Market Accessibility, Corporate Bond ETFs, and Liquidity

Jayoung Nam*
Kelley School of Business, Indiana University Bloomington

June, 2017

Abstract

I provide evidence that market accessibility *ex ante* plays an important role in how the underlying assets' liquidity changes when a basket security is introduced. First, using a multi-market version of the Kyle model, I show that if the underlying market is less accessible, then trading basket securities improves liquidity in the underlying market. In contrast, if the underlying market is more accessible, liquidity deteriorates. Second, I test the theoretical predictions using data on corporate bonds before and after the introduction of corporate bond ETFs. I find that in contrast to the stock market, the inception of corporate bond ETFs improves the liquidity of the underlying bonds. This liquidity improvement is larger for low volume, high yield, and long term bonds and for 144A bonds to which access was previously difficult for retail investors.

JEL classification: G14; G19

Keywords: Liquidity; Market accessibility; Exchange traded funds; Corporate bonds

---

*I thank Craig W. Holden, Zhenyu Wang, Veronika K. Pool, Noah Stoffman, Scott Yonker, Eitan Goldman, Azi Ben-Rephael, Charles Trzcinka, Ryan D. Israelsen, Wenyu Wang, Sreenivas Kamma, Matthew T. Billett, Batchimeg Shambalaibat, Jun Yang, Wei Yang, as well as seminar participants at Indiana University Kelley School of Business, for helpful suggestions. I thank Jonathan Reuter, Vincent Glode, Uday Rajan, and Robert Battalio for the encouragement. My paper has improved significantly with many comments from S. "Vish" Viswanathan (the discussant), Utpal Bhattarcharya, Philip Strahan and the participants at the FIRS 2017 meeting. *Contact address*: Jayoung Nam, Indiana University Bloomington, Kelley School of Business, 1309 E. 10th Street, Bloomington, IN 47405; *E-mail address*: jnam@indiana.edu.

1. Introduction

Prior literature on the introduction of basket securities shows that the liquidity of the underlying assets deteriorates because uninformed investors migrate to the basket market in order to reduce the cost of trading against informed traders. In this paper, I show that if investor participation in the underlying market is limited, the liquidity of the underlying securities can, in fact, *improve* after basket securities begin trading. To do so, I first develop a theoretical model to explain that the level of market participation *ex ante* is an important determinant of the liquidity of the underlying securities *ex post*. Second, I empirically test the predictions of the model by investigating the liquidity of corporate bonds before and after the introduction of corporate bond exchange traded funds (ETFs) using bond transaction data from TRACE.

Theories of basket securities trading typically rely on the Kyle model (Kyle, 1985), in which prices partially reflect informed investors' information. In the model, liquidity trading demand is independent of the private information associated with individual securities. Therefore, liquidity investors are concerned with their expected losses from trading due to adverse selection. Since basket securities diversify away the idiosyncratic risks of the underlying securities, these liquidity investors can reduce the expected loss of trading in basket securities. Hence discretionary liquidity investors migrate to the basket market and consequently, liquidity in the underlying asset market declines. (Subrahmanyam, 1991; Gorton and Pennacchi, 1993; Israeli, M. C. Lee, and Sridharan, 2015; Hamm, 2014; Jegadeesh and Subrahmanyam, 1993)

Largely overlooked is the fact that the basket market and the underlying market may not be equally accessible for liquidity investors. A large body of literature has shown that investors' market participation can be limited due to unfamiliarity (Merton, 1987; Huberman, 2001), home bias (Kang and Stulz, 1997; Pool, Stoffman, and Yonker, 2012), lack of information, high transaction costs or minimum investment requirements, difficulty in liability matching, or regulatory constraints (Ellul, Jotikasthira, and Lundblad, 2011). This implies that for liquidity investors, diversification can be quite costly. As a specific example, for retail investors, the transaction cost associated with trading an individual junk bond averages 110 basis points with a range from 10 basis points to more than 200 basis points (Figure 1). Furthermore, the minimum trading size for individual bonds is \$100,000 at a major brokerage firm. In sharp contrast, with the introduction of ETFs, investors can trade an entire portfolio at a cost of 1 basis point with an investment as small as \$1,000. That is, the cost of diversification is more than 100 times lower than it was before.

To derive theoretical predictions, I introduce a multi-market version of the theoretical models of Kyle (1985) and Subrahmanyam (1991) in which different markets have different levels of accessibility. I show that the introduction of basket securities in a market with limited investor participation leads to an improvement in the liquidity of the underlying securities. An important mechanism that links limited market participation to liquidity is the entrance of new uninformed investors into the basket market: they can achieve a better-diversified portfolio at the low cost available through basket securities. The high liquidity of the basket securities is spilled over to the underlying market by arbitrage trading (Holden, 1995). This liquidity spillover is greater if the influx of these new uninformed investors outweighs the departure of uninformed investors from the underlying market, which is likely to occur when the underlying market is less accessible. Conversely, if the underlying market *ex ante* is highly accessible, then the departure of uninformed investors from the underlying market exceeds the entrance of new investors and the liquidity of the underlying securities deteriorates.

Testing whether trading basket securities in a market with limited participation improves or hurts liquidity of the underlying securities requires that a proper financial instrument be available. Existing theoretical and empirical studies have focused on the equity market and unanimously find only a one-directional effect: the deterioration of liquidity of the underlying markets with the introduction of basket security trading. These results are consistent with one of the predictions of my model since the underlying equity market is already well accessible to investors.

Yet few studies have been conducted outside the equity market, for example in the bond market, due to limited availability of the instrument. One possibility would be open-end bond mutual funds, which, like bond ETFs, allow investors to achieve a diversified portfolio at a smaller cost than investing in a portfolio of individual bonds, and have long been available. However, bond mutual funds are not traded in exchanges as single securities and the price of bond mutual fund shares is available only after the market is closed. Thus, arbitrage mechanism between bond mutual funds and the underlying individual bonds is less likely to occur.

In this paper, I test the predictions of my model using TRACE intraday trading data for corporate bonds and ETFs since 2007[1]. Since corporate bonds do not have pre-trade quotes, I estimate the effective spread of corporate bonds by using the indicator-based regression methodology that was initially developed by Huang and Stoll (1997) and was later adapted to corporate bonds by Bessembinder, Maxwell, and Venkataraman (2006). Using a difference-in-difference (DD) framework, I compare the changes of the effective spread of corporate bonds before and after they are included in an ETF for the first time. My control sample consists of corporate bonds that are not included in ETFs.

I find that non-investment grade ETF bonds experience improved liquidity (i.e. a reduction in transaction costs) of 10%, while investment grade ETF bonds experience an insignificant change in trading costs. ETF bonds with low volume see improved liquidity once the ETF starts trading, while those with large volume see decreased liquidity. ETF bonds whose remaining life is greater than 2 years see improved liquidity. As an evidence for the underlying economic mechanism, I estimate the degree of arbitrage activities between the ETF market and the underlying bond markets using the creation and redemption activities of ETFs. I find that the bond liquidity improves only when the arbitrage activities are high.

[1]While equity ETFs started trading in 1993, corporate bond ETFs became available only after 2002 and have only become popular since 2007 when Vanguard entered the market by introducing ETF share classes within their bond mutual fund portfolios.

In addition, I find that the 144A bond market experiences a 20% improvement in liquidity as bond ETFs increase their holdings of 144A bonds. This further supports my theoretical predictions: if the market accessibility is more limited, then the uninformed demand from ETFs can have a more positive impact on the liquidity of the underlying bonds. 144A debt instruments are privately placed and are traded among a limited number of large institutional investors, such as insurance companies (Carey, Prowse, Rea, and Udell, 1993). The initial buyer of 144A debt can only resell the securities after the required holding period. For retail investors, 144A bonds were previously inaccessible and have only become accessible through investing in bond ETFs that holds such bonds.

To further support the causality of my results, I also use a quasi-natural experiment applied to one of the major high yield bond ETFs – the iBoxx high yield $ corporate bond ETF (HYG) – in June 2009. The index composition rules were changed and were announced only 8 days before the effective date. As a consequence, the number of newly added bonds dramatically increased to 248, while only 16 were added from January to May of 2009. Consistent with my predictions, I find an improvement in the liquidity of those newly added bonds of 3.5%.

The same natural experiment was previously used by Dannhauser (2016) to show that corporate bond ETFs improve the valuation of the underlying bonds. Dannhauser shows that the increased valuation of ETF bonds is due to the influx of institutional investors into the ETF market. She also finds that bond market liquidity is not significantly affected overall. A potential source of the discrepancy between the Dannhauser results and those in this paper stems from the differences in methodology. First, my study employs the event study method six months before and after the inclusion of ETFs, while her study utilizes a long panel regression over four year periods. Second, I use corporate bonds added for the first time to an ETF portfolio, while she studies liquidity of bonds held by all ETF portfolios. Third, I focus on the impact of the ETF inclusion, while she emphasizes the relevance of ETF ownership.

My findings contribute to our understanding of the liquidity of assets by providing new evidence about the role of market participation, the channel through which it operates, and the magnitude of its effects on the liquidity of the underlying assets. Both my theoretical results and my empirical tests using bond ETFs are unexpected, given the past evidence that trading basket securities causes the liquidity of the underlying assets to deteriorate. However, as previously noted, the existing literature focuses only on the U.S. equity market, which is a highly accessible market.

Recent theoretical work by Guedj and Huang (2015) also predicts the influx of long-term passive investors through ETFs, albeit with different economic motives. Open-end mutual funds may inherently suffer from inefficiencies, as noted by Stein (2005), since managers' long term investments can be limited by investor horizons. Frequent creations and redemptions by mutual fund investors can exert a significant price pressure on the underlying securities, especially when fund flows are correlated. By contrast, since ETF investors trade with each other, potential price impacts will be cancelled out, protecting the welfare of long term passive investors.

A large body of literature on the liquidity of the bond market focuses on the supply side of the market. The liquidity of corporate bonds is large with the use of the limit order book (Abudy and Wohl, 2015; Biais and Green, 2007) and increases with the introduction of post-trade transparency through TRACE (Bessembinder et al., 2006; Edwards, Harris, and Piwowar, 2007; Schultz, 2001), with electronic markets (Hendershott and Madhavan, 2015), and more recently with bond dealers' inventory capital commitment (Bessembinder, Jacobsen, Maxwell, and Venkataraman (2016)). My results show that fixed income ETFs benefit certain subcategories of the corporate bond market – For example, those with low trading volume, those that are non-investment grade, or have long maturities – which have been shown to be costly and hard to trade (Edwards et al., 2007; Schultz, 2001).

In addition, this paper contributes to the growing literature in ETFs by showing that long term market accessibility is an important condition on which the inclusion in ETF can have differential effects on the liquidity of the underlying assets. Bhattacharya and O'Hara (2016) shows theoretically that temporary market inaccessibility could lead to market fragility, as market makers infer the information from ETF market prices that is unrelated to the fundamental value of the underlying assets. Pan and Zeng (2017) shows that large ETF liquidity could be fragile because ETF arbitrageurs play a dual role also as bond dealers. If these bond dealers' inventory is constrained, then there is a limit to arbitrage between ETF and the underlying corporate bonds, which could hurt the liquidity of the underlying corporate bonds. In contrast, my paper investigates the net effect of ETF arbitrage in the long run on the liquidity of the underlying corporate bonds. Ben-David, Franzoni, and Moussawi (2015) show that ETF ownership increases nonfundamental market volatility. Da and Shive (2013) showed ETF increases the comovement of the asset returns.

The rest of the paper is organized as follows. In section 2, I develop a theoretical model of optimal decisions of strategic liquidity investors in a segmented bond market and derive empirical predictions. In section 3, I discuss the data and the indicator-based regression methodology. I discuss the main results in section 4, and conclude in section 5.

2. The Model

I consider a single period model of trade for uninformed investors in a market with multiple assets before and after the introduction of an ETF (Figure 2). The uninformed investors hold the market portfolio and wish to trade due to a random arrival of liquidity needs. They are discretionary in that they are not constrained to trade a particular asset. In addition, they wish to maintain a diversified portfolio. To do so, they trade the assets in a specific proportion of their portfolio, similarly to Lo and Wang (2000).

Allocating trades across assets have both positive and negative impacts on the investor's utility. Splitting trades maintains the diversification benefit of the portfolio. However, trades may become too small and may fall below a required minimum trading amount so that the trade allocation becomes infeasible. In this case, the investors either concentrate on a subset of the assets, or choose not to participate in the market. If the portfolio is large, or if a minimum trading amount required is large, then this cost of diversification increases. The market becomes less accessible.

The introduction of ETFs offers a significantly lower cost of diversification. ETFs enable investors to trade the entire portfolio at lower transaction costs and with smaller minimum trading amounts than trading portfolios of individual assets. The discretionary investors optimally choose to trade among this expanded set of assets. Arbitrageurs exploit any arbitrage opportunities between the ETF and the underlying securities, which bridges the liquidity of the two markets.

2.1. Before the introduction of ETFs

At date $t = 0$, all investors observe the values of the individual securities. At date $t = 1$, the values of the securities are v_i $(i = 1, 2)$ and are normally distributed: $v_i \sim N(\mu_i, \sigma_i^2)$. The date 1 values of the securities are correlated, $Cov(v_1, v_2) = -\rho < 0$.

I build on a standard Kyle model in each of the markets. There are k_i informed investors who specialize in the i^{th} security market, and who have private information about the idiosyncratic risk $\epsilon_i = v_i - \mu_i$ of each security i. I assume the information signal is perfect so that the informed investors observe the private information with perfect accuracy. They do not speculate in the markets for other securities. Informed investor j $(j = 1, \cdots, k_i)$ in the i^{th} security market optimizes her expected profit, $\mathbb{E}\left[x_{ij}\left(v_i - P_i\right)|\epsilon_i\right]$, from trading on her private information by choosing her optimal demand, x_{ij}. The price P_i of the security will be determined by the market makers.

There are non-discretionary liquidity investors in each of the markets, who are obligated to trade a particular security and are not allowed to trade other assets. Their demands are denoted by u_i for the individual securities. I assume u_i follow normal distributions, $N(0, \sigma_{u_i}^2)$ for $i = 1, 2$. There are Q discretionary liquidity investors who may invest in either of the assets. They have to trade the assets due to random arrival of liquidity needs. I denote their demand by a random variable $l_q \sim N(0, Var(l_q))$ for $q = 1, \cdots, Q$. However, high transaction costs and high minimum capital requirements will limit their strategic decision on allocating their portfolios.

There are market makers in each of the markets. The market makers observe the demand from all investors – informed, non-discretionary, and discretionary – and set the prices (P_i) of the securities. Competition among the market makers implies that they set the prices of the securities with zero profit. I follow the linear pricing rule as in Kyle (1985). The price of each security i is set by the market makers as $P_i = \mu_i + \lambda_i \omega_i = \mathbb{E}[v_i|\omega_i]$, where $\omega_i = x_i + u_i + w_i l_q$ is the aggregate order flow in the i^{th} security market and w_i is a specific proportion of liquidity trading implemented with the security i. If the discretionary investors prefer to trade security 1, then l_q is absent for security 2. λ_i $(i = 1, 2)$ are price impacts of the aggregate order flows by all agents and are determined endogenously within the model.

The discretionary investors' demand l_q may be split between the assets so as to maximize their diversification benefits, taking into account trading losses due to liquidity needs. I assume the uninformed trade the individual securities in proportion to the weights w_i of the portfolio. The expected loss from trading security i will be $\mathbb{E}[(P_i - v_i)w_i l_q] = \lambda_i w_i^2 Var(l_q)$, $i = 1, 2$.

I assume the discretionary uninformed investors are risk averse. For the tractability of the model, I assume a negative exponential utility, $-\frac{1}{\gamma}\exp(-\gamma W)$, where $\gamma > 0$ is the risk aversion parameter of the uninformed investor. The uninformed investors wish to maximize their expected utility,

$$\max_{w_1, w_2} \mathbb{E}\left[-\frac{1}{\gamma}\exp(-\gamma W)\right] = \max_{w_1, w_2}\left\{-\frac{1}{\gamma}\exp\left[-\gamma\mathbb{E}(W) + \frac{\gamma^2}{2}\operatorname{Var}(W)\right]\right\}. \tag{1}$$

Their wealth W depends not only on the returns on the assets of the portfolio, but also on the expected cost of trading. I assume the returns $\tilde{r}_i$ of each individual security follows a normal distribution with a mean $\overline{r}_i$ and a variance σ_i^2. I assume an initial wealth as a numeraire. Their expected wealth in the later period is $w_1\overline{r}_1 + w_2\overline{r}_2$, less the expected cost of trading, $\lambda_1 w_1^2 \operatorname{Var}(l_q)+\lambda_2 w_2^2 \operatorname{Var}(l_q)$. A variance $\operatorname{Var}(W) = w_1^2\sigma_1^2+w_2^2\sigma_2^2+2w_1w_2 \operatorname{Cov}(\tilde{r}_1, \tilde{r}_2)$. Therefore, the objective function for the discretionary investors is equivalent to the following simpler.

$$\max_{w_1,w_2} \Big[w_1\mathbb{E}(\tilde{r_1}) + w_2\mathbb{E}(\tilde{r}_2) - \frac{\gamma}{2}\left\{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2 \operatorname{Cov}(\tilde{r}_1, \tilde{r}_2)\right\} \\ -\lambda_1 w_1^2 Var(l_q) - \lambda_2 w_2^2 Var(l_q)\Big] . \tag{2}$$

The above objective function is feasible only if the optimal trading allocation decisions are feasible. I assume that the expected capital requirement in each security market must be sufficiently met with the investors' capital.

$$\mathbb{E}(w_1 l_q P_1) \geq W_1, \quad \text{for} \quad l_q \neq 0, \quad \text{and} \quad \mathbb{E}(w_2 l_q P_2) \geq W_2, \quad \text{for} \quad l_q \neq 0, \tag{3}$$

and $w_1 + w_2 = 1$. If $l_q = 0$, then no trading occurs. The strategic decision is reduced to the traditional setting for long term investment. Negative and positive l_q indicate selling or buying, respectively. I assume the minimum trading amounts W_i $(i = 1, 2)$ are exogenously provided to the model.

This strategic portfolio trading problem is constrained due to the costs associated with implementing the diversification strategy. First, the dollar amount for trading security i needs to exceed W_i $(i = 1, 2)$ in order to place a buy order. This requires that liquidity volume $Var(l_q)$ be sufficiently large and that the investors meet the capital requirement. In a frictionless world, there will be no minimum trading amount required for investment. Any portfolio trading strategy can be implemented. If, however, W_i $(i = 1, 2)$ is large, then market i becomes less accessible because only the investors with large trading volume and sufficient wealth can benefit from a diversified portfolio. If the portfolio is large, the problem only exacerbates. Second, when the trades are executed, these uninformed investors are likely to encounter a loss due to the adverse selection of trading against informed investors.

Solving the optimization problem in Eqns. (2) and (3) is nontrivial because the price impacts λ_i are affected by the trading decisions of all the investors, including the discretionary investors themselves. More specifically, I first solve simultaneously informed investors' demand decisions, and market makers' pricing decisions and find the price impact parameters endogenously within the model.

$$\lambda_i = \sqrt{\frac{k_i}{(k_i+1)^2}\frac{Var(v_i-\mu_i)}{Var(u_i)+Var(w_i l_q)}}, \quad i=1,2. \tag{4}$$

Second, I substitute (4) for the price impacts and find an optimum allocation decision w_1^* and w_2^* for the discretionary uninformed investors.

However, due to the cost of diversification, the uninformed investors may hold under-diversified portfolios, depending on the market accessibility relative to their own wealth levels. The limited market accessibility *ex ante* is critical to predict changes in the liquidity of the underlying assets *ex post* after the introduction of ETFs. I categorize the market *ex ante* into three states. (a) Any portfolio trading decisions can be implemented (Highly accessible market). (b) Trading allocation is feasible only in one of the markets (Partially accessible markets). (c) None of the desired allocation is achievable (Inaccessible market). I assume that for the cases (b) and (c), the infeasible demands are deposited in risk free assets. In Proposition 1, I formally provide the optimal decision w_i^* $(i=1,2)$ for the fully or partially accessible market. The proofs are provided in the Appendix.

Proposition 1. *Suppose the liquidity trading volume is sufficiently large, i.e.* $Var(l_q) \geq V^* := \max\{\frac{W_i}{\lambda_i(w_i^*)^2}, i=1,2\}$. *(A) if the uninformed investors' capital exceeds* $W^* = \max\{W_i, i = 1,2\}$, *then the market is highly accessible. The optimal trade allocation* w_i^* $(i=1,2)$ *in the* *individual securities market is feasible.*

$$
\begin{aligned}
w_1^* &= \frac{1}{\gamma(\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2)}\Big[\overline{r}_1-\overline{r}_2+\gamma\sigma_2^2+2\gamma\rho\sigma_1\sigma_2 \\
&\quad -Var(l_q)\left(\sqrt{A_1}+\sqrt{A_2}\right)\sqrt{\frac{1}{Var(u)+Var(l_q)}\frac{2Var(u)+Var(l_q)}{Var(u)+Var(l_q)}}\Bigg], \qquad (5)
\end{aligned}
$$

and $w_2^* = 1 - w_1^*$*, where* $A_i = Var(v_i - \mu_i)\cdot k_i/(k_i+1)^2$*,* $i = 1, 2$*.*

(B) If $\sqrt{\frac{A_1}{Var(u)+Var(l_q)}}Var(l_q) = W_1$ *and the uninformed investors' capital exceeds* W_1*, then the market is partially accessible. The investors concentrate their trading in asset 1, i.e.* $w_1^* = 1$ *and* $w_2^* = 0$*.*

2.2. After the introduction of ETFs

ETFs benefit the uninformed investors in several different ways. First, the cost of trading a diversified portfolio is significantly lower. Second, the minimum required amount of trading is reduced significantly. Third, the cost of rebalancing the portfolios is dramatically lower. These savings are greater as the portfolio consists of a larger number of securities, or as the uninformed investors rebalance their portfolios over multiple periods.

Hence, after the introduction of ETFs, the uninformed investors reconsider their strategic decisions on the optimum portfolio allocation across securities 1, 2, and ETFs. For the simplicity, I assume the entry capital required for ETF investment is nonexistent: $W_e = 0$. In addition to the price impact λ_i for individual assets, I first find the price impact λ for trading ETFs to be

$$
\lambda = \sqrt{\frac{k_1}{(k_1+1)^2}\frac{Var(v_1-\mu_1)}{Var(u_1)+Var(w_1 l_q)} + \frac{k_2}{(k_2+1)^2}\frac{Var(v_2-\mu_2)}{Var(u_2)+Var(w_2 l_q)}}.
$$

As noted in Subrahmanyam, this price impact λ of trading ETFs is smaller than the total price impacts of trading individual securities, $\lambda_1 + \lambda_2$, due to a lower adverse selection in ETFs. I then find an optimal trading allocation w_i^*, $(i = 1, 2, 3)$.

Two major migrations of uninformed investors can occur. My predictions on the liquidity *ex post* critically depend on the relative strength of the migrations, which is correlated with the accessibility of the market *ex ante* (Figure 3).[2] First, if the underlying market *ex ante* was not fully accessible, then the uninformed investors who used to meet their trading need with the risk free asset shift to the ETF market. This influx of uninformed trading provides high liquidity in the ETF market, which can spill over to the underlying security market through arbitrage activities. The liquidity of the underlying assets improves. Second, if the market *ex ante* was highly accessible, the uninformed investors in the underlying market may also prefer to move to the ETF market because of lower adverse selection cost. In this case, the liquidity of the underlying security market deteriorates. This is consistent with the prediction of Subrahmanyam. I summarize these predictions in Proposition 2. I provide the proof of Proposition 2 in the Appendix.

Proposition 2. *Let λ_i and $\tilde{\lambda}_i$ $(i = 1, 2)$ be the price impacts of trading the i-th individual securities before and after the introduction of ETFs, respectively. Assume arbitrageurs connect ETF markets and the underlying security markets. Then, after the introduction of ETFs, the liquidity of the individual securities can change in the following way.*

1. *Suppose the individual securities were highly accessible. I.e. the investors' trading demand exceeds $W_i^* = w_i^* \sqrt{\frac{A_i}{Var(u)+Var(l_q)}} Var(l_q)$, where $A_i = \frac{k_i}{(k_i+1)^2} Var(v_i - \mu_i)$, for all $i = 1, 2$, and w_i^* is defined as in (B.7) and (B.8). Then the liquidity of the individual securities decreases, i.e. $\tilde{\lambda}_i > \lambda_i$.*
2. *Suppose that the underlying securities were partially accessible with security 1 accessible, and security 2 inaccessible. In addition, assume the investors' trading demand exceeds $W_1^* = w_1^* \sqrt{\frac{A_1}{Var(u)+Var(l_q)}} Var(l_q)$, where $A_1 = \frac{k_1}{(k_1+1)^2} Var(v_1 - \mu_1)$ and w_1^* is*

[2] In this paper, I assume the supply of underlying assets is infinitely elastic. The counterargument can be made that as ETF assets under management grow, ETFs get hold of more of the underlying shares so that a reduced number of securities are available for trading. I argue that this could be a major concern if an ETF holds the entire universe of securities. If, on the other hand, an ETF has a focused portfolio, then ETF managers may cope with the limited supply of shares either by expanding the portfolios or by deviating from a common index pursued by multiple ETFs.

defined in (B.14). Then liquidity decreases for security 1, while increases for security 2. $\tilde{\lambda}_1 \geq \lambda_1$*, while* $\tilde{\lambda}_2 \leq \lambda_2$.

3. *If the underlying securities were inaccessible, then liquidity increases, i.e.* $\tilde{\lambda}_i < \lambda_i$.

3. Research Design

3.1. Regression-based model

Many of the studies on intraday trade execution costs in equity markets rely on pre-trade bid and ask quotes and construct measures such as quoted spreads or effective spreads. Alternatively, daily trading cost of securities can also be estimated from daily returns. For instance, Lesmond, Ogden, and Trzcinka (1999) developed the LOT measure, which has been applied to the global market by Fong, Holden, and Trzcinka (2016). It assumes that informed investors trade only when their expected profits outweigh the transaction costs of the securities. The daily returns themselves implicitly reflect trading costs, which can be recovered computationally. Chen, Lesmond, and Wei (2007) applied the LOT measure to corporate bonds and showed that it is correlated to their available quotes.

In contrast to equity market, pre-trade quotes for the majority of corporate bonds are not available and hence effective or quoted spreads are not viable measures.[3] To overcome this limitation, two alternative measures have been used in the literature. The first method is estimating transaction costs of corporate bonds by investigating a subset of the bonds that have both buy and sell orders on each day. The difference between the average buy and sell prices is used as a measure of the daily transaction costs of the bond. Since it is common for corporate bonds to have only buy trades or only sell trades on a given day, this method significantly reduces the number of usable data. Inferring the transaction costs for

[3]Pre-trade quotes are available for corporate bonds traded in NYSE Automated Bond Systems and MarketAxess, but are not reported to TRACE. Reuters Fixed Income Database provides the daily quotes from major bond dealers. Recently, Schestag, Schuster, and Uhrig-Homburg (2016) use the daily bond prices reported in Datastream and measure the liquidity of bonds.

other non-traded bonds from it could potentially be limited (Schultz, 2001).

A second method is the indicator-based regression model, which was originally suggested for equity markets by Huang and Stoll (1997), and then extended to corporate bonds by Bessembinder et al. (2006). This method decomposes the spreads into adverse selection costs, inventory costs, and order processing costs. This model does not require that bond trades are present both on buy and sell sides and is applicable as long as some trades occur. However, very few trades in a short time period or very sparsely traded bonds will make the cost estimates noisy.

This indicator-based regression model consists of two steps. In the first step, the unexpected order flow Q_t^* is estimated by assuming the order flows Q_t are serially correlated, following $AR(1)$ process, $Q_t = Q_{t-1} + \epsilon_t$, where Q_t is the buy-sell indicator. If the customer buys the bond at t, then $Q_t = +1$. If the customer sells the bond at t, then $Q_t = -1$. If dealers trade with other dealers, then $Q_t = 0$. In the second step, the changes of bond prices, ΔP_t, are modeled as follows:

$$\Delta P_t = \alpha S \Delta Q_t + \gamma S Q_t^* + w X_t + \delta_t, \tag{6}$$

The first term $\alpha S \Delta Q_t$ reflects the non-informational content from the order flow changes. The second term indicates the informational content from the unexpected order flow, $Q_t^* = \epsilon_t$. The third term X_t means the public information available at the time of the most recent trades. The last term δ_t is the noise. I estimate numbers, αS and γS, and a vector w. In my paper, the changes of the bond's transaction cost, αS, is of utmost interest.

Following Bessembinder et al. (2006), X_t includes three components representing public information available at the time of trading. The first is the yield, on the bond transaction day, of the on-the-run Treasury security matched for time to maturity with the corporate bond. Since the maturities of corporate bonds vary, this Treasury yield variable is indicative of changes in the interest rate term structures. The second is the daily default spread as the difference between the yield rate of aggregate BAA bonds and the risk free rate. As market sentiment leans negatively (or positively), the risk premium for aggregate BAA bonds increases (decreases). Hence, this default spread allows us to control for market wide perceptions on the overall economy. The third is the return of the common stock for the issuing firm since the last transaction date of the bond. All bonds issued by subsidiaries are considered to have the same parent company. The return on the common stock reflects firm-specific news. For instance, a positive earnings announcement increases investors' expectation of their future earnings, resulting in increases both in the firm's stock price and return.

3.2. Implementing a regression-based model

I implement the main methodology discussed in section 3.1 with two stage regressions. In the first stage, I estimate the serial correlation among the order flows Q_t using ordinary linear regression (OLS), $Q_t = \rho Q_{t-1} + \epsilon_t$. The estimated unexpected order flow is $\hat{\epsilon}_t = Q_t - \hat{\rho} Q_{t-1}$, which will be used in the second stage. I let $Q_t^* = \hat{\epsilon}_t$. The second stage is the main regression model of my paper. My goal is to investigate the changes of transaction costs of bonds with their inclusion in an ETF. To account for a potential parallel trend present in the bond market, I extend the main model (6) to the difference-in-difference framework,

$$\begin{aligned} \Delta P_t &= a + \alpha_1 S \Delta Q_t \times ETF \times Post \\ &\quad + \alpha_2 S \Delta Q_t \times ETF + \alpha_3 S \Delta Q_t \times Post + \gamma S Q_t^* + w X_t + \delta_t. \end{aligned} \tag{7}$$

ETF is a binary variable that takes 1 if the bond is included in ETF and 0 otherwise. $Post$ is a binary variable that takes 1 if the bond is traded after the inception date of ETF and 0 if the trade occurs prior to the first offer date of the ETF. The coefficient $\alpha_1 S$ is my main interest and measures changes of transaction costs of ETF bonds relative to those for non-ETF bonds, due to the inclusion in an ETF.

I estimate the second stage regression with the weighted least regression (WLS). Since recent trades are likely to be more serially correlated with today's trades than are distant trades, I use as weights the inverse of the number of days elapsed since the last transaction day. For instance, if the last trade was made two days ago, then all the trades for today are given the equal weight of 1/2.

To select the non-ETF bonds in the control group, I match on the following characteristics of the bonds, using the propensity score matching algorithm (Wooldridge, 2010): trading volume before the inclusion in ETF, credit rating at the time of trading, issue size, age, time to maturity of the bond at the time of trading. Prior studies (Schultz, 2001; Edwards et al., 2007; Goldstein, Hotchkiss, and Sirri, 2007; Hendershott and Madhavan, 2015) showed that credit rating and trading volume are robust explanatory variables for the liquidity of the bonds: bonds with higher credit rating or with large trading volume are less costly to trade. I include issue size because large issues are more frequently traded due to their availability. Trading volume also influences the spread because it allows bond dealers to manage their inventory better (Grossman and Miller, 1987). The time to maturity is relevant to the bond liquidity because long term bonds are harder to trade in general. The age of the bond matters because the majority of corporate bonds are known to end up being owned by institutional investors, who infrequently trade, within two years of issuance. Therefore, in the pre-ETF period, the two groups of bonds are similar in their liquidity. I am interested in whether their transaction costs of the bonds diverge in the post-ETF period.

I note that the standard statistical inference with the t-statistic is not valid for my analysis for the following reasons. The second stage regression has an explanatory variable $Q_t^* = \hat{\epsilon}_t$, which was estimated from the first stage. The correct inference for $\alpha_1 S$ must include the adjustment taking into account possible interactions of the noise terms between the first and second stage regressions (Wooldridge, 2010). Furthermore, some bonds are traded more than once within a day. This implies that I have multiple observations for the realized ΔP_t for the time t. The number of observations per day is random. Also, if some bonds are issued by the same parent firm, or their times to maturities are similar, then changes of bond prices are likely to be correlated. Therefore, the noise term in the second stage regression is less likely to form the standard t-distribution.

In order to estimate the true distribution of the noise term, δ_t, I use the block bootstrapping method. For each bond, since the daily trades are more likely to be correlated than inter-day trades, I sample with replacement trades from daily trades 100 times. The correlation structure between the days remains the same. But at each time, the number of trades occurring is randomly varying. The empirical distribution for the bond's transaction cost $\widehat{\alpha_1 S}$ approximates the true population distribution (Wooldridge, 2010). I use this empirical distribution and test against the null hypothesis that the true transaction cost is zero, i. e. $H_0 : \alpha_1 S = 0$. I compute the p-value as the likelihood that the estimate $\widehat{\alpha_1 S}$ has the sign opposite of the fitted value, given the null hypothesis.

3.3. Data

3.3.1. ETF and holdings

I use CRSP mutual fund database to identify ETFs using the identifier, `et_flag`. I first limit the universe of ETF into a fixed income category where the lipper asset code (`lipper_asset_cd`) is either TX or MB. I exclude the fixed income ETFs that concentrate on Treasury bonds or non-bond assets such as commodities and include ETFs that corporate bonds comprise 100% of the holdings. Although I do not know the exact timing of the acquisition or sales of the bonds, I use the portfolio holdings reported to CRSP within 6 months post the first offer date (`first_offer_dt`) of each ETF to estimate the bonds added to ETF for the first time. For ETFs offered by iShares, I manually collected the monthly holdings from the funds' websites.

For some ETFs, their portfolio holdings were reported to CRSP with a significant delay. This delayed reporting problem is more prevalent with the ETFs that were launched before 2007. For index ETFs with the reporting delay problem, I assume the identities of the portfolio holdings do not change significantly between the inception date and the earliest report date. I take the holdings on the earliest reporte date for the initial ETF portfolio.

My data on ETF bonds span 2007 to 2014. This will exclude the bonds included in the very first ETFs such as LQD (iBoxx investment grade corporate bond ETF, launched in 2002). However, the number and the assets under management of fixed income ETFs have grown rapidly (Table 1) since 2007 when Vanguard offered ETF share classes within their existing mutual funds for the first time. Hence my data are sufficient to test the theoretical predictions in Section 2.

3.3.2. Corporate bonds intraday trade data

In order to estimate the transaction costs of corporate bonds, I use the TRACE intraday trading data for corporate bonds issued by U. S. firms as my primary database. In TRACE, large dollar transaction amounts are truncated. For investment grade bonds, any trades larger than \$5 million are reported as \$5 million. For junk bonds, the maximum reported amount is \$1 million. My dataset consists of corporate bonds that were added to ETF for the first time and whose trading were reported to TRACE.

In addition to the traditional bond dealers' market, some of the bonds may have been traded in an electronic exchange such as New York Stock Exchange Automated Bond System (NYSE ABS) or in an electronic trading platform such as MarketAxess. Campbell and Taksler (2003) estimate that NYSE ABS comprises about 5% of the dollar trading volume of the bond market. Anecdotal evidence suggests that MarketAxess processes about 15%, on average, of the market trading dollar volume in the second half of 2013. The trades that occurred in these venues are reported to TRACE, as long as bond dealers take one side of the trade, although I do not know where the trades actually occurred. Still it is reasonable to assume that the majority of the bond trading occurs with bond dealers.

I exclude the transactions associated with issuing, cancelled orders, or sellers' option settlements. Inter-dealer transactions may have been reported twice by both of the dealers, among which I keep only one trade. Large changes of bond prices during the day are likely to arise due to reasons unrelated to liquidity pressure. For instance, some ETF managers may possess private information about the true value of the bond or have skills to forecast the future price changes of the bonds. These managers may prefer to include the bonds driven by their information, which may weaken the implications of my results. I therefore remove bond transactions whose prices at two consecutive time periods deviate more than 10%, following Campbell and Taksler (2003).

I obtain historical bond credit ratings from DataStream and the issue amount, the issue date and the issuer information from Bloomberg. For missing data, I use the Mergent Fixed Income Securities Database (FISD) to complement the information. The three major credit rating agencies – S&P, Moody's, Fitch – have different rating categorizations. Following the convention in the credit rating literature (Becker and Ivashina, 2015; Cornaggia and Cornaggia, 2013; Bongaerts, Cremers, and Goetzmann, 2012), I assign a numeric score for the credit rating scale from each rating agency (Becker and Ivashina, 2015) and merge the scores for each bond as follows: At any date, if three rating scores are available, I take the middle score. If two rating scores are available, I take the lower score. As the credit rating upgrades or downgrades for each bond, this combined rating score is adjusted.

I also use a separate database for the 144A bond transactions in TRACE (`trace_btds144a`) and investigate the liquidity spillover between ETFs and the underlying bond market in section 4.

3.3.3. Public information variables

Two of the public information variables – on-the-run Treasury security returns and the yield spread between BAA bonds and the risk free rate – are obtained from the Federal Reserve Statistical Release. The third, the percentage returns on the issuing firms' common stock, is obtained from the CRSP daily stock returns database.

4. Main Results

4.1. Descriptive data

I match the bonds added to ETFs for the first time with the bonds never included in ETFs portfolios using credit rating, issue size, age and time to maturity of the bonds at the inception date of ETFs, and the cumulative trading dollar volume of the bonds prior to the inclusion to ETFs. Table 2 describes static characteristics of both ETF and non-ETF bonds in the matched sample. About 64% of the bonds have the dollar issue size larger than $500M. Bonds with superior credit quality ($AA-$ and up) compose 12% of the sample. About 63% of the bonds have investment credit rating of $BBB-$ through $A+$. Non-investment grade bonds comprise 26% of the sample in both of the groups. The corporate bonds in the treatment (ETF bonds) and control (non-ETF bonds) groups match within 10% differences for large categories. However, I find it hard to identify a close match for small categories.

Trading characteristics of the bonds are shown in Table 3. The bonds are traded close to the par values. The aggregate dollar trading volume decreases for both treatment and control groups, which is consistent with Randall (2015) findings that the overall trade size of corporate bonds has decreased since 2007. However, with the inception of ETFs, the total dollar trading volume of ETF bonds decrease less and in fact increases for low volume, small trade size ETF bonds. The total number of trades increases for ETF bonds, while it decreases for non-ETF bonds. These divergent changes are most prominent for trade sizes less than $1M.

I also compare changes in transaction costs of the bonds using the two methods discussed in section 3.1. First, when the bonds have both buy and sell orders within a day, I investigate the difference between the average buy and sell prices. Changes of transaction costs are not significantly noticeable both for ETF bonds and non-ETF bonds. However this measure could be noisy due to a limited sample size. Second, I utilize the indicator-based regression method and estimate the half-way transaction costs relative to the last transaction price on the most recent transaction day. For ETF bonds, the smallest quintile based on the cumulative trading volume experiences the greatest decrease of about 30% in transaction cost – about 15 basis points. Non-ETF bonds increases transaction costs. This evidence suggests that the benefits of trading of ETFs accrue differentially on the transaction costs of the underlying bonds and that cumulative trading volume of the underlying bonds is a key state variable.

4.2. The effect of ETF inclusion of bonds

Table 4 reports the estimates of changes in transaction costs for the bonds included in an ETF for the first time using the regression-based model in equation (7). Specifically, the estimated coefficient of interest is the one on $\Delta Q \times Post \times ETF$. The coefficients are in percentages and can be interpreted in basis points.

The coefficient for Q_{t-1} is positive and statistically significantly different from zero. This implies that bond orders are positively serially correlated. The coefficient for the unexpected order flow Q_t^* is also positive significant. Since this coefficient indicates the informational content observed in the order flows of the bonds, this result implies that some of the ETFs include bonds because of private information.

The estimated coefficients on the public information variables are significant. More specifically, the coefficient for Treasury return is negative and significant. As the market interest rate risk increases, both Treasury bonds and corporate bonds respond. Stock returns of the issuing firm have a positive impact on the bond transaction costs. The impact is larger for non-investment grade bonds than for investment grade bonds. This is consistent with the finding of Hotchkiss and Ronen (2002) that both stock and bond returns respond to new information about the value of the issuing firm's underlying assets. The new information is more valuable for non-investment grade bonds, because the investors and the firms are more likely to be asymmetrically informed. The coefficient estimates for the default spread is negative and significant for investment grade bonds, and positive and significant for the non-investment grade bonds.

I estimate that the average effective half-way spread for trading individual corporate bonds is 42 basis points during my sample period. While the average cost of trading corporate bonds increases by 2.6 basis points, with the inception of ETF trading, the transaction cost of ETF bonds decreases more than that of non-ETF bonds. The difference of the difference is about 1.3 basis points on average. Hendershott and Madhavan (2015) estimated that cost of trading odd lots ($100K∼1M) for corporate bonds in 2011 was 48 basis points for the dealers' market and 14 basis points for the electronic trading platform. Our estimate of the cost of trading ETF bonds is close to that in the dealers' market. Furthermore, as an evidence for the underlying economic mechanism of my results, I estimated ETF arbitrage activities by using the intensity of ETF creations and redemptions (Table 5). I find that if ETF arbitrage activities are high, the trading costs of the underlying bonds decreases. Conversely, if ETF arbitrage activities are low, then the trading costs increases.

Table 6 reports that the underlying bond trading cost prior to the initial trading of ETF is linearly related to the cumulative trading volume, consistent with earlier findings (Edwards et al. (2007), for instance). The bonds with the smallest dollar trading volume (VOL 1) cost the most at 92 basis points. As the trading volume increases, the transaction costs decrease. The most liquid bonds with large dollar trading volume cost 41 basis points.

The initiation of ETF trading results in differential changes of trading costs of the underlying bonds. The changes are linearly related to the cumulative trading volume. The illiquid bonds (VOL 1) become less costly to trade and their transaction costs decrease by 9 basis points, i. e. about a 10% reduction. On the other hand, liquid bonds with median trading volume become more costly. The trading costs increase by 2 basis points. The most liquid bonds (VOL 5) show little impact on the liquidity due to ETF trading.

Table 7 reports transaction cost estimates arranged by the bond credit rating. Before the introduction of ETFs, non-investment grade bonds are more costly to trade at an estimated 57 basis points, while investment grade bonds cost 37 basis points. This difference of trading cost can be explained by large information asymmetry between the issuing firm of non-investment grade bonds and the investors (Holden, Mao, and Nam, 2016). With the introduction of ETF, the ETF trading from the uninformed investors has an indirect impact in the underlying junk bond market.[4] The effective spreads of the junk bonds held by ETFs decrease by 10%, i. e. 5.4 basis point. This is consistent with my theoretical predictions in Proposition 2.

Unlike bonds, fixed income ETFs have no maturity. As some incumbent bonds near maturity, ETF managers may wish to sell off the maturing bonds in order to reduce the tax burden from realizing the capital gains. This specific trading demand for these bonds may offset the positive impact of ETFs on improving the liquidity of the underlying bonds. In Table 8, I show that with the introduction of ETFs, the liquidity improves for the bonds whose time to maturity is more than two years, while the liquidity deteriorates for the maturing bonds.

4.3. Quasi-Natural Experiment for a high yield bond ETF

As an additional support for the causality of my results, I use the change in the index composition rule for iShares iBoxx $ High Yield Corporate Bond ETF (HYG). This same event has previously been used by Dannhauser (2016) to show the valuation effect of corporate bond ETFs. HYG is one of the largest high yield bond ETFs by assets under management, whose assets already reached $3.3 billion as of June 30, 2009. The HYG portfolio holds about 94% of the assets in non-investment grade corporate bonds and the remaining in cash. The expense ratio is 0.5 basis point and the portfolio turnover ratio is about 30%.

HYG is passively managed and follows the aggregate bond index managed by Markit for

[4] Dannhauser (2016) finds that the relation between ETF trading activities and the underlying bond liquidity is not robust and that ETF ownership is correlated with liquidity deterioration only for investment grade bonds. This is not in direct conflict with my results since my study and the Danhauser study explore different underlying mechanisms using different methodologies . For instance, I use an event study methodology by focusing on 6 months before and after the ETF inclusion, while Danhauser use a panel regression over 4 years of ETF ownership. I use only the bonds added for the first time to an ETF portfolio, while she uses the bonds held by all ETF portfolios. In an unreported result, I confirm my results still hold using a panel regression around the inclusion events.

their benchmark index. On June 22, 2009, Markit imposed a 3% cap on the constituent bonds and required the index to be computed as a value-weighted average, effective June 30, 2009. As a consequence, during the second half of 2009, about 248 bonds were newly added to HYG. The governing board of HYG may have known of the change in advance, and may have initiated trading the relevant bonds early. I argue that these types of trading activities are limited due to concerns about tracking errors. As indirect evidence, the numbers of the newly added bonds increased dramatically after June 2009: Only 16 were added from January to May, while 31 were added in June, and 248 were added from July to December. Therefore, I argue that these bonds were added to HYG for reasons unrelated to anticipated reduction in transaction costs. I investigate changes of the liquidity of these bonds using a difference-in-difference framework and consistently find that ETF inclusion leads to a reduction of transaction costs of the bonds by 3.5%. (Table 9)

4.4. Active and Index ETF

ETFs are diversified portfolios, but can be actively or passively managed. The extent to which the management style matters for the liquidity of the underlying bonds is unclear *ex ante*. Managers of actively managed ETFs strive to outperform the returns of a benchmark index by possessing private information about the true value of the underlying securities. Still, if the managers trade frequently, the cost of trading the securities can erode their net returns.

In Table 10, I divide the sample based on the management style of ETF. I exclude bonds held both by active and index ETFs, which reduces the overall sample size. The coefficient for ΔQ indicates that the underlying bonds of active ETFs cost more at 43 basis points, 23% larger than transaction costs of the bonds for index ETFs at 35 basis points. With the inception of trading active ETFs, the trading cost increases for the underlying bonds of active ETFs, although by a small amount, while it decreases for those of index ETFs. This result implies that my theoretical predictions in section 2 hold strongly with index ETFs. When

ETFs are known to possess private information about the underlying securities, trading such ETFs in fact increase the cost of adverse selection, reducing the liquidity of the underlying bonds.

4.5. Multivariate Analysis

In Table 11, I combine all the explanatory variables and run a multivariate analysis. First, I find that my univariate analysis results continue to hold. Second, as in column (1), bond trading volumes are the most dominant factor in determining transaction cost of the bonds. The low volume bonds benefit most from the inception of ETF trading. The bonds whose remaining life is more than 2 years had lower transaction costs by 31 basis points than those with little time to maturity. With the inception of ETF trading, these bonds near maturity exhibit even bigger transaction cost by 2.6 basis point. In column (2) and (3), I consider the multivariate analysis based on the bond credit rating. The univariate analysis from Table 7 remain unchanged.

4.6. The impact of ETF trading on 144A bonds

ETFs have increased their holdings of 144A corporate bonds since 2010 (Figure 4). ETFs allow the discretionary investors to access the 144A bond market, which has previously been restricted to a limited number of large institutional investors. My model of the liquidity in a segmented market predicts that if ETF trading volume is sufficiently large, the liquidity of these restricted securities will improve. I empirically test this prediction in this section.

144A bonds differ from the traditional corporate bonds in multiple dimensions: Unlike public debt, which requires registration and SEC oversight, firms can privately issue debt if they place it with accredited investors or a limited number of individual investors who intend to hold it for investment. This privately placed debt can be resold if the issue is registered with SEC after issuance or if investors have held it for, typically, at least two years. In 1990, Rule 144A was amended to allow investors to resell 144A securities without SEC registration or the two year holding period, as long as they are sold to qualified institutional buyers (QIB).

Although 144A securities are subject to large information asymmetry between the issuers and the investors, and illiquidity due to the required holding periods before resale, they have been popular among debt issuers and institutional investors. U.S. firms take advantage of the speedy issuance in the 144A market. As of 2013, these private placement comprises about 15% of high yield bonds issuance (Table 12).[5] Fenn (2000) showed that about 97% of the 144A high-yield securities are subsequently registered with the SEC. Among high yield 144A debt, 58% was issued by first-time bond issuers, while 29% was issued by privately owned firms. The credit risk premium for 144A high yield securities is higher than for public debt, but it gradually disappears after issuance. Furthermore, Chaplisky and Ramchand (2004) show that the credit risk premium of the 144A market is positive for investment grade debt issued by foreign firms, and for high yield debt issued by U. S. firms. Surprisingly, the weak disclosure requirement for 144A securities has no impact on the credit risk premium.

Effective February 2008, Rule 144A was further amended so that investors can resell 144A securities after only six months if the issuing firms are providing certain information to the market. This implies that non-QIBs can easily acquire 144A securities for the first time. Furthermore, a new disclosure requirement was initiated by the JOBS Act,[6] that 144A bonds transaction data start being publicly disseminated since 30 June 2014.[7] As of the third quarter of 2014, transactions of 144A securities exploded, comprising nearly 9% of the average daily volume in investment-grade corporate debt, and nearly 28% of the average daily volume in high-yield corporate debt (Table 12).

These 144A securities have low transaction costs because trades are between institutional investors and hence adverse selection cost is low (Edwards et al., 2007). As discretionary

[5] In an unreported note, I find that the composition of these privately placed debt is even higher among the most risky bonds whose credit rating is below B. As of 2013, they comprise about 50% of the issuance of the most risky bonds.

[6] The Jumpstart Our Business Startups Act (JOBS Act) directed the SEC to eliminate the long-standing prohibition against general solicitation and general advertising in offerings of securities pursuant to securities act Rule 144A and in certain other private placements. (SEC Release, No. 70345, 16 September 2013)

[7] `https://www.finra.org/newsroom/2014/finra-brings-144a-corporate-debt-transactions-light`

investors trade ETFs, I investigate whether the introduction of ETFs can affect this 144A bond market, where trading has been dominated by institutional investors, by using the TRACE 144A bond database. With the introduction of ETFs, the transaction cost of the 144A bonds owned by ETF decreases significantly by 5 basis point (Table 13). The reduction of transaction costs is larger for large volume bonds, for high-yield bonds, and for long-term bonds (not shown). This result further supports my theoretical predictions made earlier in Section 2.

4.7. Robustness

Since I match the bonds in the treatment group with those in the control group by using the bond characteristics, one potential concern would be that my results may be driven by unobservable characteristics of the bonds that are related to their issuing firms. In order to address this concern, I added firm fixed effects in my analysis. My results remain the same (Table 14).

5. Conclusion

In this paper, I show market participation plays an important role in how liquidity of the underlying securities changes when basket securities are introduced. I provide evidence that if the underlying market is less assessible, then trading basket securities improves the liquidity of the underlying assets.

To do so, I first extend the Kyle model to multi-markets where the degree of accessibility varies. The discretionary investors are uninformed and concerned about having a tradable diversified portfolio. I show that they prefer to invest in basket securities because of (1) a low expected loss of trading and (2) ease of accessibility (i.e. lower minimum investment requirement). In particular, the ease of accessibility in the basket market is critical in order to improve the liquidity of the underlying securities.

As empirical evidence, first, I use TRACE bond transaction data for traditional corporate bonds and show that in contrast to the equity market, trading bond ETFs reduces the transaction cost of the underlying bonds. The reduction of transaction costs is larger for low volume, high-yield, and long-term bonds. Transaction costs for high volume, investment grade, or short-term bonds increase. Second, 144A bonds, which were previously inaccessible to non-institutional investors, experience improved liquidity as ETFs increase their holdings of 144A bonds.

My results imply that bond investors benefit from the inclusion of the bonds in fixed income ETFs. This benefit accrues more to small retail bond investors and to junk bond investors. The aggregate savings in trading an ETF bond is as big as −\$18.5 million during the six months after inception of a fixed income ETF containing that bond.[8]

[8] $=10.259\times(-0.0881)+34.897\times(-0.0355)+78.447\times 0.0185+155.002\times(-0.0015)+844.282\times(-0.0011)$ $=-1.85$. This is equivalent to 18.5 million dollars.

Appendix A. Proof of Proposition 1

Proof. I solve the portfolio trading allocation problem using a constrained optimization method. The Lagrangian of the objective function is set up with multipliers, η_1 and η_2 as

$$\begin{aligned} U(w_1) &= w_1\overline{r}_1 + (1-w_1)\overline{r}_2 - \frac{\gamma}{2}\left[w_1^2\sigma_1^2 + w_2^2\sigma_2^2 - 2w_1(1-w_1)\rho\sigma_1\sigma_2\right] \\ &\quad -\lambda_1 w_1^2 Var(l_q) - \lambda_2(1-w_1)^2 Var(l_q) \\ &\quad -\eta_1\left[\mathbb{E}(w_1 l_q P_1) - W_1\right] - \eta_2\left[\mathbb{E}((1-w_1)l_q P_2) - W_2\right]. \end{aligned} \tag{A.1}$$

To find the first order condition (FOC), I take the first derivative of the Lagrangian.

$$\begin{aligned} \frac{dU}{dw_1} &= \overline{r}_1 - \overline{r}_2 - \gamma\left[w_1\sigma_1^2 - (1-w_1)\sigma_2^2 - (1-2w_1)\rho\sigma_1\sigma_2\right] \\ &\quad -\left[\frac{d\lambda_1}{dw_1}w_1^2 + \lambda_1\cdot 2w_1\right]Var(l_q) - \left[\frac{d\lambda_2}{dw_1}(1-w_1)^2 - \lambda_2\cdot 2(1-w_1)\right]Var(l_q) \\ &\quad -\eta_1\frac{d\left[\mathbb{E}(w_1 l_q P_1) - W_1\right]}{dw_1} - \eta_2\frac{d\left[\mathbb{E}((1-w_1)l_q P_2) - W_2\right]}{dw_1}. \\ &= 0, \end{aligned} \tag{A.2}$$

In addition, the following conditions should be met.

$$\eta_1\left[\mathbb{E}(w_1 l_q P_1) - W_1\right] = 0, \tag{A.3}$$

$$\eta_2\left[\mathbb{E}((1-w_1)l_q P_2) - W_2\right] = 0 \tag{A.4}$$

$$\eta_1 \geq 0, \quad \eta_2 \geq 0, \tag{A.5}$$

$$\mathbb{E}(w_1 l_q P_1) \geq W_1, \quad \text{and} \quad \mathbb{E}((1-w_1)l_q P_2) \geq W_2. \tag{A.6}$$

Cases 1 and 3: First, I focus on interior solutions, $\eta_i = 0$, for $i = 1, 2$ and solve the FOC for w_1. I note that $\lambda_1 = \sqrt{A_1}\left[Var(u_1) + Var(w_1 l_q)\right]^{-1/2}$, where $A_1 = Var(v_1 - \mu_1)k_1/(k_1+1)^2$

is exogenously given to the model. Furthermore,

$$\frac{d\lambda_1}{dw_1} = -\frac{\lambda_1 w_1 Var(l_q)}{Var(u_1) + Var(w_1 l_q)}, \qquad \frac{d\lambda_2}{dw_1} = -\frac{d\lambda_2}{dw_2} = \frac{\lambda_2 w_2 Var(l_q)}{Var(u_2) + Var(w_2 l_q)}.$$

The FOC is reduced to the following form. I assume $u_1 = w_1 u$.

$$\frac{d\lambda_1}{dw_1} w_1^2 + \lambda_1 \cdot 2w_1 \;=\; \lambda_1 w_1 \frac{2Var(u_1) + Var(w_1 l_q)}{Var(u_1) + Var(w_1 l_q)} = \lambda_1 w_1 \left[\frac{2Var(u) + Var(l_q)}{Var(u) + Var(l_q)}\right].$$

I note that

$$\lambda_1 = \sqrt{\frac{A_1}{Var(u) + Var(l_q)}} \frac{1}{w_1}, \quad \text{or equivalently} \quad \lambda_1 w_1 = \sqrt{\frac{A_1}{Var(u) + Var(l_q)}}.$$

Therefore, independent of w_1,

$$\frac{d\lambda_1}{dw_1} w_1^2 + \lambda_1 \cdot 2w_1 \;=\; \sqrt{\frac{A_1}{Var(u) + Var(l_q)}} \left[\frac{2Var(u) + Var(l_q)}{Var(u) + Var(l_q)}\right].$$

Similarly,

$$\begin{aligned}
\frac{d\lambda_2}{dw_1}(1-w_1)^2 - \lambda_2 \cdot 2(1-w_1) &= \frac{d\lambda_2}{dw_1} w_2^2 - \lambda_2 \cdot 2w_2 \\
&= -\lambda_2 w_2 \left[\frac{2Var(u_2) + Var(w_2 l_q)}{Var(u_2) + Var(w_2 l_q)}\right] \\
&= -\lambda_2 w_2 \left[\frac{2Var(u) + Var(l_q)}{Var(u) + Var(l_q)}\right], \\
&= \sqrt{\frac{A_2}{Var(u) + Var(l_q)}} \left[\frac{2Var(u) + Var(l_q)}{Var(u) + Var(l_q)}\right],
\end{aligned}$$

where $A_2 = k_2/(k_2+1)^2$. Putting all the relevant terms together, I solve analytically for w_1^*:

$$\begin{aligned}
w_1^* \;=\; & \frac{1}{\gamma(\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)} \Big[\overline{r}_1 - \overline{r}_2 + \gamma\sigma_2^2 + 2\gamma\rho\sigma_1\sigma_2 \\
& -Var(l_q)\left(\sqrt{A_1} + \sqrt{A_2}\right) \sqrt{\frac{1}{Var(u) + Var(l_q)} \frac{2Var(u) + Var(l_q)}{Var(u) + Var(l_q)}}\,\Bigg], \qquad \text{(A.7)}
\end{aligned}$$

and $w_2^* = 1 - w_1^*$. If the investors have sufficiently large wealth that both the optimal allocations w_1^* and w_2^* are feasible, i.e. $\mathbb{E}(w_i^* l_q P_i) \geq W_i$, then these markets are accessible. Otherwise, the investors deposit it in a risk free asset, in which case the market is inaccessible.

Case 2: Without loss of generality, I suppose $\eta_1 \neq 0$ and $\eta_2 = 0$. Using the complimentary conditions, I deduce that $\mathbb{E}(w_1 l_q P_1) = W_1$, which implies that $w_1^* = \left[\sqrt{\frac{A_1}{Var(u)+Var(l_q)}} Var(l_q)\right]^{-1} W_1$, and $w_2^* = 1 - w_1^*$. I substitute these solutions (w_i^*) to FOC, and find η_1^*. I assume $\eta_1^* > 0$. Then, let $\sqrt{\frac{A_1}{Var(u)+Var(l_q)}} Var(l_q) = W_1$. A corner solution occurs, $w_1^* = 1$ and $w_2^* = 0$. In this circumstance, the investors concentrate their portfolio trading only on the asset 1. An alternative corner solution for the case of $\eta_1 = 0$, and $\eta_2 \neq 0$ is similar, since the arguments are exactly symmetric to the above. □

Appendix B. Proof of Proposition 2

Proof. I start from each of the three equilibria based on the portfolio trading allocation *ex ante* and discuss theoretical predictions on changes of the liquidity of the underlying assets after the introduction of ETFs.

Case 1: The two underlying markets were highly accessible. The discretionary uninformed investors allocated the liquidity trading w_i^* for each market i, as defined in Proposition 1. After the introduction of an ETF, these uninformed investors reconsider their portfolio choices among security 1, security 2, and an ETF. Their objective function can be described as follows.

$$U = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}^T \begin{bmatrix} \overline{r}_1 \\ \overline{r}_2 \\ \overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2 \end{bmatrix} - \frac{\gamma}{2} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}^T \Sigma \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} - Var(l_q)\left[\lambda_1 w_1^2 + \lambda_2 w_2^2 + \lambda_3 w_3^2\right],$$

and

$$w_1 + w_2 + w_3 = 1, \quad \mathbb{E}(w_1 l_q P_1) \geq W_1, \quad \mathbb{E}(w_2 l_q P_2) \geq W_2.$$

Since minimum trading requirement of ETFs are much smaller than individual securities, I assume that ETFs do not require a minimum trading capital requirement for the simplicity. The return covariance matrix Σ is defined in the following.

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix},$$

where

$$\begin{aligned}
&\sigma_{11} = \sigma_1^2, \quad \sigma_{22} = \sigma_2^2, \\
&\sigma_{33} = Var(\overline{w}_1\tilde{r}_1 + \overline{w}_2\tilde{r}_2) = \overline{w}_1^2\sigma_1^2 + \overline{w}_2^2\sigma_2^2 - 2\rho\overline{w}_1\overline{w}_2\sigma_1\sigma_2, \\
&\sigma_{13} = \sigma_{31} = Cov(\tilde{r}_1, \overline{w}_1\tilde{r}_1 + \overline{w}_2\tilde{r}_2) = \overline{w}_1\sigma_1^2 - \overline{w}_2\rho\sigma_1\sigma_2, \\
&\sigma_{12} = \sigma_{21} = -\rho\sigma_1\sigma_2, \\
&\sigma_{23} = \sigma_{32} = Cov(\tilde{r}_2, \overline{w}_1\tilde{r}_1 + \overline{w}_2\tilde{r}_2) = -\overline{w}_1\rho\sigma_1\sigma_2 + \overline{w}_2\sigma_2^2.
\end{aligned}$$

I solve for an optimal trading allocation (w_1^*, w_2^*, w_3^*) by setting up a constrained optimization problem, similar to the method adopted in proof for Proposition 1.

$$\begin{aligned}
U(w_1, w_2) \;=\; & w_1\overline{r}_1 + w_2\overline{r}_2 + w_3(\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) \\
& -\frac{\gamma}{2}\left[w_1^2\sigma_{11} + 2w_1w_2\sigma_{12} + 2w_1w_3\sigma_{13} + w_2^2\sigma_{22} + 2w_2w_3\sigma_{23} + w_3^2\sigma_{33}\right] \\
& -\lambda_1 w_1^2 Var(l_q) - \lambda_2 w_2^2 Var(l_q) - \lambda_3 w_3^2 Var(l_q) \\
& -\eta_1\left[\mathbb{E}(w_1 l_q P_1) - W_1\right] - \eta_2\left[\mathbb{E}(w_2 l_q P_2) - W_2\right].
\end{aligned}$$

In addition, the following first order conditions should be met.

$$\frac{\partial U}{\partial w_1} = 0, \quad \frac{\partial U}{\partial w_2} = 0, \tag{B.1}$$

$$\eta_1 \left[\mathbb{E}(w_1 l_q P_1) - W_1\right] = 0, \quad \eta_2 \left[\mathbb{E}(w_2 l_q P_2) - W_2\right] = 0, \tag{B.2}$$

$$\eta_1 \geq 0, \quad \eta_2 \geq 0, \tag{B.3}$$

$$\mathbb{E}(w_1 l_q P_1) \geq W_1, \quad \mathbb{E}(w_2 l_q P_2) \geq W_2. \tag{B.4}$$

First, I focus on interior solutions: $\eta_1 = 0$, and $\eta_2 = 0$. Then,

$$\begin{aligned}
\frac{\partial U}{\partial w_1} &= \overline{r}_1 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - \frac{\gamma}{2}\left[2w_1\sigma_{11} + 2w_2\sigma_{12} + 2\sigma_{13}(w_3 - w_1) - 2w_2\sigma_{23} - 2w_3\sigma_{33}\right] \\
&\quad -Var(l_q)\left[\frac{\partial \lambda_1}{\partial w_1} w_1^2 + 2\lambda_1 w_1\right] - Var(l_q)\left[\frac{\partial \lambda_3}{\partial w_1} w_3^2 - 2\lambda_3 w_3\right] \\
&= 0.
\end{aligned}$$

I note that $\lambda_1 = \sqrt{\frac{k_1}{(k_1+1)^2}\frac{Var(v_1-\mu_1)}{Var(u_1)+Var(w_1 l_q)}} = \sqrt{A_1}(Var(u_1) + Var(w_1 l_q))^{-1/2}$, where $A_1 = \frac{k_1}{(k_1+1)^2} Var(v_1 - \mu_1)$. The first partial derivative $\partial\lambda_1/\partial w_1 = -\lambda_1 w_1 Var(l_q)/(Var(u_1) + Var(w_1 l_q))$. In addition, I assume $u_1 = w_1 u$.

$$\begin{aligned}
\frac{\partial \lambda_1}{\partial w_1} w_1^2 + 2\lambda_1 w_1 &= \sqrt{\frac{A_1}{Var(u) + Var(l_q)}}\left[\frac{2Var(u) + Var(l_q)}{Var(u) + Var(l_q)}\right], \text{ independent of } w_1, \\
&:= C_1.
\end{aligned}$$

Similarly, $\lambda_3 = \sqrt{\frac{1}{Var(u_3)+Var(w_3 l_q)}\sum_{i=1}^{2}\frac{k_i}{(k_i+1)^2}\overline{w}_i^2 Var(v_i - \mu_i)} = \sqrt{A}(Var(u_3)+Var(w_3 l_q))^{-1/2}$, where $A = \sum_{i=1}^{2}\frac{k_i}{(k_i+1)^2}\overline{w}_i^2 Var(v_i - \mu_i)$. The first order partial derivative $\partial\lambda_3/\partial w_1 = -\partial\lambda_3/w_3 = \lambda_3 w_3 Var(l_q)/(Var(u_3) + Var(w_3 l_q))$. I assume $u_3 = w_3 u$. Hence,

$$\begin{aligned}
\frac{\partial \lambda_3}{\partial w_1} w_3^2 - 2\lambda_3 w_3 &= -\sqrt{\frac{A}{Var(u) + Var(l_q)}}\left[\frac{2Var(u) + Var(l_q)}{Var(u) + Var(l_q)}\right], \text{ independent of } w_1, \\
&:= C.
\end{aligned}$$

Plugging all of these intermediate calculations to the first FOC, I obtain the following.

$$\begin{aligned} &\gamma(\sigma_{11}-\sigma_{13})w_1+\gamma(\sigma_{12}-\sigma_{23})w_2+\gamma(\sigma_{13}-\sigma_{33})w_3 \\ &=\overline{r}_1-(\overline{w}_1\overline{r}_1+\overline{w}_2\overline{r}_2)-Var(l_q)(C_1+C). \end{aligned} \tag{B.5}$$

Second, the FOC with respect to w_2:

$$\begin{aligned} \frac{\partial U}{\partial w_2} &= \overline{r}_2-(\overline{w}_1\overline{r}_1+\overline{w}_2\overline{r}_2)-\frac{\gamma}{2}\left[2w_1\sigma_{12}-2w_1\sigma_{13}+2w_2\sigma_{22}+2(w_3-w_2)\sigma_{23}-2w_3\sigma_{33}\right] \\ &\quad -Var(l_q)\left[\frac{\partial\lambda_2}{\partial w_2}w_2^2+2\lambda_2 w_2\right]-Var(l_q)\left[\frac{\partial\lambda_3}{\partial w_2}w_3^2-2\lambda_3 w_3\right] \\ &= 0. \end{aligned}$$

I follow the similar steps as before and find the following.

$$\gamma(\sigma_{12}-\sigma_{13})w_1+\gamma(\sigma_{22}-\sigma_{23})w_2+\gamma(\sigma_{23}-\sigma_{33})w_3=\overline{r}_2-(\overline{w}_1\overline{r}_1+\overline{w}_2\overline{r}_2)-Var(l_q)(C_2+C), \tag{B.6}$$

where

$$C_2:=\sqrt{\frac{A_2}{Var(u)+Var(l_q)}}\left[\frac{2Var(u)+Var(l_q)}{Var(u)+Var(l_q)}\right],\quad A_2=\frac{k_2}{(k_2+1)^2}Var(v_2-\mu_2).$$

Along with the condition $w_1+w_2+w_3=1$, and the two equations (B.5) and (B.6), I obtain the optimum allocation decision.

$$\begin{bmatrix} 1 & 1 & 1 \\ \gamma(\sigma_{11}-\sigma_{13}) & \gamma(\sigma_{12}-\sigma_{23}) & \gamma(\sigma_{13}-\sigma_{33}) \\ \gamma(\sigma_{12}-\sigma_{13}) & \gamma(\sigma_{22}-\sigma_{23}) & \gamma(\sigma_{23}-\sigma_{33}) \end{bmatrix}\begin{bmatrix} w_1^* \\ w_2^* \\ w_3^* \end{bmatrix}=\begin{bmatrix} 1 \\ \overline{r}_1-(\overline{w}_1\overline{r}_1+\overline{w}_2\overline{r}_2)-Var(l_q)(C_1+C) \\ \overline{r}_2-(\overline{w}_1\overline{r}_1+\overline{w}_2\overline{r}_2)-Var(l_q)(C_2+C) \end{bmatrix}.$$

I reduce the system of 3 equations to a system of 2 equations by substituting w_3 with

$1 - w_1 - w_2$ as follows.

$$\gamma \begin{bmatrix} (\sigma_{11} - \sigma_{13}) - (\sigma_{13} - \sigma_{33}) & (\sigma_{12} - \sigma_{23}) - (\sigma_{13} - \sigma_{33}) \\ (\sigma_{12} - \sigma_{13}) - (\sigma_{23} - \sigma_{33}) & (\sigma_{22} - \sigma_{23}) - (\sigma_{23} - \sigma_{33}) \end{bmatrix} \begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix}$$
$$= \begin{bmatrix} \overline{r}_1 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - Var(l_q)(C_1 + C) - \gamma(\sigma_{13} - \sigma_{33}) \\ \overline{r}_2 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - Var(l_q)(C_2 + C) - \gamma(\sigma_{23} - \sigma_{33}) \end{bmatrix}.$$

I find the following solution:

$$\begin{bmatrix} w_1^* \\ w_2^* \end{bmatrix} = \frac{1}{\gamma} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}^{-1} \begin{bmatrix} \overline{r}_1 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - Var(l_q)(C_1 + C) - \gamma(\sigma_{13} - \sigma_{33}) \\ \overline{r}_2 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - Var(l_q)(C_2 + C) - \gamma(\sigma_{23} - \sigma_{33}) \end{bmatrix}.$$

where

$$B_{11} = (\sigma_{11} - \sigma_{13}) - (\sigma_{13} - \sigma_{33}), \quad B_{12} = (\sigma_{12} - \sigma_{23}) - (\sigma_{13} - \sigma_{33}),$$
$$B_{21} = (\sigma_{12} - \sigma_{13}) - (\sigma_{23} - \sigma_{33}), \quad B_{22} = (\sigma_{22} - \sigma_{23}) - (\sigma_{23} - \sigma_{33}).$$

Equivalently,

$$w_1^* = \frac{1}{\gamma D} \begin{pmatrix} B_{22}\left[\overline{r}_1 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - Var(l_q)(C_1 + C) - \gamma(\sigma_{13} - \sigma_{33})\right] \\ -B_{12}\left[\overline{r}_2 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - Var(l_q)(C_2 + C) - \gamma(\sigma_{23} - \sigma_{33})\right] \end{pmatrix}, \quad \text{(B.7)}$$
$$w_2^* = \frac{1}{\gamma D} \begin{pmatrix} B_{21}\left[\overline{r}_1 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - Var(l_q)(C_1 + C) - \gamma(\sigma_{13} - \sigma_{33})\right], \\ +B_{11}\left[\overline{r}_2 - (\overline{w}_1\overline{r}_1 + \overline{w}_2\overline{r}_2) - Var(l_q)(C_2 + C) - \gamma(\sigma_{23} - \sigma_{33})\right] \end{pmatrix}, \quad \text{(B.8)}$$
$$w_3^* = 1 - w_1^* - w_2^*, \quad \text{(B.9)}$$

where $D = B_{11}B_{12} - B_{12}B_{21}$. If the allocations w_i^* $(i = 1, 2)$ in (B.7) and (B.8) satisfy the condition (B.4), then the investors achieve the optimum utility.

Importantly, I note that to the extent that $w_3^* > 0$, the allocation to the assets 1 and 2 are reduced. This implies that after the introduction of ETFs, the investors migrate to ETFs, and reduce the allocation to individual assets 1 and 2: $\tilde{w}_i^* < w_i^*$, $(i = 1, 2)$. Furthermore, the corner solutions or no solutions are not feasible.

Furthermore, I denote the liquidity of the underlying assets before and after the introduction of ETFs by λ_i, $\tilde{\lambda}_i$, respectively. Then,

$$\lambda_i = \sqrt{\frac{1}{Var(w_i^* u) + Var(w_i^* l_q)} \frac{k_i}{(k_i+1)^2} Var(v_i - \mu_i)},$$

and

$$\tilde{\lambda}_i = \sqrt{\frac{1}{Var(\tilde{w}_i^* u) + Var(\tilde{w}_i^* l_q)} \frac{k_i}{(k_i+1)^2} Var(v_i - \mu_i)}.$$

Since $\tilde{w}_i^* < w_i^*$, the liquidity of the underlying assets decrease after the introduction of ETFs: $\tilde{\lambda}_i^* > \lambda_i^*$. If all the discretionary investors $(q = 1, \cdots, Q)$ migrate to ETF market simultaneously, then the improvement of liquidity will be larger.

Case 2: Only the market 1 was accessible. The investors consider the portfolio allocation between the security 1 and an ETF. I repeat the similar analysis as in Case 1. I first set up the Lagrangian of the objective function for a constrained optimization problem.

$$\begin{aligned} U(w_1, w_3) &= w_1\bar{r}_1 + w_3(\overline{w}_1\bar{r}_1 + \overline{w}_2\bar{r}_2) \\ &\quad -\frac{\gamma}{2}\left[w_1^2\sigma_{11} + 2w_1w_3\sigma_{13} + w_3^2\sigma_{33}\right] \\ &\quad -\lambda_1 w_1^2 Var(l_q) - \lambda_3 w_3^2 Var(l_q) \\ &\quad -\eta_1\left[\mathbb{E}(w_1 l_q P_1) - W_1\right]. \end{aligned}$$

In addition to $w_1 + w_3 = 1$, the following complementarity conditions should be met.

$$\frac{\partial U}{\partial w_1} = 0, \quad \frac{\partial U}{\partial w_3} = 0, \tag{B.10}$$

$$\eta_1\left[\mathbb{E}(w_1 l_q P_1) - W_1\right] = 0, \tag{B.11}$$

$$\eta_1 \geq 0, \quad \eta_2 \geq 0, \tag{B.12}$$

$$\mathbb{E}(w_1 l_q P_1) \geq W_1. \tag{B.13}$$

Similar to Case 1, I solve for the interior solutions, w_1^* and w_3^*.

$$\begin{aligned} w_1^* &= \frac{1}{\gamma(\sigma_{11} - 2\sigma_{13} + \sigma_{33})} \left[\overline{r}_1 - (\overline{w}_1\overline{r}_1 + \overline{w}_3\overline{r}_3) - Var(l_q)(C_1 + C) \right. \\ &\qquad \left. -\gamma(\sigma_{13} - \sigma_{33})\right], \qquad \text{(B.14)} \\ w_3^* &= 1 - w_1^*. \end{aligned}$$

I assume these allocations are feasible, i.e. $\mathbb{E}(w_1^* l_q P_1) \geq W_1$. This implies that flows of new investors invest and trade in the ETF market. Since $\tilde{w}_1^* < w_1^*$, the corner solutions are not feasible.

I note that these new influx of investors in ETF markets may have differential impact on the liquidity of the underlying assets. For security 2, as arbitrageurs connect ETF market and individual markets, additional uninformed trading have been supplied. Therefore, the price impact in the security 2 market becomes

$$\tilde{\lambda}_2 = \sqrt{\frac{1}{Var(u_2) + Var(\theta^*) + Cov(\theta^*, u_2)} \frac{k_2}{(k_2+1)^2} Var(v_2 - \mu_2)}, \tag{B.15}$$

while $\lambda_2 = \sqrt{\frac{1}{Var(u_2)} \frac{k_2}{(k_2+1)^2} Var(v_2 - \mu_2)}$. Since $\tilde{\lambda}_2 \leq \lambda_2$, the liquidity of security 2 improves.

On the other hand, the liquidity of security 1 is ambiguous. It is because the existing uninformed investors migrate to ETF, which will decrease the liquidity of security 1. At the same time new uninformed investors flow into ETF, which will bring in uninformed investors back to security 1 by arbitrageurs. Formally, before the introduction of ETF, $\lambda_1 = \sqrt{\frac{1}{Var(u_1)+Var(w_1^* l_q)} \frac{k_1}{(k_1+1)^2} Var(v_1 - \mu_1)}$. After the introduction of ETF, $\tilde{\lambda}_1 = \sqrt{\frac{1}{Var(u_1)+Var(\tilde{w}_1^* l_q)+Var(\theta^*)+Cov(\theta^*,u_1)} \frac{k_1}{(k_1+1)^2} Var(v_1 - \mu_1)}$. However, $\tilde{w}_1^* < w_1^*$, and $\theta^* > 0$. The predictions as between $\tilde{\lambda}_1$ and λ_1 are uncertain.

<u>Case 3:</u> The two underlying markets were inaccessible. The investors consider the portfolio

allocation to the ETF. The objective function will be described as

$$U(w_3) \quad = \quad w_3\overline{r}_3 - \frac{\gamma}{2}w_3^2\sigma_3^2 - \lambda_3 w_3^2 Var(l_q).$$

I repeat the same procedure as before and find that

$$w_3^* = \frac{1}{\gamma\sigma_3^2}\left[\overline{r}_3 - CVar(l_q)\right], \tag{B.16}$$

where

$$C = \sqrt{\frac{A}{Var(u)+Var(l_q)}}\left[\frac{2Var(u)+Var(l_q)}{Var(u)+Var(l_q)}\right], \quad A = \sum_{i=1}^{2}\frac{k_i}{(k_i+1)^2}\overline{w}_i^2 Var(v_i-\mu_i).$$

This implies ETF market has the new influx of uninformed trading by $w_3^* l_q$. The arbitrageurs connect the ETF market and the underlying market. I denote their optimal trading activity by θ^*. The liquidity after the introduction of ETF become

$$\tilde{\lambda}_i = \sqrt{\frac{k_i}{(k_i+1)^2}\frac{Var(v_i-\mu_i)}{Var(u_i)+Var(w_3^* l_q)+Cov(w_3^* l_q, \theta^*)}}. \tag{B.17}$$

In contrast, before the introduction of ETFs, discretionary investors were nonexistent. The liquidity of the underlying assets are

$$\lambda_i = \sqrt{\frac{k_i}{(k_i+1)^2}\frac{Var(v_i-\mu_i)}{Var(u_i)}}. \tag{B.18}$$

By comparing the liquidity before and after the introduction of ETFs, (B.17) and (B.18), I conclude $\lambda_i > \tilde{\lambda}_i$. Hence, in the inaccessible market, the introduction of ETF improves the liquidity of the underlying securities. □

References

Abudy, M., Wohl, A., 2015. Corporate Bond Trading by Retail Investors in a Limit Order Book Exchange. Tel Aviv University Working paper .

Becker, B., Ivashina, V., 2015. Reaching for Yield in the Bond Market. Journal of Finance 70, 1863–1902.

Ben-David, I., Franzoni, F., Moussawi, R., 2015. Do ETFs Increase Volatility? Ohio State University Working Paper .

Bessembinder, H., Jacobsen, S., Maxwell, W., Venkataraman, K., 2016. Capital Commitment and Illiquidity in Corporate Bonds. Arizona State University and Southern Methodist University Working Paper .

Bessembinder, H., Maxwell, W., Venkataraman, K., 2006. Market transparency, liquidity externalities, and institutional trading costs in corporate bonds. Journal of Financial Economics 82, 251–288.

Bhattacharya, A., O'Hara, M., 2016. Can etfs Increase Market Fragility? Effect of Information Linkages in ETF Markets. Cornell University Working Paper .

Biais, B., Green, R., 2007. The Microstructure of the Bond Market in the 20th Century. Carnegie Mellon University Working Paper .

Bongaerts, D., Cremers, M., Goetzmann, W. N., 2012. Tiebreaker: Certification and Multiple Credit Ratings. Journal of Finance 67, 113–152.

Campbell, J., Taksler, G., 2003. Equity Volatility and Corporate Bond Yields. Journal of Finance 58, 2321–2349.

Carey, M., Prowse, S., Rea, J., Udell, G., 1993. The Economics of the Private Placement Market. The Staff Report by the Board of Governors pp. 1–111.

Chaplisky, S., Ramchand, L., 2004. The Borrowing Costs of International Issuers: SEC Rule 144A. Journal of Business 77, 1073–1097.

Chen, L., Lesmond, D. A., Wei, J., 2007. Corporate yield spreads and bond liquidity. Journal of Finance 62, 119–149.

Cornaggia, J., Cornaggia, K. J., 2013. Estimating the Costs of Issuer-paid Credit Ratings. Review of Financial Studies 26, 2229–2269.

Da, Z., Shive, S., 2013. When the Bellwether Dances to Noise: Evidence from Exchange-Traded-Funds. University of Notre Dame Working Paper .

Dannhauser, C. D., 2016. The Impact of Innovation: Evidence from Corporate Bond ETFs. Journal of Financial Economics, forthcoming .

Edwards, A. K., Harris, L. E., Piwowar, M. S., 2007. Corporate Bond Market Transaction Costs and Transparency. Journal of Finance 62, 1421–1451.

Ellul, A., Jotikasthira, P., Lundblad, C. T., 2011. Regulatory pressure and fire sales in the corporate bond market. Journal of Financial Economics 101, 596–620.

Fenn, G. W., 2000. Speed of issuance and the adequacy of disclosure in the 144a high-yield debt market. Journal of Financial Economics 56, 383–405.

Fong, K. Y. L., Holden, C. W., Trzcinka, C. A., 2016. What Are the Best Liquidity Proxies for Global Research? Indiana University Working Paper .

Goldstein, M. A., Hotchkiss, E. S., Sirri, E. R., 2007. Transparency and liquidity: A Conrolled Experiment on Corporate Bonds. Review of Financial Studies 20, 235–273.

Gorton, G. B., Pennacchi, G. G., 1993. Security Baskets and Index-linked Securities. The Journal of Business 66, 1–27.

Grossman, S. J., Miller, M. H., 1987. Liquidity and market structure. Journal of Finance 43, 617–633.

Guedj, I., Huang, J., 2015. Are ETFs Replacing Index Mutual Funds? University of Texas at Austin Working Paper .

Hamm, S. J. W., 2014. The Effect of ETFs on Stock Liquidity. the Ohio State University Working paper .

Hendershott, T., Madhavan, A., 2015. Click or Call? Auction versus Search in the Over-the-Counter Market. Journal of Finance 70, 419–447.

Holden, C. W., 1995. Index arbitrage as cross-sectional market making. Journal of Futures Markets 15, 423–455.

Holden, C. W., Mao, Y., Nam, J., 2016. Price Discovery in the Stock and Corporate Bond Markets. Cornell University and Indiana Unviersity Working Paper .

Hotchkiss, E., Ronen, T., 2002. Informational efficiency of the corporate bond market: an intraday analysis. Review of Financial Studies 15, 1325–1354.

Huang, R., Stoll, H., 1997. The components of the bid-ask spread: a general approach. Review of Financial Studies 10, 995–1034.

Huberman, G., 2001. Familiarity Breeds Investment. Review of Financial Studies 14, 659–680.

Israeli, D., M. C. Lee, C., Sridharan, S., 2015. Is there a Dark Side to Exchange Traded Funds (ETF)? An Information Perspective. Working paper .

Jegadeesh, N., Subrahmanyam, A., 1993. Liquidity Effects of the Introduction of the S&P 500 Index Futures Contract on the Underlying Stocks. Journal of Business 66, 171–187.

Kang, J.-K., Stulz, R. A. M., 1997. Why is There a Home Bias? an Analysis of Foreign Portfolio Equity Ownership in Japan. Journal of Financial Economics 46, 3–28.

Kyle, A. S., 1985. Continuous Auctions and Insider Trading. Econometrica 53, 1315–1336.

Lesmond, D. A., Ogden, J., Trzcinka, C. A., 1999. A New Estimate of Transaction Costs. Review of Financial Studies 12, 1113–1141.

Lo, W. A., Wang, J., 2000. Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory. Review of Financial Studies 13, 257–300.

Merton, R. C., 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information. The Journal of Finance 42, 483–510.

Pan, K., Zeng, Y., 2017. ETF Arbitrage under Liquidity Mismatch. Harvard University and University of Washington Working Paper .

Pool, V. K., Stoffman, N., Yonker, S. E., 2012. No Place Like Home: Familiarity in Mutual Fund Manager Portfolio Choice. Review of Financial Studies 25, 2563–2599.

Randall, O., 2015. Pricing and Liquidity in Over-The-Counter Markets. Emory University Working Paper .

Schestag, R., Schuster, P., Uhrig-Homburg, M., 2016. Measuring Liquidity in Bond Markets. Forthcoming in Review of Financial Studies .

Schultz, P., 2001. Corporate bond trading costs: a peek behind the curtain. Journal of Finance 56, 677–698.

Stein, J. C., 2005. Why are most Funds Open-end? Competition and the Limits of Arbitrage. The Quarterly Journal of Economics 120, 247–272.

Subrahmanyam, A., 1991. A Theory of Trading in Stock Index Futures. Review of Financial Studies 4, 17–51.

Wooldridge, J. M., 2010. Econometric Analysis of Cross Sectional and Panel Data. The MIT Press.

Figure 1. **Distribution of transaction costs and minimum trading quantities for individual junk bonds.** These bonds are held by the junk bond ETF portfolio (JNK) as of March 30, 2016. For each bond, the bid-ask spreads is divided by their mid point price. The average bid-ask spread is 110 basis points. The minimum trading quantities vary with bid or ask orders. The average minimum bid quantity $133,000, while the average minimum ask quantity is $116,000. The median bid or ask quantity is $100,000. Data source: Fidelity corporate bond database ("Depth of Book").

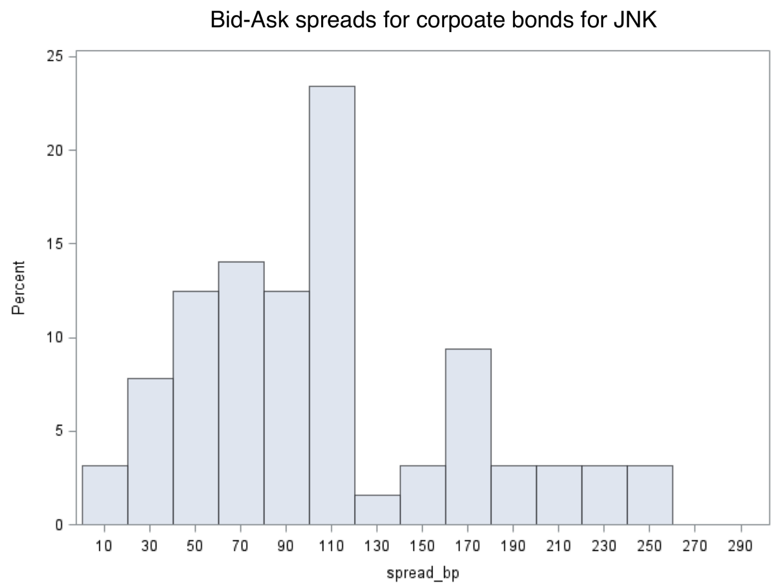

Figure 2. **Model overview for less accessible markets.** (A) Before the introduction of ETFs, investors' portfolios may have limited access to some market segments due to a minimum trading requirement, high transaction costs and other reasons. The cost of having a diversified portfolio is high. These diversification costs increases as the portfolio is large or the investors trade over multiple periods. The investors tend to have an under-diversified portfolio. (B) ETFs significantly lower the diversification cost than a portfolio of individual securities. First, investors flow into ETFs due to low cost of diversification. If the market was less accessible *ex ante*, then the influx of new investors to ETFs is greater. Second, ETFs are liquid. Third, large ETF liquidity is spilled over to the underlying securities through arbitrage mechanism. The liquidity of the underlying securities improves in the previously less accessible market.

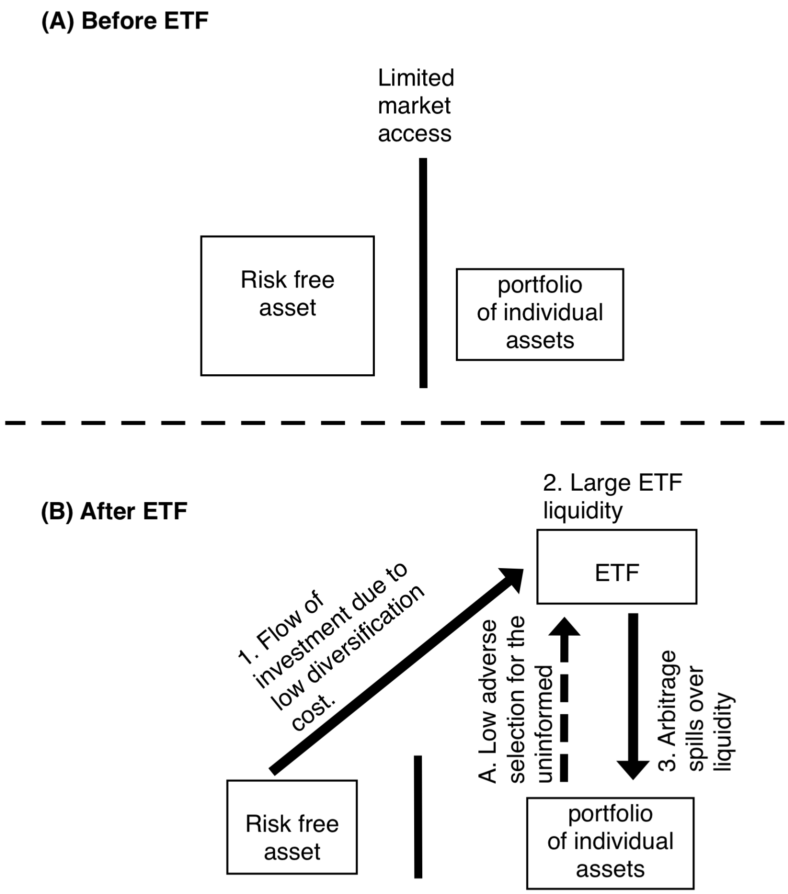

Figure 3. **Theoretical predictions after the introduction of ETFs.** (A) If the market was highly accessible *ex ante*, discretionary uninformed investors can easily trade a portfolio of individual assets. With the introduction of ETFs, they prefer to trade ETFs in order to reduce further the adverse selection cost. This migration of the uninformed investors causes liquidity of underlying securities to decrease. (B) Suppose the market was partially accessible, i.e. only security 1. Then the uninformed investors would prefer to migrate to ETF, which will reduce the liquidity of security 1. At the same time, the uninformed investors who hold the risk free assets would prefer to move investment from the risk free asset to ETF. The large ETF liquidity will spill over the underlying securities 1 and 2. Hence, the liquidity of security 1 is ambiguous, while the liquidity of security 2 improves. (C) If the market is inaccessible, then the liquidities of both security 1 and security 2 improve by applying similar logic.

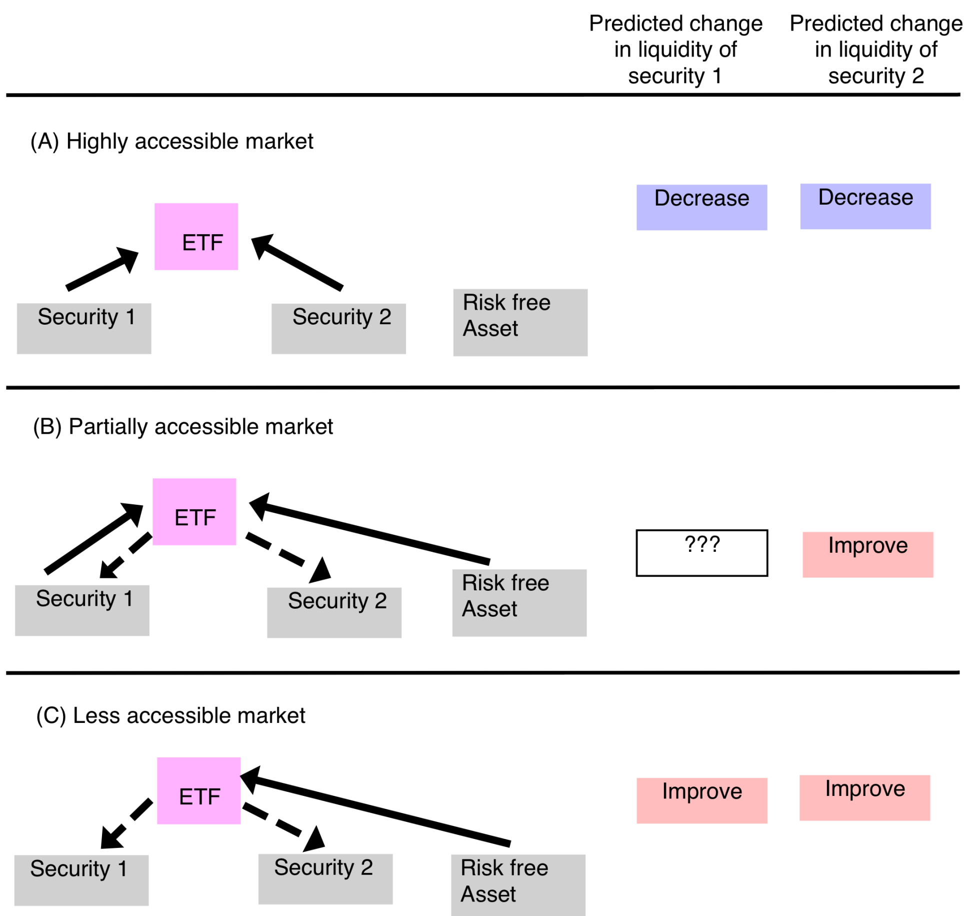

Figure 4. **Number of ETF bonds whose trading was reported to TRACE.** The number of the incumbent bonds of corporate bond ETFs whose trading was reported to TRACE has grown by more than 10 times since 2008. As of 2014Q4, 1253 distinct corporate bonds were traded and reported to TRACE. 144A bonds that were only available to institutional investors through private placement are increasingly included in ETFs since 2013. ETF holdings are obtained from the CRSP portfolio holdings. The transactions of regular or 144A corporate bonds are obtained from TRACE.

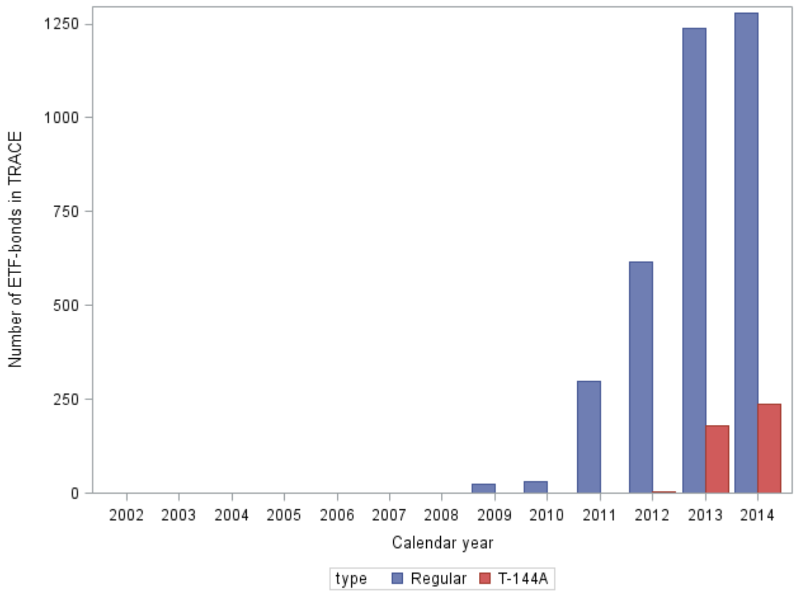

Table 1: **ETF summary statistics.** (1) N is the number of the existent ETF in each year. (2) AUM is the assets under management in million dollars among the existent ETFs in the 4^{th} quarter of each year. (3) N(Hold) is the average number of holdings in each quarter per ETF. (4) %(Hold) is the total TNA reported in CRSP mutual fund holdings, relative to AUM. (5) N is the number of the newly launched ETFs in each year whose holdings are reported within 6 months since the inception date. (6) The average number of holdings of the new ETFs. (7) The aggregate % of the total net assets for the new ETFs. (1), (2), (5) are from CRSP mutual fund database. (3), (4) and (6) and (7) are based on the CRSP portfolio holdings database.

	All ETF				New ETF		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
yr	N	AUM	N(Hold)	%(Hold)	N	N(Hold)	%(Hold)
2002					1		
2003	2	2,520			1		
2004	2	3,510					
2005	2	5,312			1		
2006	2	7,632			7		
2007	34	15,267	24	71	36	22	71
2008	52	33,832			24		
2009	60	70,346	1	4	24	1	0
2010	107	103,411	235	102	28	262	100
2011	157	144,663	244	105	40	57	107
2012	184	204,786	281	107	31	83	110
2013	209	213,094	318	111	36	114	108
2014	243	254,493	329	104	38	95	94

Table 2: **Bond descriptive statistics.** For credit rating, I assign a numeric score for each credit rating by each of the three rating agencies – S& P, Moody's, Fitch – following Becker and Ivashina (2015). At each given time, I merge the three credit rating scores as follows: If the three ratings are available, I take the middle rating. If two ratings are available, I take the lower. Non-ETF bonds are matched with ETF bonds using issue amount, credit rating, maturity at issuance, age at the inception date of ETF, and cumulative trading volume during 6 months prior to the inception date of ETF.

	ETF bonds		non-ETF bonds	
Total number of bond issues	2,637		2,637	
Average issue size ($M)	1,007		1,032	
$\leq$ 100M	16	1%	517	20%
100-500M	726	28%	692	26%
$>$ 500M	1,895	72%	1,428	54%
Average credit quality	19.37		19.07	
$AA-$ and up	283	11%	351	13%
$BBB-$ thru $A+$	1,729	66%	1,554	59%
below $BBB-$	625	24%	732	28%
Average maturity at issue (in years)	7.37		7.87	
$<$ 2 years	21	1%	105	4%
2-5 years	465	18%	597	23%
5-20 years	2,048	78%	1,845	70%
$>$ 20 years	103	4%	90	3%
Age at the inception date of ETF (in years)	1.39		1.60	
$<$ 2 years	1,955	74%	1,959	74%
2-5 years	663	25%	617	23%
$\geq$ 5 years	19	1%	61	2%

Table 3: **Bond trading characteristics.** Transaction price, transaction volume, transaction frequency. To compute average transaction cost per day, I first use a subset of corporate bonds whose buy and sell transactions are observed for each day. I compute the average sell price subtracted by the average buy price. To estimate the average trading cost per transaction, I use the last trading price on the most recent transaction day as a benchmark price and examine the changes of trading prices on the current day relative to the benchmark price, following the insight of Huang and Stoll (1997); Hendershott and Madhavan (2015). Transaction costs are further examined in details based on the cumulative trading volume during 6 months prior to the inception date of ETF.

	ETF bonds		non-ETF bonds	
	before	after	before	after
Average trade price(% of par value)	101.11	101.01	100.40	99.93
Total trading volume ($B)	1341.73	1122.89	1009.62	788.94
Total number of trades (1000)	2304.62	2364.67	2143.66	1913.70
< $100K	1291.74	1419.42	1280.89	1194.49
$100K-1M	547.10	565.50	466.48	403.93
$1-5M	351.69	291.44	321.79	261.44
>$5M	114.10	88.31	74.50	53.84
Average two-way transaction cost per day (bp)	65.72	67.21	75.42	76.74
Average transaction cost per transaction (bp)				
VOL 1 (small)	47.69	32.78	42.75	52.44
VOL 2	27.11	22.54	27.64	30.85
VOL 3	21.82	22.52	25.08	24.20
VOL 4	20.18	20.00	24.15	28.78
VOL 5 (large)	20.15	21.45	17.81	19.72

Table 4: **Full sample analysis.** The effective half spreads of ETF-bonds are compared with those of non-ETF bonds. Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is $+1$ for the buy order, -1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with $Post$ and ETF. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include three public information variables, each of which is measured on the date of the most recent transaction on a day prior to the current transaction. The changes in the interest rate for on-the-run Treasury security is matched to the corporate bond based on maturity. The percentage return on the issuing firm's common stock. The change in the default spread between long-term indexes of BAA-rated bonds and U. S. Treasury securities. I interact these variables with investment- and non-investment-grade indicator variables to account for potential differences in sensitivity based on the bond's risk. I estimate the first and second stage regressions using the weighted least squares, where the weight is inversely related to the number of days elapsed between Q_t and Q_{t-1}. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	coeff	p-value
First stage regression		
Intercept	0.1137***	0.000
Q_{t-1}	0.0678***	0.000
Second stage regression		
Intercept	0.0276*	0.060
Q^*	0.0316***	0.000
Treasury return	−0.0119***	0.000
Stock return × invgrd	0.0176***	0.000
Stock return × noninvgrd	0.0745***	0.000
Default spread × invgrd	−0.0074***	0.000
Default spread × noninvgrd	0.0386***	0.000
ΔQ	0.4221***	0.000
$\Delta Q\times$ Post	0.0262***	0.000
$\Delta Q\times$ ETF	−0.0537***	0.000
$\Delta Q\times$ Post × ETF	−0.0132***	0.000
R^2	0.1804	
RMSE	0.7324	
N of block boostrapping	50	

Table 5: **By Arbitrage activities** To estimate the arbitrage activities of each ETF, I add the magnitude of the daily changes of the ETF shares outstanding in the post event period, relative to the average ETF shares outstanding. If the arbitrage activities fall below (or above) the 40th percentile, i.e. 125%, I consider it low (or high) arbitrage ETFs. The effective half spreads of ETF-bonds are compared with those of non-ETF bonds. Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is +1 for the buy order, −1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a+\gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with $Post$ and ETF. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include three public information variables X_t as discussed in 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	coeff	p-value
First stage regression		
Intercept	0.1074***	0.000
Q_{t-1}	0.0685***	0.000
Second stage regression		
Intercept	0.0372***	0.000
Q^*	0.0325***	0.000
Treasury return	−0.0157***	0.000
Stock return × invgrd	0.0262***	0.000
Stock return × noninvgrd	0.0780***	0.000
Default spread × invgrd	−0.0086***	0.000
Default spread × noninvgrd	0.0478***	0.000
Arb H × ΔQ	0.4266***	0.000
Arb H × ΔQ× Post	0.0227***	0.000
Arb H × ΔQ× ETF	0.0227***	0.000
Arb H × ΔQ× Post × ETF	−0.0300***	0.000
Arb L × ΔQ	0.4143***	0.000
Arb L × ΔQ× Post	0.0210***	0.000
Arb L × ΔQ× ETF	−0.0486***	0.000
Arb L × ΔQ× Post × ETF	0.0119***	0.000
R^2	0.1874	
N of block bootstrapping	50	

Table 6: **Changes of the effective half spreads by cumulative trading volume.** The bonds are segmented into the quintiles based on their cumulative trading volume during 6 months prior to ETFs. Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is $+1$ for the buy order, -1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with $Post$ and ETF. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include the three public information variables X_t as discussed in section 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	coeff	p-value
First stage regression		
Intercept	0.1137***	0.000
Q_{t-1}	0.0678***	0.000
Second stage regression		
Intercept	0.0268***	0.000
Q^*	0.0316***	0.000
Treasury return	−0.0116***	0.000
Stock return × invgrd	0.0175***	0.000
Stock return × noninvgrd	0.0745***	0.000
Default spread × invgrd	−0.0075***	0.000
Default spread × noninvgrd	0.0389**	0.040
VOL 1 (Small) $\times\Delta Q$	0.9263***	0.000
VOL 1 $\times\Delta Q\times$ Post	−0.0009	0.300
VOL 1 $\times\Delta Q\times$ ETF	−0.4695***	0.000
VOL 1 $\times\Delta Q\times$ Post × ETF	−0.0881***	0.000
VOL 2 $\times\Delta Q$	0.8714***	0.000
VOL 2 $\times\Delta Q\times$ Post	0.0141**	0.020
VOL 2 $\times\Delta Q\times$ ETF	−0.5017***	0.000
VOL 2 $\times\Delta Q\times$ Post × ETF	−0.0355***	0.000
VOL 3 $\times\Delta Q$	0.7360***	0.000
VOL 3 $\times\Delta Q\times$ Post	−0.0384***	0.000
VOL 3 $\times\Delta Q\times$ ETF	−0.3180***	0.000
VOL 3 $\times\Delta Q\times$ Post × ETF	0.0185***	0.000
VOL 4 $\times\Delta Q$	0.4440***	0.000
VOL 4 $\times\Delta Q\times$ Post	0.0394***	0.000
VOL 4 $\times\Delta Q\times$ ETF	−0.1036***	0.000
VOL 4 $\times\Delta Q\times$ Post × ETF	−0.0015	0.240
VOL 5 (Large) $\times\Delta Q$	0.4159***	0.000
VOL 5 $\times\Delta Q\times$ Post	0.0237***	0.000
VOL 5 $\times\Delta Q\times$ ETF	−0.0528***	0.000
VOL 5 $\times\Delta Q\times$ Post × ETF	−0.0011	0.180
R^2	0.1822	
RMSE	0.7316	
N of block bootstrapping	100	

Table 7: **Changes of the effective half spreads by credit rating of the bonds.** The bonds are divided into two groups based on the bond credit ratings. I assign numeric scores to the bond credit ratings from S&P, Moody's, and Fitch, following Bo and Ivashina (2012). For each bond, if two credit ratings are available, I take the lower. If three credit ratings are available, I take the middle. Investment grades bonds have credit ratings $BBB-$ or above based on S&P or Fitch (or equivalently Baa3 from Moody's). Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is +1 for the buy order, −1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with $Post$ and ETF. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include the three public information variables X_t as discussed in section 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	Investment grade		Non-investment grade	
	coeff	p-value	coeff	p-value
First stage regression				
Intercept	0.1190***	0.000	0.1003***	0.000
Q_{t-1}	0.0672***	0.000	0.0684***	0.000
Second stage regression				
Intercept	0.0072***	0.000	0.1053***	0.000
Q^*	0.0172***	0.000	0.0842***	0.000
Treasury return	−0.0101***	0.000	−0.0270***	0.000
Stock return × invgrd	0.0178***			
Stock return × noninvgrd			0.0747***	0.000
Default spread × invgrd	0.0067***			
Default spread × noninvgrd			0.0341***	0.000
ΔQ	0.3728***	0.000	0.5695***	0.000
$\Delta Q\times$ Post	0.0194***	0.000	0.0214***	0.000
$\Delta Q\times$ ETF	−0.0347***	0.000	−0.0200***	0.000
$\Delta Q\times$ Post × ETF	0.0001	0.200	−0.0542***	0.000
R^2	0.1603		0.2291	
RMSE	0.5493		0.9049	
N of block bootstrapping	100		100	

Table 8: **Changes of the effective half spreads of bonds by the time to maturity.** The time to maturity is measured as the number of days to the maturity on the first offer date of ETF. The bonds are segmented into two groups with a threshold of 2 years i.e. 730 days. Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is +1 for the buy order, −1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with *Post* and *ETF*. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include the three public information variables X_t as discussed in section 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	Life remaining ≤ 2 years		Life remaining ≥ 2 years	
	coeff	*p*-value	coeff	*p*-value
First stage regression				
Intercept	−0.0529***	0.000	0.1298***	0.000
Q_{t-1}	0.0354***	0.000	0.0667***	0.000
Second stage regression				
Intercept	−0.0043	0.290	0.0354***	0.000
Q^*	0.0006***	0.000	0.0385***	0.000
Treasury return	0.0160***	0.000	−0.0144***	0.000
Stock return × invgrd	0.0041***	0.000	0.0194***	0.000
Stock return × noninvgrd	0.0713***	0.000	0.0744***	0.000
Default spread × invgrd	−0.0028***	0.000	−0.0084***	0.000
Default spread × noninvgrd	−0.0661***	0.000	0.0393***	0.000
ΔQ	0.1591**	0.040	0.4547***	0.000
$\Delta Q \times$ Post	−0.0246***	0.000	0.0353***	0.000
$\Delta Q \times$ ETF	−0.0539***	0.000	−0.0861***	0.000
$\Delta Q \times$ Post × ETF	0.0989***	0.000	−0.0190***	0.000
R^2	0.0762		0.1936	
RMSE	0.3026		0.7270	
N of block bootstrapping	100		100	

Table 9: **Quasi-natural experiment for HYG.** I investigate changes of the liquidity of the constituent bonds for HYG around the index rule change during 2009. The effective half spreads of ETF-bonds are compared with those of non-ETF bonds. Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is +1 for the buy order, −1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$.To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with *Post* and *ETF*. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include the three public information variables X_t as discussed in section 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	coeff	p-value
First stage regression		
Intercept	0.0485***	0.00
Q_{t-1}	0.0686***	0.00
Second stage regression		
Intercept	0.1090***	0.00
Q^*	0.0964***	0.00
Treasury return	−0.0052***	0.00
Stock return × invgrd	0.0054***	0.00
Stock return × noninvgrd	0.0025***	0.00
Default spread × invgrd	0.0195***	0.00
Default spread × noninvgrd	0.1943***	0.00
ΔQ	0.4318***	0.00
$\Delta Q \times$ Post	0.0686***	0.00
$\Delta Q \times$ ETF	0.2083***	0.00
$\Delta Q \times$ Post × ETF	−0.0150***	0.00
R^2	0.1469	
RMSE	0.9876	
N of block boostrapping	100	

Table 10: **ETF by management styles.** ETFs can be actively or passively managed. I separate the bonds included by active or index ETFs. The bonds can be owned by both styles of the ETFs, which I exclude from the analysis. The effective half spreads of ETF-bonds are compared with those of non-ETF bonds. Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is +1 for the buy order, −1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with $Post$ and ETF. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include the three public information variables X_t as discussed in section 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	Active ETF		Index ETF	
	coeff	p-value	coeff	p-value
First stage regression				
Intercept	0.1490***	0.000	0.0857***	0.000
Q_{t-1}	0.0640***	0.000	0.0473***	0.000
Second stage regression				
Intercept	0.0092***	0.000	−0.0169***	0.000
Q^*	0.0143***	0.000	0.0518***	0.000
Tresury return	−0.0084***	0.000	0.0240***	0.000
Stock return × Invgrd	0.0150***	0.000	0.0350***	0.000
Stock return × Noninvgrd	0.0500***	0.000	0.0890***	0.000
Default spread × Invgrd	−0.0124***	0.000	0.0024***	0.000
Default spread × Noninvgrd	0.0271***	0.000	0.0490***	0.000
ΔQ	0.4310***	0.000	0.3510***	0.000
$\Delta Q \times Post$	0.0086***	0.000	0.0639***	0.000
$\Delta Q \times ETF$	−0.0530***	0.000	0.0083***	0.000
$\Delta Q \times Post \times ETF$	0.0080***	0.000	−0.1028***	0.000
R^2	0.1813		0.1851	
RMSE	0.6364		0.6762	
N of block bootstrapping	100		100	

Table 11: **Multivariate analysis.** I evaluate the relative strength of the explanatory variables – cumulative trading volume, bond credit rating, age, ETF styles – to the liquidity of the underlying bonds. The effective half spreads of ETF-bonds are compared with those of non-ETF bonds. Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is +1 for the buy order, −1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with $Post$ and ETF. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include the three public information variables X_t as discussed in section 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	All		Invgrad		Non-invgrd	
	coeff	p-value	coeff	p-value	coeff	p-value
First stage regression						
Intercept	0.1137***	0.000	0.1190***	0.000	0.1003***	0.000
Q_{t-1}	0.0678***	0.000	0.0672***	0.000	0.0684***	0.000
Second stage regression						
Intercept	0.0225***	0.000	0.0024***	0.000	0.1034***	0.000
Q^*	0.0360***	0.000	0.0213***	0.000	0.0858***	0.000
Treasury return	−0.0088***	0.000	−0.0070***	0.000	−0.0259***	0.000
Stock return × Invgrd	0.0176***	0.000	0.0178***	0.000		
Stock return × Noninvgrd	0.0742***	0.000			0.0746***	0.000
Default spread × Invgrd	−0.0079***	0.000	−0.0073***	0.000		
Default spread × Noninvgrd	0.0383***	0.000			0.0336***	0.000
VOL 1 (Small) $\times \Delta Q$	0.9692***	0.000	0.9814***	0.000	0.6823***	0.000
VOL 1 $\times \Delta Q \times Post$	0.0056***	0.000	0.0087***	0.000	0.0604***	0.000
VOL 1 $\times \Delta Q \times ETF$	−0.5092***	0.000	−0.5623***	0.000	−0.1966***	0.000
VOL 1 $\times \Delta Q \times Post \times ETF$	−0.0902***	0.000	−0.0482***	0.000	−0.1979***	0.000
VOL 3 $\times \Delta Q$	0.7490***	0.000	0.9482***	0.000	0.3548***	0.000
VOL 3 $\times \Delta Q \times Post$	−0.0271***	0.000	0.0028***	0.000	0.0070***	0.000
VOL 3 $\times \Delta Q \times ETF$	−0.3178***	0.000	−0.6446***	0.000	−0.3601***	0.000
VOL 3 $\times \Delta Q \times Post \times ETF$	−0.0150***	0.000	0.0266***	0.000	−0.1078***	0.000
VOL 5 (Large) $\times \Delta Q$	0.4504***	0.000	0.4019***	0.000	0.5844***	0.000
VOL 5 $\times \Delta Q \times Post$	0.0333***	0.000	0.0276***	0.000	0.0289***	0.000
VOL 5 $\times \Delta Q \times ETF$	−0.0895***	0.000	−0.0386***	0.000	−0.0803***	0.000
VOL 5 $\times \Delta Q \times Post \times ETF$	0.0091	0.000	−0.0039	0.000	0.2300***	0.000
Life $\leq$ 2yr $\times \Delta Q$	−0.3139***	0.000	−0.2789***	0.000	−0.0421***	0.000
Life $\leq$ 2yr $\times \Delta Q \times Post$	−0.0589***	0.000	−0.0548***	0.000	−0.1876***	0.000
Life $\leq$ 2yr $\times \Delta Q \times ETF$	0.0062***	0.000	0.0326***	0.000	0.2370***	0.000
Life $\leq$ 2yr $\times \Delta Q \times Post \times ETF$	0.1351***	0.000	0.0839***	0.000	0.3524***	0.000
R^2	0.1910		0.1755		0.2310	
RMSE	0.6729		0.7774		1.0202	
N of block bootstrapping	100		100		100	

Table 12: **144A vs. public bonds trading characteristics.** 144A and public bonds comparison of average daily trading volume in million dollars and issue amount in billion dollars. Investment grade bonds have credit ratings $BBB-$ and above. The average daily trading volume exploded during the 3rd quarter of 2014. This was prevalent both in investment grade and non-investment grade bonds.

		Average daily trading volume									Issue Amounts					
		Full			Inv grade			Noninv grd			Inv grd			Noninv grd		
Year	Qtr.	144A	Public		144A	Public		144A	Public		144A	Public		144A	Public	
2008	4	1	9262	0%	0	0		1	3386	0%	1092	16636	8%	1168	6373	18%
2010	4	8	11742	0%	20	7467	0%	2	4274	0%	1358	23941	7%	1371	6286	22%
2011	4	3	11455	0%	3	7137	0%	0	4318		676	18573	5%	1021	5371	19%
2012	4	1	9483	0%	0	6042		1	4088	0%	597	5304	8%	1038	11633	9%
2013	4	8	10417	0%	4	4993	0%	8	5611	0%	188	1269	15%	658	4807	14%
2014	1	4	12306	0%	4	5360	0%	4	7153	0%						
2014	2	57	11339	1%	29	4338	1%	53	7250	1%						
2014	3	2440	10638	23%	366	3771	10%	2109	7184	29%						
2014	4	2558	11817	22%	352	4067	9%	2261	7953	28%						
2015	1	2846	13481	21%	351	4549	8%	2531	9330	27%						
2015	2	2896	12855	23%	292	3779	8%	2675	9219	29%						
2015	3	2735	11494	24%	253	3135	8%	2524	8617	29%						
2015	4	2765	11096	25%	239	2892	8%	2561	8377	31%						

Table 13: **Liquidity of 144A bonds.** Half-spreads by all, volume and credit rating with the 144A bonds. The first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is +1 for the buy order, −1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-144A bonds relative to the changes of effective spreads of non-ETF 144A bonds, I interact ΔQ_t with $Post$ and ETF. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include the three public information variables X_t as discussed in section 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	Full		By volume		By credit rating			
			High/Low Vol		Investment grade		Non-investment grade	
	coeff	p-value	coeff	p-value	coeff	p-value	coeff	p-value
First stage regression								
Intercept	-0.0389^{***}	0.000	-0.0389^{***}	0.000	-0.0670^{**}	0.000	-0.0168^{*}	0.090
Q_{t-1}	0.0299^{**}	0.010	0.0299^{**}	0.010	0.0181^{**}	0.100	0.0357^{**}	0.010
Second stage regression								
Intercept	-0.0408^{**}	0.050	-0.0477^{**}	0.030	-0.0974^{***}	0.000	-0.0148	0.036
Q^{*}	0.1004^{**}	0.000	0.0976^{***}	0.000	0.0247^{*}	0.060	0.1115^{***}	0.000
Treasury return	-0.0005	0.540	0.0045	0.310	0.0584^{***}	0.000	-0.0119	0.430
Stock return × invgrd	0.0025	0.480	-0.0059^{**}	0.020	-0.0039^{**}	0.020		
Stock return × noninvgrd	0.0597^{***}	0.000	0.0572^{***}	0.000			0.0600^{***}	0.000
Default spread × invgrd	-1.5579^{***}	0.000	-1.3569^{***}	0.000	-1.2386^{***}	0.000		
Default spread × noninvgrd	0.0675	0.130	0.0032	0.140			0.0881	0.230
ΔQ	0.0231^{*}	0.070			0.0321^{***}	0.000	0.0202	0.110
$\Delta Q\times$ Post	0.0267	0.100			0.0306^{***}	0.000	0.0266	0.140
$\Delta Q\times$ ETF	0.0716^{**}	0.000			0.0052	0.230	0.1297^{***}	0.000
$\Delta Q\times$ Post × ETF	-0.0492^{**}	0.020			-0.0351^{**}	0.010	-0.0971^{***}	0.000
VOL L $\times\Delta Q$			0.1828^{***}	0.000				
VOL L $\times\Delta Q\times$ Post			-0.0114	0.360				
VOL L $\times\Delta Q\times$ ETF			-0.1291^{***}	0.000				
VOL L $\times\Delta Q\times$ Post × ETF			0.0879^{**}	0.010				
VOL H $\times\Delta Q$			-0.0277	0.110				
VOL H $\times\Delta Q\times$ Post			0.0401^{**}	0.010				
VOL H $\times\Delta Q\times$ ETF			0.1349^{***}	0.000				
VOL H $\times\Delta Q\times$ Post × ETF			-0.1224^{***}	0.000				
R^2	0.1191		0.1406		0.2186		0.1245	
N of block boostrapping	100		100		100		100	

Table 14: **Control for firm fixed effect.** To control for unobservable bond characteristics related to the bond issuing firms, I include firm fixed effects in the analysis. The effective half spreads of ETF-bonds are compared with those of non-ETF bonds. Following the methodology of Bessembinder et al. (2006), the first stage is estimated as $Q_t = a + bQ_{t-1} + \epsilon_t$, where Q_t is $+1$ for the buy order, -1 for the sell order, and 0 for the inter-dealer orders. I write the unexpected order flow ϵ_t as Q_t^*. The second stage is estimated as $\Delta P_t = a + \gamma S Q_t^* + \alpha S \Delta Q_t + w X_t + \omega_t$. To assess the changes of effective spreads of ETF-bonds relative to the changes of effective spreads of non-ETF bonds, I interact ΔQ_t with *Post* and *ETF*. The non-informational component αS is divided into the changes of the effective spreads in the treatment group and the changes of the effective spreads in the control group: $\alpha_1 \Delta Q_t + \alpha_2 S \Delta Q_t \times Post + \alpha_3 S \Delta Q_t \times ETF + \alpha_4 S \Delta Q_t \times Post \times ETF$. Of my main focus is the coefficient $\alpha_4 S$. I include the three public information variables X_t as discussed in section 3.1. The statistical significance is measured the probability from block bootstrapping, following Bessembinder et al. (2006).

	coeff	p-value
First stage regression		
Intercept	0.1137***	0.000
Q_{t-1}	0.0678***	0.000
Second stage regression		
Q^*	0.0455***	0.000
Treasury return	−0.0138***	0.000
Stock return × invgrd	0.0175***	0.000
Stock return × noninvgrd	0.0740***	0.000
Default spread × invgrd	−0.0069***	0.000
Default spread × noninvgrd	0.0383***	0.000
ΔQ	0.4160***	0.000
$\Delta Q\times$ Post	0.0252***	0.000
$\Delta Q\times$ ETF	−0.0531***	0.000
$\Delta Q\times$ Post × ETF	−0.0122***	0.000
Firm Fixed Effect	Y	
R^2	0.1880	
N of block bootstrapping	50	