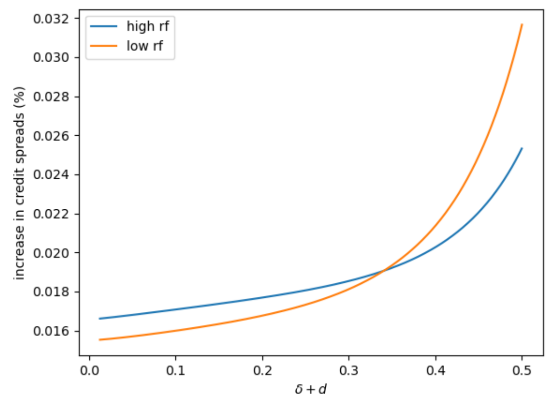
[28]The rationale is that the regulation change has been implemented for a while, and the market probably has adjusted to the new equilibrium.

Figure 9: Counter-factual Exercises

(a)



(b)



Using calibrated parameter values, we increase ξ_1 by 1% in panel (a). We allow investors to re-sort and recalculate the coupon payment. In panel (b), we increase the seller-buyer ratio in each sub-market exogenously by 1%, fixing investor composition and coupon rate. The figure plots the increase in credit spreads for each bond $\delta_i + d_i$. The blue line feeds in market conditions in 2005, and the orange line feeds in the market conditions in 2019.

Having more short-term investors could either amplify or alleviate the impact of a change in dealer regulations on credit conditions. On the one hand, with more short-term investors, credit spreads are more sensitive to secondary market frictions. This channel suggests we should see a larger impact of regulation with more short-term investors. On the other hand, short-term investors provide liquidity. Investors in the secondary market now rely less on the dealer sector to intermediate trades (lower seller-buyer ratio in equilibrium). This channel implies that a change in dealer regulations should have a smaller impact with more short-term investors in the market. In our calibration, we find that the latter channel dominates. Compared with 2005, the effect of an increase in ξ_1 on credit spreads is much smaller in 2019. Since today's market is populated with more short-term investors, the reliance on dealers to intermediate the trades has decreased. As a result, even if the dealer sector experiences an increase in cost, the effect on credit spreads is quite small. In other words, the entrance of mutual funds has masked the effects of changes in dealer regulation to some extent.

In the second exercise, we examine what happens when the selling pressure in all sub-markets increases exogenously. One can think of this experiment as similar to an aggregate liquidity shock hitting all investors, inducing higher selling pressure in the market. To capture this, we increase the seller-buyer ratio by 1%, keeping the investor composition in all sub-markets fixed as in either 2005 or 2019.[29] The results are illustrated in Figure 9b. As before, having more short-term investors has two counteracting effects; on the one hand, the markets have a lower seller-buyer ratio in steady state, but credit spreads are more sensitive to secondary market frictions. For bonds held by short-term investors (high $\delta + d$), the negative effect dominates. Cross-sectionally, bonds held by short-term investors experience a larger disruption in response to aggregate liquidity shocks. Over time, bonds held by short-term investors become even more sensitive to aggregate shocks.

7 Conclusion

In this paper, we connect two important trends happening in the corporate bond market. The first trend is the massive growth of mutual fund shares, especially in the post-crisis period. The second trend is new: the sensitivity of credit yields to bid-ask spreads increased fourfold in the last 15 years, leading to an increase in the liquidity component from 5% in 2005 to around 25% in 2019, for both investment-grade and high-yield bonds.

We then build a model linking the growth of mutual fund shares in the bond market with the long-run decline in risk-free rates around the world. Our model is able to match the aggregate trends in the data jointly. In particular, the model predicts that as the risk-free rate declines, more short-term investors enter the bond market, searching for higher yields. This creates more liquidity and lowers the bid-ask spreads. However, because of the higher trading needs and shortened waiting time, the credit yields become more sensitive to bid-ask spreads. As a result, the liquidity component increases. We then test the model predictions in the cross section and find consistent evidence. Finally, we calibrate the model and conduct counterfactual analysis.

Our results indicate secondary market frictions have become a more important factor in pricing liquidity into credit spreads and ultimately in determining firms' borrowing cost. We contribute to the heated debate about the importance of liquidity in the corporate bond market. Although others have studied how dealer regulations alter the level of liquidity in the secondary market, we point out that credit yields are more sensitive to underlying

[29] Here, we keep all the coupon rates fixed. Since this is a shock in the secondary market, the coupon rate, which is determined in the primary market, does not respond instantaneously.

frictions today than before because of the change in investor composition.

Both the current COVID-19 crisis and the 2008 financial crisis have induced high selling pressure in the bond market. Our model implies that the disruptive impact of aggregate liquidity shocks is much higher today than before because of the differences in investor composition. We also find that the entry of mutual funds is potentially masking the negative impact of changes in dealer regulation on the corporate bond market.

It would be interesting to apply our framework to evaluate the effectiveness of different quantitative easing (QE) programs. Past QE measures in the U.S. have focused on purchasing long-term government bonds. In the recent COVID-19 crisis, the Federal Reserve has started purchasing investment-grade corporate bonds. Our framework can analyze the effectiveness of these two measures in lowering credit spreads, taking into account the difference in the investor base for different bonds.

Finally, different from past crises, investment-grade bonds have suffered a larger dislocation relative to high-yield bonds in the current COVID-19 crisis. Market participants have argued that this phenomenon is due to the large outflow of corporate bond mutual funds. Given that investor composition today is quite different from that in the past, it would be interesting to investigate the role of investor heterogeneity in explaining the different behaviors across bonds. One can extend the model to a full dynamic setup, to evaluate the movements of investor types in different bond markets and how the economy's reaction to aggregate shocks depends on the initial distribution of investor composition.

References

Adrian, T., N. Boyarchenko, and O. Shachar (2017): "Dealer balance sheets and bond liquidity provision," *Journal of Monetary Economics*, 89(C), 92–109.

Adrian, T., M. Fleming, O. Shachar, and E. Vogt (2017): "Market liquidity after the financial crisis," *Annual Review of Financial Economics*, 9(1), 43–83.

Amihud, Y., and H. Mendelson (1980): "Dealership market: Market-making with inventory," *Journal of Financial Economics*, 8(1), 31 – 53.

——— (1986): "Asset pricing and the bid-ask spread," *Journal of financial Economics*, 17(2), 223–249.

Anderson, M., and R. M. Stulz (2017): "Is post-crisis bond liquidity lower?," Discussion paper, National Bureau of Economic Research.

Bai, J., and P. Collin-Dufresne (2019): "The CDS-bond basis," *Financial Management*, 48(2), 417–439.

Bao, J., J. Pan, and J. Wang (2011): "The illiquidity of corporate bonds," *The Journal of Finance*, 66(3), 911–936.

Beber, A., M. W. Brandt, and K. A. Kavajecz (2009): "Flight-to-quality or Flight-to-liquidity? Evidence from the Euro-area bond market," *Review of Financial Studies*, 22(3), 925–957.

Bernanke, B. S., C. Bertaut, L. P. DeMarco, and S. Kamin (2011): "International capital flows and returns to safe assets in the United States, 2003-2007," *International Finance Discussion Papers*.

Bretscher, L., L. Schmid, I. Sen, and V. Sharma (2020): "Review of Financial Studies," *Swiss Finance Institute Research Paper*, (No. 21-07).

Chen, H., R. Cui, Z. He, and K. Milbradt (2017): "Quantifying liquidity and default risks of corporate bonds over the business cycle," *The Review of Financial Studies*, 31(3), 852–897.

Chen, X., J.-Z. Huang, Z. Sun, T. Yao, and T. Yu (2020): "Liquidity premium in the eye of the beholder: An analysis of the clientele effect in the corporate bond market," *Management Science*, (2), 932–957.

Choi, J., and O. Shachar (2014): "Did liquidity providers become liquidity seekers? Evidence from the CDS-bond basis during the 2008 financial crisis," *FRB of New York Staff Report*, (650).

Choi, J., O. Shachar, and S. S. Shin (2019): "Dealer liquidity provision and the break-down of the law of one price: Evidence from the CDS–bond basis," *Management Science*, 65(9), 4100–4122.

Di Maggio, M., A. Kermani, and Z. Song (2017): "The value of trading relations in turbulent times," *Journal of Financial Economics*, 124(2), 266 – 284.

Dick-Nielsen, J. (2014): "How to clean enhanced TRACE data," *Available at SSRN 2337908*.

Dick-Nielsen, J., P. Feldhütter, and D. Lando (2012): "Corporate bond liquidity before and after the onset of the subprime crisis," *Journal of Financial Economics*, 103(3), 471–492.

Dick-Nielsen, J., and M. Rossi (2018): "The cost of immediacy for corporate bonds," *The Review of Financial Studies*, 32(1), 1–41.

Duffie, D. (2012): "Market making under the proposed Volcker Rule," Stanford University Working Paper.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005): "Over-the-counter markets," *Econometrica*, 73(6), 1815–1847.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2007): "Valuation in over-the-counter markets," *The Review of Financial Studies*, 20(6), 1865–1900.

Eeckhout, J., and P. Kircher (2010): "Sorting and decentralized price competition," *Econometrica*, 78(2), 539–574.

Ericsson, J., and O. Renault (2006): "Liquidity and credit risk," *The Journal of Finance*, 61(5), 2219–2250.

Feldhütter, P. (2012): "The same bond at different prices: Identifying search frictions and selling Pressures," *The Review of Financial Studies*, 25(4), 1155–1206.

Feroli, M., A. K. Kashyap, K. L. Schoenholtz, and H. S. Shin (2014): "Market tantrums and monetary policy," *Chicago Booth Research Paper*, (No. 14-09).

Friewald, N., R. Jankowitsch, and M. G. Subrahmanyam (2013): "Illiquidity or credit deterioration: A study of liquidity in the US corporate bond market during financial crises," in *Managing and Measuring Risk: Emerging Global Standards and Regulations After the Financial Crisis*, pp. 159–200. World Scientific.

Friewald, N., and F. Nagler (2019): "Over-the-counter market frictions and yield spread changes," *The Journal of Finance*, 74(6), 3217–3257.

Gagnon, E., B. K. Johannsen, and D. Lopez-Salido (2016): "Understanding the new normal: The role of demographics," *Finance and Economics Discussion Series, Washington: Board of Governors of the Federal Reserve System.*

Goldstein, I., H. Jiang, and D. T. Ng (2017): "Investor flows and fragility in corporate bond funds," *Journal of Financial Economics*, 126(3), 592 – 613.

Goldstein, M. A., E. S. Hotchkiss, and D. J. Pedersen (2019): "Secondary market liquidity and primary market pricing of corporate bonds," *Journal of Risk and Financial Management*, 12(2), 86.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007): "The US Treasury yield curve: 1961 to the present," *Journal of monetary Economics*, 54(8), 2291–2304.

Hamilton, J. D., E. S. Harris, J. Hatzius, and K. D. West (2016): "The equilibrium real funds rate: past, present, and future," *IMF Economic Review*, 64(4), 660–707.

He, Z., P. Khorrami, and Z. Song (2020): "Commonality in credit spread changes: dealer inventory and intermediary distress," *Working Paper*.

He, Z., and K. Milbradt (2014): "Endogenous liquidity and defaultable bonds," *Econometrica*, 82(4), 1443–1508.

Ho, T. S. Y., and H. R. Stoll (1983): "The dynamics of dealer markets under competition," *The Journal of Finance*, 38(4), 1053–1074.

Hugonnier, J., B. Lester, and P.-O. Weill (2019): "Frictional intermediation in over-the-counter markets," *The Review of Economic Studies*, 87(3), 1432–1469.

Kargar, M., S. Liu, B. Lester, P.-O. Weill, D. Lindsay, and D. Zúñiga (2020): "The corporate bond liquidity during the COVID-19 crisis," NBER Working Paper.

Lagos, R., and G. Rocheteau (2007): "Search in asset markets: Market structure, liquidity, and welfare," *American Economic Review*, 97(2), 198–202.

Lagos, R., G. Rocheteau, and P.-O. Weill (2011): "Crises and liquidity in over-the-counter markets," *Journal of Economic Theory*, 146(6), 2169 – 2205.

Longstaff, F. A., S. Mithal, and E. Neis (2005): "Corporate yield spreads: default risk or liquidity? New evidence from the credit default swap market," *The Journal of Finance*, 60(5), 2213–2253.

Ma, Y., K. Xiao, and Y. Zeng (2021): "Mutual Fund Liquidity Transformation and Reverse Flight to Liquidity," *Review of Financial Studies*, Forthcoming.

Musto, D., G. Nini, and K. Schwarz (2018): "Notes on bonds: Illiquidity feedback during the financial crisis," *The Review of Financial Studies*, 31(8), 2983–3018.

Oehmke, M., and A. Zawadowski (2016): "The anatomy of the CDS market," *The Review of Financial Studies*, 30(1), 80–119.

O'Hara, M., and X. Zhou (2021): "Anatomy of a liquidity crisis: Corporate bonds in the COVID-19 crisis," *Journal of Financial Economics*, 142(1), 46–68.

Rachel, L., and T. D. Smith (2017): "Are low real interest rates here to stay?," *International Journal of Central Banking*, 13(3), 1–42.

Schwarz, K. (2018): "Mind the gap: disentangling credit and liquidity in risk spreads*," *Review of Finance*, 23(3), 557–597.

Vayanos, D., and T. Wang (2007): "Search and endogenous concentration of liquidity in asset markets," *Journal of Economic Theory*, 136(1), 66 – 104.

Wu, B. (2020): "Increasing corporate bond liquidity premium and post-crisis regulations," NYU Stern School of Business Working Paper.

Appendix A Derivations and Proofs

A.1 Proof for Lemma 1

In this proof, we omit the subscript j in order to present the proof in the simplest way. The proof applies for all j:

$$\rho V_s = r - \Delta + d(s - V_s) + \delta(1 - V_s) + \mu_s(p_s - \xi - V_s)$$
$$\rho V_b = \mu_b(V_0 - V_b - p_b)$$
$$\rho V_0 = r + d(s - V_0) + \delta(1 - V_0) + \theta(V_s - V_0)$$

and $p_b - p_s = \xi$. From the above four equation, we have

$$V_0 - V_b - V_s - \xi = \frac{\Delta + \mu_s(\frac{\rho V_b}{\mu_b} + \xi)}{\delta + d + \rho + \theta + \mu_s} - \xi$$

when ξ is small, we have $V_0 - V_b - V_s - \xi \geq 0$.

A.2 Proof for Lemma 2

The first differential equation is derived from $\theta' = \frac{\alpha_s}{\alpha_b \lambda}$, where α_s is the measure of sellers and α_b is the measure of buyers. Denote α_h as the measure of patient bond holders. In steady state,

$$\theta' m_I f(\theta) - \mu_b \alpha_b - \tilde{\delta} = 0, \qquad -\mu_s \alpha_s - \tilde{\delta}\alpha_s + \theta\alpha_h = 0, \qquad -\theta\alpha_h + \mu_b\alpha_b - \tilde{\delta}\alpha_h + \tilde{\delta} = 0.$$

With these conditions, we can derive

$$\frac{\alpha_s}{\alpha_b \lambda} = \frac{\theta m_I f(\theta)\mu_b + \tilde{\delta}\lambda(\mu_s + \tilde{\delta})(\theta + \tilde{\delta})}{(\mu_s + \tilde{\delta})(\tilde{\delta} + \theta) m_I f(\theta)\lambda}$$

To derive the second differential equation, consider the seller's problem,

$$U_s = \max_{\theta,\lambda} \quad \frac{\frac{\Delta}{\rho+\delta+d+\theta} - U(\theta)(1 + \frac{\rho}{\mu_b}) - \xi(\lambda)}{\frac{1}{\mu_s} + \frac{1}{\rho+\delta+d+\theta}}$$

Taking first order condition with respect to θ and λ,

$$[\theta]: \qquad -\frac{\left[\frac{\Delta}{\rho+\delta+d+\theta}+U'(\theta)(1+\frac{\rho}{\mu_b})\right](\frac{1}{\mu_s}+\frac{1}{\rho+\delta+d+\theta})+\frac{U(\theta)(1+\frac{\rho}{\mu_b})+\xi}{(\rho+\delta+d+\theta)^2}}{(\frac{1}{\mu_s}+\frac{1}{\rho+\delta+d+\theta})^2}=0 \tag{A.1}$$

$$[\lambda]: \qquad \frac{(\frac{U(\theta)\rho}{\mu_b^2}\mu_b'-\xi')(\frac{1}{\mu_s}+\frac{1}{\rho+\delta+d+\theta})+\frac{1}{\mu_s^2}\mu_s'\left[\frac{\Delta}{\rho+\delta+d+\theta}-U(\theta)(1+\frac{\rho}{\mu_b})-\xi(\lambda)\right]}{(\frac{1}{\mu_s}+\frac{1}{\rho+\delta+d+\theta})^2}=0 \tag{A.2}$$

from equation A.2,

$$U(\theta)=\frac{\xi'(\frac{1}{\mu_s}+\frac{1}{\rho+\delta+d+\theta})-\frac{\mu_s'}{\mu_s^2}(\frac{\Delta}{\rho+\delta+d+\theta}-\xi)}{\frac{\rho}{\mu_b^2}\mu_b'(\frac{1}{\mu_s}+\frac{1}{\rho+\delta+d+\theta})-\frac{\mu_s'}{\mu_s^2}(1+\frac{\rho}{\mu_b})} \tag{A.3}$$

plug this into equation A.1, we get

$$U'(\theta)=-\frac{\frac{U(\theta)(1+\frac{\rho}{\mu_b})+\xi}{(\rho+\delta+d+\theta)^2(\frac{1}{\mu_s}+\frac{1}{\rho+\delta+d+\theta})}+\frac{\Delta}{\rho+\delta+d+\theta}}{1+\frac{\rho}{\mu_b}}<0 \tag{A.4}$$

here we can see that long term buyers get higher value from participating in the bond market, which is consistent with our one bond value function.

Next we total-differentiate equation A.2 with respect to θ, and plug in $U'(\theta)$, rearrange terms we get

$$\lambda'(\theta)=-\frac{\rho\gamma\lambda U'\theta'}{\rho\gamma U-\xi''\eta\lambda^\gamma-\xi'\eta\gamma\lambda^{\gamma-1}}+\frac{(1-\gamma)\left[\frac{\Delta}{\mu_s}-U(1+\frac{\rho}{\mu_b})-\xi\right]}{[\rho\gamma U-\xi''\eta\lambda^\gamma-\xi'\eta\gamma\lambda^{\gamma-1}]\left[\frac{\rho+\delta+d+\theta}{\mu_s}+1\right]} \tag{A.5}$$

A.3 Proof for Proposition 1

To satisfy second order condition, denote the hessian as $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, where

$$a_{11} = -\frac{\left[-\frac{\Delta}{(\rho+\delta+d+\theta)^2} + U''(\theta)(1+\frac{\rho}{\mu_b})\right](\frac{1}{\mu_s} + \frac{1}{\rho+\delta+d+\theta}) - \frac{\frac{\Delta}{\rho+\delta+d+\theta} + U'(\theta)(1+\frac{\rho}{\mu_b})}{(\rho+\delta+d+\theta)^2} + \frac{U'(\theta)(1+\frac{\rho}{\mu_b})+\xi}{(\rho+\delta+d+\theta)^2} - \frac{U'(\theta)(1+\frac{\rho}{\mu_b})}{(\rho+\delta+d+\theta)^2} - \frac{2U(\theta)(1+\frac{\rho}{\mu_b})+\xi}{(\rho+\delta+d+\theta)^3}}{(\frac{1}{\mu_s} + \frac{1}{\rho+\delta+d+\theta})^2}$$

$$a_{12} = a_{21} = -\frac{-U'(\theta)\frac{\rho}{\mu_b^2}\mu_b'(\frac{1}{\mu_s} + \frac{1}{\rho+\delta+d+\theta}) - \frac{\left[\frac{\Delta}{\rho+\delta+d+\theta} + U'(\theta)(1+\frac{\rho}{\mu_b})\right]}{\mu_s^2}\mu_s' + \frac{U(\theta)\frac{\rho}{\mu_b^2}\mu_b' + \xi'}{(\rho+\delta+d+\theta)^2}}{(\frac{1}{\mu_s} + \frac{1}{\rho+\delta+d+\theta})^2}$$

$$a_{22} = -\frac{\frac{\mu_s'}{\mu_s^2}}{\frac{1}{\mu_s} + \frac{1}{\rho+\delta+d+\theta}}\left[\rho\gamma U - \xi''\eta\lambda^\gamma - \xi'\eta\gamma\lambda^{\gamma-1}\right]$$

If second order condition is satisfied, $a_{22} < 0$, i.e.

$$\rho\gamma U - \xi''\eta\lambda^\gamma - \xi'\eta\gamma\lambda^{\gamma-1} < 0. \tag{A.6}$$

In addition, $a_{11}a_{22} - a_{12}^2 \geq 0$. Totally differentiate equation A.1, to get U'', plug it into a_{11}, we can get

$$a_{11}a_{22} - a_{12}^2 \geq 0 \Leftrightarrow \frac{\frac{\Delta}{\mu_s} - U(\theta)(1+\frac{\rho}{\mu_b}) - \xi}{\theta'} \geq 0 \tag{A.7}$$

For positive assortative matching to be true, we need

$$\frac{\Delta}{\mu_s} - U(\theta)(1+\frac{\rho}{\mu_b}) - \xi \geq 0$$

when $\gamma \to 1$ and ξ_0 and ξ_1 are small relative to Δ, this condition is satisfied.

A.4 Proof for Corollary 1

From equation (A.5), we have

$$\lambda'(\theta) = -\frac{\rho\gamma\lambda U'\theta'}{\rho\gamma U - \xi''\eta\lambda^\gamma - \xi'\eta\gamma\lambda^{\gamma-1}} + \frac{(1-\gamma)\left[\frac{\Delta}{\mu_s} - U(1+\frac{\rho}{\mu_b}) - \xi\right]}{\left[\rho\gamma U - \xi''\eta\lambda^\gamma - \xi'\eta\gamma\lambda^{\gamma-1}\right]\left[\frac{\rho+\delta+d+\theta}{\mu_s} + 1\right]}$$

Given that we have positive assortative matching, we have $\frac{\Delta}{\mu_s} - U(1 + \frac{\rho}{\mu_b}) - \xi \geq 0$ (from equation (A.7)). From the other second order condition equation (A.6), we know

$$\rho\gamma U - \xi''\eta\lambda^{\gamma} - \xi'\eta\gamma\lambda^{\gamma-1} < 0$$

Lastly, we proved $U' < 0$ in equation (A.4). As a result, both terms in λ' are negative, i.e. $\lambda' < 0$.

A.5 Proof for Proposition 2

As risk-free rate declines, by implicit function theorem, it is straightforward that $\frac{d\bar{\theta}}{dr_f} < 0$. For each bond, we show this using a finite difference method, discretize the two ODEs to

$$\theta_{i+1} - \theta_i + f_1(\theta_{i+1}, \lambda_{i+1})t_\theta = 0$$
$$\lambda_{i+1} - \lambda_i + f_2(\theta_{i+1}, \lambda_{i+1})t_\lambda = 0$$

where t_θ is the grid size of θ and t_λ is the grid size of λ. Order the variables as $X = (\theta_N, \theta_{N-1}, ..., \theta_1, \lambda_N, \lambda_{N-1}, ..., \lambda_2)$. λ_1 can be found using θ_2 and λ_2, I leave it out of the system here. So the system of equation can be defined as

$$GX = \mathbf{0}$$

where

$$G = \begin{pmatrix} 1 & 0 & 0 & ... & 0 & 0 & 0 & ... & 0 \\ 1+a & -1 & 0 & ... & 0 & b & 0 & ... & 0 \\ 0 & 1+a & -1 & ... & 0 & 0 & b & ... & 0 \\ & & & ... & & & & ... & \\ 0 & 0 & .... & 1+a & -1 & 0 & 0 & ... & b \\ & & & & & & & & \\ c & 0 & ... & 0 & 0 & 1+d & -1 & ... & 0 \\ 0 & c & ... & 0 & 0 & 0 & 1+d & -1 & ... \\ & & ... & & & & ... & & \\ 0 & 0 & .... & c & 0 & 0 & ... & 0 & 1+d \end{pmatrix}$$

$$a = -t_\theta \frac{df_1(\theta, \lambda)}{d\theta} \qquad b = -t_\theta \frac{df_1(\theta, \lambda)}{d\lambda} \qquad c = -t_\lambda \frac{df_2(\theta, \lambda)}{d\theta} \qquad d = -t_\lambda \frac{df_2(\theta, \lambda)}{d\lambda}$$

The upper left block is $N \times N$, denote it as A; the upper right block is $N \times (N-1)$, denote is as B; the lower left block is $(N-1) \times N$, denote it as C; the lower right block is D, $(N-1) \times (N-1)$.

To show that $\overline{\theta}$ is decreasing in r_f, consider

$$\frac{dX}{d\overline{\theta}} = -G^{-1} \begin{pmatrix} -1 \\ 0 \\ ... \\ 0 \end{pmatrix}.$$

We just have to show the first column of G is positive. By block inversion, we need to show first columns of $(A - BD^{-1}C)^{-1}$ and $-D^{-1}C(A - BD^{-1}C)^{-1}$ are positive.

$$D^{-1} = \begin{pmatrix} \frac{1}{1+d} & \frac{1}{(1+d)^2} & ... & 0 \\ 0 & \frac{1}{1+d} & ... & 0 \\ & & ... & \\ 0 & 0 & ... & \frac{1}{1+d} \end{pmatrix}, \qquad D^{-1}C = \begin{pmatrix} \frac{c}{1+d} & \frac{c}{(1+d)^2} & ... & 0 & 0 \\ 0 & \frac{c}{1+d} & ... & 0 & 0 \\ & & ... & & 0 \\ 0 & 0 & ... & \frac{c}{1+d} & 0 \end{pmatrix}$$

$$BD^{-1}C = \begin{pmatrix} 0 & 0 & ... & 0 & 0 \\ \frac{bc}{1+d} & \frac{bc}{(1+d)^2} & ... & 0 & 0 \\ 0 & \frac{bc}{1+d} & ... & 0 & 0 \\ & & ... & & 0 \\ 0 & 0 & ... & \frac{bc}{1+d} & 0 \end{pmatrix} \quad A - BD^{-1}C = \begin{pmatrix} 1 & 0 & ... & 0 & 0 \\ 1 + a - \frac{bc}{1+d} & -1 - \frac{bc}{(1+d)^2} & ... & 0 & 0 \\ 0 & 1 + a - \frac{bc}{1+d} & ... & 0 & 0 \\ & & ... & & 0 \\ 0 & 0 & ... & 1 + a - \frac{bc}{1+d} & -1 \end{pmatrix}$$

the first column of its inverse is $\begin{pmatrix} 1 & x_2 & x_3 & ... & x_N \end{pmatrix}^\top$, where

$$1 + a - \frac{bc}{1+d} - (1 + \frac{bc}{(1+d)^2})x_2 = 0, \qquad (1 + a - \frac{bc}{1+d})x_{N-1} - x_N = 0$$
$$(1 + a - \frac{bc}{1+d})x_i - (1 + \frac{bc}{(1+d)^2})x_{i+1} \qquad \text{for} \quad i = 2..., N-2.$$

a, b, c, d are small, so $x_i > 0$ for $2 \leq i \leq N$. Next,

$$-D^{-1}C(A - BD^{-1}C)^{-1} = -\begin{pmatrix} \frac{c}{1+d} & \frac{c}{(1+d)^2} & \dots & 0 & 0 \\ 0 & \frac{c}{1+d} & \dots & 0 & 0 \\ & & \dots & & 0 \\ 0 & 0 & \dots & \frac{c}{1+d} & 0 \end{pmatrix} \begin{pmatrix} 1 & \dots & \dots \\ x_2 & \dots & \dots \\ & \dots & \\ x_N & \dots & \dots \end{pmatrix}.$$

we just need to show $c < 0$, where $c = -\frac{df_2(\theta,\lambda)}{d\theta}$. One can show that when Δ is large, this is indeed the case.

A.6 Proof for Lemma 3

Interest rate is pinned down by $V_b = V_0 - 1$, where $V_b = U(\theta)$ and

$$\begin{aligned} \rho V_s &= r - \Delta + \delta(1 - V_s) + d(s - V_s) + U_s \\ \rho V_0 &= r + \delta(1 - V_0) + d(s - V_0) + \theta(V_s - V_0) \\ r &= \rho + d(1 - s) + (\rho + \delta + d)U + \theta \frac{\Delta - U_s}{\rho + \delta + d + \theta} \\ &= \rho + d(1 - s) + (\rho + \delta + d)U + \frac{\mu_s \theta \left[\frac{\Delta}{\mu_s} + U(1 + \frac{\rho}{\mu_b}) + \xi\right]}{\mu_s + \rho + \delta + d + \theta} \end{aligned}$$

A.7 Proof for Proposition 3

To look at $\frac{dr}{d\xi_0}$,

$$\begin{aligned} \frac{dr}{d\xi_0} =& \frac{\mu_s \theta}{\mu_s + \rho + \delta + d + \theta} + \left[\rho + \delta + d + \frac{\mu_s \theta (1 + \frac{\rho}{\mu_b})}{\mu_s + \rho + \delta + d + \theta}\right] \frac{dU}{d\xi_0} \\ =& \frac{\mu_s \theta}{\mu_s + \rho + \delta + d + \theta} + \left[\rho + \delta + d + \frac{\mu_s \theta (1 + \frac{\rho}{\mu_b})}{\mu_s + \rho + \delta + d + \theta}\right] \frac{\frac{\mu_s'}{\mu_s^2}}{\frac{\rho}{\mu_b^2}\mu_b'(\frac{1}{\mu_s} + \frac{1}{\rho+\delta+d+\theta}) - \frac{\mu_s'}{\mu_s^2}(1 + \frac{\rho}{\mu_b})} \end{aligned}$$

as $\gamma \to 1$, $\mu_s' \to 0$, $\frac{dr}{d\xi_0} \to \frac{\mu_s \theta}{\mu_s + \rho + \delta + d + \theta}$

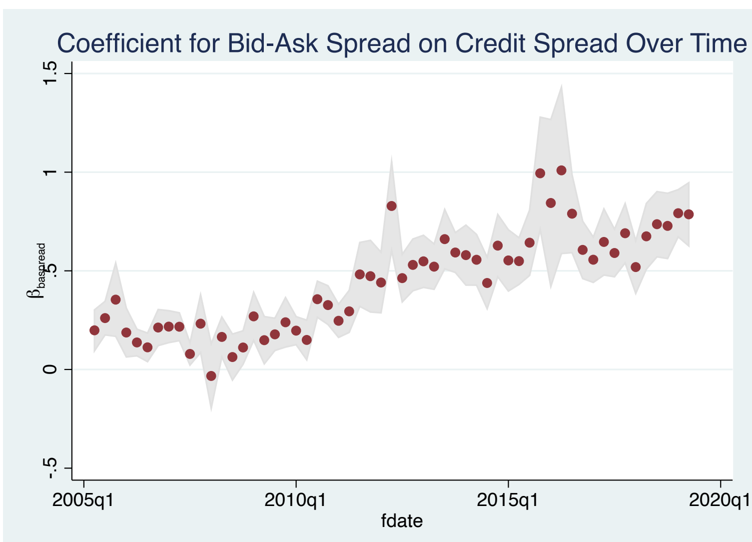
Appendix B Controlling for Credit Default Swaps

In this appendix, we use credit default swaps (CDS) to address the concern that change in default risk and risk appetite may be driving our results. We show that both our time series result and cross-sectional result survive even if we subset to the bonds with CDS traded and control for CDS spreads. This cuts our sample by half.

The CDS data is from Markit and we only select CDS spreads for US dollar contract, refers to senior unsecured debt and with documentation clause type "No Restructuring". We use linear interpolation the bond's CDS based on issuer's CDS of 0.5, 1, 2, 3, 4, 5, 7, 10, 15, 20 and 30 years.

We first show the aggregate trend in Section 2 survives using the subset of the corporate bonds that have a corresponding CDS. In addition to the firm and bond characteristics in Equation (2.1), we also control for CDS spreads to better approximate the default component. The results are shown in Figure 10.

Figure 10: Regression Coefficients of Bid/Ask Spread Over-time



We regress credit spread on bid-ask spreads quarter by quarter for bonds with CDS, controlling for bond and firm characteristics, as well as the CDS spread. We then plot the coefficients over time in red. The shaded region indicates the 95% confidence interval. Data comes from WRDS Bond Return and Mergent FISD

In addition, our cross-sectional result in Section 5 also holds once we control for CDS spreads. We repeat the analysis in Equation 5.4, adding CDS spreads as a control. The results are shown in Table 10. We still find the coefficient for the interaction of investor turnover and bid-ask spreads positively significant, consistent with the results in the main text.

Table 10: Credit Spread Regression for Corporate Bonds with CDS

	Credit spread	Credit spread	Credit spread
CDS	0.694***	0.689***	0.680***
	(0.0392)	(0.0399)	(0.0401)
Bid-ask Spread	0.112*	0.109*	-0.142*
	(0.0462)	(0.0478)	(0.0576)
Investor Comp		-0.00286	-0.0181***
		(0.00414)	(0.00472)
Bid-ask Spread × Investor Comp			3.995***
			(0.708)
Coupon	0.000282**	0.000281*	0.000250*
	(0.000103)	(0.000106)	(0.000107)
Offering amount	3.89e-10*	4.02e-10*	4.18e-10*
	(1.90e-10)	(1.94e-10)	(1.94e-10)
N	78354	75612	75612
adj. R^2	0.827	0.827	0.829

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table presents the regression results for equation 5.4 for the subsample of corporate bonds with valid CDS. "Investor Comp" is the measure of investor composition, higher value indicates the investor have higher frequency of liquidity shock, i.e. shorter term investors. We also control for bond age, firm leverage, fraction of long-term debt, equity volatility and profitability, where equity volatility is the standard deviation of the equity return of the issuer in that quarter.

Internet Appendix

Appendix C Micro-founding Dealer Sector

This appendix micro-founds the dealer sector and verifies that our main results go through. We model the dealer sector explicitly: there are two sets of directed search markets: one in which the dealers and sellers trade and the second one in which the dealers and buyers trades. The two markets are running parallel with each other. For simplicity, we assume default rates are 0 for all bonds, and only consider the heterogeneity in maturity rate. It is easy to extend the framework to consider heterogeneous default rates.

C.1 Seller-Dealer Markets

Denote the measure of sellers of bond j: measure $\alpha_s(\delta_j)$, value function $V_{s,j}$. Similar to before, the value function of sellers is

$$\rho V_{s,j} = r_j - \Delta + \delta_j(1 - V_{s,j}) + \max_{\lambda_1, p_1} \mu_s(\lambda_1)(p_1 - V_{s,j})$$

where r_j is the interest rate paid on the bond and δ_j is the maturity rate. The sellers problem is to choose a sub-market with tightness λ_1 and price p_1 to maximize expected selling value.

Dealers: Each period, measure D units of dealers flow into the economy. Measure g_j dealers enter the market with bond j each period, hence

$$\int_j g_j dj = D$$

Each dealer can at most hold one unit of bond. Dealers who do not hold any bond participate in the market with sellers; after obtaining the bond, the dealers participate in the market with buyers. We denote the dealers with no bond holdings, who choose to participate in the market of bond j, by $d_{0,j}$; and those with bond holdings as $d_{1,j}$. Dealers incur a flow cost c_j when holding bond j on their balance sheets. The corresponding measures are $\alpha_{d_0}(\delta_j)$ and

$\alpha_{d_1}(\delta_j)$ respectively; and the value functions are denoted by $V_{d_0,j}$ and $V_{d_1,j}$ respectively.

$$\rho V_{d_0,j} = \max_{(p_1,\lambda_1,j)\in G_1} \mu_b(\lambda_1)(V_{d_1,j} - V_{d_0,j} - p_1)$$

In equilibrium, $V_{d_0,j}$ needs to be equalized across j, so that dealers have no incentive to flow to other submarkets. We denote equilibrium dealer value by $\rho V_{d_0,j} = u$.

Value function for dealer holding bond j is $V_{d_1,j}$,

$$\rho V_{d_1,j} = r_j - c_j + \delta_j(1 - V_{d_1,j}) + \max_{\lambda_2,p_2} \mu_s(\lambda_2)(p_2 - V_{d_1,j})$$

For the baseline case, we assume c_j is the same across j.

On the seller-dealer markets, all the dealers are homogeneous

$$\lambda_1 = \frac{\alpha_s}{\alpha_{d_0}} \tag{C.1}$$

The first order condition from seller's maximization problem is

$$\begin{aligned} 0 = &-\frac{\mu_s'(\lambda_1)}{\rho+\delta+\mu_s(\lambda_1)}\left[(\rho+\delta)\left(\frac{r-c+\delta+\mu_s(\lambda_2)p_2}{\rho+\delta+\mu_s(\lambda_2)} - \frac{\rho+\mu_b(\lambda_1)}{\rho\mu_b(\lambda_1)}u\right) - (r-\Delta+\delta)\right] \\ &-\mu_s(\lambda_1)\left(\frac{\mu_b'(\lambda_1)}{\mu_b^2(\lambda_1)}u + \frac{\mu_s(\lambda_2)}{\rho+\delta+\mu_s(\lambda_2)}\frac{dp_2}{d\lambda_1}\right) \end{aligned} \tag{C.2}$$

C.2 Buyer-Dealer Markets

Similar as in the main text, value function for end buyer of type θ is $V_b(\theta)$,

$$\rho V_b(\theta) = \max_{(p_2,\lambda_2,j)\in G_2} \mu_b(\lambda_2)(V_{0,j}(\theta) - V_b(\theta) - p_2)$$

Investor type θ holding bond j, value function $V_{0,j}(\theta)$,

$$\rho V_{0,j}(\theta) = r_j + \delta_j(1 - V_{0,j}) + \theta(V_{s,j} - V_{0,j})$$

The buyer-dealer market is closely related to the seller-buyer market we analyze in the main text. The first ODE is

$$\theta'(\delta) = \frac{\alpha_{d_1}}{\alpha_b \lambda_2} \tag{C.3}$$

To derive the second ODE, consider the dealer's problem

$$\max_{\lambda_2, \theta} F(\delta, \theta, \lambda_2) - G(\delta, \theta, \lambda_2) h(\theta) \tag{C.4}$$

where

$$\begin{aligned}
G &= \frac{(\rho + \delta + \theta)\frac{\mu_b(\lambda_2)+\rho}{\rho\mu_b(\lambda_2)}}{Q} \\
Q &= \frac{\theta}{\rho + \delta + \mu_s(\lambda_1)} + 1 + \frac{\theta + \rho + \delta}{\mu_s(\lambda_2)} \\
F &= \frac{c + \frac{\theta}{\rho+\delta+\mu_s(\lambda_1)}[c - \Delta - u\mu_s(\lambda_1)(\frac{1}{\mu_b(\lambda_1)} + \frac{1}{\rho})]}{Q}
\end{aligned}$$

First order conditions from the maximization problem are

$$[\lambda_2]: \quad F_3 - G_3 h(\theta) = 0 \tag{C.5}$$

$$[\theta]: \quad F_2 - G_2 h - G h' = 0 \tag{C.6}$$

Second order condition

$$\begin{pmatrix} F_{33} - G_{33}h & F_{32} - G_{32}h - G_3 h' \\ F_{32} - G_{32}h - G_3 h' & F_{22} - G_{22}h - 2G_2 h' - G h'' \end{pmatrix}$$

Total differentiate equation C.6,

$$\begin{aligned}
&F_{21} + F_{23}\lambda_2' - G_{21}h - G_{23}h\lambda_2' - G_3 h'\lambda_2' + (F_{22} - G_{22}h - 2G_2 h' - G h'')\theta' = 0 \\
&F_{22} - G_{22}h - 2G_2 h' - G h'' = -\frac{1}{\theta'}[F_{21} + F_{23}\lambda_2' - G_{21}h - G_{23}h\lambda_2' - G_3 h'\lambda_2']
\end{aligned}$$

Total differentiate equation C.5 and reorganize the terms we get the second ODE,

$$\lambda_2' = -\frac{F_{31} - G_{31}\frac{F_3}{G_3} + (F_{32} - G_{32}h - G_3 h')\theta'}{F_{33} - G_{33}h} \tag{C.7}$$

Together with the boundary conditions

$$\theta(\underline{\delta}) = \underline{\theta} \qquad \theta(\overline{\delta}) = \overline{\theta} \qquad V_b(\theta) = V_f \tag{C.8}$$

we can solve the equilibrium in the second set of markets.

Lastly, the dynamics are characterized by the following system of equations,

$$\begin{aligned}
\dot{\alpha}_s &= -\mu_s(\lambda_1)\alpha_s - \delta_j\alpha_s + \theta\alpha_h \\
\dot{\alpha}_h &= -(\theta + \delta_j)\alpha_h + \mu_b(\lambda_2)\alpha_b + \delta_j \\
\dot{\alpha}_b &= m_I f(\theta)\theta' - \delta_j - \mu_b(\lambda_2)\alpha_b \\
\dot{\alpha}_{d_0} &= g(\delta) - \mu_b(\lambda_1)\alpha_{d_0} \\
\dot{\alpha}_{d_1} &= -(\mu_s(\lambda_2) + \delta_j)\alpha_{d_1} + \mu_b(\lambda_1)\alpha_{d_0}
\end{aligned}$$

where α_h is the measure of patient investors holding bonds. We focus on the stationary equilibrium, where the measures of different types of investors and dealers are constant.

To endogenize the offering yield r, participating in the secondary market and primary market must be the same

$$V_0(\theta) - 1 = V_b(\theta)$$

We define bid-ask spreads in percentage term

$$BA(\delta) = \frac{p_2 - p_1}{\frac{p_1 + p_2}{2}}$$

C.3 Results

Using Equation (C.1) and (C.2) from the seller-dealer market, and Equation (C.3) and (C.7) from the buyer-dealer market, together with the boundary conditions in (C.8), we can solve the system numerically. We verify the following three results still hold

- In equilibrium, the matching between investors and bonds features positive assortative matching $\theta'(\delta) > 0$
- When risk free rate drops, more short-term investors enter the market $\frac{d\bar{\theta}}{dr_f} < 0$
- As short-term investors enter the market, the bid-ask spreads measured in percentage term decrease. Moreover, when we regress credit spreads on bid-ask spreads controlling for bond maturity rate, the coefficient in front of bid-ask spreads increases as risk free rate decreases.

Appendix D Alternative Measure of Illiquidity

In this appendix, we construct an alternative measurement of illiquidity and repeat our empirical analysis in Subsection 5.1. We first compute Amihud measure, Roll measure, imputed round-trip cost (IRC), number of zero trading days for the firm and the bond, average bid-ask spreads, standard deviation of Amihud measure, IRC and bid-ask spreads for each bond at the quarterly level. Then we conduct principle component analysis (PCA).

The principle component analysis of Table 11 indicates that the first component, which mainly consists of the price measures, explains nearly half of the variation in liquidity, while the second and third components are closer to quantity measure, and are almost orthogonal to the first component. Following the literature, we take the first component as the measure of illiquidity for bonds. Moreover, as Table 12 shows, the price measures are highly correlated. The reason for conducting PCA is to get a more accurate measure of the underlying frictions. Lastly, all the liquidity measures are standardized over time.[30] We denote this measure as $\lambda_{i,t}$. Higher value of $\lambda_{i,t}$ means less liquid. Figure 11 plots our measure of illiquidity over time, the illiquidity peaks during the financial crisis. Overall, bonds with higher ratings are more liquid.

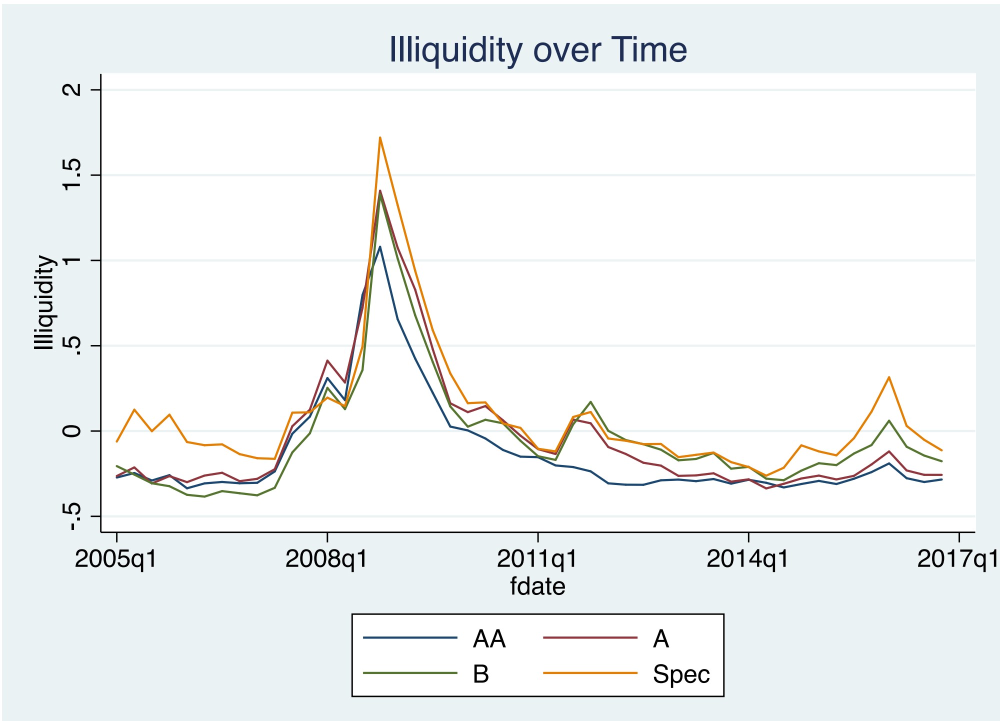
In terms of aggregate trends, Figure 12 shows the liquidity component for investment grade and speculative grade bonds over time. Despite high volatility, the liquidity components as a fraction of credit spreads have been increasing, at least for the post-crisis period.

Next, we move on to the cross-sectional analysis. First of all, Table 13 shows the sorting result of investor types and the sub-markets. The column of time to maturity is the same as before. We replace the bid-ask spreads with the new measure of illiquidity. We find the same result as before: sub-markets with more short-term investors have higher liquidity. Figure 13 is the counterpart of Figure 7a, replacing the bid-ask spread the alternative measure of illiquidity. We sort the bonds by their investor composition into five groups, the summary statistics for each group of bonds is shown in Table 14. We then regress the credit spreads on the new measure of illiquidity group by group, controlling for bond and firm characteristics. We obtain the same result as before, i.e. the sensitivity of credit spreads to illiquidity is higher for bonds with more short-term investors. We present the regression analysis in Table 15. The key point is that the interaction term of illiquidity and investor composition is highly

[30] This means we subtract away the sample mean and divide by the sample standard deviation.

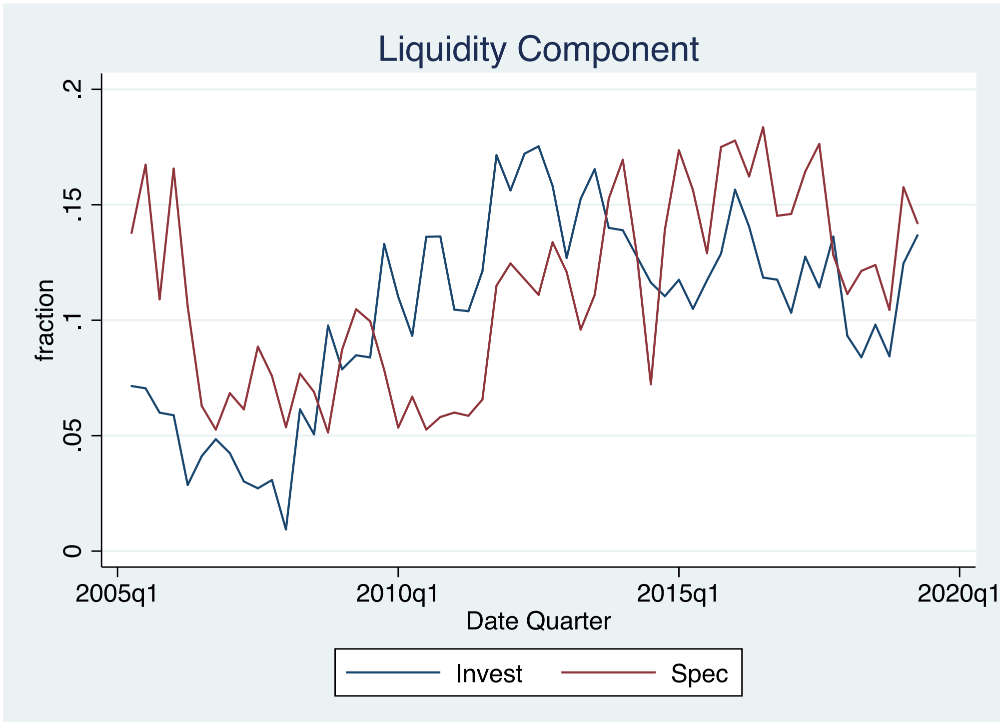
significant in determining the credit spreads. We repeat the analysis for offering yields in Table 16.

Figure 11: Illiquidity by Rating over Time



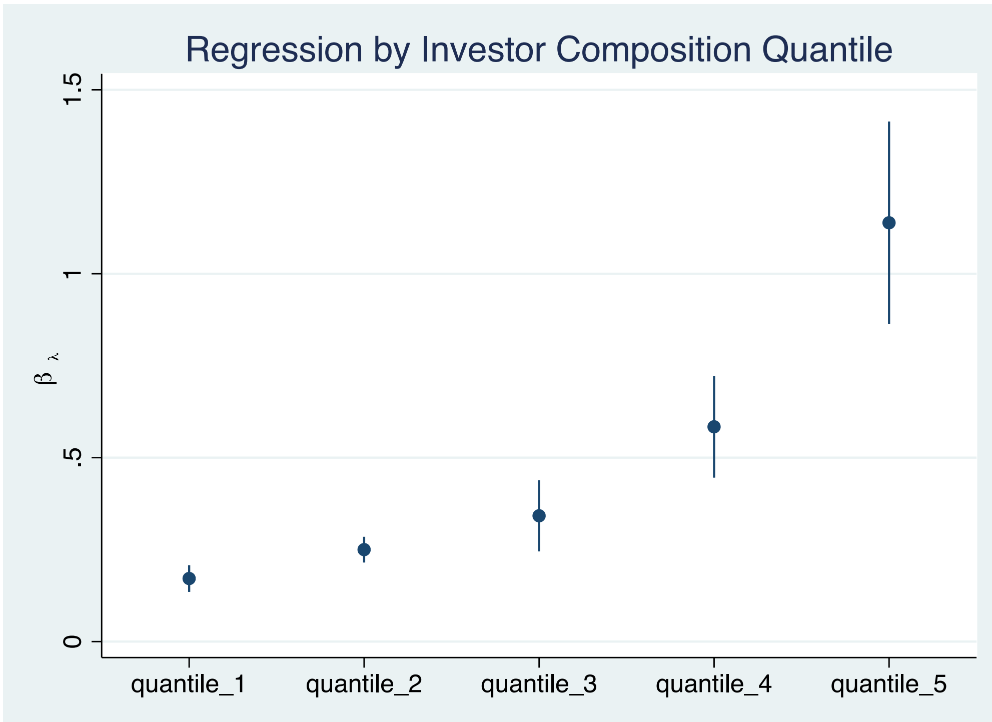
The figure plots quarterly measure of illiquidity from 2005Q1 and 2016Q4. We construct the measure of illiquidity by taking the first principle component of the common liquidity measures in the literature. The corporate bond price data is from TRACE and rating data is from Mergent FISD.

Figure 12: Liquidity Component as a Fraction of Credit Spreads



The figure plots the liquidity component of investment grade and speculative grade bonds over time. For each quarter t, first run 5.4 to get the coefficient of the alternative illiquidity measure β. Bond characteristic data is from Mergent FISD and WRDS Bond Returns, the price data is from TRACE. Next for each bond, define $liquidity\_component_{i,t} = \frac{\beta_t \times illiquidity_{it}}{credit\_spread_{it}}$. The figure plots the median liquidity component of investment grade and speculative grade bonds over time.

Figure 13: Sensitivity of Credit Yields to Alternative Measure of Illiquidity by Investor Quantile



We sort the bonds (in each quarter from 2005Q2 to 2019Q2) into five groups, by investor base turnover. Group 1 contains bonds whose investors have the lowest turnover rate and group 5 contains bonds whose investors have the highest turnover rate. The figure plots the regression coefficients and 1% confidence interval of credit spread regressed on the alternative illiquidity measure, controlling for bond and firm characteristics, group by group.

Table 11: Principle Component Loadings on the Liquidity Variables

	Comp 1	Comp 2	Comp 3
Roll's measure	.306816	.1680553	.1158815
Amihud measure	.3537716	.007733	-.221156
Amihud std	.3772206	-.0198852	-.1215953
IRC	.4141384	.0365575	.2021284
IRC std	.3914108	-.0754453	.1529368
Bid-ask spread	.3995355	.0639768	-.155747
Bid-ask spread std	.388697	-.0459457	-.0664872
Firm with zero trading days	-.007644	.6181877	.3801652
Bond with zero trading days	-.0283332	.6705573	.0729346
Turnover	.0461081	-.35534	.8246865
Cum Explained	0.467	0.644	0.744

The table reports the results of Principle Component Analysis (PCA) of the common liquidity measure used in Dick-Nielsen, Feldhütter, and Lando (2012). We construct each measures at the quarter level using enhanced TRACE then conduct the PCA. The sample period is 2005Q1 to 2015Q4

Table 13: Sorting of Investors

	Time to maturity	Illiquidity
Investor comp	-49.57***	-1.666***
	(3.884)	(0.133)
Age	-0.592***	0.0230***
	(0.0591)	(0.00165)
Coupon	2.126***	-0.0358***
	(0.154)	(0.00471)
Offering amount	-0.000000431	-5.14e-08***
	(0.000000254)	(6.58e-09)
N	179965	170957
adj. R^2	0.200	0.345

Standard errors in parentheses: sym* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table presents correlation between investor composition and the maturity of bonds they hold, and the illiquidity in each sub-market. "Investor comp" is the measure of investor composition, higher value indicates the investor have higher frequency of liquidity shock, i.e. shorter term investors. We also control for bond age, firm leverage, fraction of long-term debt, equity volatility and profitability, where equity volatility is the standard deviation of the equity return of the issuer in that quarter.

Table 12: Correlation of the Iliquidity Measures

	Illiquidity	BA spread	BA spread std	Amihud measure	Amihud measure std	IRC	IRC std	Roll's measure	0 trading days (firm)	0 trading days (bond)	Turnover
Illiquidity	1										
BA spread	0.809***	1									
BA spread std	0.824***	0.783***	1								
Amihud measure	0.720***	0.584***	0.497***	1							
Amihud measure std	0.811***	0.552***	0.600***	0.569***	1						
IRC	0.852***	0.593***	0.582***	0.618***	0.626***	1					
IRC std	0.817***	0.517***	0.580***	0.438***	0.657***	0.853***	1				
Roll's measure	0.607***	0.424***	0.440***	0.315***	0.356***	0.430***	0.386***	1			
0 trading days (firm)	0.0166***	0.0101***	-0.0889***	0.0567***	-0.0556***	0.0780***	-0.0479***	0.0940***	1		
0 trading days (bond)	0.0238***	0.0384***	-0.0964***	0.0934***	-0.0341***	0.0373***	-0.0877***	0.122***	0.573***	1	
Turnover	-0.0263***	-0.0772***	0.0491***	-0.123***	-0.0312***	0.0505***	0.0870***	-0.0522***	-0.217***	-0.484***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The table reports the correlation matrix of various illiquidity measures used in Dick-Nielsen, Feldhütter, and Lando (2012). We construct each measures at the quarter level using enhanced TRACE then conduct the PCA. The sample period is 2005Q1 to 2015Q4

Table 14: Summary Statistics by Illiquidity Quantile

	(1)		(2)		(3)		(4)		(5)	
	mean	std	mean	std	mean	std	mean	std	mean	std
Investor comp	0.091	0.050	0.094	0.050	0.088	0.048	0.081	0.045	0.076	0.044
Bid-ask spread	0.001	0.001	0.002	0.001	0.003	0.002	0.005	0.003	0.009	0.007
Alternative measure of illiquidity	-0.543	0.133	-0.366	0.189	-0.187	0.265	0.081	0.382	0.879	0.902
Credit spread	0.017	0.020	0.021	0.021	0.021	0.021	0.023	0.024	0.031	0.033
Quarterly transaction volume (million USD)	111	202.5	160.7	299.4	180.1	304.2	163.7	265.0	130.7	251.0
Time to Maturity (30/360 Convention)	4.776	4.906	7.351	6.808	10.352	9.207	11.845	9.719	13.110	10.543
Interest rate (%)	5.350	2.171	5.541	2.043	5.507	1.871	5.479	1.759	5.742	1.672
Fraction of AAA-A	0.332	0.471	0.309	0.462	0.366	0.482	0.382	0.486	0.319	0.466
Fraction of BBB	0.458	0.498	0.385	0.487	0.376	0.484	0.412	0.492	0.453	0.498
Fraction of Speculative Grade	0.210	0.407	0.306	0.461	0.257	0.437	0.207	0.405	0.228	0.420
Observations	35870		37942		38262		37647		34404	

Bond-Quarter level summary statistics by illiquidity (λ) quantile. For each quarter, we sort the bond into five groups based on their liquidity. Group 1 are most liquid bonds, group 5 are most illiquid. A-AAA is a dummy indicating the bond has rating A-AAA. Similar for BBB and Speculative.

Table 15: Credit Spread Regression on Alternative Measure of Illiquidity

	Credit spread	Credit spread	Credit spread
Alternative measure of illiquidity	0.00356***	0.00349***	-0.00202***
	(0.000444)	(0.000457)	(0.000419)
Investor comp		-0.0185*	-0.0138
		(0.00712)	(0.00899)
Illiquidity × Investor comp			0.0804***
			(0.00724)
Time to maturity	0.0000687***	0.0000568**	0.0000616**
	(0.0000190)	(0.0000208)	(0.0000205)
Age	-0.0000108	-0.0000546	-0.0000233
	(0.0000589)	(0.0000487)	(0.0000481)
Coupon	0.000900***	0.000895***	0.000829***
	(0.000173)	(0.000174)	(0.000174)
Offering amount	6.01e-10	6.24e-10	3.76e-10
	(3.17e-10)	(3.13e-10)	(2.84e-10)
N	118058	110060	110060
adj. R^2	0.764	0.768	0.779

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table presents the regression results for equation 5.4. We use the illiquidity measure as in Dick-Nielsen, Feldhütter, and Lando (2012)and "Investor Comp" is the measure of investor composition, higher value indicates the investor have higher frequency of liquidity shock, i.e. shorter term investors. We also control for bond age, firm leverage, fraction of long-term debt, equity volatility and profitability, where equity volatility is the standard deviation of the equity return of the issuer in that quarter.

Table 16: Regression of Offering Spreads on Alternative Measure of Illiquidity

	Offering yield	Offering yield	Offering yield	Bond maturity
Alternative measure of illiquidity	-0.000252	-0.000114	-0.000898	2.429***
	(0.000445)	(0.000409)	(0.000704)	(0.456)
Investor comp		0.0295***	0.0314***	-28.94***
		(0.00237)	(0.00272)	(4.175)
Illiquidity × Investor comp			0.00872	
			(0.00756)	
Coupon	0.00641***	0.00666***	0.00663***	4.219***
	(0.000323)	(0.000291)	(0.000299)	(0.271)
Offering amount	-5.67e-11	-2.59e-10*	-2.59e-10*	-0.000000195
	(1.10e-10)	(1.26e-10)	(1.25e-10)	(0.000000168)
N	5480	5307	5307	6085
adj. R^2	0.905	0.911	0.911	0.383

Standard errors in parentheses: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression results for Equation 5.4 where the left hand variable is replaced by offering spreads. Higher value of investor composition indicates the investor have higher frequency of liquidity shock. Equity volatility is the standard deviation of the equity return of the issuer in that quarter.