Steering a Ship in Illiquid Waters: Active Management of Passive Funds

Naz Koont

Yiming Ma

Ľuboš Pástor

Yao Zeng*

May 4, 2022

Abstract

Exchange-traded funds (ETFs) are typically viewed as passive index trackers. In contrast, we show that corporate bond ETFs actively manage their portfolios, trading off index tracking against liquidity transformation. In our model, ETFs optimally choose creation and redemption baskets that include cash and only a subset of index assets, especially if those assets are illiquid. Our evidence supports the model's predictions. We find that ETFs dynamically adjust their baskets to correct portfolio imbalances while facilitating ETF arbitrage. Basket inclusion improves bond liquidity, except in periods of large imbalance between ETF creations and redemptions, such as the COVID-19 crisis of 2020.

JEL classifications: G12, G23

Keywords: Exchange-traded funds, ETFs, liquidity transformation, active management

---

*Koont and Ma are at Columbia Business School. Pástor is at the University of Chicago Booth School of Business, NBER, CEPR, and the National Bank of Slovakia. Zeng is at the Wharton School of the University of Pennsylvania. The views in this paper are the responsibility of the authors, not the institutions they are affiliated with.

We are grateful for comments from Livia Amato, Matthew Bartolini, Itzhak Ben-David, Davidson Heath, Wei Jiang, Thomas Marta, Stefan Nagel, Christopher Reilly, Stijn Van Nieuwerburgh, Laura Veldkamp, and seminar participants at Chicago, Columbia, Indiana, Notre Dame, SAIF, and USC. We thank market participants at BlackRock and Vanguard for helpful conversations regarding ETF portfolio management. We thank Livia Amato for excellent research assistance. This research was funded in part by the Fama-Miller Center for Research in Finance at the University of Chicago Booth School of Business.

1 Introduction

Exchange-traded funds (ETFs) are among the most important financial intermediaries. Their assets under management have grown quickly since their first appearance in 1993, reaching $7.2 trillion by the end of 2021 in the U.S. alone. This amount exceeds total assets of U.S. fixed income mutual funds, and it is about half of total assets of U.S. equity mutual funds. Given their large and growing footprint, ETFs seem relatively understudied.

The vast majority of ETFs track passive indexes. To manage index deviations, ETFs rely on authorized participants (APs) to conduct arbitrage trades, in which APs create and redeem ETF shares in exchange for baskets of securities called the "creation basket" and the "redemption basket," respectively. These baskets are chosen by the ETF.

We study how ETFs use creation and redemption baskets to manage their portfolios. We find that, despite their passive image, ETFs are remarkably active in their portfolio management. They often use baskets that deviate substantially from the underlying index, and they adjust those baskets dynamically. By analyzing ETF baskets and their dynamics, we gain new insights into the economics of ETFs.

The main insight is that ETFs are active to facilitate liquidity transformation. ETF shares tend to be more liquid than the underlying securities, in part because APs' arbitrage trades tend to absorb shocks to investors' demand for ETF shares. When investors sell ETF shares, APs can buy and redeem them; when investors buy ETF shares, APs can create and sell them. By absorbing the trades of ETF investors, APs reduce the price impact of those trades. APs' arbitrage trading thus provides liquidity to investors who must trade ETFs at short notice. This liquidity provision requires APs to trade basket securities and thus incur transaction costs. To help reduce those costs, ETFs adjust their baskets to make them cheaper to trade. While these basket adjustments facilitate liquidity transformation, they also constrain the ETF's index-tracking capacity. We argue that ETFs are active because they care not only about index tracking but also about liquidity transformation, and because only active basket management allows them to balance both objectives.

Our empirical analysis focuses on U.S. corporate bond ETFs. We choose this category to emphasize the role of liquidity transformation, which is enhanced by the relative illiquidity of corporate bonds. Corporate bond ETFs have been growing fast, with assets tripling over the past six years before reaching over $342 billion at the end of 2021. We analyze the baskets, portfolios, and indexes of bond ETFs, along with the characteristics of the underlying bonds. Our analysis uses both proprietary and non-proprietary data from 2017 to 2020.

We begin by establishing two simple facts about bond ETF baskets. First, these baskets include a fair amount of cash. The average creation (redemption) basket contains 4.6% (7.8%) of its assets in cash, based on the baskets pre-announced by the ETF at the start of a trading day. The cash proportions are even larger, 11.6% (8.2%) for creation (redemption) baskets, based on realized baskets imputed from ETF holdings. Second, ETF baskets are concentrated—they include only a small subset of the bonds that appear in the underlying index. Both facts are costly to the ETF in terms of index tracking.

To rationalize these facts, we build a model that highlights ETFs' dual role of index tracking and liquidity transformation. The model features two types of investors, patient and impatient, and an AP. The impatient investors face a liquidity shock that forces them to either consume or save at short notice. Investors who need to consume sell ETF shares; those who need to save buy them. Any imbalance between the buys and sells is exploited by the AP via an arbitrage trade. When there are more sells than buys, the AP buys ETF shares in the market and redeems them in exchange for the redemption basket. When there are more buys than sells, the AP sells ETF shares after creating them in exchange for the creation basket. Either way, the AP improves the ETF's liquidity by mitigating the price impact of impatient investors. This liquidity improvement comes at a cost to the AP, who incurs transaction costs by trading basket securities. To reduce those costs, the ETF finds it optimal to use baskets that include some cash and only a subset of index securities, as we observe in the data. However, such baskets deviate from the index, which is generally costly to patient investors. Therefore, when designing its basket, the ETF balances the benefits of liquidity transformation against the costs of imperfect index-tracking.

Our model also predicts that ETFs have more basket cash and larger tracking errors when their securities are less liquid. Intuitively, when liquidity transformation is more costly, ETFs find it optimal to reduce this cost by tolerating looser index-tracking. We test these predictions empirically. We find that ETFs investing in less liquid corporate bonds indeed have more cash in their baskets and larger tracking errors, consistent with the model.

We make two sets of predictions about how an ETF should manage its portfolio over time. First, the ETF should dynamically adjust its creation and redemption baskets to steer its portfolio toward the index. Second, these adjustments should be smaller when liquidity transformation is more costly. For example, suppose the ETF's portfolio currently overweights a given security relative to the index. The ETF should add some of this security to the redemption basket and remove it from the creation basket. Similarly, if a security is underweighted, it should be added to the creation basket and removed from the redemption basket. Either way, these basket modifications should be attenuated when the ETF invests in less liquid securities. The same ideas apply to the cash balance. If the ETF currently holds more cash than usual, it should add cash to the redemption basket and remove it from the creation basket, but less so if the underlying index is less liquid.

We find empirical support for these predictions in our sample of corporate bond ETFs. One estimation challenge is that the over/underweighting of securities in ETF portfolios can be driven by unobserved fundamentals. To address this challenge, we make use of monthly index rebalancing. Bond indexes rebalance at month-ends to reflect changes in bond characteristics during the month. Bond index weights jump at the rebalancing dates, providing plausibly exogenous variation in over/underweighting that allows us to trace out the ETFs' response. We find that when index rebalancing makes a bond more overweighted (underweighted) in an ETF's portfolio, the ETF includes more of that bond in its redemption (creation) basket. ETFs thus revise their baskets to steer their portfolios toward the index, as predicted. We also find these basket adjustments are less pronounced for ETFs tracking less liquid indexes, again consistent with our predictions.

Having analyzed the nature and causes of ETFs' active portfolio management, we study its consequences for bond market liquidity. We find that ETFs' active efforts to improve the liquidity of their shares have significant effects on the liquidity of the underlying bonds. Specifically, we find that a bond's inclusion in an ETF basket makes the bond more liquid in normal times. Basket inclusion makes the bond less liquid, however, in periods of large imbalance between creations and redemptions.

We reach these conclusions after estimating the effect of basket inclusion on liquidity in two ways. First, we relate bond liquidity to three measures of basket inclusion in the presence of numerous controls and fixed effects. Second, we exploit the monthly bond index rebalancing once again to construct a novel instrument for basket inclusion. For each bond, the instrument measures the jump in portfolio overweighting at the rebalancing date added up across all ETFs experiencing creations or redemptions. The instrumented inclusion of that bond in an ETF basket should then reflect a plausibly exogenous shock. Repeating our estimation using two-stage least squares, we find that a bond's inclusion in an ETF basket, creation or redemption, improves the bond's liquidity in the full sample.

This positive effect of basket inclusion on bond liquidity holds in normal times, when investors' liquidity shocks are largely idiosyncratic. A random mix of creations and redemptions across ETFs increases the trading activity in the bonds included in ETF baskets, improving their liquidity. That is not the case, however, in periods when liquidity shocks are systematic, resulting in imbalances between creations and redemptions. For example, bond

ETFs experienced systematic redemptions in March 2020, after the onset of the COVID-19 pandemic (Falato et al., 2020). These redemptions moved many redemption-basket bonds to APs' balance sheets. The APs, who also act as market makers in these bonds, thus became reluctant to purchase more of the same bonds, reducing their liquidity. Indeed, when we rerun our baseline regression for the subperiod of the spring of 2020, we find that being included in a redemption basket worsens rather than improves a bond's liquidity. Similarly, when we interact basket inclusion with basket imbalance, we find that the inclusion's effect on bond liquidity is negative when the imbalance is sufficiently large. We conclude that the effect of ETFs' active portfolio management on bond liquidity is state-dependent.

Related Literature. We spotlight ETFs' liquidity transformation. The literature has traditionally examined equity ETFs (see Ben-David et al., 2017, Lettau and Madhavan, 2018, for surveys), for which liquidity transformation is modest, whereas we analyze bond ETFs. We show that bond ETFs' role in liquidity transformation turns passive ETFs into active managers of their portfolios. ETFs balance this role against their better-known index-tracking role, with strong implications for the underlying securities.

We merge data on ETF creation and redemption baskets, ETF holdings, index holdings, and the underlying bonds, assembling the most comprehensive ETF database to date, to our knowledge. ETF basket data are rarely analyzed. Our work complements two ongoing studies that do examine bond ETF baskets. Shim and Todorov (2021) show that these baskets tend to cover only a small fraction of ETF holdings. They explore the implications of concentrated baskets for ETF premiums and discounts, a focus different from ours. Reilly (2021) finds that APs engaging in ETF arbitrage tend to deliver underperforming bonds in creation baskets. His focus on the information asymmetry between ETFs and APs is also different from ours. We follow these studies in imputing ETF baskets from ETF holdings. We depart from them by using also data on reported baskets, by focusing on ETF activeness, and by presenting a different set of theoretical and empirical results.

We are not the first to note that ETFs are somewhat active. Some ETFs track non-traditional indexes, while others exhibit large index deviations (Akey et al., 2021, Ben-David et al., 2021, Brogaard et al., 2021, Easley et al., 2021). Some ETFs engage in securities lending or cross-subsidization of affiliated financial institutions (Cheng et al., 2019). We contribute to this literature in two ways. First, while the above studies analyze equity ETFs, we focus on bond ETFs and the related liquidity transformation. Second, and more important, we emphasize a different notion of ETFs' activeness—their active basket management, including the dynamic management of their cash balances and individual bonds. By examining ETFs' active portfolio management and its underlying tradeoffs, we also relate to the literature on fund activeness (e.g., Kacperczyk et al., 2008, Cremers and Petajisto, 2009) and the tradeoffs among active funds' characteristics (Pastor et al., 2020).

Our work is also related to the literature on the asset pricing implications of ETF ownership. For equity ETFs, this literature shows that ETF ownership has positive effects on stock volatility (Ben-David et al., 2018), return co-movement (Da and Shive, 2018), as well as liquidity comovement (Agarwal et al., 2018), but an unclear effect on stock liquidity (Israeli et al., 2017, Saglam et al., 2019). For bond ETFs, Dannhauser and Hoseinzade (2022) find that the selling pressure on bond ETFs induced by the 2013 taper tantrum was transmitted to the bonds owned by those ETFs. The evidence on the effect of ETF ownership on the liquidity of the constituent bonds is mixed. While Dannhauser (2017) finds that ETF ownership has a weak or even negative effect on bond liquidity, Holden and Nam (2019) and Marta (2019) find positive effects. Instead of ETF ownership, we analyze ETF creation and redemption baskets, which offer a more granular perspective on ETF activity. We show that the basket concentration generated by the active management of bond ETFs has significant state-dependent effects on the liquidity of the underlying bonds.

Finally, we contribute to the broad literature on liquidity transformation by financial intermediaries. Starting with Diamond and Dybvig (1983), this literature has traditionally focused on banks. More recently, this literature has come to emphasize the liquidity provision by shadow banks (e.g., Kacperczyk and Schnabl, 2013, Sunderam, 2015), including mutual funds (e.g., Chen et al., 2010, Goldstein et al., 2017, Chernenko and Sunderam, 2017, Ma et al., 2019, Choi et al., 2020, Jin et al., 2022). We analyze the liquidity transformation by ETFs, the fastest-growing financial intermediary engaging in such transformation.

The paper is organized as follows. Section 2 describes our data. Section 3 presents the stylized facts that motivate our model, which we present in Section 4. Section 5 tests the model's predictions, offering evidence of active basket management by ETFs. The effect of this management on bond liquidity is analyzed in Section 6. Section 7 concludes.

2 Data

To construct our sample, we identify passive U.S. corporate bond ETFs in the ETF Global database. We exclude ETFs that use active strategies, total bond market ETFs, and ETFs that primarily invest in Treasuries, mortgage-backed securities, international bonds, munis, and loans. Our sample includes 118 ETFs in January 1, 2017 to December 31, 2020.

For each ETF, we obtain daily data on fund portfolio holdings, shares, and prices from

ETF Global. We also obtain index holding data from Bloomberg for about half of the ETFs in our sample. In the Appendix, we report the summary statistics for all ETFs in our sample, as well as for the subset of ETFs whose index holding data are available. The two sets of ETFs look similar based on most observable characteristics. For example, the average ETF in our sample holds 753 bonds in its portfolio. The average portfolio size is slightly larger, 835, when computed only across the ETFs with index holding data. For the latter ETFs, the average number of bonds in the index is 1,153.

Before describing our data on ETF baskets, it seems useful to provide some institutional background. ETFs are investment funds that issue shares backed by a portfolio of securities. ETF shares trade in the secondary market. In the primary market, investors known as APs can create and redeem ETF shares in-kind for the underlying securities.[1] The APs can profit from arbitrage trading across the primary and secondary markets. When ETF shares are relatively cheap in the secondary market, APs can buy them, redeem them in-kind for a basket of securities called the redemption basket, and sell these securities at a profit. When ETF shares are expensive, APs can deliver securities in the creation basket to the ETF issuer and sell the newly created ETF shares. Figure 1 illustrates the process.

ETFs disclose their desired creation and redemption baskets ahead of each trading day. During the day, an AP can request basket modifications, such as the omission of hard-to-locate securities or the inclusion of securities that are of interest to the AP. The final basket composition is a result of negotiations between the AP and the ETF. The use of custom baskets that are not fully representative of ETF portfolio holdings is permitted under the SEC's rule 6c-11 as long as ETF issuers adopt and enforce policies governing the construction of custom baskets and act in the best interests of the ETF and its shareholders. Custom baskets were common also before the introduction of rule 6c-11 in 2019, due to a widespread use of exemptive orders that allowed non-pro-rata baskets. We do not see any significant change in basket representativeness following the adoption of rule 6c-11, which suggests that prior restrictions on basket composition were not binding.

2.1 Realized ETF Baskets

We impute ETFs' realized creation (CR) and redemption (RD) baskets from changes in ETF holdings on days with CR or RD activity (Shim and Todorov, 2021). We identify CR (RD) days as those on which there is a positive (negative) change in the number of ETF shares. We then use daily changes in the number of bonds held to determine the composition of the

[1] For bond ETFs, APs are mostly large dealer banks such as JP Morgan Chase, Goldman Sachs, and Morgan Stanley, which also operate bond trading desks and act as market makers in the bond market.

average ETF basket on that day. This imputation assumes that ETFs' portfolio changes reflect only APs' CR and RD activities. This assumption could in principle be violated in several ways. First, ETFs could trade in the secondary market instead of relying solely on CR and RD activities to manage their portfolios. However, such trades are uncommon as they could create tax liabilities for investors, unlike the in-kind exchanges that occur during CR and RD events. Second, different APs may negotiate different baskets with an ETF on the same day. Third, there may be RD baskets on days with net creations and CR baskets on days with net redemptions. For these reasons, each imputed/realized basket is best interpreted as the average net basket for the given ETF on the given day.

On most CR/RD days, realized baskets coincide with changes in the number of ETF shares. However, the daily portfolio change recorded in the ETF Global database occasionally leads or lags the corresponding change in ETF shares by one day. Such instances are easy to spot because of the large number of zero-share-change days in the sample. To correct these apparent data errors, we shift the dates of such realized baskets by one day to align them with the dates of ETF share changes. We do not perform such shifts, and do not impute any baskets, on days when the ETF portfolio changes are very small, affecting only one or two bonds. Such small portfolio changes on days with no changes in the number of ETF shares can result from occasional secondary-market transactions by the ETF, or from corporate events by bond issuers leading to changes in bond identifiers. The average ETF in our sample has 104 (147) bonds in the realized CR (RD) basket.

2.2 Reported ETF Baskets

We obtain proprietary data on ETFs' reported baskets from the Depository Trust and Clearing Corporation (DTCC). Pursuant to the Securities Exchange Act of 1934, Rule 19(b), ETF issuers announce the CR and RD baskets to all APs at the close of each trading day for use on the following trading day. These baskets represent specific portfolios—names and quantities—of securities that ETF issuers intend to exchange for ETF shares. ETF issuers distribute this information via the DTCC's National Securities Clearing Corporation, a clearinghouse service that automates the CR/RD process.

The main advantage of this dataset is that baskets are reported on the eve of each trading day, including days with no CR or RD events. On any given day, there are many ETFs with no CR or RD activity, and thus no realized baskets. The reported basket data are therefore particularly useful for cross-sectional analysis comparing ETF baskets on a given day.

The main disadvantage is that reported baskets may differ from the actual baskets used in the CR/RD process. APs can negotiate with the ETF issuer about basket composition, arriving at a realized basket that may differ from the pre-announced reported basket. We thus rely mostly on realized baskets for our time-series analysis, while using reported baskets mostly to establish the stylized facts in the cross section.[2]

The average ETF in our sample has 424 bonds in the reported basket. This number exceeds the average numbers of bonds in realized CR and RD baskets (104 and 147, respectively), indicating that reported baskets tend to be larger than realized baskets. In contrast, reported baskets tend to be smaller than ETF portfolios as well as index portfolios, which shows that even intended ETF baskets are not designed to fully replicate the underlying index. The summary statistics are tabulated in the Appendix.

2.3 Individual Bonds

We obtain bond price and trading data from TRACE and bond-level characteristics from Mergent FISD. We use the TRACE data to calculate three daily measures of market illiquidity for each bond: the effective tick size (*Tick*; see Holden (2009), Goyenko et al. (2009)), the imputed roundtrip cost (*IRC*; see Feldhutter (2012)), and the interquartile range (*IQR*; see (Song and Zhou, 2007, Pu, 2009)). We describe these measures and their construction in detail in the Appendix. The Appendix also provides bond-level summary statistics for the three measures as well as other bond characteristics. When we tabulate our results later in the paper, we use IL1, IL2, and IL3 to denote *Tick*, *IRC*, and *IQR*, respectively. Unless we note otherwise, we winsorize all fund- and bond-level variables at the 1% level.

3 Stylized Facts

In this section, we present stylized facts related to cash holdings and basket concentration.

3.1 Cash in ETF Baskets and Portfolios

The first fact is that corporate bond ETF baskets contain substantial amounts of cash. To identify cash and cash equivalents in ETF baskets, we manually classify the securities in realized baskets by their asset class and security description in ETF Global. We include any cash-like money market instruments, such as securities labeled as "Cash," "Currency," "US

[2]Vanguard ETFs do not disclose daily portfolio holdings, precluding the imputation of realized baskets. We proxy for Vanguard ETFs' realized baskets by reported baskets on days with ETF share changes.

Dollars," as well as money market funds and short-term Treasury ETFs.[3]

We compute the proportions of cash in both realized and reported baskets. For each realized basket, we divide the imputed amount of cash by total basket value on the same day. For each reported basket, we divide the amount of cash reported to DTCC by total basket value. We compute time-series averages of these "cash ratios" fund by fund and show their empirical distributions in Table 1.

Table 1 shows that cash accounts for a significant proportion of realized ETF baskets: 11.6% of CR baskets and 8.2% of RD baskets, on average. The average proportion of cash in reported baskets is a bit smaller, ranging from 4.6% to 7.8% depending on whether we average across all days, days with RD activity, or days with CR activity.

These averages mask large dispersion in the cash ratios across ETF baskets. The cross-sectional distributions of the basket cash ratios have long right tails. Their medians are dwarfed by the means and the 90th percentiles of imputed cash ratios exceed 34%. These results show that cash plays a significant role in ETF baskets.[4]

Next, we turn our attention from ETF baskets to ETF holdings. Table 1 shows that the average ETF holds 1.7% of its portfolio in cash. ETFs typically keep their cash in money market sweep vehicles. Most of the cash holdings seem discretionary. Only a small part comes from receipts of coupon payments that have not yet been reinvested. The correlation between cash from coupon payments and cash holdings at the ETF-day level is only 1%, presumably because coupon payments can be fully anticipated, and thus easily managed, by the ETF. Proceeds from matured bonds also matter little because most bonds leave the ETF portfolio before they mature. When we add proceeds from matured bonds to coupon payments and recompute the above correlation, it still rounds to 1%. Even though ETFs' cash holdings are nontrivial, it is clear from Table 1 that ETFs hold substantially more cash in their baskets than in their portfolios. The disproportionate use of cash in ETF baskets is consistent with ETFs' desire to incentivize the CR and RD activity by APs.

ETFs do not need to hold cash to meet withdrawals, unlike banks and mutual funds, as ETFs do not engage in cash transactions with investors. Furthermore, the indexes tracked by bond ETFs include little or no cash. Any index cash results from intra-month coupon

[3]The most popular Treasury ETFs are "BlackRock Cash Funds: Treasury, SL Agency" and "Invesco Premier U.S. Government Money Portfolio". We do not include Treasury bonds among cash equivalents.

[4]About a quarter of realized RD baskets, and a smaller fraction of realized CR baskets, have negative amounts of cash. Negative cash can result from the use of credit lines, failed trades, accounting errors, and basket price adjustments, as revealed to us in private communication with ETF managers. Some of the negative amounts can also result from the imputation errors described earlier.

and principal payments that are not reinvested between monthly rebalance dates. Index cash builds up during the month as the payments are received, before dropping to zero at the next month-end rebalancing date. The amounts are small. For example, for the ICE BoFA US Corporate Bond Index, the average daily cash value in 2021 is 11 bps, and the daily cash levels range from zero to 29 bps. The numbers are only slightly larger for indexes of high-yield bonds, for which coupon payments account for a larger fraction of the return. For example, cash in the ICE BoFA US High Yield Master II Index has a mean of 14 bps and a maximum of 38 bps in 2021.[5] These index cash amounts are at least an order of magnitude smaller than the cash amounts in ETF baskets reported in Table 1. As a result, any nontrivial cash in ETF portfolios is costly in terms of index-tracking.

The use of cash by ETFs is reminiscent of its use by mutual funds for the purpose of liquidity transformation (Chernenko and Sunderam, 2017). Although mutual funds and ETFs use cash differently—mutual funds hold it as a liquid buffer to meet investor redemptions, whereas ETFs use it in baskets to incentivize AP arbitrage—both transform portfolios of relatively illiquid assets into more liquid claims. One of the main messages of our paper is that ETFs include cash in their baskets to facilitate liquidity transformation.

3.2 Basket Concentration

Our second stylized fact is that ETF baskets are highly concentrated compared to index portfolios. This fact may seem surprising, just like cash in baskets, given the stated objective of passive ETFs to track the underlying index. A natural way for an ETF to ensure index-tracking would be to use CR and RD baskets representative of the index. Instead, ETFs use baskets that contain fewer securities, resulting in imperfect index-tracking.

To assess a given basket's concentration on a given day, we divide the number of bonds in the basket by the number of bonds in the underlying index. We compute these "basket ratios" for both CR and RD baskets, reported as well as realized. We compute time-series averages of these ratios and report their cross-sectional distributions in Table 2.

Table 2 shows that realized baskets are highly concentrated relative to their indexes. The average realized basket ratios for CR and RD baskets are 24.6% and 29.8%, respectively. The medians are even lower, at 19.4% and 22.0%. These findings are consistent with those of Shim and Todorov (2021), who also examine realized baskets. Reported baskets, which are not examined by Shim and Todorov (2021), are less concentrated. The average reported basket ratio is about 76%—much larger than the averages for realized baskets, but much

[5]We are grateful to Matthew Bartolini of State Street for his helpful guidance in this matter.

smaller than one. We thus see two dimensions of ETFs' active basket choice. First, ETFs pre-announce baskets that differ substantially from the index. Second, ETFs allow the APs to use realized baskets that differ substantially from the pre-announced baskets.

Table 2 also reports the cross-sectional distribution of analogous ratios that capture ETF portfolio concentration rather than basket concentration. These are ratios of the number of bonds in the ETF portfolio to the number of bonds in the index. The mean and median of these ratios are 81.3% and 88.0%, respectively. These values are larger than their counterparts for any of the previously discussed baskets, suggesting that ETFs actively update their baskets to keep their portfolios fairly close to the index.[6]

When ETFs choose individual bonds for inclusion in their baskets, they do not sample index bonds evenly. Rather, some bonds are persistently included, whereas others rarely appear. To show this fact, we calculate the basket inclusion probability for each bond-ETF pair as the number of times the bond appears in the ETF's CR or RD basket divided by the total number of baskets in which this bond could have appeared. We find that basket inclusion probabilities differ substantially across bonds. The median bond has only an 8% (13%) likelihood of appearing in its ETF's CR (RD) basket, whereas for the 99th percentile bond, this likelihood is 72% (88%). The Appendix offers more details.

Just like cash in baskets, concentrated baskets help ETFs incentivize AP arbitrage. Through this channel, the use of concentrated baskets helps ETFs achieve their objective of liquidity transformation. We formalize these ideas in the following section.

4 Model

Motivated by the stylized facts uncovered in Section 3, we develop a simple model to shed light on the underlying economic mechanism. In the model, an ETF's optimal basket choice reflects a tradeoff between the ETF's dual objectives of index tracking and liquidity transformation. Besides predicting the two main facts from Section 3, the model makes additional predictions regarding ETFs' cash and basket management. We test those predictions empirically in Section 5. All the proofs are in the Appendix.

[6]For some ETFs, basket ratios exceed one. About 90% of the securities that are held by ETFs but absent from the index are fixed-income securities. The remaining securities are mostly cash equivalents.

4.1 Model Setup

The economy has three dates, $t = 0, 1, 2$, with no time discount. A unit measure of ex-ante identical infinitesimal agents is born at $t = 0$, each endowed with one unit of a consumption good called cash. Cash is riskless, liquid, and it serves as the numeraire. The agents jointly form a representative ETF at $t = 0$, each holding an equal share of it. ETF shares trade in a competitive market at $t = 1$ at the market price p_E, which is determined in equilibrium. The ETF matures at $t = 2$ at the value v_E, also determined in equilibrium. Beyond the ETF, agents cannot access the underlying security market or any other investment technology to transfer wealth across time. This assumption captures the difficulty investors face in accessing illiquid security markets on their own. All agents have the same mean-variance utility function, $u(c) = \mathrm{E}(c) - (\rho/2)\mathrm{Var}(c)$, over their lifetime consumption c.[7]

Each agent privately learns their preferences at $t = 1$, becoming one of three types:

1. Impatient consumer: With probability π_c, the agent is subject to an idiosyncratic consumption shock at $t = 1$. The agent sells their ETF share at $t = 1$ and consumes the proceeds, obtaining utility $u(p_E)$ at that time.

2. Impatient saver: With probability π_s, the agent is subject to an idiosyncratic saving shock at $t = 1$. The agent buys an additional ETF share at $t = 1$, paying the price p_E per share. The agent holds two ETF shares until $t = 2$, at which time the agent consumes the ETFs' matured value $2v_E$, enjoying total utility $u(2v_E - p_E)$.

3. Patient investor: With probability $1 - \pi_c - \pi_s$, the agent faces no shock at $t = 1$. The agent consumes the matured value of one ETF share at $t = 2$, receiving utility $u(v_E)$.

The modeling of both consumption and saving shocks allows us to consider ETF creations and redemptions in a unified way, highlighting liquidity transformation in both cases. Consumption shocks correspond to liquidity shocks in Diamond and Dybvig (1983). Saving shocks are similar because they, too, reflect agents' idiosyncratic and immediate demand for liquidity (e.g., due to receiving a windfall that is not immediately consumed).

In the knife-edge case of $\pi_c = \pi_s$, the masses of impatient consumers and savers are equal, so the two groups trade with each other: impatient consumers sell their ETF shares to impatient savers. When $\pi_c \neq \pi_s$, though, the imbalance between the demand and supply for ETF shares in the secondary market must be met by their creation or redemption in the

[7]The mean-variance assumption simplifies the exposition and facilitates analytical comparative statics. In the Appendix, we show that the model's key predictions remain unchanged when we replace mean-variance with constant-absolute-risk-aversion (CARA) utility and solve the model numerically.

primary market. For this purpose, we introduce a deep-pocketed, risk-neutral, representative arbitrageur, or "AP." When $\pi_c < \pi_s$, the AP meets the excess demand for ETF shares by creating new shares at $t = 1$ and immediately selling them to impatient savers at the price p_E. When $\pi_c > \pi_s$, the AP meets the excess supply of ETF shares by buying them from impatient consumers for p_E and immediately redeeming them at $t = 1$. The AP creates and redeems ETF shares in kind, by exchanging those shares for a basket of securities chosen by the ETF at $t = 0$. Our modeling of the AP's activity closely mirrors the real-world CR/RD process, which is unique to ETFs and only available to authorized participants.

By meeting excess demand and supply in the secondary market for ETF shares, the AP makes these shares more liquid, supporting the liquidity transformation function of ETFs. The AP's presence in the market effectively reduces the price impact of the liquidity-demanding trades of impatient agents. The extent of the AP's liquidity provision can be inferred from the ETF's equilibrium market price p_E, as we explain later.

At $t = 0$, the newly formed ETF is endowed with an equal-weighted portfolio of N risky securities, which are identical ex ante. We normalize the N securities' values at $t = 0$ and $t = 1$ to unity. Each security's value at $t = 2$ is distributed as i.i.d. normal, $N(\mu, \sigma^2)$, with $\mu > 1$. We label the equal-weighted portfolio of the N securities as the "index."[8]

The N securities are not only risky but also illiquid, in that sourcing and trading them is costly to the AP at $t = 1$. There are two types of costs, fixed and variable. First, sourcing a unique security entails a fixed cost $\lambda > 0$, so that a basket with I unique securities entails a total cost of λI. This fixed cost proxies for the search-and-matching costs typical of over-the-counter security markets. Second, regardless of the basket count, the AP incurs a variable cost of $\frac{1}{2}\phi s^2$, where $\phi > 0$, when holding s security units as a result of ETF creations or redemptions. The value of s can be positive or negative, capturing long or short security positions after redemptions or creations, respectively. This variable cost captures the AP's transaction costs associated with liquidating the inventory associated with creations or redemptions, as well as any balance-sheet costs of carrying that inventory.

The ETF manages its portfolio by designing the basket of securities that the AP can exchange for ETF shares. The same basket is used for both creations and redemptions. The ETF cannot trade securities directly; it can change its portfolio composition only through basket design. Consistent with this assumption, real-world ETFs rarely trade on their own because in-kind exchange of ETF shares for security baskets is more tax-efficient.

[8]In the Appendix, we prove all of our theoretical results in a more general setting, in which the ETF's endowment portfolio can include not only the index portfolio but also some amount of cash.

The ETF's basket design decision at $t = 0$ involves two choices. First, the ETF chooses α, the basket cash weight, where $0 \leq \alpha \leq 1$. The ETF basket holds α in cash and $1 - \alpha$ in risky securities. Second, the ETF chooses I, the number of risky securities in the basket, where $I \leq N$. Since all securities are identical ex ante, the choice among them boils down to the choice of I. We refer to I as the basket count. The ETF chooses α and I to maximize agents' expected aggregate welfare:

$$\max_{\alpha,I} \mathrm{E}[W] \;\equiv\; \underbrace{\pi_c\, u(p_E)}_{\substack{\text{impatient}\\ \text{consumers}}} + \underbrace{\pi_s\, E[u(2v_E - p_E)]}_{\substack{\text{impatient}\\ \text{savers}}} + \underbrace{(1-\pi_c-\pi_s)\,\mathrm{E}[u(v_E)]}_{\substack{\text{patient}\\ \text{investors}}}\,. \tag{1}$$

By maximizing the welfare of its shareholders, the ETF maximizes its appeal to clients, which should enhance its fee revenue. We do not model management fees explicitly. Abstracting from agency frictions allows us to sharpen our focus on ETF's liquidity transformation.

4.2 Equilibrium ETF Pricing and Liquidity Transformation

We first solve for the equilibrium ETF price at $t = 1$, p_E, taking the ETF's basket cash weight α and basket count I as given. We solve for α and I later, in Section 4.3.

The ETF's market-clearing price ensures that the AP is willing to clear the imbalance between the demand and supply of ETF shares by the impatient agents. If there is excess demand for ETF shares (i.e., $\pi_s - \pi_c > 0$), the AP must create $\pi_s - \pi_c$ new ETF shares and sell them to impatient savers. If there is excess supply (i.e., $\pi_c - \pi_s > 0$), the AP must redeem $\pi_c - \pi_s$ shares after buying them from impatient consumers. Let x denote the number of ETF shares bought and redeemed by the AP at $t = 1$ (where $x > 0$ for redemptions and $x < 0$ for creations). The AP chooses x to maximize its profit function, $\Gamma(x)$:

$$\max_{x} \Gamma(x) \;\equiv\; [\alpha + (1-\alpha)\mu - p_E]\, x \;-\; \frac{1}{2}\phi\, [(1-\alpha)x]^2 \;-\; \lambda I\,, \tag{2}$$

subject to the participation constraint $\Gamma(x) \geq 0$. The first term in equation (2) captures the AP's expected arbitrage profit. Suppose $x > 0$. After buying x ETF shares from impatient consumers at the price of p_E per share, the AP exchanges these shares for x units of the ETF basket. This basket contains α units of cash and $\frac{1-\alpha}{I}$ units of each of I risky securities, so its expected payoff is $\alpha+(1-\alpha)\mu$. The AP's profit per ETF share is therefore $\alpha+(1-\alpha)\mu-p_E$, yielding the first term. If $x < 0$, the same logic applies, except that the AP creates ETF shares and sells them for p_E. The second term in equation (2) reflects the AP's variable costs associated with carrying and transacting $1 - \alpha$ units of risky basket securities for each of the x ETF shares created or redeemed. Finally, the last term represents the fixed costs associated with sourcing I distinct basket securities in the market.

From the first-order condition corresponding to equation (2), we obtain

$$x^* = \frac{\alpha + (1-\alpha)\mu - p_E}{\phi(1-\alpha)^2} \,. \tag{3}$$

When ETF shares are cheap relative to the basket, so that $p_E < \alpha + (1-\alpha)\mu$, the AP buys x^* ETF shares, redeems them, and sells the basket securities. When the ETF is expensive relative to the basket, so that $p_E > \alpha + (1-\alpha)\mu$, the AP purchases the basket, creates x^* ETF shares, and sells them. Either way, the AP makes money, before transaction costs. The AP will execute this trade only if the profit net of transaction costs is also positive.

The ETF's equilibrium price equates supply and demand for ETF shares. We obtain this price by combining equation (3) with the market-clearing condition $x = \pi_c - \pi_s$.

Proposition 1. *The ETF's market-clearing price at time 1 is given by*

$$p_E \;=\; \underbrace{\alpha + (1-\alpha)\mu}_{\text{basket NAV}} + \underbrace{\phi(1-\alpha)^2(\pi_s - \pi_c)}_{\text{premium/discount}} \,. \tag{4}$$

The ETF market clears at time 1 if and only if $I \leq \bar{I}$, where $\bar{I}$ is the basket count that makes the AP break even (i.e., for which the supremum of $\Gamma(x)$ in equation (2) is zero).

The first term in equation (4) is the basket's expected payoff, which is also its net asset value from the perspective of the risk-neutral AP. The second term is the ETF premium or discount, as perceived by the AP. From the AP's viewpoint, the ETF trades at a premium if and only if $\pi_s > \pi_c$, which indicates excess demand for ETF shares in the secondary market and leads to net ETF creation in the primary market. Similarly, the ETF trades at a discount if $\pi_s < \pi_c$, which indicates excess supply of ETF shares and net redemption.

We interpret the ETF premium/discount as the equilibrium price impact of impatient traders. A discount means that excess supply of ETF shares by impatient consumers has pushed p_E down, reducing those consumers' utility, $u(p_E)$. A premium means that excess demand by impatient savers has pushed p_E up, increasing their disutility. The magnitude of the premium/discount thus captures the degree of liquidity transformation performed by the AP. The lower the magnitude, the lower is the price impact of impatient traders in the secondary ETF market, and the higher is the ETF liquidity transformation.

From equation (4), the magnitude of the ETF premium/discount increases with ϕ but decreases with α. Both effects are intuitive under the price impact interpretation. A larger cash weight α indicates a more liquid basket, which reduces the AP's variable transaction costs. Given the lower cost of providing liquidity, the AP chooses to provide more of it, resulting in a smaller price impact in the ETF market. In contrast, a larger ϕ means that the risky securities are more costly for the AP to trade. The AP then provides less liquidity and the price impact is larger. This model prediction is consistent with the empirical finding that ETF premiums and discounts tend to be larger for ETFs holding less liquid securities (e.g., Petajisto, 2017, Pan and Zeng, 2019, Shim and Todorov, 2021).

The price impact is infinite, and no liquidity transformation takes place, if the AP's participation constraint is not satisfied. In that case, any excess demand or supply from impatient agents remains unmet and the ETF market fails to clear. As noted in Proposition 1, the basket count must be sufficiently low to incentivize the AP to trade. We thus see that not only a larger cash weight but also a lower basket count facilitate liquidity transformation. However, they come at a cost, as we explain in the following section.

4.3 Tradeoff between Liquidity Transformation and Index-Tracking

Recall that the ETF's decision at $t = 0$ involves choosing the basket cash weight, α, and basket count, I. The ETF faces a tradeoff: increasing α or decreasing I improves liquidity transformation, as explained earlier, but it comes at the expense of index-tracking. To understand this, recall that the ETF is endowed with the index portfolio containing all N securities. If the basket includes only a subset of those securities (i.e., $I < N$), then the ETF's portfolio after creations or redemptions will deviate from the index. Similarly, if the basket includes cash, so will the ETF's post-CR/RD portfolio, again deviating from the (cash-free) index. While index-tracking per se does not enter agents' preferences, it is related to the welfare of patient investors, as we explain below.

We solve the ETF's optimal basket design problem at $t = 0$. Substituting the results from Section 4.2 into equation (1), we re-express the objective function as

$$\max_{\alpha, I} \mathrm{E}[W] \ = \ (\pi_c - \pi_s)p_E \ + \ V(\mu_E, \sigma_E^2)\,, \tag{5}$$

where p_E is in equation (4), μ_E and σ_E^2 are the mean and variance of the ETF portfolio's time-2 payoff, and $V(\mu_E, \sigma_E^2)$ is a function increasing in μ_E and decreasing in σ_E^2. Explicit formulas for μ_E, σ_E^2, and $V(\mu_E, \sigma_E^2)$ are in the Appendix.

Equation (5) sheds more light on the tradeoff between liquidity transformation and index-tracking. The first term captures liquidity transformation; the second index-tracking. The first term, $(\pi_c - \pi_s)p_E$, echoes our prior discussion of price impact. When there is excess demand for ETF shares (i.e., $\pi_c - \pi_s < 0$), the ETF wants to push p_E down to reduce the price impact of the impatient savers' buying pressure. When there is excess supply (i.e.,

$\pi_c - \pi_s > 0$), the ETF wants to push p_E up to counter the price impact of the selling pressure. The value of p_E is determined by the amount of liquidity that the AP is willing to provide. The ETF can incentivize the AP to provide more liquidity by using more cash or fewer securities in the basket (i.e., higher α or lower I).

However, higher α or lower I imply a larger index deviation at $t = 2$, as explained earlier. This larger index deviation affects the second term in equation (5). In particular, σ_E^2 increases when I decreases, holding α constant. With fewer securities in the basket, the ETF portfolio will have a larger variance than the equal-weighted index at $t = 2$, due to reduced diversification. While I has no effect on μ_E, α affects both μ_E and σ_E^2. A larger α implies more (less) cash in the ETF portfolio at $t = 2$ after creations (redemptions). More cash in the portfolio implies a lower expected payoff, μ_E, because risky securities have higher expected returns than cash ($\mu > 1$), but also lower variance, σ_E^2, because cash is risk-free. The net effect of α on the welfare of mean-variance investors is thus ambiguous. However, the poorer index-tracking caused by a lower I unambiguously reduces the welfare of patient investors because a decrease in I increases σ_E^2, which in turn reduces $V(\mu_E, \sigma_E^2)$.

Solving the ETF's basket design problem in equation (5), we obtain the following results.

Proposition **2.** *If the illiquidity parameters ϕ and λ are sufficiently large, then*

(i) The ETF optimally chooses a positive basket cash weight, $\alpha^* > 0$.

(ii) The ETF optimally chooses a concentrated basket, $I^* < N$.

(iii) The optimal basket cash weight, α^*, increases with ϕ.

(iv) The index-tracking error of the ETF's equilibrium security portfolio increases with λ.

Parts (i) and (ii) of Proposition 2 rationalize the two stylized facts presented in Section 3. Parts (iii) and (iv) show that the tension between liquidity transformation and index-tracking is exacerbated by the illiquidity of the underlying securities. This illiquidity makes it costlier for the AP to provide liquidity in the ETF market. The result is lower ETF liquidity, which increases the price impact of impatient investors, reducing their welfare. To restore the balance between the welfare of patient and impatient investors, the ETF incentivizes more liquidity provision by the AP. Specifically, larger variable transaction costs (ϕ) lead the ETF to raise its basket cash weight, whereas larger fixed transaction costs (λ) induce it to use a more concentrated basket, resulting in a larger tracking error.

Parts (iii) and (iv) of Proposition 2 generate additional testable predictions that go beyond the stylized facts from Section 3. Part (iii) directly motivates the following hypothesis.

Hypothesis **1.** *The proportion of cash in an ETF's baskets increases when the underlying* *securities are less liquid.*

Part (iv) of Proposition 2 directly motivates the following hypothesis.

HYPOTHESIS **2.** *The tracking error of an ETF's security portfolio is larger when the underlying securities are less liquid.*

A key feature of our model is a wedge between an ETF's portfolio and its basket. In reality, baskets often deviate from portfolios (see Section 3). Many basket-portfolio deviations are driven by the ETF's desire to adjust its portfolio. ETF portfolios often deviate from the portfolios that ETFs would like to hold, for reasons such as index rebalancing and APs' demands for custom baskets. We hypothesize that when an ETF's portfolio differs from its target, the ETF favors baskets whose use pushes the portfolio back to the target.

Suppose an ETF holds too much of a given security relative to the underlying index. The ETF could in principle sell some of this security in the secondary market, but such transactions are not commonly done in practice because they trigger tax liabilities. Instead, it can be more efficient for the ETF to increase the security's weight in the RD basket. Upon delivering this RD basket to an AP in exchange for ETF shares, the ETF reduces its holdings of this security in a tax-free manner (because the exchange is in-kind). The same objective can be achieved by reducing the weight of this security in the CR basket, because new creations then bring less of this security to the ETF's portfolio.

The same logic, in reverse, applies when an ETF holds too little of a given security relative to the index. The ETF can then increase the security's weight in the CR basket, and reduce it in the RD basket, to steer the ETF portfolio back to the index. The logic applies not only to securities but also to cash holdings. We thus obtain the following hypotheses about the active management of ETF baskets over time.

HYPOTHESIS **3.** *When cash is overweighted in an ETF's portfolio, the ETF increases its weight in the RD basket and decreases it in the CR basket. Such adjustments are attenuated when the underlying securities are less liquid.*

HYPOTHESIS **4.** *When a given security is overweighted in an ETF's portfolio, the ETF increases its weight in the RD basket and reduces it in the CR basket. Such adjustments are attenuated when the underlying securities are less liquid.*

The liquidity-related parts of both hypotheses are motivated by our model. The model does not feature dynamic basket adjustments, but it does explain how liquidity transformation comes at the expense of index-tracking. Proposition 2 shows that ETFs tolerate larger index deviations when index securities are less liquid. Suppose an ETF's portfolio is currently overweighting certain securities, or cash, relative to the target. To improve index-tracking, the ETF will adjust its baskets to reduce this overweighting. However, these basket adjustments will be smaller if the underlying securities are costlier to trade, because the ETF needs the AP to be willing to trade those securities to exchange them for the baskets. Instead of making aggressive basket adjustments that would correct index deviations immediately, the ETF tolerates some tracking error to deliver more liquidity transformation.

5 Cash and Basket Management

In this section, we test the four hypotheses developed in Section 4. We test Hypothesis 1 in Section 5.1, Hypothesis 2 in Section 5.2, etc. Our evidence supports all four hypotheses. We find that ETFs tracking less liquid indexes use more cash and have larger tracking errors. We also find that ETFs actively steer their portfolios toward their indexes, and that their capacity to do so is constrained by the illiquidity of their holdings.

5.1 Cash and Liquidity

According to Hypothesis 1, ETFs tracking less liquid indexes should have more cash in their baskets. Intuitively, trading less liquid bonds imposes larger transaction costs on APs, so that more basket cash is needed to encourage arbitrage trading by APs.

To test the hypothesis, we relate $Cash_{jt}$, the proportion of cash in the reported basket of ETF j on day t (i.e., the cash ratio), to $Illiquidity_{jt}$, the average illiquidity of the non-cash securities in the ETF's portfolio, weighted by the number of bonds held. We use portfolio illiquidity to proxy for index illiquidity for our cross-sectional analysis because index data are available for only about half of our ETFs, as noted earlier. We include only ETF-days on which the ETF experienced no creations or redemptions to better capture the ETF's equilibrium cash exposure. We exclude ETF-days with cash ratios exceeding 10%, to suppress outliers. We drop ETFs that track indexes with target maturity dates because such ETFs tend to increase their cash holdings as the maturity date approaches.

Figure 2 shows scatterplots of $Cash_{jt}$ against $Illiquidity_{jt}$ in Panels A, C, and E. The three panels correspond to our three measures of liquidity (see Section 2.3). To reduce noise in the data, we sort observations by $Illiquidity_{jt}$ into 20 equal-size bins. We then compute the average values of $Cash_{jt}$ and $Illiquidity_{jt}$ within each bin and plot one against the other. In each panel, the relation between $Cash_{jt}$ and $Illiquidity_{jt}$ is clearly positive, indicating that ETFs tracking less liquid indexes tend to have larger cash ratios.

Each panel of Figure 2 also plots the line of best fit from the regression

$$Cash_{jt} = \beta \, Illiquidity_{jt} + \omega_{It} + \epsilon_{jt} \,, \tag{6}$$

where ω_{It} is an issuer-time fixed effect. By including these fixed effects, we remove time-varying differences in cash management across ETF issuers. We are effectively estimating a cross-sectional correlation, comparing ETFs from the same issuer on the same day.

The β estimates—slopes of the lines plotted in Panels A, C, and E—range from 0.22 to 0.48. These estimates imply that a one-standard-deviation increase in $Illiquidity_{jt}$, which is standardized to unit variance, is associated with an increase in the cash ratio of 22 to 48 basis points (bps), depending on the liquidity measure. These estimates are economically significant relative to the 109 bps median cash ratio for reported baskets. All three estimates are also statistically significant at the 1% level, as we show in the Appendix.

We prioritize reported baskets over realized ones in this test because reported baskets are available every day, whereas realized baskets can be imputed only on days with CR or RD activity. Such activity is often induced by buying or selling pressure, which could indicate a temporary deviation from the equilibrium. In addition, imputed cash ratios exhibit larger outliers. Nevertheless, we also estimate regression (6) using cash in realized baskets. We exclude realized baskets containing fewer than 10 securities, both here and in subsequent analysis, because such small baskets are more likely to be driven by imputation errors or trades occurring outside the AP process. We winsorize cash ratios at the 5% level. We find results similar to those based on reported baskets, but with weaker statistical significance. The β estimates are positive in all six regression specifications (three liquidity measures times two baskets, CR and RD), but only two of them are statistically significant, both at the 1% level. See the Appendix for details. Overall, we find that ETFs holding less liquid bonds use more cash in their baskets, as predicted by Hypothesis 1.

While Hypothesis 1 applies to basket cash, it also carries implications for cash in ETFs' portfolio holdings. In periods of sustained net ETF share creation, more cash in baskets translates into more cash in holdings. Therefore, in such periods, ETFs tracking less liquid indexes should hold more cash not only in their baskets but also in their portfolios. In our sample period of 2017 through 2020, bond ETFs have indeed experienced a large amount of net creation, as we document later in Section 6.3. Therefore, Hypothesis 1 seems relevant also to cash in ETF portfolios, albeit to a lesser extent.

To examine this version of Hypothesis 1, we repeat the above analysis replacing cash in baskets by cash in holdings, so that $Cash_{jt}$ is the proportion of cash in the portfolio of ETF j on day t. The results are reported in Panels B, D, and F of Figure 2. The β estimates range from 0.11 to 0.19, indicating that a one-standard-deviation increase in $Illiquidity_{jt}$ is associated with an increase in the cash ratio of 11 to 19 bps. These estimates are not as large as those in Panels A, C, and E (visually, the slopes in the right-hand panels are not as steep as those in the left-hand panels), but they are sizable relative to the median cash ratio of 83 bps. Moreover, all three estimates are significant at the 1% level. ETFs holding less liquid bonds hold more cash in their portfolios, as predicted.

5.2 Tracking Error and Liquidity

According to Hypothesis 2, ETFs tracking less liquid indexes should track less closely. As trading illiquid bonds is costly to APs, ETFs holding such bonds tolerate larger tracking errors to accommodate APs' demands and thereby incentivize APs' arbitrage activity.

For each ETF j and day t, we compute $TrackingError_{jt}$ as the standard deviation of the difference between the ETF's daily return and the index return over the past month. As before, we sort ETF-day observations into 20 equal-size bins by $Illiquidity_{jt}$. We compute the average values of $TrackingError_{jt}$ and $Illiquidity_{jt}$ within each bin and plot one against the other in Panels A, C, and E of Figure 3. The three panels correspond to the three liquidity measures used before. Each panel also plots the line of best fit from the regression

$$TrackingError_{jt} = \beta\, Illiquidity_{jt} + \omega_{It} + \epsilon_{jt}\,, \tag{7}$$

where ω_{It} is an issuer-time fixed effect.[9] As in equation (6), by including ω_{It}, we are conducting cross-sectional comparisons of ETFs from the same issuer on the same day.

Figure 3 shows a strongly positive relation between $TrackingError_{jt}$ and $Illiquidity_{jt}$. In each panel, the fitted regression line slopes upward, indicating larger index deviations for ETFs tracking less liquid indexes. All three β estimates are positive and statistically significant at the 1% level, as we show in the Appendix.

One potential concern is the role of stale bond prices. $TrackingError_{jt}$ is calculated from the difference between ETF returns and index returns. While ETF returns are computed

[9]We are recycling some of the notation from equation (6)—β for the slope, ω_{It} for issuer-day fixed effects, and ϵ_{jt} for the error term—even though there is no direct relation between equations (6) and (7). To economize on notation, we also engage in similar recycling in all subsequent regression equations. Throughout the paper, we use β as generic notation for a regression slope, ω as generic notation for a fixed effect, and ϵ as generic notation for the error term, with various subscripts or superscripts. The various β's, ω's, and ϵ's are not related across regression equations.

from ETF prices, which change frequently, index returns are computed from the underlying bond prices, which change infrequently. To remove this mismatch, we construct an alternative measure of tracking error that computes both ETF and index returns from the same bond-level price data. Specifically, we construct daily returns on the ETF portfolio (index) by weight-averaging the returns on the individual bonds in the portfolio (index) and compute the monthly standard deviation of the difference between the resulting return series. In this approach, stale pricing may still be present, but it affects both ETF and index returns symmetrically. Panels B, D, and F of Figure 3 present results for this alternative measure of tracking error. The relation of interest remains positive and significant at the 1% level. A one-standard-deviation increase in $Illiquidity_{jt}$ is associated with an ETF tracking error that is 32 to 38 percentage points of a standard deviation larger. Overall, ETFs investing in less liquid bonds deviate more from their indexes, consistent with Hypothesis 2.

5.3 Active Cash Management

In this section, we analyze how ETFs actively manage their basket cash over time. According to Hypothesis 3, ETFs dynamically adjust their baskets to revert to a certain desired level of cash, which may or may not be zero. ETFs holding too much cash respond by adding cash to RD baskets and removing it from CR baskets, and vice versa. This active cash management is constrained by the illiquidity of the underlying securities.

To test Hypothesis 3, we construct three variables. $BasketCash_{jt}$ is the proportion of cash in a CR or RD basket of ETF j on day t. We use realized baskets because they exhibit much more time variation than reported baskets. As before, we exclude baskets with fewer than 10 securities and ETFs with target maturity dates. $\Delta Cash_{jt-1}$ is the difference between the fraction of cash in the ETF's portfolio on day $t-1$ and the average of those fractions over the prior month. Interpreting this average as the equilibrium level of cash, $\Delta Cash_{jt-1}$ captures the ETF's over- or underweighting of cash. Finally, $Illiquidity_j$ is ETF j's average index illiquidity, which is computed by first averaging the illiquidity of the ETF's index constituents and then averaging over time. We estimate how $BasketCash_{jt}$ varies with $\Delta Cash_{jt-1}$ and its interaction with $Illiquidity_j$:

$$BasketCash_{jt} = \beta_1 \, \Delta Cash_{jt-1} + \beta_2 \, \Delta Cash_{jt-1} \times Illiquidity_j + \omega_j + \epsilon_{jt} \, , \tag{8}$$

where ω_j is an ETF fixed effect. By including this fixed effect, we are effectively running a series of time-series regressions, one for each ETF.

Table 3 reports the estimates of β_1 and β_2. We find that $\beta_1 < 0$ for CR baskets and $\beta_1 > 0$ for RD baskets. These results indicate that when cash is overweighted in an ETF's portfolio (i.e., $\Delta Cash_{jt-1} > 0$), the ETF tends to remove cash from CR baskets and add it to RD baskets. Such basket adjustments help ETFs steer their cash holdings toward their desired long-term average. We also find that $\beta_2 > 0$ for CR baskets and $\beta_2 < 0$ for RD baskets, indicating that the above basket adjustments are attenuated for ETFs tracking less liquid indexes. All of these results are consistent with Hypothesis 3.

The estimates of β_1 and β_2 are statistically significant for both CR and RD baskets and for all three measures of liquidity. They are also economically significant. A one-standard-deviation increase in $\Delta Cash_{jt-1}$ is associated with 3.3 to 3.6 percentage points (pps) less cash in CR baskets and 3.1 to 3.2 pps more cash in RD baskets for an ETF with the average level of illiquidity. When the illiquidity is one standard deviation larger, the effect is mitigated by 1.3 to 1.7 pps for CR baskets and 1.1 to 1.4 pps for RD baskets.

5.4 Active Basket Management

Finally, we test Hypothesis 4, according to which ETFs actively manage the composition of their baskets to steer their portfolios toward the index. ETFs holding too much of a given bond respond by adding this bond to RD baskets and removing it from CR baskets, and vice versa. These adjustments are smaller when the bonds are less liquid.

To test this hypothesis, we use index rebalancing as a source of plausibly exogenous variation in the over- or underweighting of bonds in ETF portfolios. Using such variation helps alleviate the concern that a bond's over- or underweighting could be driven by unobserved characteristics that could also cause the bond's basket inclusion. We take advantage of the fact that fixed-income indexes rebalance on the last day of each month. While changes in the bond universe and bond characteristics occur throughout the month, index composition is not updated until the month-end to reflect those changes. The monthly jump in index portfolio weights therefore constitutes a plausibly exogenous shock to the over- or underweighting of bonds in ETF portfolios relative to the index. Conversations with ETF managers indicate that the changes in index weights are not fully predictable. Even though some of them are predictable, ETFs incur tracking error if they make anticipatory portfolio adjustments before the month-end. Therefore, ETFs often postpone such adjustments until the month-end. Indeed, the difference between ETF portfolio weights and index weights tends to spike at the monthly index rebalancing dates, as we show in the Appendix. The Appendix also contains a more detailed discussion of the rebalancing of fixed-income indexes.

Let $Deviation_{ijt}$ denote the difference between bond i's weight in ETF j's portfolio and the bond's weight in the index on day t. We refer to this difference as the bond's "overweighting" in the ETF's portfolio relative to the index, with the understanding that negative values represent underweighting.[10] For each index rebalancing day h, we compute each bond's overweighting in excess of its average overweighting over the previous week:

$$\Delta Deviation_{ijh} \equiv Deviation_{ijh} - \frac{1}{5}\sum_{k=1}^{5} Deviation_{ijh-k}\,. \tag{9}$$

We interpret $\Delta Deviation_{ijh}$ as a rebalancing-induced shock to the bond's overweighting. We relate this shock to the bond's basket inclusion by estimating the regression models

$$Basket_{ijt}^{CR} = \beta_1^{CR}\,\Delta Deviation_{ijh} + \beta_2^{CR}\,\Delta Deviation_{ijh} \times Illiquidity_j + \omega_{jt} + \epsilon_{ijt} \tag{10}$$

$$Basket_{ijt}^{RD} = \beta_1^{RD}\,\Delta Deviation_{ijh} + \beta_2^{RD}\,\Delta Deviation_{ijh} \times Illiquidity_j + \omega_{jt} + \epsilon_{ijt}\,, \tag{11}$$

where $Basket_{ijt}^{CR}$ ($Basket_{ijt}^{RD}$) is a dummy variable equal to one when bond i is included in ETF j's CR (RD) basket on day t, $Illiquidity_j$ is the ETF's index illiquidity, as before, and ω_{jt} represent ETF-day fixed effects. Given these fixed effects, we exploit cross-sectional variation in basket inclusion across bonds within the same ETF on the same day.

For each ETF, we include in the sample the first ten baskets after each monthly rebalancing date. We use the first few baskets after each index rebalancing because they are the most likely to respond to the rebalancing. We do not use more than ten baskets because the first ten baskets are likely to close much of the gap between the portfolio and the index, and we do not want to use baskets unaffected by the rebalancing. If fewer than ten baskets are available for the given month, we use all available baskets in that month.

Panel A of Table 4 reports the regression estimates. We see that $\beta_1^{CR} < 0$ and $\beta_1^{RD} > 0$. These estimates show that when a bond becomes more overweighted in an ETF's portfolio after the rebalancing of the ETF's index, the bond is less likely to be included in CR baskets but more likely to be included in RD baskets. These results are consistent with ETFs steering their portfolios back to the index after sudden deviations caused by index rebalancing. We also find that $\beta_2^{CR} > 0$ and $\beta_2^{RD} < 0$, indicating that the aforementioned basket adjustments are attenuated when the bonds held by the ETF are less liquid. All of these results are statistically significant at the 1% level and fully consistent with Hypothesis 4.

Panel B of Table 4 reports the results from an analogous estimation in which the dummy variables $Basket_{ijt}^{CR}$ and $Basket_{ijt}^{RD}$ are replaced by quantities, namely, the log of one plus the number of shares of bond i in ETF j's CR and RD baskets, respectively. Given the heterogeneity in issuance across bonds, we add a control for the log number of shares outstanding of

[10] The value of $Deviation_{ijt}$ is missing if bond i is held by neither ETF j nor its index on day t. If the bond is not held by the ETF but is in the index, $Deviation_{ijt}$ is negative to indicate underweighting. If the bond is held by the ETF but is not in the index, $Deviation_{ijt}$ is positive to indicate overweighting.

bond i. The results are very similar to those in Panel A in that $\beta_1^{CR} < 0$, $\beta_1^{RD} > 0$, $\beta_2^{CR} > 0$, and $\beta_2^{RD} < 0$. As before, all the estimates are statistically significant at the 1% level. They are also economically significant. A one-standard-deviation increase in rebalancing-induced overweighting for a given bond is associated with a 33.2% to 34.2% decrease of that bond's presence in CR baskets for an ETF of average index illiquidity. The effect is attenuated by 6.1% to 6.8% when the index illiquidity is one standard deviation higher. The same one-standard-deviation increase corresponds to a 20.9% to 23.8% increase of the bond's quantity in RD baskets for an ETF of average illiquidity. That effect is attenuated by 3.9% to 6.9% when the illiquidity is one standard deviation higher.

6 The Effect of ETF Basket Inclusion on Bond Liquidity

Recall that the active management of passive ETFs is motivated by the desire to boost the liquidity of ETF shares. Since ETF shares are connected to the underlying securities through an arbitrage mechanism, it seems natural to ask whether this boost to ETF-level liquidity spills over to the security level. With this question in mind, we now examine the implications of ETFs' active management for the liquidity of the underlying bonds.

We estimate the effect of basket inclusion on individual bond liquidity in two ways. In Section 6.1, we estimate the relation between basket inclusion and subsequent bond liquidity in the presence of controls and fixed effects. In Section 6.2, we use an instrument for basket inclusion to estimate its causal effect on bond liquidity by two-stage least squares. Using both approaches, we find that a bond's inclusion in an ETF basket makes the bond more liquid. However, basket inclusion makes the bond less liquid in periods of large imbalance between creations and redemptions, as we show in Section 6.3.

6.1 Baseline Analysis

Our baseline analysis relates an individual bond's illiquidity on a given day to the bond's basket inclusion on the previous day. Let $Illiquidity_{it+1}$ denote the illiquidity of bond i on day $t+1$, for any of our three illiquidity measures. Let $Basket_{it}^{CR}$ denote the extent of bond i's inclusion in ETFs' realized CR baskets on day t. We consider two measures of basket inclusion (and a third in the Appendix). The first measure is the number of ETFs that include this bond in their CR baskets on this day. The second measure is the log of one plus the number of shares of this bond that are included in ETFs' CR baskets on this day.

$Basket_{it}^{RD}$ is defined analogously, except for RD baskets. We estimate the model

$$Illiquidity_{it+1} = \beta^{CR}\, Basket_{it}^{CR} + \beta^{RD}\, Basket_{it}^{RD} + Controls_{it} + \omega_{Ft} + \omega_{mt} + \epsilon_{it}\,. \qquad (12)$$

The controls include the average $Illiquidity_{it}$ over the prior week to address the possibility that bonds selected into baskets may be more liquid ex ante. We further control for the average basket size of ETFs holding bond i because, as noted earlier, baskets containing few bonds may reflect imputation errors or adjustments that do not operate through ETF arbitrage. Finally, when using the second basket inclusion measure, we also control for the log number of bond i's outstanding shares on day t. We include firm-day fixed effects, ω_{Ft}, and maturity-day fixed effects, ω_{mt}. The former fixed effects control for time-varying fundamental shocks at the firm level, whereas the latter account for time-varying yield-curve effects. By including these fixed effects, we are effectively comparing near-identical bonds (of the same maturity, issued by the same firm) on the same day.

Our sample includes only bonds that are held in at least one ETF portfolio. We remove observations in which a bond appears in CR baskets for the first time as well as those in which a bond appears in RD baskets for the last time. We do this to remove mechanical basket inclusions because a bond newly added to the index can only appear in the ETF portfolio by being part of the CR basket and a bond newly excluded from the index can only leave the ETF portfolio by being part of the RD basket. The remaining variation in basket inclusion better represents ETFs' discretionary basket management.

Table 5 shows that the estimates of β^{CR} and β^{RD} are both negative, for both basket inclusion measures and all three liquidity measures. All 12 of these estimates are statistically significant (11 of them at the 1% level, one at the 5% level). These results are consistent with the hypothesis that inclusion in ETF baskets improves a bond's liquidity.

To interpret the magnitudes of the estimates, we standardize $Illiquidity_{it}$ by its mean and standard deviation, both computed across all of its bond-day observations, and multiply its values by 100 to improve the readability of the coefficients. The estimates imply that being included in a CR (RD) basket of one more ETF is associated with a decrease in the bond's next-day illiquidity by 1.0% to 3.7% (1.8% to 4.3%) of a standard deviation. In addition, a 10% increase in the number of a bond's shares included in CR (RD) baskets is associated with a drop in the bond's next-day illiquidity by 0.5% to 1.8% (1.0% to 1.8%) of a standard deviation. These relations are likely to be causal given that we are comparing a bond to near-identical bonds on the same day and controlling for the bond's recent illiquidity.

6.2 The Instrumental Variables Approach

In Section 6.1, we address potential concerns regarding the endogeneity of a bond's basket inclusion by controlling for the bond's recent illiquidity and using firm-time and maturity-time fixed effects. Nonetheless, readers may still wonder what the determinants of the remaining variation in basket inclusion are and whether they correlate with liquidity within the same firm and maturity bracket. One determinant, uncovered earlier in this paper, is the bond's over- or underweighting in the ETF's portfolio relative to the index. Recall from Section 5.4 that ETFs are more likely to include underweighted bonds in CR baskets and overweighted bonds in RD baskets. In this section, we use plausibly exogenous variation in this over- or underweighting to construct an instrument for basket inclusion.

We construct this instrument by recognizing, again, that fixed-income indexes rebalance monthly. Even though changes in bond characteristics that trigger index inclusions and exclusions happen throughout the month, index composition is not updated until the monthly rebalancing day. Any effects of changes in bond characteristics on bond liquidity should be incorporated in asset markets at the time when those characteristics are changing, not later when the index is rebalanced. Therefore, shocks to bond over- or underweighting on index rebalancing days should affect ETF basket inclusion without being confounded by changes in unobservable bond characteristics that also influence bond liquidity.

To instrument for the basket inclusion variables $Basket_{it}^{CR}$ and $Basket_{it}^{RD}$ in equation (12), we construct the variables $CRInstr_{it}$ and $RDInstr_{it}$, respectively, as follows:

$$CRInstr_{it} \equiv \sum_{j \in J_t^{CR}} \Delta Deviation_{ijh} \tag{13}$$

$$RDInstr_{it} \equiv \sum_{j \in J_t^{RD}} \Delta Deviation_{ijh} \, , \tag{14}$$

where $\Delta Deviation_{ijh}$, which is defined in equation (9), is a shock to bond i's overweighting in ETF j's portfolio induced by the rebalancing of the ETF's index on day h. For any day t following day h, we compute bond i's instrument $CRInstr_{it}$ by adding up the bond's overweighting shocks over the set of all ETFs that have CR baskets, J_t^{CR}. Similarly, we compute $RDInstr_{it}$ by summing the same shocks over the set J_t^{RD} of ETFs with RD baskets.[11] As before, we use only the first ten baskets per ETF after each rebalancing.

Intuitively, bonds with high values of $CRInstr_{it}$ or $RDInstr_{it}$ have become more over-

[11]Note that both sets J_t^{CR} and J_t^{RD} can include only ETFs for which we have index composition data, because such data are needed to compute $\Delta Deviation_{ijh}$. Also note that an ETF does not have to hold bond i in order to be included in the set J_t^{CR} or J_t^{RD}; it only has to have a CR or RD basket on day t.

weighted in ETF portfolios as a result of index rebalancing. Given our results from Section 5.4, we expect such bonds to be disproportionately included in RD baskets and excluded from CR baskets. We expect $CRInstr_{it}$ to better capture exclusion from CR baskets because the set J_t^{CR} over which $CRInstr_{it}$ is computed contains ETFs that have CR baskets but need not have RD baskets. Similarly, we expect $RDInstr_{it}$ to better capture inclusion in RD baskets because the set J_t^{RD} contains ETFs that have RD baskets.

The first-stage specifications in our 2SLS estimation are

$$\begin{aligned} Basket_{it}^{RD} &= \beta_1^{RD}\, RDInstr_{it} + \beta_2^{RD}\, CRInstr_{it} + Controls_{it} + \omega_{Ft} + \omega_{mt} + \epsilon_{it} \quad (15) \\ Basket_{it}^{CR} &= \beta_1^{CR}\, RDInstr_{it} + \beta_2^{CR}\, CRInstr_{it} + Controls_{it} + \omega_{Ft} + \omega_{mt} + \epsilon_{it}\,. \quad (16) \end{aligned}$$

We expect to find $\beta_1^{RD} > 0$ and $\beta_2^{CR} < 0$, as explained in the previous paragraph.

The first-stage results are reported in Table 6. As before, we consider two measures of bond i's basket inclusion on day t, $Basket_{it}^{CR}$. First, we measure it by the number of ETFs that include the bond in their CR baskets on that day. Second, we use the log of one plus the number of shares of this bond that are included in ETFs' CR baskets on that day. The results for the first measure are in Panel A of Table 6; those for the second measure are in Panel B. We measure $Basket_{it}^{RD}$ analogously based on RD baskets. The results for $Basket_{it}^{RD}$ are in columns 1 to 3; those for $Basket_{it}^{CR}$ are in columns 4 to 6.

Table 6 shows that $\beta_1^{RD} > 0$ and $\beta_2^{CR} < 0$, as predicted. Both results hold for both basket inclusion measures and all three measures of liquidity, which is included among the controls. All of these estimates are significant at the 1% level.[12] We thus see that bonds overweighted after index-rebalancing shocks tend to be included in RD baskets but excluded from CR baskets. These bond-level results are not surprising given the ETF-bond-level results reported in Table 4, which we used to motivate our instruments. The cross-effects of $CRInstr_{it}$ on RD basket inclusion (β_2^{RD}) and $RDInstr_{it}$ on CR basket inclusion (β_1^{CR}) are weaker, as expected. As explained earlier, it is primarily overweighting at ETFs with CR (RD) baskets that affects a bond's CR (RD) basket inclusion.

From the first-stage regressions (15) and (16), we obtain the predicted values $\widehat{Basket}_{it}^{CR}$ and $\widehat{Basket}_{it}^{RD}$. In the second stage, we use these predicted values to instrument for $Basket_{it}^{CR}$ and $Basket_{it}^{RD}$, respectively, in equation (12). The second-stage slopes on these predicted values estimate the causal effect of basket inclusion on bond illiquidity.

Table 7 shows the second-stage results. The slope estimates on instrumented basket

[12] In addition, all of the Cragg and Donald (1993) F-statistics from our first-stage regressions easily pass the Stock and Yogo (2005) test for weak instruments.

inclusion are always negative. This is true in all 12 specifications: two measures of basket inclusion, two types of baskets (CR and RD), and three liquidity measures. Among the 12 negative estimates, five are significant at the 1% level, four at the 5% level, and two at the 10% level. We conclude that basket inclusion causes an improvement in bond liquidity.

This effect is significant not only statistically but also economically. When a bond is included in one additional RD basket, its illiquidity decreases by 7.3% to 15.3% of a standard deviation. When the number of a bond's shares included in RD baskets increases by 10%, the bond's illiquidity decreases by 5.5% to 14.1% of a standard deviation. These magnitudes are substantial despite the presence of firm-time and maturity-time fixed effects, which absorb some of the variation of interest. The results for CR baskets are similar.

6.3 The Role of Basket Imbalance

In Sections 6.1 and 6.2, we show that basket inclusion improves bond liquidity in our full sample. On most days in our sample, the shocks that give rise to buying or selling pressure in ETF shares are idiosyncratic, leading APs to alternate between CR and RD activity. Bonds used in CR or RD baskets thus tend to move back and forth between the asset and liability sides of the APs' balance sheets. Their frequent presence on both sides of AP inventory makes basket bonds more liquid, because the APs also tend to act as market makers in these bonds. However, when the shocks in the ETF market are asymmetric, the resulting CR-RD imbalances could potentially make basket bonds less liquid. In particular, systematic CR-RD imbalances generate imbalances in the market maker's inventory, with basket bonds concentrated on one side of the inventory, which could harm the bonds' liquidity. In this section, we examine the role of CR-RD imbalances in the basket-liquidity relation.

Panel A of Figure 4 plots the time series of the number of ETFs in our sample that create or redeem shares between January 2017 and December 2020. We see that ETFs engaged in creations generally outnumber those engaged in redemptions. On a typical day, about 20 ETFs create new shares, but only five or six ETFs redeem. Panel B shows that creations tend to outweigh redemptions also in dollar terms. However, net redemption volume exhibits substantial week-to-week variation, and there are about 40 weeks in which redemptions outweigh creations in dollar terms. Our sample period is thus characterized mostly by net creation, but with a fair amount of redemption occurring as well.

The only time when bond ETFs experienced large and persistent redemptions was during the COVID-19 crisis. In March 2020, we saw sharp increases in the number of ETFs engaged in redemption as well as in the volume of net redemptions. In the week of March 11 to 18 alone, total net redemptions by the ETFs in our sample reached \$3.2 billion. This spike in redemptions is clearly visible in both panels of Figure 4.

To examine how COVID-induced redemptions affect the basket-liquidity relation, we estimate regression (12) over the "COVID subperiod" of March 2 to April 15, 2020. Starting the week of March 2, corporate bond spreads began to surge (Haddad et al., 2020). March 2 also marks the beginning of systematic redemptions from corporate bond ETFs (Figure 4). Redemptions began to decline following the Federal Reserve's market interventions, and they largely returned to their pre-COVID levels by mid-April.[13] Our choice of April 15 as the end of the COVID subperiod is close to the dates chosen by others for similar purposes. For example, Haddad et al. (2020) mark April 16 as the end of the recovery period in the corporate bond basis, while He et al. (2020) end their sample period on April 13.

Table 8 shows that the slopes from regression (12) estimated in the COVID subperiod are very different from those obtained in the full period. Recall that, in the full period, all 12 estimates of the slopes on basket inclusion are negative and significant (Table 5). In contrast, nine of the 12 estimates are positive in the COVID subperiod. The results for RD baskets are especially different. The magnitudes of the estimated slopes on RD basket inclusion are similar to those in Table 5, but the signs are opposite. All six of these RD basket slopes are positive and five of them are statistically significant (two at the 1% level, one at the 5% level, and two at the 10% level). We thus see that inclusion in RD baskets decreases rather than increases bond-level liquidity during the COVID subperiod.

In this exercise, we cannot follow our instrumental variables approach due to the short length of the COVID subperiod, which includes only one index rebalancing date. Moreover, several index providers canceled their March 2020 rebalancing due to extreme market conditions. However, the results in Table 8 are based on same-day comparisons of bonds of the same maturity issued by the same firm, controlling for recent bond-level illiquidity, all of which alleviate endogeneity concerns. Moreover, the conclusions from Table 8 are robust to the inclusion of additional controls and fixed effects (see the Appendix).

The COVID subperiod provides preliminary evidence that inclusion in RD baskets can hurt a bond's liquidity when redemptions are systematic and persistent. In this subperiod, many investors experienced liquidity shocks that led them to sell ETF shares. This selling pressure was met by APs who purchased many ETF shares from investors, redeemed them,

[13]The key interventions in the corporate bond market took place on March 23, 2020, when the Fed announced that it would buy investment-grade corporate bonds, both through ETFs and directly, and on April 9, 2020, when it announced that it would buy non-investment-grade bond ETFs. These interventions also stemmed the decline in corporate bond liquidity (Kargar et al., 2021, O'Hara and Zhou, 2021).

and then tried to sell the bonds acquired through RD baskets. Bonds heavily represented in RD baskets thus became heavily represented in APs' inventory. Given their balance sheet constraints, APs became reluctant to purchase even more of the same bonds in their role as market makers. Bonds present in RD baskets thus lost their most natural buyers. When its own market makers do not want to buy it, a security can become quite illiquid.

We now explore the role of basket imbalance more generally. To measure basket imbalance for a given bond, we first let N_{it}^{CR} and N_{it}^{RD} denote the numbers of CR baskets and RD baskets, respectively, in which bond i appears anytime during the week immediately preceding day t. We then define two basket imbalance variables:

$$Imbal_{it}^{CR} \equiv \max\left(N_{it}^{CR} - N_{it}^{RD}, 0\right) \tag{17}$$

$$Imbal_{it}^{RD} \equiv \max\left(N_{it}^{RD} - N_{it}^{CR}, 0\right) . \tag{18}$$

That is, when a bond is included in more CR baskets than RD baskets (i.e., $N_{it}^{CR} > N_{it}^{RD}$), we set $Imbal_{it}^{CR}$ equal to the positive difference and $Imbal_{it}^{RD}$ to zero. And vice versa. We also let $Basket_{it}$ denote the total number of baskets, CR or RD, that contain bond i in the week preceding day t. We then estimate a modified version of regression (12):

$$Illiquidity_{it+1} = \beta_1\, Basket_{it} + \beta_2^{CR}\, Basket_{it} \times Imbal_{it}^{CR} + \beta_2^{RD}\, Basket_{it} \times Imbal_{it}^{RD} + e_{it}\,, \tag{19}$$

where $e_{it} \equiv Controls_{it} + \omega_{Ft} + \omega_{mt} + \epsilon_{it}$. As before, the controls include the bond's average illiquidity over the prior week and average basket size, and we also include a firm-day fixed effect, ω_{Ft}, and a maturity-day fixed effect, ω_{mt}.

Table 9 reports the regression estimates. We find that $\beta_1 < 0$, indicating that a bond's liquidity improves after the bond is included in a larger number of baskets. This result is similar to that observed in Table 5 for CR and RD baskets separately. However, we also find that $\beta_2^{CR} > 0$ and $\beta_2^{RD} > 0$, both highly significant, indicating that basket imbalance of either kind, CR or RD, weakens the favorable effect of basket inclusion on bond liquidity. This effect can become unfavorable if basket imbalance is sufficiently large.

To get a sense of the magnitudes, consider a bond included in five CR baskets and five RD baskets (i.e., $Imbal_{it}^{RD} = 0$). If this bond were instead included in ten CR baskets and no RD baskets (i.e., $Imbal_{it}^{CR} = 10$), the bond's liquidity would be lower by 17.0% to 19.0% of a standard deviation. Alternatively, if this bond were included in ten RD baskets and no CR baskets (i.e., $Imbal_{it}^{RD} = 10$), its liquidity would be lower by 14.8% to 29.8% of a standard deviation. To put these numbers into context, at the height of the COVID-19 crisis, 32 ETFs experienced redemptions while only five ETFs had creations. The effect of basket imbalance on the basket-liquidity relation can therefore be substantial.

Overall, we find that basket inclusion need not always improve a bond's liquidity. When there are systematic creations or redemptions across many ETFs, basket bonds can become less liquid as a result of being held one-sidedly by their market makers. This scenario, in the form of systematic redemptions, materialized during the COVID-19 crisis.

6.4 Robustness

In this subsection, we show that our results in Section 6 are robust to a number of alternative specifications. We summarize our findings here and show the details in the Appendix.

Sections 6.1 through 6.3 show that our results on the effect of basket inclusion on bond liquidity hold for two measures of basket inclusion and three measures of liquidity, and they survive the inclusion of multiple fixed effects and controls. The results also hold for a third measure of basket inclusion, a dummy variable equal to one if the bond is included in at least one ETF basket, and they survive the addition of a bond fixed effect, which isolates the time variation in a given bond's basket inclusion. The results also survive a control for ETF ownership, which is explored in prior studies and correlated with basket inclusion. We measure a bond's ETF ownership by the fraction of the bond's shares outstanding that are held by ETFs. We find that ETF ownership is positively related to future bond liquidity in normal times and negatively in the COVID subperiod, but adding this control does not affect any of our conclusions regarding the effects of basket inclusion.

We also explore the role of correlated trading of non-ETF investors. For example, if an ETF tracks the same index as an index mutual fund, index inclusion could affect bond liquidity not only through ETF basket inclusion but also through mutual fund trading. To address this issue, we add controls for bond inclusion in the baskets of ETFs that share indexes with index mutual funds, ensuring that the control matches the measure of basket inclusion (number of baskets or basket shares) used as the independent variable. These controls are nonzero for 5% of bond-day observations for which bonds are included in baskets. We find that the inclusion of these controls has a negligible effect on our results.

Our measures of bond liquidity are sometimes missing because not all bonds trade every day. However, we find very similar results when we repeat our tests using bond liquidity averaged over the first three trading days with observed bond trades after basket inclusion. We further control for the number of zero trading days, that is, the number of days that a bond has gone without trading over the past month. Again, all of our results remain, confirming that our findings are not driven by infrequent trading of a subset of bonds.

Finally, we consider a simpler measure of basket imbalance:

$$Imbal_{it} \equiv |N_{it}^{CR} - N_{it}^{RD}|, \tag{20}$$

that is, the absolute value of the difference between the numbers of CR and RD baskets in which the bond appears. We then run the regression

$$Illiquidity_{it+1} = \beta_1\, Basket_{it} + \beta_2\, Basket_{it} \times Imbal_{it} + Controls_{it} + \omega_{Ft} + \omega_{mt} + \epsilon_{it}\,. \tag{21}$$

The results mirror those in Table 9: $\beta_1 < 0$ and $\beta_2 > 0$, both significant at the 1% level. Overall, we find robust evidence that a bond's inclusion in ETF baskets improves the bond's liquidity, and that this effect is attenuated in times of high CR/RD imbalance.

7 Conclusion

We show that passive ETFs actively manage their portfolios, balancing index-tracking against liquidity transformation. ETFs update their baskets frequently to steer their portfolios toward the index. For example, when an ETF's portfolio underweights a security relative to the index, the ETF tends to add this security to the CR basket and remove it from the RD basket. These basket adjustments are attenuated for ETFs holding less liquid securities because ETFs also aim to incentivize arbitrage activity. Arbitrageurs boost the liquidity of ETF shares by absorbing other investors' trades, mitigating their price impact. To help reduce the arbitrageurs' transaction costs, ETFs adjust their baskets by including cash and excluding some of the index securities. The use of cash and basket concentration is more common when the index securities are less liquid, and it comes at the expense of index-tracking. We capture the tradeoff between index-tracking and liquidity transformation in a theoretical model. While we model ETFs in general, our evidence is based on corporate bond ETFs, for which liquidity transformation is particularly important.

To offer an analogy, imagine steering a ship powered by a team of rowers to a harbor on the other side of the lake. If the lake water were clear, you would chart a straight path and track it closely by quickly correcting all random deviations. However, the lake is full of mud. Rowing through mud is more exhausting. How do you adjust the ship's path in light of the mud? Instead of always tracking the straight path, you may want to avoid the muddiest parts. By steering clear of mud, you make it easier for the rowers to move the ship forward, which helps you get to the harbor more effectively. An astute reader already knows that the ship is a metaphor for an ETF portfolio, you are the ETF's manager, and the rowers symbolize the arbitrageurs. Clear water represents liquidity and mud illiquidity.

ETFs' efforts to improve the liquidity of their shares have consequences for the liquidity of the underlying securities. We find that a bond's inclusion in an ETF basket has a significant state-dependent effect on the bond's liquidity. This effect is positive in normal times but negative in periods of large imbalance between creations and redemptions. A salient example occurred in the spring of 2020. The COVID-19 crisis witnessed acute selling pressure in the bond market, which led to net redemptions from bond ETFs, which in turn strained the liquidity of the bonds concentrated in RD baskets. Given the growing role of ETFs in liquidity transformation, future episodes of ETF-induced liquidity strains seem likely.

Future research can examine additional consequences of ETFs' active basket management. One promising step in this direction is Reilly (2021), which studies the performance of bonds included in CR baskets. It would also be useful to analyze the security-level determinants of basket inclusion. Our imperfect understanding of many aspects of ETFs' liquidity transformation presents numerous opportunities for future work.

Figure 1: ETF Arbitrage

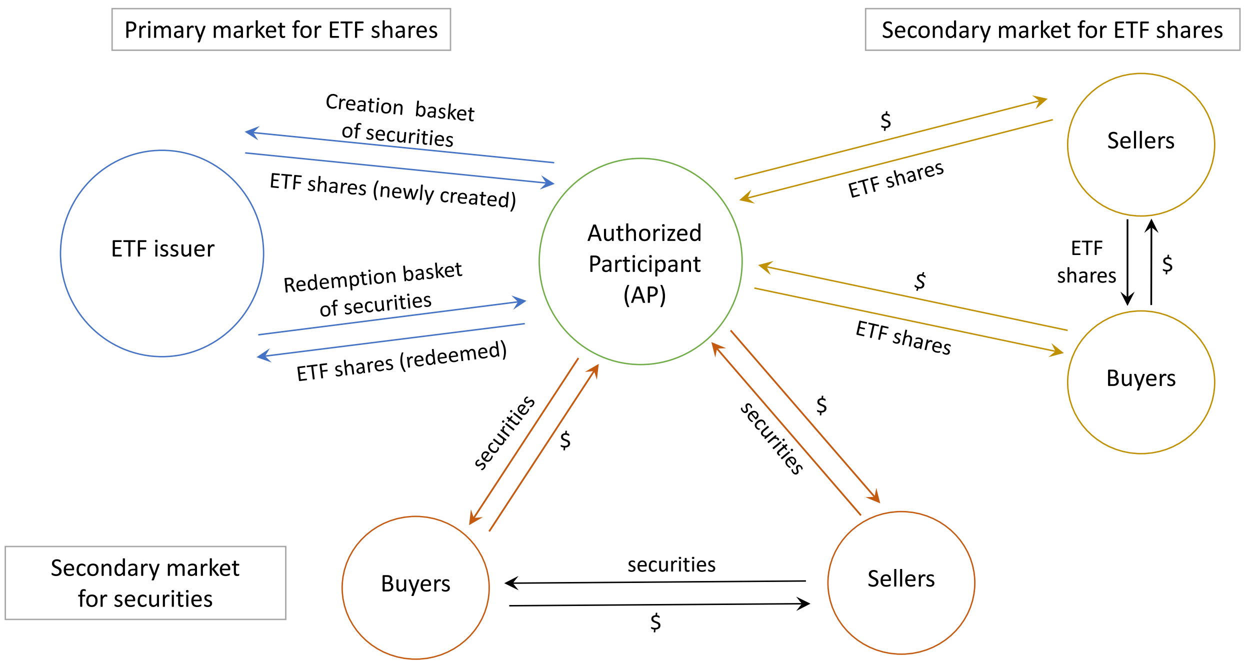
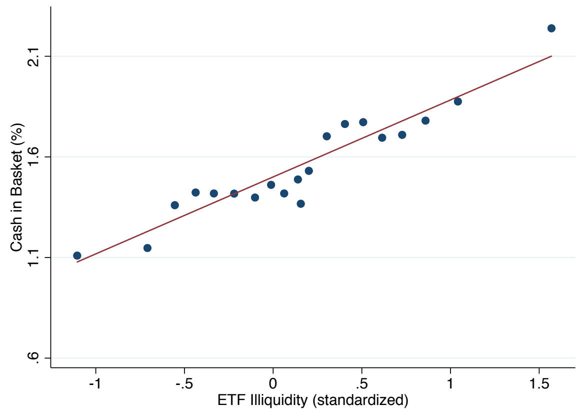


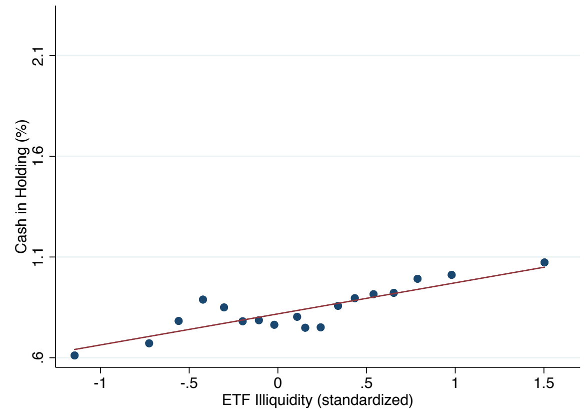
Figure 2: Cash and Liquidity

This figure shows binned scatterplots of the proportion of cash in reported ETF baskets (Panels A, C, and E) and ETF portfolio holdings (Panels B, D, and F) against the average illiquidity of the non-cash securities in the ETF's portfolio. The three measures of illiquidity are the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). Each panel also plots the line of best fit from a linear regression with issuer-time fixed effects.

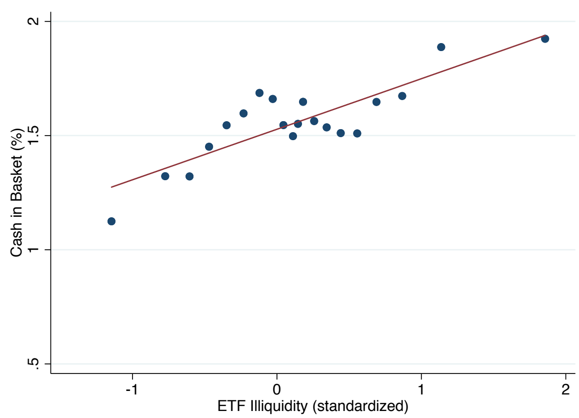
(a) Basket Cash and Illiquidity IL1



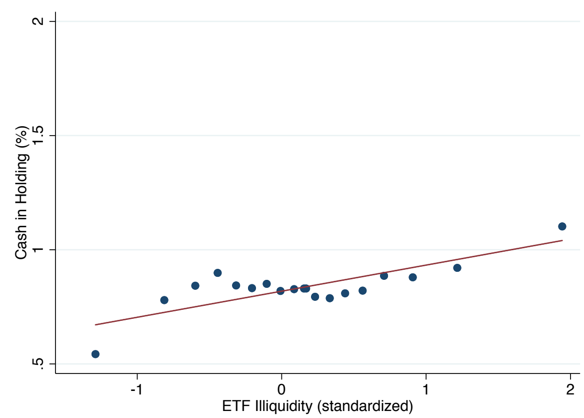
(b) Holdings Cash and Illiquidity IL1



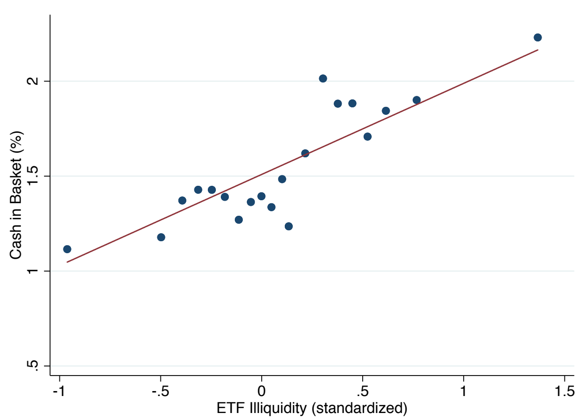
(c) Basket Cash and Illiquidity IL2



(d) Holdings Cash and Illiquidity IL2



(e) Basket Cash and Illiquidity IL3



(f) Holdings Cash and Illiquidity IL3

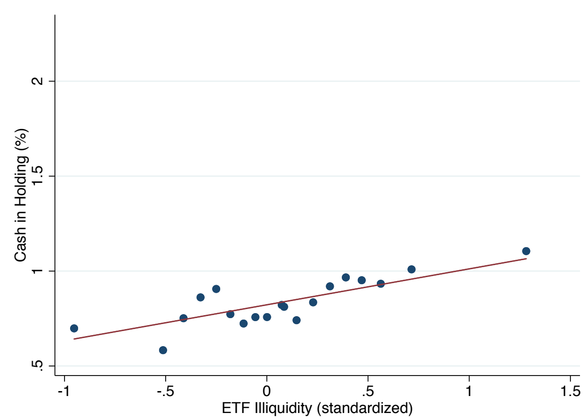
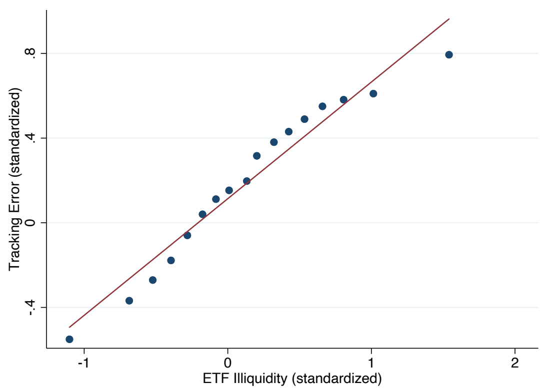


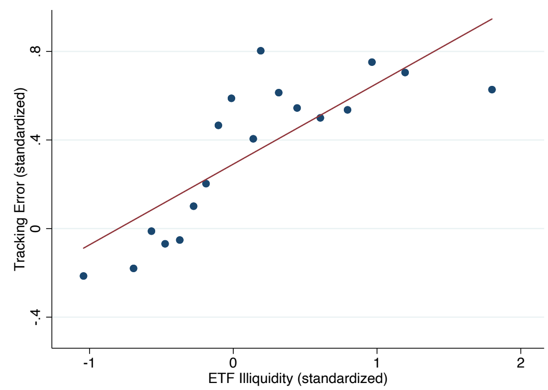
Figure 3: Tracking Error and Liquidity

This figure shows binned scatterplots of ETF tracking error against the average illiquidity of the non-cash securities in the ETF's portfolio. The three measures of illiquidity are the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). Both measures of tracking error are monthly standard deviations of daily differences between ETF returns and index returns. For Tracking Error 1, ETF returns are computed from ETF share prices and index returns come from Bloomberg. For Tracking Error 2, both ETF and index returns are computed from the prices of the underlying bonds. Each panel plots the line of best fit from a regression with issuer-time fixed effects.

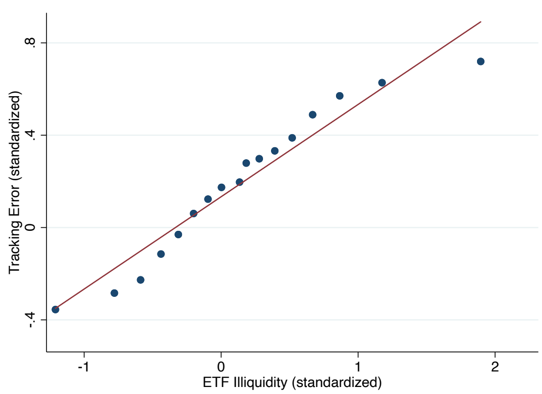
(a) Tracking Error 1 and Illiquidity IL1



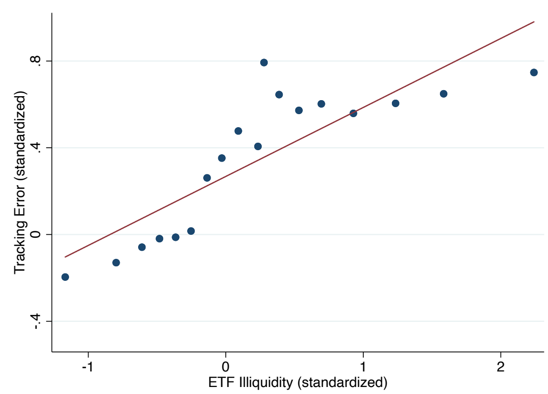
(b) Tracking Error 2 and Illiquidity IL1



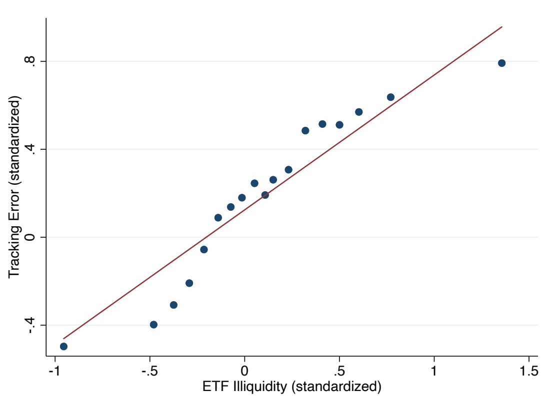
(c) Tracking Error 1 and Illiquidity IL2



(d) Tracking Error 2 and Illiquidity IL2



(e) Tracking Error 1 and Illiquidity IL3



(f) Tracking Error 2 and Illiquidity IL3

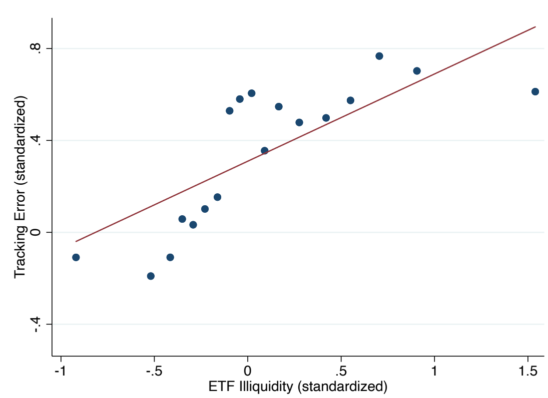
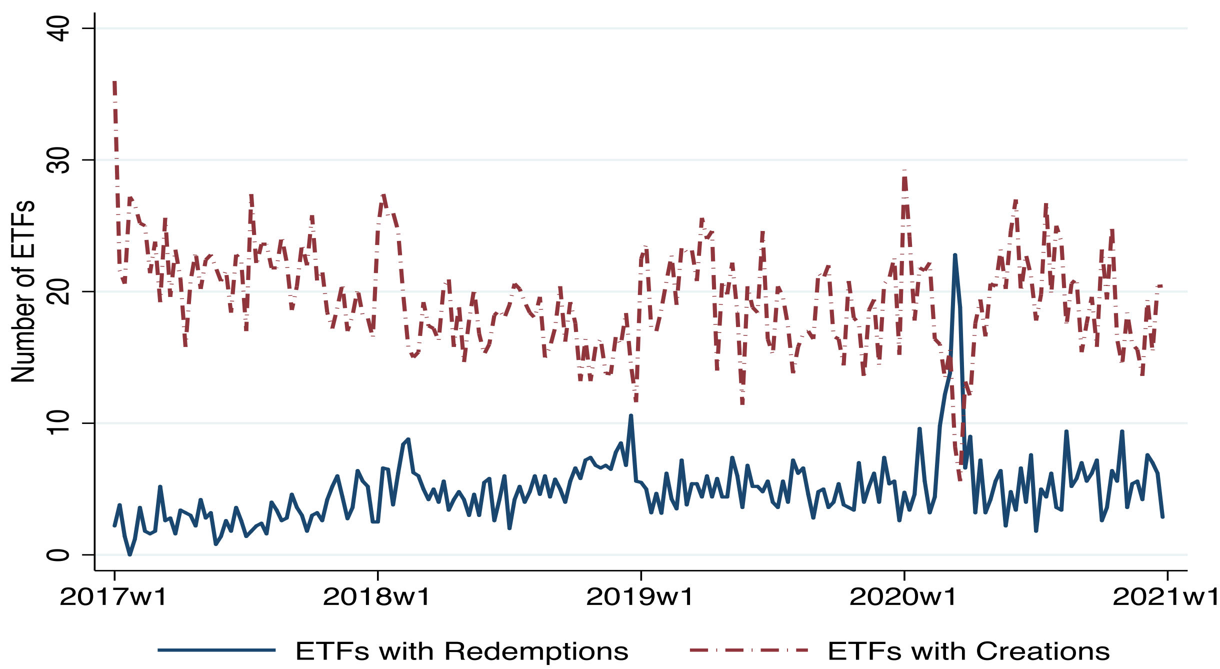


Figure 4: Creations and Redemptions over Time

Panel A plots the number of corporate bond ETFs that create or redeem shares on a given day, averaged weekly. Panel B plots the weekly volume of net redemptions (i.e., redemptions minus creations) by corporate bond ETFs, in billions of dollars.

(a) Number of Funds that Redeem and Create (Weekly Average)



(b) Net Redemption Volume (Weekly Sum)

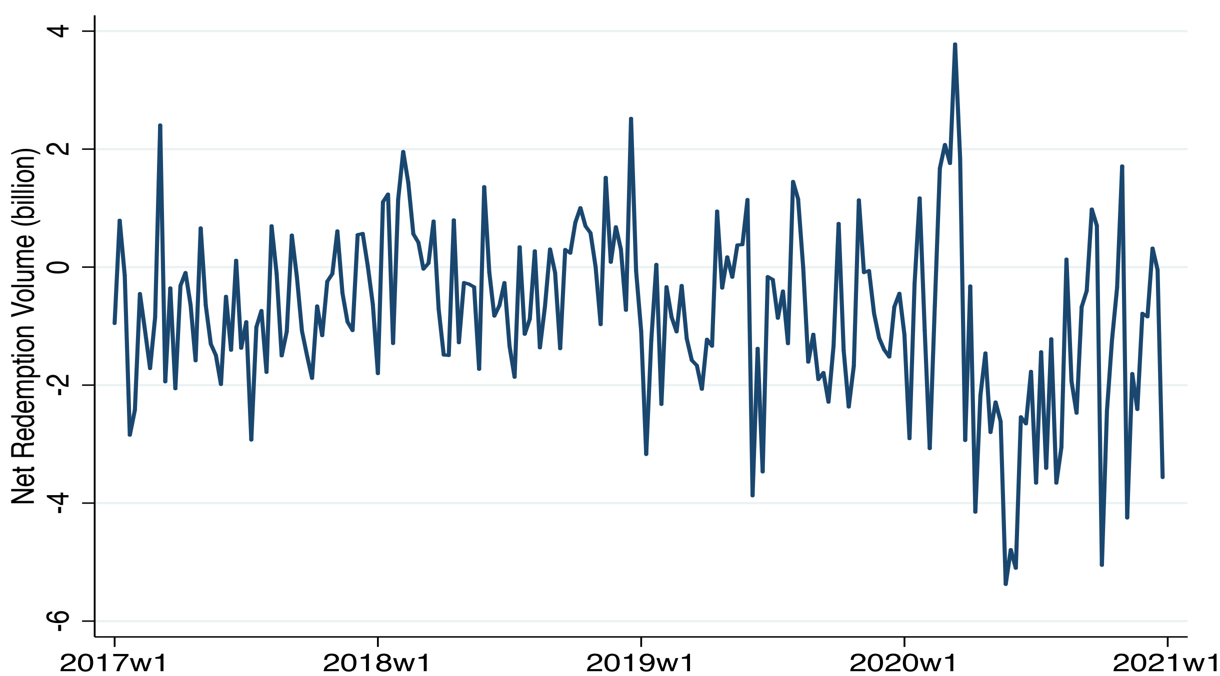


Table 1: Cash in ETF Baskets and Holdings

This table shows the cross-sectional distributions of cash ratios in ETF baskets and portfolios. Cash ratios are computed by dividing the amount of cash in the basket (portfolio) by total basket (portfolio) value on the same day. We first calculate the time-series average of a cash ratio at the ETF level and then report two averages and the 10th, 25th, 50th, 75th, and 90th percentiles of its cross-sectional distribution. The 'unweighted' average is equal-weighted; the 'weighted' average is weighted by assets under management. We report the distributions for realized RD baskets, realized CR baskets, reported baskets, and ETF portfolio holdings. The distributions of reported cash ratios are computed across three sets of days: days with CR activity, days with RD activity, and all days, which include days with CR or RD activity as well as no activity.

	Average		Distribution				
	Unweighted	Weighted	p10	p25	p50	p75	p90
Realized creation baskets	11.60	7.27	-0.44	0.28	6.25	18.96	34.26
Realized redemption baskets	8.18	3.68	-2.05	-0.00	0.73	6.43	35.19
Reported baskets (All days)	5.39	2.03	0.19	0.54	1.09	2.27	9.28
Reported baskets (CR days)	4.58	1.89	0.11	0.34	0.97	2.23	5.94
Reported baskets (RD days)	7.78	2.45	0.01	0.24	0.80	2.69	10.89
Portfolio holdings	1.70	0.85	0.00	0.44	0.83	1.57	2.47

Table 2: Concentration of ETF Baskets and Holdings

This table shows the cross-sectional distributions of concentration ratios for ETF baskets and portfolios. The concentration ratios for ETF baskets are computed by dividing the number of bonds in the basket by the number of bonds in the underlying index on the same day. The concentration ratios for ETF portfolios are computed by dividing the number of bonds in the ETF portfolio by the number of bonds in the index. We first calculate the time-series average of the given ratio at the ETF level and then report two averages and the 10th, 25th, 50th, 75th, and 90th percentiles of its cross-sectional distribution. The 'unweighted' average is equal-weighted; the 'weighted' average is weighted by assets under management. We report the distributions for realized RD baskets, realized CR baskets, reported baskets, and ETF portfolio holdings. The distributions of reported basket ratios are computed across three sets of days: days with CR activity, days with RD activity, and all days, which include days with CR or RD activity as well as no activity.

	Average		Distribution				
	Unweighted	Weighted	p10	p25	p50	p75	p90
Realized creation baskets	24.56	18.13	4.22	10.21	19.44	37.44	52.75
Realized redemption baskets	29.80	18.00	3.77	9.64	22.01	38.72	87.12
Reported baskets (All days)	76.06	57.51	44.65	64.28	84.09	94.88	97.35
Reported baskets (CR days)	77.56	59.53	45.15	63.98	85.42	95.13	97.56
Reported baskets (RD days)	78.40	56.22	45.73	69.80	86.75	96.72	100.00
Portfolio holdings	81.28	91.39	45.64	64.76	88.03	96.70	102.39

Table 3: Active Cash Management

This table reports the slope estimates from the regressions of the proportion of cash in a CR or RD basket on one-day-lagged values of $\Delta Cash$ and their interactions with the illiquidity of the ETF's index. $\Delta Cash$ is the difference between the fraction of cash in the ETF's portfolio and the average of those fractions over the prior month. The three measures of index illiquidity are the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). Observations are at the ETF-day level. An ETF fixed effect is included in all specifications. Standard errors are reported in parentheses. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
ΔCash	3.07***	3.16***	3.16***	-3.38***	-3.57***	-3.33***
	(0.74)	(0.74)	(0.74)	(0.46)	(0.47)	(0.46)
ΔCash × IL	-1.42**	-1.18*	-1.12**	1.66***	1.58***	1.28***
	(0.63)	(0.61)	(0.55)	(0.44)	(0.41)	(0.43)
Observations	2,272	2,272	2,272	5,108	5,108	5,108
Adjusted R^2	0.13	0.13	0.13	0.09	0.09	0.09

Table 4: Active Basket Management

This table reports the slope estimates from the regressions of a bond's basket inclusion on $\Delta Deviation$ and its interaction with the illiquidity of the ETF's index. Basket inclusion is measured either by a dummy variable, which is equal to one if the bond is included in the basket and zero otherwise (Panel A), or by the log of one plus the number of bond shares in the basket (Panel B). $\Delta Deviation$ for a given bond held by a given ETF is an index-rebalancing-induced shock to the bond's overweighting in the ETF's portfolio relative to the index. The three measures of index illiquidity are the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). Observations are at the ETF-bond-day level. An ETF-day fixed effect is included in all specifications. We multiply both dependent variables by 100 to improve readability of the coefficients. Standard errors, which are clustered at the bond and day level, are reported in parentheses. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% level, respectively.

(a) Basket Inclusion Dummy

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
ΔDeviation	1.98***	1.74***	1.99***	-2.96***	-2.88***	-2.97***
	(0.15)	(0.14)	(0.15)	(0.09)	(0.09)	(0.09)
ΔDeviation × IL	-0.59***	-0.32***	-0.57***	0.58***	0.54***	0.53***
	(0.11)	(0.11)	(0.10)	(0.06)	(0.06)	(0.06)
Observations	2726592	2726592	2726592	7803001	7803001	7803001
Adjusted R^2	0.41	0.41	0.41	0.30	0.30	0.30

(b) Basket Shares

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
ΔDeviation	23.70***	20.94***	23.77***	-34.13***	-33.18***	-34.16***
	(1.84)	(1.69)	(1.83)	(1.08)	(1.06)	(1.08)
ΔDeviation × IL	-6.94***	-3.88***	-6.54***	6.84***	6.51***	6.14***
	(1.33)	(1.34)	(1.29)	(0.76)	(0.75)	(0.77)
Amount Outstanding	60.08***	60.19***	60.08***	54.06***	54.02***	54.05***
	(3.08)	(3.08)	(3.08)	(1.53)	(1.53)	(1.53)
Observations	2639566	2639566	2639566	7579237	7579237	7579237
Adjusted R^2	0.45	0.45	0.45	0.35	0.35	0.35

Table 5: The Effect of Basket Inclusion on Bond Liquidity

This table reports the slope estimates from the regressions of an individual bond's illiquidity on the bond's basket inclusion on the previous day. There are two measures of basket inclusion: the number of ETFs that include this bond in their baskets (columns 1 to 3) and the log of one plus the number of the bond's shares that are included in ETFs' baskets (columns 4 to 6). There are three measures of bond illiquidity: the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). We standardize illiquidity by its mean and standard deviation and multiply it by 100 to improve readability of the coefficients. We control for the average bond illiquidity over the prior week, the average basket size of ETFs holding the bond, and, in the last three specifications, the log number of the bond's outstanding shares. Observations are at the bond-day level. All specifications include firm-day and maturity-day fixed effects. Standard errors, which are clustered at the bond and day level, are reported in parentheses. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-4.31***	-1.84***	-2.31***	-0.18***	-0.11***	-0.10***
	(0.26)	(0.27)	(0.25)	(0.02)	(0.03)	(0.02)
CR	-3.68***	-1.02***	-1.96***	-0.18***	-0.05**	-0.12***
	(0.14)	(0.14)	(0.14)	(0.02)	(0.02)	(0.02)
Bond IL	12.77***	11.81***	19.71***	11.42***	11.64***	19.21***
	(0.19)	(0.19)	(0.36)	(0.18)	(0.19)	(0.35)
Avg Basket Size	-0.84***	-1.42***	-2.17***	-1.38***	-1.56***	-2.47***
	(0.27)	(0.28)	(0.27)	(0.24)	(0.27)	(0.25)
Amount Outstanding				-12.37***	-4.60***	-6.14***
				(0.30)	(0.34)	(0.32)
Observations	3254055	2831031	2901286	3254055	2831031	2901286
Adjusted R^2	0.23	0.13	0.44	0.23	0.13	0.44

Table 6: First Stage: The Effect of Index Rebalancing on Basket Inclusion

This table reports the slope estimates from the first-stage regression of basket inclusion on our instruments, $CRInstr_{it}$ ("CR Instrument") and $RDInstr_{it}$ ("RD Instrument"). Basket inclusion is measured either by the number of ETFs that include the bond in their baskets (Panel A) or by the log of one plus the number of bond shares in the baskets (Panel B). We control for the average bond illiquidity over the prior week, the average basket size of ETFs holding the bond, and, in Panel B, the log amount of the bond outstanding. The three measures of bond illiquidity are the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). Observations are at the bond-day level. All specifications include firm-day and maturity-day fixed effects. We multiply both dependent variables by 100 to improve readability of the coefficients. Standard errors, which are clustered at the bond and day level, are reported in parentheses. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% level, respectively.

(a) Number of Baskets

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	1.15***	1.17***	1.15***	-0.08	-0.08	-0.08
	(0.10)	(0.11)	(0.11)	(0.05)	(0.05)	(0.05)
CR Instrument	-0.46***	-0.46***	-0.46***	-2.50***	-2.51***	-2.51***
	(0.05)	(0.05)	(0.05)	(0.13)	(0.13)	(0.13)
Bond IL	-0.90***	-0.50***	-0.66***	-1.87***	-0.75***	-1.48***
	(0.06)	(0.04)	(0.08)	(0.10)	(0.06)	(0.14)
Avg Basket Size	8.26***	8.22***	8.32***	10.34***	10.28***	10.43***
	(0.72)	(0.72)	(0.73)	(0.81)	(0.81)	(0.81)
Observations	3391931	3304918	3230624	3391931	3304918	3230624
Adjusted R^2	0.47	0.47	0.48	0.39	0.39	0.39

(b) Basket Shares

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	10.93***	11.06***	10.85***	-1.03**	-1.06**	-1.03**
	(1.10)	(1.10)	(1.10)	(0.41)	(0.41)	(0.41)
CR Instrument	-6.24***	-6.17***	-6.25***	-19.05***	-18.99***	-19.01***
	(0.57)	(0.56)	(0.57)	(0.89)	(0.89)	(0.89)
Bond IL	-4.01***	-3.08***	-3.15***	-4.76***	-2.71***	-4.57***
	(0.43)	(0.38)	(0.56)	(0.53)	(0.40)	(0.86)
Avg Basket Size	65.51***	65.24***	66.11***	73.76***	73.11***	74.21***
	(6.32)	(6.33)	(6.38)	(5.97)	(5.96)	(6.03)
Amount Outstanding	88.03***	88.57***	88.60***	138.98***	140.29***	139.39***
	(3.58)	(3.59)	(3.58)	(3.26)	(3.27)	(3.26)
Observations	3391931	3304918	3230624	3391931	3304918	3230624
Adjusted R^2	0.41	0.41	0.41	0.36	0.36	0.36

Table 7: Second Stage: The Effect of Instrumented Basket Inclusion on Bond Liquidity

This table reports the slope estimates from the second-stage regression of next-day bond illiquidity on instrumented basket inclusion. There are two measures of basket inclusion: the number of ETFs that include the bond in their baskets (columns 1 to 3) and the log of one plus the number of bond shares included in ETF baskets (columns 4 to 6). There are three measures of bond illiquidity: the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). We standardize illiquidity by its mean and standard deviation and multiply it by 100 to improve readability of the coefficients. We control for the average bond illiquidity over the prior week, the average basket size of ETFs holding the bond, and, in the last three columns, the log number of the bond's outstanding shares. Observations are at the bond-day level. All specifications include firm-day and maturity-day fixed effects. Standard errors, which are clustered at the bond and day level, are reported in parentheses. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-15.25***	-15.08***	-7.30*	-1.07**	-1.41**	-0.55
	(4.93)	(5.50)	(4.18)	(0.52)	(0.58)	(0.45)
CR	-3.80*	-5.38**	-8.58***	-1.32***	-0.92**	-1.55***
	(2.26)	(2.58)	(1.90)	(0.33)	(0.37)	(0.28)
Bond IL	11.78***	11.16***	18.06***	10.46***	11.03***	17.70***
	(0.12)	(0.09)	(0.11)	(0.10)	(0.09)	(0.10)
Avg Basket Size	0.54	1.33**	-0.16	0.52	1.12**	-0.19
	(0.47)	(0.54)	(0.41)	(0.40)	(0.45)	(0.34)
Amount Outstanding				-10.20***	-2.34***	-3.43***
				(0.46)	(0.55)	(0.40)
Observations	2020546	1753639	1803581	2020546	1753639	1803581
Adjusted R^2	0.01	0.01	0.02	0.01	0.01	0.02

Table 8: The Effect of Basket Inclusion on Bond Liquidity during COVID-19

This table is the counterpart of Table 5 estimated over the COVID subperiod. Instead of using the full sample of 2017 to 2020, this table uses the subperiod of March 2 to April 15, 2020.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	0.64	3.98***	4.03***	0.20*	0.35**	0.39*
	(0.96)	(1.13)	(1.43)	(0.11)	(0.13)	(0.20)
CR	-1.45	-0.55	2.69	0.10	-0.12	0.31*
	(1.29)	(0.86)	(1.82)	(0.13)	(0.10)	(0.18)
Bond IL	9.52***	8.04***	15.39***	9.50***	7.96***	15.07***
	(0.62)	(0.54)	(0.74)	(0.62)	(0.54)	(0.71)
Avg Basket Size	0.22	-1.34	-0.39	-0.35	-1.12	-0.41
	(1.19)	(0.95)	(2.09)	(1.16)	(0.97)	(1.99)
Amount Outstanding				-1.83	4.27***	10.02***
				(1.70)	(1.20)	(2.87)
Observations	111707	100263	101341	111707	100263	101341
Adjusted R^2	0.26	0.07	0.36	0.26	0.07	0.36

Table 9: Interactions with Basket Imbalance

This table reports the slope estimates from the regressions of an individual bond's next-day illiquidity on the number of baskets that contain the bond in the prior week and its interactions with two measures of the bond's basket imbalance. These imbalance measures are defined in equations (17) and (18). There are three measures of bond illiquidity: the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). We standardize illiquidity by its mean and standard deviation and multiply it by 100 to improve readability of the coefficients. We control for the average bond illiquidity over the prior week and the average basket size of ETFs holding the bond. Observations are at the bond-day level. All specifications include firm-day and maturity-day fixed effects. Standard errors, which are clustered at the bond and day level, are reported in parentheses. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	(1) IL1	(2) IL2	(3) IL3
Num Baskets	-10.98***	-4.27***	-6.65***
	(0.40)	(0.40)	(0.40)
Num Baskets $\times$ CR Imbal	1.89***	1.90***	1.70***
	(0.24)	(0.23)	(0.21)
Num Baskets $\times$ RD Imbal	2.98***	1.48**	2.64***
	(0.51)	(0.70)	(0.61)
Bond IL	12.52***	11.78***	19.59***
	(0.19)	(0.19)	(0.35)
Avg Basket Size	-0.45*	-1.28***	-1.91***
	(0.26)	(0.27)	(0.26)
Observations	3254055	2831031	2901286
Adjusted R^2	0.23	0.13	0.44

References

AGARWAL, VIKAS, PAUL HANOUNA, RABIH MOUSSAWI and CHRISTOF W. STAHEL (2018). "Do ETFs increase the commonality in liquidity of underlying stocks?" Working Paper.

AKEY, PAT, ADRIANA Z. ROBERTSON and MIKHAIL SIMUTIN (2021). "Closet active management of passive funds." Working Paper.

BEN-DAVID, ITZHAK, FRANCESCO FRANZONI and RABIH MOUSSAWI (2017). "Exchange-traded funds." *Annual Review of Financial Economics* 9: 169–189.

BEN-DAVID, ITZHAK, FRANCESCO FRANZONI and RABIH MOUSSAWI (2018). "Do ETFs increase volatility?" *Journal of Finance* 73: 2471–2535.

BEN-DAVID, ITZHAK, FRANCESCO FRANZONI, BYUNGWOOK KIM and RABIH MOUSSAWI (2021). "Competition for attention in the ETF space" Working Paper.

BROGAARD, JONATHAN, DAVIDSON HEATH and DA HUANG (2021). "The heterogeneous effects of passive investing on asset markets" Working Paper.

CHEN, QI, ITAY GOLDSTEIN and WEI JIANG (2010). "Payoff complementarities and financial fragility: Evidence from mutual fund outflows." *Journal of Financial Economics* 97: 239–262.

CHENG, SI, MASSIMO MASSA and HONG ZHANG (2019). "The unexpected activeness of passive investors: a worldwide analysis of ETFs" *Review of Asset Pricing Studies* 9: 296–355.

CHERNENKO, SERGEY and ADI SUNDERAM (2017). "Liquidity transformation in asset management: Evidence from the cash holdings of mutual funds." Working Paper.

CHOI, JAEWON, SAEID HOSEINZADE, SEAN SEUNGHUN SHIN and HASSAN TEHRANIAN (2020). "Corporate bond mutual funds and asset fire sales." *Journal of Financial Economics* 138: 432–457.

CRAGG, JOHN G. and STEPHEN G. DONALD (1993). "Testing Identifiability and Specification in Instrumental Variable Models." *Econometric Theory* 9, 222-240.

CREMERS, MARTIJN and ANTTI PETAJISTO (2009). "How active is your fund manager? A new measure that predicts performance." *Review of Financial Studies* 22: 3329–3365.

DA, ZHI and SOPHIE SHIVE (2018). "Exchange traded funds and asset return correlations." *European Financial Management* 24: 136–168.

DANNHAUSER, CAITLIN (2017). "The impact of innovation: Evidence from corporate bond ETFs." *Journal of Financial Economics* 125: 182–199.

DANNHAUSER, CAITLIN and SAEID HOSEINZADE (2022). "The unintended consequences of corporate bond ETFs: Evidence from the taper tantrum." *Review of Financial Studies* 35: 51–90.

DIAMOND, DOUGLAS and PHILIP DYBVIG (1983). "Bank runs, deposit insurance, and liquidity." *Journal of Political Economy* 91: 401–419.

EASLEY, DAVID, DAVID MICHAYLUK, MAUREEN O'HARA and TALIS PUTNINS (2021). "The active world of passive investing." *Review of Finance*, forthcoming.

Falato, Antonio, Itay Goldstein and Ali Hortacsu (2020). "Financial fragility in the COVID-19 crisis: The case of investment funds in corporate bond markets." *Journal of Monetary Economics*, forthcoming.

Feldhutter, Peter (2012). "The same bond at different prices: Identifying search frictions and selling pressures." *Review of Financial Studies* 25: 1155–1206.

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

Goyenko, Ruslan Y, Craig W. Holden and Charles A. Trzcinka (2009). "Do liquidity measures measure liquidity." *Journal of Financial Economics* 92: 153–181.

Haddad, Valentin, Alan Moirera and Tyler Muir (2020). "When selling becomes viral: Disruptions in debt markets in the Covid-19 crisis and the Fed's response." *Review of Financial Studies*, forthcoming.

He, Zhiguo, Stefan Nagel and Zhaogang Song (2021). "Treasury inconvenience yields during the COVID-19 crisis." *Journal of Financial Economics*, forthcoming.

Holden, Craig W. (2009). "New low-frequency spread measures." *Journal of Financial Markets* 12: 778–813.

Holden, Craig W. and Jayoung Nam (2019). "Market accessibility, corporate bond ETFs, and liquidity." Working Paper.

Israeli, Doron, Charles Lee and Suhas Sridharan (2017). "Is there a dark side to exchange traded funds? An information perspective." *Review of Accounting Studies* 22: 1048–1083.

Jin, Dunhong, Marcin Kacperczyk, Bige Kahraman and Felix Suntheim (2022). "Swing pricing and fragility in open-end mutual funds." *Review of Financial Studies* 35: 1–50.

Kacperczyk, Marcin, Clemens Sialm and Lu Zheng (2008). "Unobserved actions of mutual funds." *Review of Financial Studies* 21: 2379–2416.

Kacperczyk, Marcin and Philipp Schnabl (2013). "How safe are money market funds?" *Quarterly Journal of Economics* 128: 1073–1122.

Kargar, Mahyar, Benjamin Lester, David Lindsay, Shuo Liu, Pierre-Olivier Weill, and Diego Zuniga (2021). "Corporate bond liquidity during the COVID-19 crisis." *Review of Financial Studies* 34: 5352–5401.

Lettau, Martin and Ananth Madhavan (2018). "Exchange-traded funds 101 for economists." *Journal of Economic Perspectives* 32: 135–154.

Li, Sida (2021). "Should passive investors actively manage their trades?" Working Paper.

Ma, Yiming, Kairong Xiao and Yao Zeng (2019). "Bank debt versus mutual fund equity in liquidity provision." Working Paper.

Marta, Thomas (2019). "Fixed income ETFs, bond liquidity, and stressed markets." Working Paper.

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

Pan, Kevin and Yao Zeng (2019). "ETF Arbitrage under Liquidity Mismatch." Working Paper.

Pastor, Lubos, Robert F. Stambaugh and Lucian A. Taylor (2020). "Fund tradeoffs." *Journal of Financial Economics* 138: 614–634.

Petajisto, Antti (2017). "Inefficiencies in the pricing of exchange-traded funds." *Financial Analysts Journal* 73: 24–54.

Pu, Xiaoling (2009). "Liquidity commonality across the bond and CDS markets" *Journal of Fixed Income* 19: 26–39.

Reilly, Christopher (2021). "The hidden cost of corporate bond ETFs." Working Paper.

Saglam, Mehmet, Tugkan Tuzun and Russ Wermers (2019). "Do ETFs increase liquidity?" Working Paper.

Shim, John J. and Karamfil Todorov (2021). "ETFs, illiquid assets, and fire sales." Working Paper.

Stock, James H. and Motohiro Yogo (2005). "Testing for Weak Instruments in Linear IV Regression" *Identification and Inference for Econometric Models: Essays in Honor of Thomas J. Rothenberg.* Cambridge University Press.

Song, Han and Hao Zhou (2007). "Nondefault bond spread and market trading liquidity." Working Paper.

Sunderam, Adi (2015). "Money creation and the shadow banking system." *Review of Financial Studies* 28: 939–977.

Appendix

Contents

A.1 Measures of Bond Liquidity

We measure bond liquidity in three ways. All three proxies take larger values when liquidity is lower, so they should be interpreted as proxies for illiquidity rather than liquidity. We calculate these measures for each bond on a daily basis by using TRACE data.

First, we use the Effective Tick Size (*Tick*), which infers the effective bid-ask spread from the clustering of trade prices on round price increments (Holden, 2009, Goyenko et al., 2009). To illustrate, Goyenko et al. (2009) explain that "we assume that price clustering is completely determined by spread size. For example, if the spread is $\$\frac{1}{4}$, the model assumes that the bid and ask prices employ only even quarters. The quote could be $\$25\frac{1}{4}$ bid, $\$25\frac{1}{2}$ offered, but never $\$25\frac{3}{8}$ bid, $\$25\frac{5}{8}$ offered. Thus, if odd-eighth transaction prices are observed, one infers that the spread must be $\$\frac{1}{8}$. This implies that the simple frequency with which closing prices occur in particular price clusters can be used to estimate the spread probabilities."

Second, we use the Imputed Roundtrip Cost (*IRC*), a standard proxy for bid-ask spread in fixed-income markets. Developed by Feldhutter (2012), this measure is based on the differences between the highest and lowest prices for a given bond that are likely part of the same round-trip trade. Round-trip trades are imputed for a given bond on a given day if there are two or three trades of the same volume within 15 minutes. The highest trade price is assumed to be an investor buying from a dealer, the lowest price an investor selling to a dealer, and the investor round-trip cost to be the highest minus the lowest price. A bond's IRC on a given day is the average round-trip cost for that bond on that day.

Finally, we use the Inter-Quartile Range (*IQR*), which is the inter-quartile range of a bond's prices on a given day. For each bond on each day, we first compute three metrics from the distribution of the bond's trade prices on that day: the average price, $\bar{p}$, and the 25th and 75th percentiles, p_{25} and p_{75}. The *IQR* is then given by $100 \times (p_{75} - p_{25})/\bar{p}$. Developed by Song and Zhou (2007) and Pu (2009), *IQR* reduces the effect of outliers on measured illiquidity compared to other measures. Nevertheless, it may also underestimate illiquidity by excluding the tails of the bond price distribution.

A.2 Proofs

We prove all of our theoretical results in a general setting, in which the ETF's endowment portfolio can include not only the index portfolio but also some amount of cash. Specifically, let α_0 denote the fraction of the ETF's initial portfolio that is invested in cash, so that the remaining fraction, $1 - \alpha_0$, is invested in the equal-weighted index portfolio of the N securities. Our results and their proofs hold for any $0 \leq \alpha_0 \leq 1$.

Proof of Proposition 1. Consider the AP's profit maximization problem (2). Suppose the AP's participation constraint is not binding. The first-order condition is given by

$$\alpha + (1-\alpha)\mu - p_E - \phi(1-\alpha)^2 x = 0\,. \tag{A1}$$

Solving for x, we immediately obtain equation (3). Combining equation (3) with the market clearing condition $x = \pi_c - \pi_s$ immediately yields the equilibrium ETF price in equation (4).

At the market-clearing ETF price p_E, the AP's participation constraint is given by

$$\Gamma = \frac{1}{2}\phi(1-\alpha)^2(\pi_c - \pi_s)^2 - \lambda I \geq 0\,,$$

which gives the cutoff basket count at which the AP breaks even:

$$\bar{I} = \frac{\phi(1-\alpha)^2(\pi_c - \pi_s)^2}{2\lambda}\,, \tag{A2}$$

concluding the proof. □

Example. We now present an example to illustrate the evolution of the ETF portfolio with a given ETF basket. Suppose $N = 2$, $\alpha_0 = 0.2$, $I = 1$, $\alpha = 0.6$, and $\pi_c - \pi_s = -0.5$. The following table summarizes the evolution of the ETF portfolio over time.

Date		ETF shares	Cash	Security 1	Security 2
$t = 0$	Initial ETF portfolio	1	0.2	0.4	0.4
$t = 1$	Portfolio delivered for creations	0.5	0.3	0.2	0
$t = 2$	Resulting ETF portfolio	1.5	0.5	0.6	0.4

In words, row 1 shows that at $t = 0$, the ETF holds a portfolio whose weight on cash is 0.2 and whose weights on both of the unique security names, security 1 and security 2, are 0.4. The ETF issues a total of 1 ETF share at $t = 0$.

Row 2 shows that at $t = 1$, there is a net demand for 0.5 ETF shares, so the AP creates 0.5 ETF shares by delivering a portfolio to the ETF according to the creation basket. The creation basket contains cash and security 1 only, with a portfolio weight of 0.6 on cash and 0.4 on security 1. This implies that the AP delivers 0.3 units of cash and 0.2 units of security 1 to the ETF, in exchange for 0.5 ETF shares to be created.

Row 3 describes the situation after the in-kind creation. The ETF portfolio, which corresponds to a total of $1.5 = 1 + 0.5$ ETF shares, includes $0.5 = 0.2 + 0.3$ units of of cash, $0.6 = 0.2 + 0.4$ units of of security 1, and $0.4 = 0.4 + 0$ units of security 2. Scaling the portfolio weights to add up to one, these weights are given by a 3×1 vector:

$$w_E = \left(\frac{1}{3}, \ \frac{2}{5}, \ \frac{4}{15}\right) .$$

The three weights correspond to cash, security 1, and security 2, respectively.

Proof of statements following equation (5). Denote $\delta_\pi \equiv \pi_c - \pi_s$. Generalizing the above example, the date-2 ETF portfolio weights are given by the $(N+1) \times 1$ vector

$$w_E = \left(\underbrace{\frac{\alpha_0 - \delta_\pi \alpha}{1-\delta_\pi}}_{\text{cash}} ; \underbrace{\frac{(1-\alpha_0)I - \delta_\pi(1-\alpha)N}{(1-\delta_\pi)IN}, \ldots, \frac{(1-\alpha_0)I - \delta_\pi(1-\alpha)N}{(1-\delta_\pi)IN}}_{I \text{ basket securities}} ; \underbrace{\frac{1-\alpha_0}{(1-\delta_\pi)N}, \ldots, \frac{1-\alpha_0}{(1-\delta_\pi)N}}_{N-I \text{ non-basket securities}} \right) . \tag{A3}$$

Thus, the expected value and variance of the date-2 ETF portfolio is given by

$$\mu_E = \frac{\alpha_0 - \delta_\pi \alpha}{1-\delta_\pi} + \frac{1-\alpha_0-\delta_\pi(1-\alpha)}{1-\delta_\pi}\,\mu\,, \tag{A4}$$

$$\sigma_E^2 = \frac{(1-\alpha_0)(1-2\delta_\pi(1-\alpha)-\alpha_0)I + \delta_\pi^2(1-\alpha)^2 N}{(1-\delta_\pi)^2 IN}\,\sigma^2\,. \tag{A5}$$

Straightforward calculation shows that

$$\frac{\partial \mu_E}{\partial \alpha} = \frac{\delta_\pi(\mu - 1)}{1-\delta_\pi}\,,$$

which is negative (positive) when $\delta_\pi < 0$ ($\delta_\pi > 0$), that is, when creations (redemptions) happen. In other words, more cash in the basket reduces the ETF portfolio's expected time-2 payoff after creations, but increases it after redemptions. We also have

$$\frac{\partial \sigma_E^2}{\partial I} = -\frac{(1-\alpha)^2\delta_\pi^2\sigma^2}{(1-\delta_\pi)^2 I^2} < 0\,. \tag{A6}$$

Consider the ETF's ex-ante welfare maximization problem (1) and re-arrange terms:

$$\begin{aligned} \mathrm{E}[W] &= \pi_c u(p_E) + \pi_s E[u(2v_E - p_E)] + (1-\pi_c-\pi_s)E[u(v_E)] \\ &= \delta_\pi p_E + V(\mu_E, \sigma_E^2)\,, \end{aligned} \tag{A7}$$

where

$$V(\mu_E, \sigma_E^2) = (1-\pi_c-\pi_s)\left(\mu_E - \frac{\rho\sigma_E^2}{2}\right) + \pi_s\left(2\mu_E - 2\rho\sigma_E^2\right)\,, \tag{A8}$$

which gives the closed-form expressions for the re-expressed objective function (5), and μ_E and σ_E^2 are given by (A4) and (A5), respectively. It is straightforward that $V(\mu_E, \sigma_E^2)$ increases with μ_E and decreases with σ_E^2. Note that p_E depends on transaction costs ϕ and λ, whereas $V(\mu_E, \sigma_E^2)$ does not.

Proof of Proposition 2. First, we consider the optimal choice of α. The ETF chooses α to maximize $\mathrm{E}[W]$ in equation (A7), where $V(\mu_E, \sigma_E^2)$ comes from equation (A8):

$$\begin{aligned} \mathrm{E}[W] &= \delta_\pi p_E + V(\mu_E, \sigma_E^2) \\ &= \delta_\pi p_E + (1-\pi_c-\pi_s)\left(\mu_E - \frac{\rho\sigma_E^2}{2}\right) + \pi_s\left(2\mu_E - 2\rho\sigma_E^2\right). \end{aligned} \tag{A9}$$

We substitute for p_E, μ_E, and σ_E^2 from equations (4), (A4), and (A5), respectively. All of them are functions of α. We then differentiate $\mathrm{E}[W]$ in equation (A9) with respect to α, obtaining the first-order condition

$$\frac{\delta_\pi\left[(1-\pi_c+3\pi_s)((1-\alpha)\delta_\pi N-(1-\alpha_0)I)\rho\sigma^2+2(1-\alpha)\delta_\pi(1-\delta_\pi)^2\phi IN\right]}{(1-\delta_\pi)^2 IN}=0\,, \tag{A10}$$

from which we obtain a closed-form solution,

$$\alpha^* = 1 - \frac{(1-\pi_c+3\pi_s)(1-\alpha_0)I\rho\sigma^2}{(1-\pi_c+3\pi_s)\delta_\pi N\rho\sigma^2 + 2\delta_\pi(1-\delta_\pi)^2\phi IN}\,. \tag{A11}$$

This is a valid solution as long as $\delta_\pi > 0$. If $\delta_\pi \leq 0$ then we obtain $\alpha^* = 1$, given the restriction $0 \leq \alpha \leq 1$. From equation (A11), we immediately see that

$$\lim_{\phi\to+\infty} \alpha^* = 1\,.$$

When the illiquidity parameter ϕ approaches infinity, α^* approaches one. Therefore, $\alpha^* > 0$ when ϕ is sufficiently large, as stated in part (i) of the proposition.

To prove part (iii), we differentiate α^* in equation (A11) with respect to ϕ:

$$\frac{\mathrm{d}\alpha^*}{\mathrm{d}\phi} = \frac{2(1-\pi_c+3\pi_s)(1-\alpha_0)(1-\delta_\pi)^2 I^2\rho\sigma^2}{\delta_\pi\left(2(1-\delta_\pi)^2 I\phi + (1-\pi_c+3\pi_s)\rho\sigma^2\right)^2 N} > 0\,,$$

as long as $\delta_\pi > 0$. If $\delta_\pi \leq 0$ then α^* does not depend on ϕ. Therefore, α^* is weakly increasing in ϕ. Another way to see the result is to use the implicit function theorem. Denoting the left-hand side of equation (A10) by W_α, we obtain

$$\begin{aligned} \frac{\mathrm{d}\alpha^*}{\mathrm{d}\phi} &= -\frac{\partial W_\alpha}{\partial\phi}\left(\frac{\partial W_\alpha}{\partial\alpha}\right)^{-1} \\ &= \frac{2(1-\delta_\pi)^2(1-\alpha)I}{(1-\pi_c+3\pi_s)\rho\sigma^2 + 2(1-\delta_\pi)^2\phi I} \geq 0\,, \end{aligned}$$

when $\alpha \leq 1$.

Next, we consider the optimal basket count. From Proposition 1, the AP's participation constraint is not binding if and only if $I \leq \bar{I}$, where $\bar{I}$ is in equation (A2). When $I \leq \bar{I}$, equation (4) implies that p_E does not depend on I. Therefore, based on equation (A7), $\mathrm{E}[W]$ depends on I only through $V(\mu_E, \sigma_E^2)$, which is increasing in I, according to equation (A6). As a result, the ETF chooses an I as large as possible subject to the AP's participation constraint $I \leq \bar{I}$ and the feasibility constraint $I \leq N$, implying that

$$I^* = \min\{\bar{I}, N\}\,.$$

Note that I^* decreases in λ because $\bar{I}$ decreases in λ. In addition,

$$\lim_{\lambda \to +\infty} I^* = \lim_{\lambda \to +\infty} \bar{I} = 0\,.$$

Therefore, $I^* < N$ when λ is sufficiently large, as stated in part (ii) of the proposition.

Finally, we turn to part (iv) of the proposition. Let $w_{E,i}$ and $w_{B,i}$ denote the portfolio weights of security i in the ETF portfolio and the benchmark index portfolio, respectively. Since the returns of the N securities are i.i.d. with volatility σ^2, the tracking error of the ETF's security portfolio simplifies to

$$\Delta = \mathrm{Var}\left(\sum_{i=1}^{N}(w_{E,i} - w_{B,i})\tilde{r}_i\right) = \sigma^2 \sum_{i}^{N} (w_{E,i} - w_{B,i})^2 \;,$$

where $\tilde{r}_i$ is the return on security i. Recognizing that $w_{B,i} = 1/N$ and substituting for $w_{E,i}$ from equation (A3), we obtain

$$\frac{\partial \Delta}{\partial I} = -\frac{(1-\alpha)^2 \delta_\pi^2 \sigma^2}{(1-\delta_\pi)^2 I^2} < 0\,.$$

Because I^* decreases in λ and Δ depends on λ only through I, we have that Δ is increasing in λ, concluding the proof of part (iv) of the proposition. □

A.3 Model Extension: CARA Utility

In our baseline model, we use a mean-variance utility function to obtain analytical solutions. In this section, we replace mean-variance utility with CARA utility, $u(c) = -\exp(-\rho c)$, where c represents the agent's total lifetime consumption. We maintain all other assumptions of the baseline model, setting $\alpha_0 = 0$ to simplify the exposition. We solve the model numerically and show that its main results continue to hold.

The results regarding the optimal basket count hold automatically because, according to the proof of Proposition 2, the ETF's optimal basket count choice depends on the AP's participation constraint but not on the agents' utility function.

The results regarding the optimal basket cash share also continue to hold under CARA utility, as we show numerically. After solving the model for many combinations of plausible parameter values, we find that the equilibrium basket cash share is always strictly between 0 and 1, and that this share is increasing with security illiquidity. In addition, we obtain the same results in the special case of $\sigma = 0$, in which the risk-based motive for holding cash is absent and ETF cash holdings are driven solely by liquidity transformation.

Under CARA utility, the ETF is maximizing an objective function given by

$$\max_{\alpha} \pi_c \left(-e^{-\rho p_E}\right) + \pi_s \left(-e^{-\rho\left(2\mu_E - 2\rho\sigma_E^2 - p_E\right)}\right) + (1-\pi_c-\pi_s)\left(-e^{-\rho\left(\mu_E - \frac{1}{2}\rho\sigma_E^2\right)}\right), \tag{A12}$$

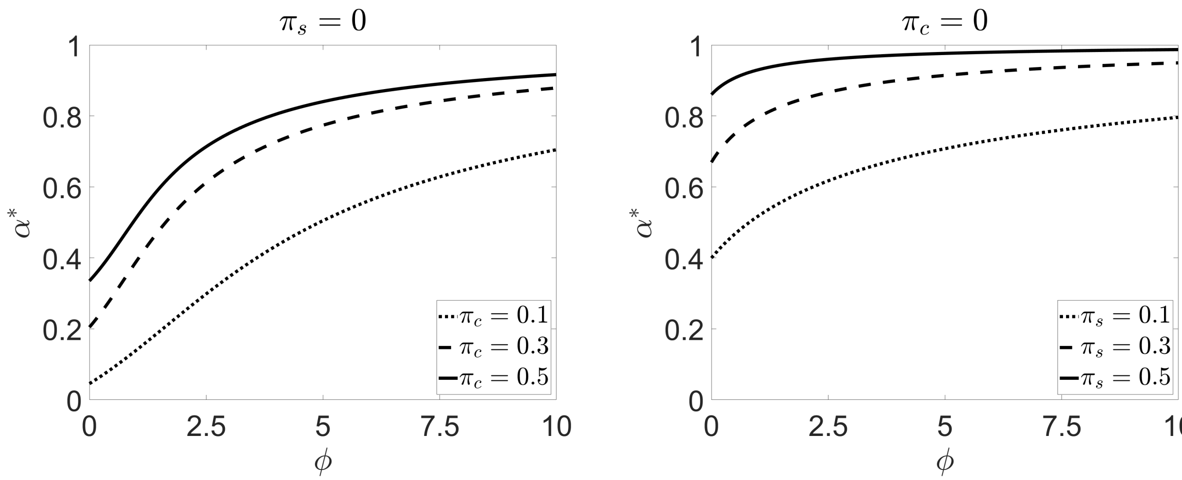
where we highlight the focus on optimizing the basket cash share, α. In equation (A12), the values of p_E, μ_E, and σ_E^2 are the same as in the baseline model (i.e., they are given in equations (4), (A4), and (A5), respectively). Denote $\delta_\pi = \pi_c - \pi_s$, as before. Differentiating the expression in equation (A12) with respect to α, we obtain the first-order condition

$$\begin{aligned}
0 \;=\; & \pi_c(1-\mu+2(1-\alpha)\phi\delta_\pi)\exp\left(-\rho\left(\alpha+\mu-\alpha\mu-(1-\alpha)^2\phi\delta_\pi\right)\right) \\
+\; & \pi_s\left(\mu-1-2(1-\alpha)\phi\delta_\pi+\frac{2(\mu-1)\delta_\pi}{1-\delta_\pi}+\frac{4\rho\sigma^2\delta_\pi((1-\alpha)N\delta_\pi-I)}{IN(1-\delta_\pi)^2}\right) \\
& \exp\left(-\rho\left(\alpha\mu-\mu-\alpha+(1-\alpha)^2\phi\delta_\pi+\frac{2(\mu(1-(1-\alpha)\delta_\pi-\alpha\delta_\pi))}{1-\delta_\pi}-\frac{2\rho\sigma^2\left((1-\alpha)^2N\delta_\pi^2-2(1-\alpha)\delta_\pi I+I\right)}{IN(1-\delta_\pi)^2}\right)\right) \\
+\; & (1-\pi_c-\pi_s)\left(\frac{(\mu-1)\delta_\pi}{1-\delta_\pi}+\frac{\rho\sigma^2\delta_\pi((1-\alpha)N\delta_\pi-I)}{IN(1-\delta_\pi)^2}\right) \\
& \exp\left(-\rho\left(\frac{\mu(1-\delta_\pi(1-\alpha))-\alpha\delta_\pi}{1-\delta_\pi}-\frac{\rho\sigma^2\left((1-\alpha)^2N\delta_\pi^2-2(1-\alpha)\delta_\pi I+I\right)}{2IN(1-\delta_\pi)^2}\right)\right).
\end{aligned} \tag{A13}$$

This equation cannot be solved analytically. We solve it numerically for many sets of parameter values. To illustrate the results, we use the parameters $N = 10, I = 1, \rho = 1, \mu = 2, \sigma = 1$. We plot the optimal basket cash weight α^* against the illiquidity parameter ϕ and vary both π_c and π_s to show the comparative statics. Figure A.1 displays the results.

The left-hand panel of Figure A.1 considers the case in which the ETF expects redemptions (i.e., $\pi_c > \pi_s$); the right-hand panel considers creations (i.e., $\pi_c < \pi_s$). In both cases, the ETF optimally chooses a basket cash share strictly between 0 and 1. Moreover, the optimal basket cash share is higher when the demand for liquidity transformation is higher in either of two ways: (i) when the security is more illiquid (i.e., when ϕ is larger) or (ii) when the imbalance in the secondary market for ETF shares is larger (i.e., when $|\pi_c - \pi_s|$

Figure A.1: ETF optimal basket cash share with CARA utility



Parameters: $N = 10, I = 1, \rho = 1, \mu = 2, \sigma = 1$

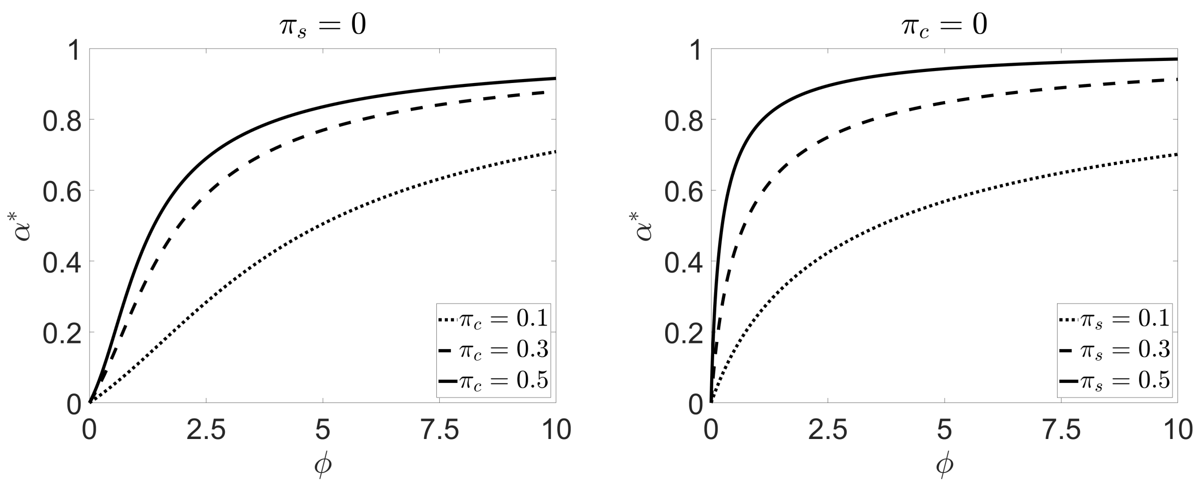
is larger). We find the same patterns for a wide range of parameter values for which the first-order condition is numerically solvable. Even though we have not been able to derive analytical proofs, our extensive parameter search has not found a counterexample.

In addition to demonstrating the robustness of our theoretical results to an alternative utility function, the CARA utility framework provides useful insights into the special case of $\sigma = 0$. In this case, the standard risk-based motive for holding cash is shut down and ETF cash holdings are driven solely by liquidity transformation. (In the baseline model with mean-variance utility, the optimal basket cash share is invariant to ϕ when $\sigma = 0$, precluding the analysis of how the cash share responds to a higher need for liquidity transformation.) Conveniently, when $\sigma = 0$, the first-order condition in equation (A13) does not depend on N or I. This first-order condition quickly yields the following result.

Proposition **3.** *A zero basket cash share, $\alpha^* = 0$, is optimal when $\phi = 0$ and $\sigma = 0$.*

The proposition states that if the N underlying securities are risk-free and liquidity transformation is of no concern, it is optimal for the ETF to include zero cash in its CR/RD basket. Figure A.2 illustrates this result, plotting α^* against ϕ for the same parameters as before, except for $\sigma = 0$. (Figure A.2 is the counterpart of Figure A.1, except that it has $\sigma = 0$.) The figure shows that the ETF's optimal basket includes no cash when the underlying securities are perfectly liquid (i.e., $\alpha^* = 0$ when $\phi = 0$, as in Proposition 3). In addition, the basket includes more cash when the underlying securities are less liquid, just like in Figure A.1 (i.e., α^* is increasing in ϕ). Given that $\sigma = 0$, the presence of cash in the basket is completely driven by the ETF's desire to transform liquidity. When there is no need for liquidity transformation, there is no need for cash in the basket, and a higher need for liquidity transformation calls for more basket cash.

Figure A.2: ETF optimal basket cash share with CARA utility and $\sigma = 0$



Parameters: $\rho = 1, \mu = 2, \sigma = 0$

A.4 Index Rebalancing

The vast majority of fixed-income indexes rebalance on a monthly basis, at month-ends. Monthly rebalancing applies to all of the indexes that we use in our analysis, as we confirm by manually checking the prospectuses of ETFs tracking those indexes. The only exception is WisdomTree indexes, which remove downgraded bonds monthly but do the rest of their rebalancing quarterly. We treat the monthly bond removals by WisdomTree as rebalancing for our purposes because they, too, cause jumps in index portfolio weights, albeit of smaller magnitudes compared to regular index rebalancing.

To illustrate the effects of index rebalancing on index composition, we turn to one of the largest bond ETFs, the iShares 1-5 Year Investment Grade Corporate Bond ETF (IGSB). This ETF, which has over $22 billion under management as of February 2022, tracks the ICE BofA US Corporate (1-5Y) index. We calculate daily changes in index composition by first computing the first difference in each bond's daily index weights and then summing the absolute values of these differences across bonds. Figure A.4 plots the time series of these daily changes (black dash-dot line). We observe that changes in index composition tend to be zero, except for positive spikes at monthly intervals.

Figure A.4 also plots the time series of the ETF's deviations from the index (blue solid line). To compute this deviation on a given day, we first calculate the difference between the ETF portfolio weights and index weights for each bond, and then sum the absolute values of these differences across all bonds in the ETF's portfolio. Comparing the solid and dash-dot lines, we see that the ETF deviations from the index tend to spike at about the same time as do the changes in index composition.

The presence of monthly spikes in ETFs' index deviations alleviates the concern that index rebalancing is fully anticipated by ETFs and incorporated into their basket management strategies before the rebalancing date. Conversations with fund managers confirm that index rebalancing cannot be fully predicted ex ante. Furthermore, even if the rebalancing were fully anticipated, the ETF would presumably not want to adjust its portfolio too much ahead of the rebalancing date, because such forward-looking portfolio adjustments create tracking error. Consistent with this argument, Li (2021) reports that equity ETFs mostly rebalance on the reconstitution day. While the levels of index deviations vary across ETFs, their patterns largely resemble those in Figure A.4.

In our index data, month-end rebalancing is often reported on the first day of the following month. There are also instances of delayed reporting when the rebalancing date is shifted by a day or two. To address these occasional reporting discrepancies, we take the day with the largest change in index composition in each month as that month's rebalancing date.

A.5 Appendix Figures and Tables

Figures

Tables

Bond Fixed Effects

Basket Inclusion Indicator Variable

3 Day Bond Liquidity

ETF Ownership Control

Shared Indexes Control

Zero Trading Days Control

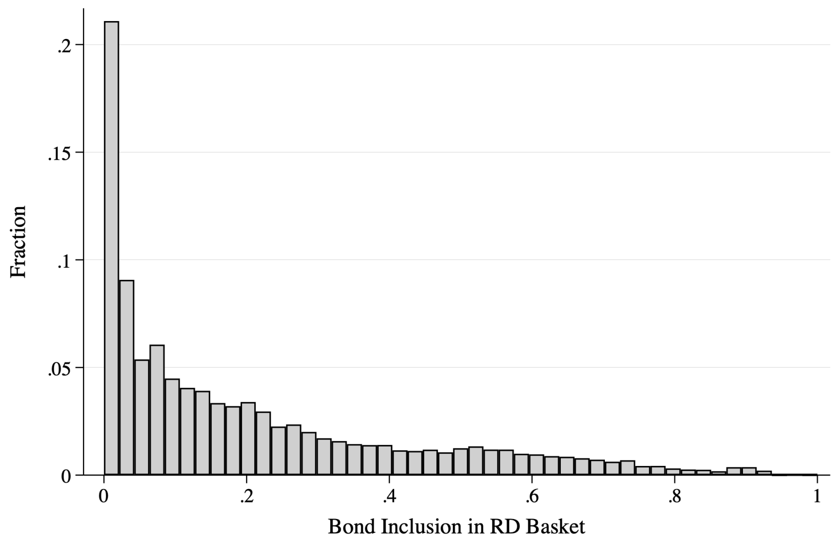
All Controls

Alternate Basket Imbalance Measure

Figure A.3: Distributions of Basket Inclusion Probabilities

This figure shows the empirical distribution of basket inclusion probabilities for all bonds held by corporate bond ETFs. The basket inclusion probability for a given bond held by a given ETF is calculated as the number of times this bond appears in this ETF's CR (RD) basket divided by the total number of the ETF's CR (RD) baskets. This figure supports the discussion in Section 3.2.

(a) Redemption Basket



(b) Creation Basket

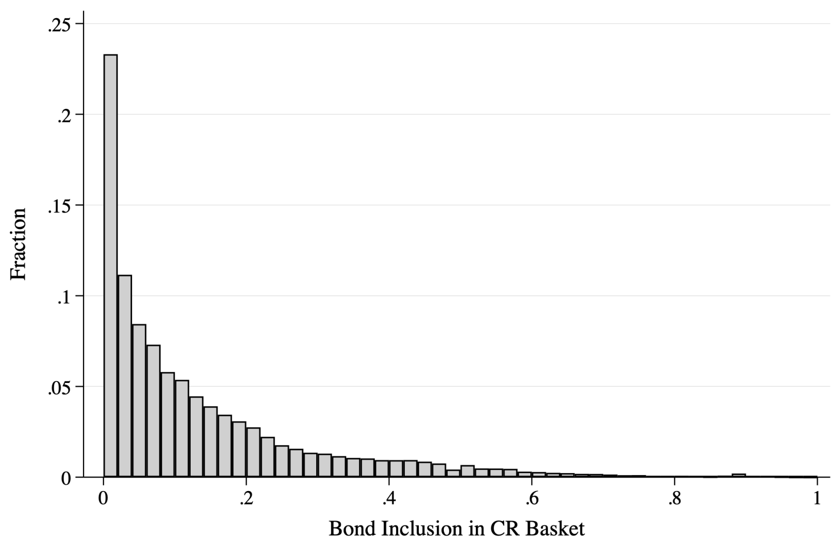


Figure A.4: Changes in Index Composition and ETF Index Deviations

This figure shows the time series of the changes in index composition (black dash-dot line) and ETF index deviations (blue solid line) for the iShares 1-5 Year Investment Grade Corporate Bond ETF (IGSB). To compute daily changes in index composition, we first compute the first difference in each bond's daily index weights and then sum the absolute values of these differences across bonds. To compute the ETF's index deviation on a given day, we first calculate the difference between the ETF portfolio weights and index weights for each bond, and then sum the absolute values of these differences across all bonds in the ETF's portfolio.

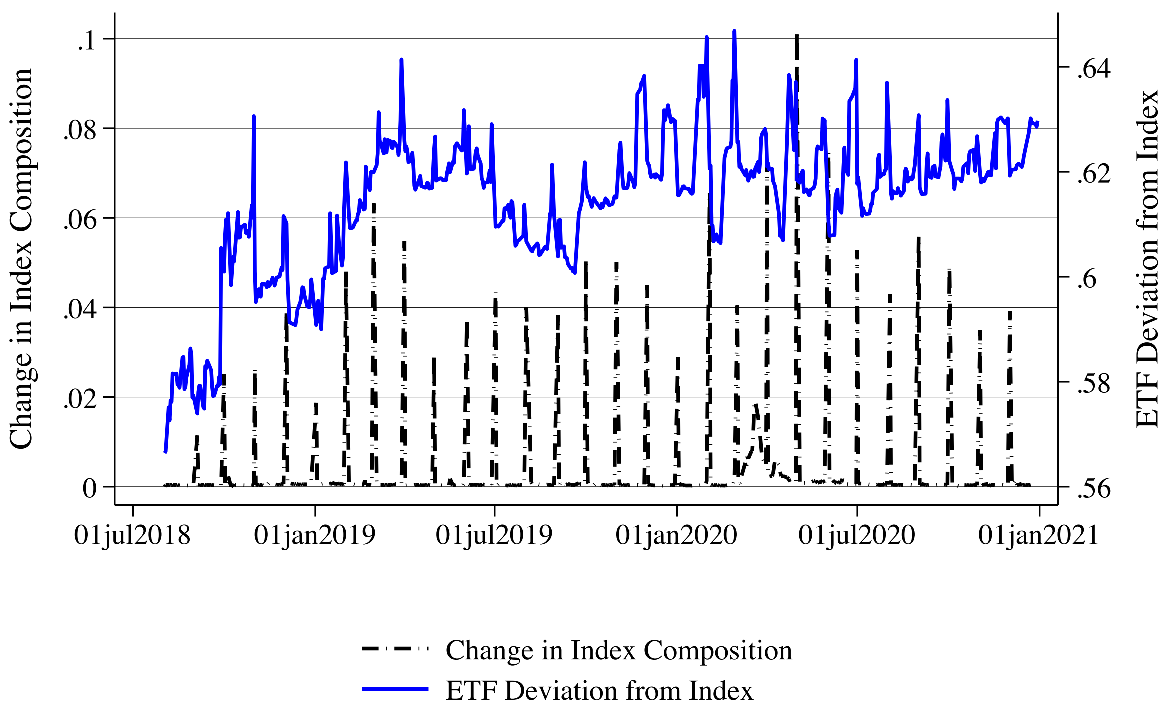


Table A.1: Summary Statistics

This table shows summary statistics for the full ETF-level sample (panel (a)), the ETF-level sample for which index data is available (panel (b)), and the bond-level sample (panel (c)). The bond-level sample includes all bonds that appear in ETF portfolios.

(a) ETF-level (Full Sample)

	Mean	Std Dev	p25	p50	p75
AUM (Million $)	2127	5547	77	340	1346
ETF Days	748	251	547	823	969
Bonds in Portfolio	753	1008	186	323	1004
Bonds in Reported Basket	424	485	121	223	556
Bonds in Realized RD Basket	147	137	52	97	205
Bonds in Realized CR Basket	104	90	51	78	119
Proportion of RD Days	0.058	0.096	0.003	0.012	0.058
Proportion of CR Days	0.169	0.169	0.022	0.107	0.274
Proportion of No Change Days	0.773	0.240	0.679	0.847	0.966
Observations	118				

(b) ETF-level (Index-merged Sample)

	Mean	Std Dev	p25	p50	p75
AUM (Million $)	2702	5645	118	576	1723
ETF Days	585	289	364	575	790
Bonds in Index	1153	1217	297	536	1847
Bonds in Portfolio	835	849	256	414	1126
Bonds in Reported Basket	696	683	170	523	924
Bonds in Realized RD Basket	138	123	54	91	242
Bonds in Realized CR Basket	128	96	74	97	172
Proportion of RD Days	0.078	0.112	0.006	0.016	0.139
Proportion of CR Days	0.208	0.190	0.044	0.155	0.342
Proportion of No Change Days	0.715	0.269	0.635	0.806	0.931
Observations	57				

(c) Bond-level

	Mean	Std Dev	p25	p50	p75
Days Held by ETFs	615	383	281	572	987
Days in RD Baskets	57	93	1	14	72
Days in CR Baskets	92	121	4	36	141
RD Basket Share	0.208	0.130	0.114	0.188	0.282
CR Basket Share	0.181	0.120	0.096	0.148	0.244
Effective Tick	0.003	0.001	0.002	0.002	0.003
IRC	0.193	0.148	0.095	0.149	0.247
IQR	0.003	0.002	0.002	0.003	0.004
Observations	18,746				

Table A.2: Cash and ETF Liquidity

This table reports the slope estimates from the regressions of ETF cash ratios on the average illiquidity of the non-cash securities in the ETF's portfolio. The three measures of index illiquidity are the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). The three types of cash ratios are the proportion of cash in ETF portfolio holdings (panel (a)), in reported ETF baskets (panel (b)), and in realized baskets (panel (c)) Observations are at the ETF-day level. An issuer-time fixed effect is included. Standard errors are reported in parentheses. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% level, respectively.

(a) Cash in Holdings

	IL1	IL2	IL3
IL	0.15***	0.11***	0.19***
	(0.01)	(0.01)	(0.01)
Observations	30,780	30,762	30,741
Adjusted R2	0.19	0.19	0.19

(b) Cash in Reported Baskets

	(1)	(2)	(3)
Illiquidity IL1	0.38***		
	(0.02)		
Illiquidity IL2		0.22***	
		(0.02)	
Illiquidity IL3			0.48***
			(0.02)
Observations	36,078	36,072	36,070
Adjusted R2	0.44	0.44	0.44

(c) Cash in Realized Baskets

	RD Basket			CR Basket		
	(1)	(2)	(3)	(4)	(5)	(6)
Illiquidity IL1	1.14			1.79***		
	(0.95)			(0.61)		
Illiquidity IL2		1.46			0.51	
		(1.01)			(0.63)	
Illiquidity IL3			0.68			2.10***
			(0.99)			(0.75)
Observations	1,849	1,849	1,849	5,568	5,568	5,568
Adjusted R2	0.14	0.14	0.14	0.18	0.17	0.18

Table A.3: Tracking Error and ETF Liquidity

This table reports the slope estimates from the regressions of ETF tracking error on the average illiquidity of the non-cash securities in the ETF's portfolio. The three measures of index illiquidity are the effective tick size (IL1), imputed roundtrip cost (IL2), and interquartile range (IL3). Both measures of tracking error are monthly standard deviations of daily differences between ETF returns and index returns. For Tracking Error 1, ETF returns are computed from ETF share prices and index returns come from Bloomberg. For Tracking Error 2, both ETF and index returns are computed from the prices of the underlying bonds. Observations are at the ETF-day level. An issuer-time fixed effect is included. Standard errors are reported in parentheses. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	Tracking Error 1			Tracking Error 2		
	(1)	(2)	(3)	(4)	(5)	(6)
Illiquidity IL1	0.55***			0.36***		
	(0.01)			(0.02)		
Illiquidity IL2		0.40***			0.32***	
		(0.01)			(0.01)	
Illiquidity IL3			0.61***			0.38***
			(0.01)			(0.02)
Observations	24,855	24,845	24,826	5,075	5,070	5,072
Adjusted R^2	0.67	0.61	0.65	0.41	0.42	0.39

A.5.1 Bond Fixed Effect

Table A.4: OLS Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 5 in the paper, except that it also includes bond fixed effects.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-0.94***	-0.55**	-0.18	-0.12***	-0.08***	-0.03*
	(0.20)	(0.23)	(0.19)	(0.02)	(0.02)	(0.02)
CR	-0.90***	-0.48***	-0.54***	-0.10***	-0.06***	-0.06***
	(0.11)	(0.12)	(0.11)	(0.02)	(0.02)	(0.01)
Bond IL	6.92***	6.10***	11.78***	6.91***	6.09***	11.78***
	(0.14)	(0.14)	(0.26)	(0.14)	(0.14)	(0.26)
Avg Basket Size	-0.24	-0.29	-0.07	-0.24	-0.27	-0.08
	(0.19)	(0.21)	(0.20)	(0.19)	(0.21)	(0.20)
Amount Outstanding				-7.26***	-2.47	-1.77
				(1.60)	(1.54)	(1.71)
Observations	3253720	2830696	2900900	3253720	2830696	2900900
Adjusted R^2	0.24	0.15	0.46	0.24	0.15	0.46

Table A.5: First Stage Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 6 in the paper, except that it also includes bond fixed effects.

(a) Number of Baskets

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	1.10***	1.11***	1.10***	-0.19***	-0.20***	-0.20***
	(0.10)	(0.10)	(0.10)	(0.05)	(0.05)	(0.05)
CR Instrument	-0.50***	-0.49***	-0.50***	-2.57***	-2.58***	-2.58***
	(0.04)	(0.04)	(0.05)	(0.12)	(0.13)	(0.13)
Bond IL	-0.20***	-0.16***	-0.06	-0.37***	-0.24***	-0.34***
	(0.04)	(0.03)	(0.05)	(0.05)	(0.04)	(0.07)
Avg Basket Size	7.38***	7.38***	7.51***	9.29***	9.25***	9.41***
	(0.68)	(0.68)	(0.69)	(0.77)	(0.77)	(0.78)
Observations	3391879	3304865	3230572	3391879	3304865	3230572
Adjusted R^2	0.50	0.50	0.50	0.43	0.43	0.43

(b) Basket Shares

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	10.91***	11.04***	10.83***	-1.18***	-1.22***	-1.18***
	(1.08)	(1.09)	(1.08)	(0.39)	(0.39)	(0.40)
CR Instrument	-5.80***	-5.71***	-5.82***	-18.18***	-18.11***	-18.15***
	(0.53)	(0.52)	(0.53)	(0.86)	(0.85)	(0.85)
Bond IL	-2.67***	-2.09***	-1.20***	-2.50***	-2.10***	-1.95***
	(0.37)	(0.30)	(0.45)	(0.42)	(0.31)	(0.68)
Avg Basket Size	57.01***	57.05***	58.05***	63.33***	62.83***	63.80***
	(5.85)	(5.86)	(5.92)	(5.63)	(5.62)	(5.69)
Amount Outstanding	65.23***	65.67***	66.07***	119.85***	119.58***	121.69***
	(8.44)	(8.35)	(8.53)	(10.76)	(10.70)	(10.95)
Observations	3391879	3304865	3230572	3391879	3304865	3230572
Adjusted R^2	0.43	0.43	0.43	0.38	0.38	0.38

Table A.6: IV Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 7 in the paper, except that it also includes bond fixed effects.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-10.50**	-12.18**	-4.11	-1.04**	-1.23**	-0.40
	(5.03)	(5.61)	(4.25)	(0.51)	(0.57)	(0.44)
CR	-7.50***	-5.89**	-7.16***	-1.07***	-0.82**	-1.07***
	(2.17)	(2.46)	(1.81)	(0.34)	(0.38)	(0.28)
Bond IL	6.48***	5.72***	11.10***	6.45***	5.71***	11.09***
	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)
Avg Basket Size	1.04**	1.36***	0.55	0.80**	1.09**	0.43
	(0.47)	(0.52)	(0.40)	(0.39)	(0.44)	(0.34)
Amount Outstanding				-6.87***	-1.41	-1.14
				(1.05)	(1.16)	(0.88)
Observations	2020227	1753297	1803180	2020227	1753297	1803180
Adjusted R^2	-0.00	-0.00	0.00	-0.00	-0.00	0.00

Table A.7: Basket Inclusion on COVID Bond Liquidity

This table is the counterpart of Table 8 in the paper, except that it also includes bond fixed effects.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	1.27	2.11*	2.16	0.25**	0.27**	0.40**
	(1.04)	(1.15)	(1.64)	(0.12)	(0.12)	(0.19)
CR	-1.86*	-1.87**	-1.06	-0.05	-0.13	-0.01
	(0.95)	(0.90)	(1.24)	(0.12)	(0.11)	(0.17)
Bond IL	-7.96***	-7.59***	-6.05***	-7.96***	-7.59***	-6.05***
	(1.50)	(1.47)	(1.98)	(1.50)	(1.47)	(1.98)
Avg Basket Size	-0.86	1.26	1.19	-1.15	1.21	0.95
	(1.44)	(1.37)	(2.48)	(1.42)	(1.34)	(2.48)
Amount Outstanding				-24.24	-14.29	27.27
				(19.58)	(14.81)	(44.19)
Observations	111277	99710	100816	111277	99710	100816
Adjusted R^2	0.28	0.08	0.38	0.28	0.08	0.38

Table A.8: Basket Imbalance on Bond Liquidity

This table is the counterpart of Table 9 in the paper, except that it also includes bond fixed effects.

	(1) IL1	(2) IL2	(3) IL3
Num Baskets	-2.42***	-1.13***	-0.98***
	(0.28)	(0.30)	(0.28)
Num Baskets × CR Imbal	0.29	0.29	-0.14
	(0.20)	(0.19)	(0.18)
Num Baskets × RD Imbal	0.93*	1.11*	1.79***
	(0.50)	(0.66)	(0.61)
Bond IL	6.91***	6.09***	11.77***
	(0.14)	(0.14)	(0.26)
Avg Basket Size	-0.30	-0.35*	-0.11
	(0.19)	(0.20)	(0.20)
Observations	3253720	2830696	2900900
Adjusted R^2	0.24	0.15	0.46

A.5.2 Indicator Variable for Basket Inclusion

Table A.9: OLS Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 5 in the paper, except that it uses a different measure of basket inclusion and considers specifications both with and without bond fixed effects. Basket inclusion is measured by a dummy variable that is equal to one if the bond is included in the CR or RD basket of at least one ETF.

	IL1		IL2		IL3	
	(1)	(2)	(3)	(4)	(5)	(6)
RD Basket Inclusion	-5.30***	-1.28***	-2.61***	-0.98***	-2.89***	-0.33
	(0.30)	(0.24)	(0.32)	(0.26)	(0.27)	(0.21)
CR Basket Inclusion	-5.02***	-1.15***	-1.70***	-0.70***	-2.84***	-0.68***
	(0.21)	(0.17)	(0.21)	(0.18)	(0.21)	(0.16)
Bond IL	12.77***	6.92***	11.80***	6.09***	19.71***	11.78***
	(0.19)	(0.14)	(0.19)	(0.14)	(0.36)	(0.26)
Avg Basket Size	-0.77***	-0.22	-1.32***	-0.25	-2.11***	-0.06
	(0.27)	(0.19)	(0.28)	(0.21)	(0.27)	(0.20)
Issuer-Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Bond FE	No	Yes	No	Yes	No	Yes
Observations	3254055	3253720	2831031	2830696	2901286	2900900
Adjusted R^2	0.23	0.24	0.13	0.15	0.44	0.46

Table A.10: First Stage Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 6 in the paper, except that it uses a different measure of basket inclusion and considers specifications both with and without bond fixed effects. Basket inclusion is measured by a dummy variable that is equal to one if the bond is included in the CR or RD basket of at least one ETF.

(a) Firm-Time and Maturity-Time Fixed Effects

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	0.95***	0.96***	0.94***	-0.04	-0.04	-0.04
	(0.09)	(0.09)	(0.09)	(0.04)	(0.04)	(0.04)
CR Instrument	-0.42***	-0.42***	-0.42***	-1.46***	-1.46***	-1.46***
	(0.04)	(0.04)	(0.05)	(0.07)	(0.07)	(0.07)
Bond IL	-0.80***	-0.47***	-0.61***	-1.20***	-0.60***	-1.01***
	(0.05)	(0.04)	(0.07)	(0.06)	(0.05)	(0.11)
Avg Basket Size	6.11***	6.09***	6.16***	7.20***	7.14***	7.22***
	(0.56)	(0.57)	(0.57)	(0.55)	(0.55)	(0.55)
Observations	3391931	3304918	3230624	3391931	3304918	3230624
Adjusted R^2	0.42	0.42	0.42	0.35	0.35	0.35

(b) Issuer-Time, Maturity-Time, and Bond Fixed Effects

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	0.91***	0.92***	0.90***	-0.12***	-0.12***	-0.12***
	(0.09)	(0.09)	(0.09)	(0.03)	(0.03)	(0.03)
CR Instrument	-0.45***	-0.45***	-0.46***	-1.51***	-1.50***	-1.50***
	(0.04)	(0.04)	(0.04)	(0.07)	(0.07)	(0.07)
Bond IL	-0.20***	-0.16***	-0.08**	-0.23***	-0.18***	-0.20***
	(0.03)	(0.02)	(0.04)	(0.04)	(0.03)	(0.05)
Avg Basket Size	5.52***	5.52***	5.62***	6.29***	6.25***	6.33***
	(0.55)	(0.55)	(0.56)	(0.52)	(0.52)	(0.53)
Observations	3391879	3304865	3230572	3391879	3304865	3230572
Adjusted R^2	0.45	0.45	0.45	0.39	0.39	0.39

Table A.11: IV Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 7 in the paper, except that it uses a different measure of basket inclusion and considers specifications both with and without bond fixed effects. Basket inclusion is measured by a dummy variable that is equal to one if the bond is included in the CR or RD basket of at least one ETF.

	IL1		IL2		IL3	
	(1)	(2)	(3)	(4)	(5)	(6)
RD Basket Inclusion	-18.89***	-13.07**	-18.83***	-15.27**	-9.22*	-5.30
	(6.11)	(6.18)	(6.83)	(6.92)	(5.28)	(5.33)
CR Basket Inclusion	-6.35	-13.26***	-9.23*	-10.24**	-15.63***	-13.06***
	(4.21)	(4.03)	(4.84)	(4.59)	(3.60)	(3.41)
Bond IL	11.75***	6.46***	11.12***	5.72***	17.99***	11.09***
	(0.13)	(0.09)	(0.10)	(0.09)	(0.11)	(0.10)
Avg Basket Size	0.44	1.05**	1.28**	1.34***	-0.05	0.63*
	(0.45)	(0.44)	(0.51)	(0.49)	(0.39)	(0.38)
Issuer-Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Bond FE	No	Yes	No	Yes	No	Yes
Observations	2020546	2020227	1753639	1753297	1803581	1803180
Adjusted R^2	0.01	-0.00	0.01	-0.00	0.02	0.00

Table A.12: Basket Inclusion on COVID Bond Liquidity

This table is the counterpart of Table 8 in the paper, except that it uses a different measure of basket inclusion and considers specifications both with and without bond fixed effects. Basket inclusion is measured by a dummy variable that is equal to one if the bond is included in the CR or RD basket of at least one ETF.

	IL1		IL2		IL3	
	(1)	(2)	(3)	(4)	(5)	(6)
RD Basket Inclusion	2.08	2.91**	5.07***	3.10**	6.65***	4.49*
	(1.35)	(1.40)	(1.48)	(1.35)	(2.37)	(2.22)
CR Basket Inclusion	0.65	-0.64	-0.68	-2.05*	6.23**	-0.22
	(1.78)	(1.42)	(1.11)	(1.18)	(2.63)	(1.99)
Bond IL	9.50***	-7.95***	8.05***	-7.59***	15.34***	-6.05***
	(0.62)	(1.50)	(0.54)	(1.47)	(0.74)	(1.98)
Avg Basket Size	-0.27	-1.19	-1.29	1.25	-0.96	0.90
	(1.15)	(1.43)	(0.95)	(1.33)	(2.04)	(2.49)
Issuer-Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Maturity-Day FE	Yes	Yes	Yes	Yes	Yes	Yes
Bond FE	No	Yes	No	Yes	No	Yes
Observations	111707	111277	100263	99710	101341	100816
Adjusted R^2	0.26	0.28	0.07	0.08	0.36	0.38

A.5.3 Effect of Basket Inclusion on 3-Day Bond Liquidity

Table A.13: OLS Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 5 in the paper for 3-day-ahead bond liquidity, calculated as the average of $t + 1$, $t + 2$, and $t + 3$ bond liquidity.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-5.00***	-2.48***	-2.47***	-0.20***	-0.15***	-0.11***
	(0.28)	(0.28)	(0.29)	(0.02)	(0.02)	(0.02)
CR	-4.35***	-1.20***	-2.45***	-0.20***	-0.03*	-0.15***
	(0.16)	(0.14)	(0.15)	(0.02)	(0.02)	(0.02)
Bond IL	12.20***	12.96***	17.70***	10.99***	12.76***	17.20***
	(0.19)	(0.23)	(0.35)	(0.17)	(0.22)	(0.34)
Avg Basket Size	-1.20***	-1.92***	-2.67***	-1.50***	-2.00***	-2.83***
	(0.30)	(0.30)	(0.29)	(0.26)	(0.29)	(0.27)
Amount Outstanding				-14.08***	-6.13***	-7.33***
				(0.37)	(0.41)	(0.38)
Observations	4551644	4254774	4196240	4551644	4254774	4196240
Adjusted R^2	0.30	0.18	0.48	0.30	0.18	0.48

Table A.14: IV Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 7 in the paper for 3-day-ahead bond liquidity, calculated as the average of $t+1$, $t+2$, and $t+3$ bond liquidity.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-8.66**	-5.40	-10.29***	-0.34	-0.32	-0.83**
	(3.80)	(4.14)	(3.25)	(0.40)	(0.43)	(0.34)
CR	-7.15***	-6.24***	-10.74***	-1.86***	-1.23***	-1.84***
	(1.82)	(2.02)	(1.54)	(0.26)	(0.29)	(0.22)
Bond IL	11.18***	12.25***	16.18***	10.00***	12.06***	15.85***
	(0.08)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)
Avg Basket Size	0.12	0.33	0.30	0.39	0.41	0.28
	(0.36)	(0.40)	(0.31)	(0.31)	(0.34)	(0.27)
Amount Outstanding				-11.69***	-4.42***	-3.88***
				(0.37)	(0.42)	(0.32)
Observations	2807592	2626760	2589840	2807592	2626760	2589840
Adjusted R^2	0.01	0.01	0.02	0.01	0.01	0.01

Table A.15: Basket Inclusion on COVID Bond Liquidity

This table is the counterpart of Table 8 in the paper for 3-day-ahead bond liquidity, calculated as the average of $t+1$, $t+2$, and $t+3$ bond liquidity.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	1.56	4.69***	6.05***	0.11	0.30***	0.34**
	(0.94)	(0.96)	(1.46)	(0.10)	(0.10)	(0.17)
CR	0.34	-0.00	3.60*	0.17	-0.03	0.32*
	(1.30)	(0.83)	(2.08)	(0.10)	(0.08)	(0.18)
Bond IL	9.21***	7.26***	13.57***	9.17***	7.21***	13.11***
	(0.58)	(0.53)	(0.68)	(0.59)	(0.53)	(0.65)
Avg Basket Size	1.53	-1.46	0.87	1.25	-1.30	0.91
	(1.19)	(0.93)	(1.87)	(1.12)	(0.96)	(1.65)
Amount Outstanding				2.10	4.35***	14.85***
				(2.08)	(1.35)	(3.17)
Observations	152571	145920	142237	152571	145920	142237
Adjusted R^2	0.32	0.11	0.40	0.32	0.11	0.41

Table A.16: Basket Imbalance on Bond Liquidity

This table is the counterpart of Table 9 in the paper for 3-day-ahead bond liquidity, calculated as the average of $t+1$, $t+2$, and $t+3$ bond liquidity.

	(1) IL1	(2) IL2	(3) IL3
Num Baskets	-13.03***	-5.83***	-7.90***
	(0.49)	(0.47)	(0.49)
Num Baskets × CR Imbal	2.23***	2.68***	2.13***
	(0.28)	(0.27)	(0.25)
Num Baskets × RD Imbal	4.04***	1.98***	3.61***
	(0.57)	(0.72)	(0.69)
Bond IL	11.96***	12.92***	17.58***
	(0.18)	(0.22)	(0.34)
Avg Basket Size	-0.71**	-1.69***	-2.33***
	(0.28)	(0.29)	(0.27)
Observations	4551644	4254774	4196240
Adjusted R^2	0.30	0.18	0.48

A.5.4 ETF Ownership

Table A.17: OLS Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 5 in the paper, with the additional ETF Ownership control. ETF Ownership is measured as the standardized daily proportion of the bond's shares held by ETFs, relative to shares outstanding.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-3.18***	-1.37***	-1.48***	-0.12***	-0.09***	-0.06***
	(0.24)	(0.26)	(0.23)	(0.02)	(0.03)	(0.02)
CR	-2.92***	-0.70***	-1.41***	-0.13***	-0.03	-0.08***
	(0.13)	(0.14)	(0.13)	(0.02)	(0.02)	(0.02)
Bond IL	12.60***	11.78***	19.58***	11.36***	11.63***	19.14***
	(0.19)	(0.19)	(0.35)	(0.17)	(0.19)	(0.35)
Avg Basket Size	-0.13	-1.14***	-1.67***	-0.93***	-1.40***	-2.09***
	(0.26)	(0.28)	(0.26)	(0.23)	(0.27)	(0.25)
ETF Ownership	-3.91***	-1.56***	-2.88***	-2.33***	-0.88***	-2.09***
	(0.23)	(0.20)	(0.22)	(0.21)	(0.20)	(0.21)
Amount Outstanding				-11.98***	-4.44***	-5.80***
				(0.30)	(0.34)	(0.32)
Observations	3254005	2830996	2901243	3254005	2830996	2901243
Adjusted R^2	0.23	0.13	0.44	0.23	0.13	0.44

Table A.18: First Stage Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 6 in the paper, with the additional ETF Ownership control. ETF Ownership is measured as the standardized daily proportion of the bond's shares held by ETFs, relative to shares outstanding.

(a) Number of Baskets

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	1.09***	1.10***	1.09***	-0.19***	-0.19***	-0.19***
	(0.10)	(0.10)	(0.10)	(0.05)	(0.05)	(0.05)
CR Instrument	-0.57***	-0.56***	-0.57***	-2.68***	-2.70***	-2.69***
	(0.05)	(0.05)	(0.05)	(0.13)	(0.13)	(0.13)
Bond IL	-0.70***	-0.37***	-0.44***	-1.51***	-0.51***	-1.07***
	(0.05)	(0.04)	(0.07)	(0.08)	(0.05)	(0.11)
Avg Basket Size	7.71***	7.68***	7.78***	9.36***	9.30***	9.46***
	(0.72)	(0.72)	(0.73)	(0.79)	(0.79)	(0.80)
ETF Ownership	7.17***	7.20***	7.23***	12.82***	12.97***	12.91***
	(0.32)	(0.32)	(0.32)	(0.38)	(0.38)	(0.38)
Observations	3391872	3304870	3230573	3391872	3304870	3230573
Adjusted R^2	0.48	0.48	0.49	0.40	0.40	0.40

(b) Basket Shares

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	10.48***	10.60***	10.39***	-1.65***	-1.69***	-1.66***
	(1.09)	(1.10)	(1.10)	(0.41)	(0.41)	(0.41)
CR Instrument	-6.99***	-6.92***	-7.02***	-20.07***	-20.03***	-20.05***
	(0.57)	(0.56)	(0.57)	(0.88)	(0.89)	(0.88)
Bond IL	-3.09***	-2.34***	-1.83***	-3.49***	-1.68***	-2.76***
	(0.40)	(0.35)	(0.49)	(0.48)	(0.37)	(0.75)
Avg Basket Size	61.17***	60.96***	61.80***	67.77***	67.16***	68.30***
	(6.34)	(6.35)	(6.40)	(5.96)	(5.95)	(6.03)
Amount Outstanding	76.96***	77.36***	77.58***	123.73***	124.76***	124.29***
	(3.22)	(3.23)	(3.23)	(2.95)	(2.95)	(2.94)
ETF Ownership	59.05***	59.00***	59.58***	81.30***	81.63***	81.64***
	(2.60)	(2.59)	(2.61)	(2.09)	(2.09)	(2.09)
Observations	3391872	3304870	3230573	3391872	3304870	3230573
Adjusted R^2	0.42	0.42	0.42	0.37	0.37	0.37

Table A.19: IV Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 7 in the paper, with the additional ETF Ownership control. ETF Ownership is measured as the standardized daily proportion of the bond's shares held by ETFs, relative to shares outstanding.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-13.85***	-14.91***	-6.66	-1.03*	-1.42**	-0.52
	(5.10)	(5.70)	(4.32)	(0.53)	(0.59)	(0.45)
CR	-5.32**	-5.48**	-9.31***	-1.37***	-0.90**	-1.59***
	(2.16)	(2.46)	(1.82)	(0.32)	(0.36)	(0.27)
Bond IL	11.67***	11.15***	18.01***	10.44***	11.03***	17.68***
	(0.11)	(0.09)	(0.10)	(0.09)	(0.09)	(0.09)
Avg Basket Size	0.71	1.34***	-0.08	0.57	1.09***	-0.16
	(0.45)	(0.51)	(0.38)	(0.38)	(0.42)	(0.33)
ETF Ownership	-2.13***	-0.21	-1.09***	-0.56*	0.20	-0.45
	(0.38)	(0.44)	(0.33)	(0.32)	(0.37)	(0.28)
Amount Outstanding				-10.08***	-2.39***	-3.34***
				(0.40)	(0.47)	(0.35)
Observations	2020511	1753613	1803553	2020511	1753613	1803553
Adjusted R^2	0.01	0.01	0.02	0.01	0.01	0.02

Table A.20: Basket Inclusion on COVID Bond Liquidity

This table is the counterpart of Table 8 in the paper, with the additional ETF Ownership control. ETF Ownership is measured as the standardized daily proportion of the bond's shares held by ETFs, relative to shares outstanding.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	0.03	3.50***	3.00**	0.15	0.31**	0.31
	(0.96)	(1.12)	(1.42)	(0.11)	(0.13)	(0.20)
CR	-1.98	-0.96	1.82	0.05	-0.16	0.24
	(1.27)	(0.86)	(1.79)	(0.13)	(0.10)	(0.19)
Bond IL	9.45***	8.02***	15.31***	9.44***	7.95***	15.02***
	(0.62)	(0.54)	(0.74)	(0.61)	(0.53)	(0.72)
Avg Basket Size	-0.06	-1.55	-0.85	-0.61	-1.30	-0.76
	(1.20)	(0.94)	(2.05)	(1.18)	(0.96)	(1.97)
ETF Ownership	2.61**	2.00***	4.52***	2.42**	1.64**	3.41**
	(0.98)	(0.67)	(1.49)	(0.96)	(0.63)	(1.39)
Amount Outstanding				-2.12	4.06***	9.65***
				(1.69)	(1.18)	(2.83)
Observations	111690	100255	101326	111690	100255	101326
Adjusted R^2	0.26	0.07	0.36	0.26	0.07	0.36

Table A.21: Basket Imbalance on Bond Liquidity

This table is the counterpart of Table 9 in the paper, with the additional ETF Ownership control. ETF Ownership is measured as the standardized daily proportion of the bond's shares held by ETFs, relative to shares outstanding.

	(1) IL1	(2) IL2	(3) IL3
Num Baskets	-8.96***	-3.34***	-4.99***
	(0.37)	(0.39)	(0.37)
Num Baskets × CR Imbal	1.42***	1.69***	1.32***
	(0.23)	(0.22)	(0.20)
Num Baskets × RD Imbal	2.74***	1.37**	2.45***
	(0.51)	(0.69)	(0.61)
Bond IL	12.45***	11.77***	19.52***
	(0.18)	(0.19)	(0.35)
Avg Basket Size	-0.06	-1.11***	-1.60***
	(0.25)	(0.27)	(0.26)
ETF Ownership	-2.69***	-1.21***	-2.23***
	(0.23)	(0.20)	(0.21)
Observations	3254005	2830996	2901243
Adjusted R^2	0.23	0.13	0.44

A.5.5 Shared Indexes

Table A.22: OLS Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 5 in the paper, with the additional Shared Index control. This variable controls for the number of baskets or amount of shares of each bond that are included in baskets of ETFs that share their indexes with index mutual funds, where the control corresponds to the measure of basket inclusion used in a given specification.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-4.34***	-1.91***	-2.38***	-0.17***	-0.12***	-0.11***
	(0.27)	(0.28)	(0.26)	(0.02)	(0.03)	(0.02)
CR	-3.67***	-1.08***	-1.97***	-0.17***	-0.05***	-0.12***
	(0.15)	(0.14)	(0.14)	(0.02)	(0.02)	(0.02)
Bond IL	12.72***	11.81***	19.69***	11.36***	11.64***	19.18***
	(0.19)	(0.19)	(0.35)	(0.17)	(0.19)	(0.34)
Avg Basket Size	-0.79***	-1.53***	-2.05***	-1.30***	-1.66***	-2.34***
	(0.29)	(0.28)	(0.27)	(0.25)	(0.27)	(0.26)
Shared Index	1.09**	2.08***	1.12**	0.05	0.19***	0.07
	(0.54)	(0.52)	(0.49)	(0.05)	(0.05)	(0.04)
Amount Outstanding				-12.52***	-4.65***	-6.22***
				(0.31)	(0.34)	(0.33)
Observations	3255898	2832964	2902611	3255898	2832964	2902611
Adjusted R^2	0.23	0.13	0.44	0.23	0.13	0.44

Table A.23: First Stage Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 6 in the paper, with the additional Shared Index control. This variable controls for the number of baskets or amount of shares of each bond that are included in baskets of ETFs that share their indexes with index mutual funds, where the control corresponds to the measure of basket inclusion used in a given specification.

(a) Number of Baskets

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	1.08***	1.10***	1.08***	-0.19***	-0.19***	-0.19***
	(0.10)	(0.10)	(0.10)	(0.06)	(0.06)	(0.06)
CR Instrument	-0.40***	-0.39***	-0.40***	-2.38***	-2.39***	-2.39***
	(0.05)	(0.05)	(0.05)	(0.12)	(0.13)	(0.13)
Bond IL	-0.85***	-0.49***	-0.62***	-1.78***	-0.73***	-1.40***
	(0.06)	(0.04)	(0.08)	(0.10)	(0.06)	(0.14)
Avg Basket Size	7.05***	7.00***	7.10***	8.35***	8.28***	8.40***
	(0.70)	(0.70)	(0.71)	(0.77)	(0.77)	(0.78)
Shared Index	41.33***	41.41***	41.24***	67.68***	67.90***	67.74***
	(2.81)	(2.82)	(2.81)	(2.97)	(2.98)	(2.97)
Observations	3392882	3305958	3231632	3392882	3305958	3231632
Adjusted R^2	0.49	0.49	0.49	0.40	0.40	0.40

(b) Basket Shares

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	10.30***	10.41***	10.20***	-1.77***	-1.80***	-1.78***
	(1.05)	(1.05)	(1.05)	(0.44)	(0.44)	(0.44)
CR Instrument	-5.53***	-5.44***	-5.53***	-18.14***	-18.07***	-18.09***
	(0.59)	(0.58)	(0.59)	(0.86)	(0.86)	(0.86)
Bond IL	-3.75***	-3.08***	-2.87***	-4.43***	-2.68***	-4.20***
	(0.43)	(0.38)	(0.57)	(0.53)	(0.40)	(0.87)
Avg Basket Size	55.81***	55.43***	56.31***	62.55***	61.86***	62.83***
	(6.14)	(6.14)	(6.21)	(5.66)	(5.65)	(5.72)
Amount Outstanding	84.62***	85.05***	85.10***	135.52***	136.74***	135.85***
	(3.59)	(3.60)	(3.59)	(3.27)	(3.28)	(3.27)
Shared Index	34.96***	35.00***	34.83***	39.69***	39.66***	39.52***
	(2.41)	(2.42)	(2.41)	(1.84)	(1.84)	(1.84)
Observations	3392882	3305958	3231632	3392882	3305958	3231632
Adjusted R^2	0.42	0.42	0.42	0.37	0.37	0.37

Table A.24: IV Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 7 in the paper, with the additional Shared Index control. This variable controls for the number of baskets or amount of shares of each bond that are included in baskets of ETFs that share their indexes with index mutual funds, where the control corresponds to the measure of basket inclusion used in a given specification.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-15.60***	-16.52***	-7.53*	-1.17**	-1.58***	-0.59
	(5.19)	(5.77)	(4.42)	(0.54)	(0.60)	(0.47)
CR	-4.50*	-6.13**	-8.98***	-1.44***	-1.03***	-1.63***
	(2.31)	(2.64)	(1.94)	(0.33)	(0.38)	(0.28)
Bond IL	11.76***	11.15***	18.08***	10.43***	11.02***	17.71***
	(0.12)	(0.10)	(0.11)	(0.10)	(0.09)	(0.10)
Avg Basket Size	0.59	1.29***	-0.26	0.57	1.12***	-0.25
	(0.44)	(0.50)	(0.38)	(0.38)	(0.42)	(0.33)
Shared Index	6.71***	10.88***	7.21***	0.90***	1.02***	0.73***
	(2.34)	(2.68)	(2.03)	(0.19)	(0.21)	(0.16)
Amount Outstanding				-10.19***	-2.21***	-3.45***
				(0.49)	(0.57)	(0.43)
Observations	2020911	1754120	1803915	2020911	1754120	1803915
Adjusted R^2	0.01	0.01	0.02	0.01	0.01	0.01

Table A.25: Basket Inclusion on COVID Bond Liquidity

This table is the counterpart of Table 8 in the paper, with the additional Shared Index control. This variable controls for the number of baskets or amount of shares of each bond that are included in baskets of ETFs that share their indexes with index mutual funds, where the control corresponds to the measure of basket inclusion used in a given specification.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	1.76*	4.20***	5.75***	0.35***	0.34**	0.59***
	(1.03)	(1.19)	(1.57)	(0.12)	(0.14)	(0.21)
CR	-1.10	-0.29	3.37*	0.13	-0.10	0.37*
	(1.30)	(0.81)	(1.84)	(0.13)	(0.09)	(0.18)
Bond IL	9.28***	7.99***	15.39***	9.25***	7.91***	15.06***
	(0.63)	(0.56)	(0.74)	(0.63)	(0.56)	(0.71)
Avg Basket Size	0.37	-1.78*	-0.78	-0.20	-1.52	-0.82
	(1.16)	(0.99)	(2.11)	(1.12)	(1.04)	(1.99)
Shared Index	-11.16***	-1.32	-17.53***	-1.05***	-0.08	-1.57***
	(2.82)	(2.75)	(3.50)	(0.24)	(0.23)	(0.34)
Amount Outstanding				-2.13	4.26***	9.89***
				(1.76)	(1.21)	(2.95)
Observations	111720	100319	101388	111720	100319	101388
Adjusted R^2	0.26	0.07	0.36	0.26	0.07	0.36

Table A.26: Basket Imbalance on Bond Liquidity

This table is the counterpart of Table 9 in the paper, with the additional Shared Index control. This variable controls for the number of baskets of each bond that are included in baskets of ETFs that share their indexes with index mutual funds.

	(1) IL1	(2) IL2	(3) IL3
Num Baskets	-10.83***	-4.22***	-6.52***
	(0.41)	(0.41)	(0.41)
Num Baskets × CR Imbal	1.82***	1.82***	1.58***
	(0.23)	(0.23)	(0.21)
Num Baskets × RD Imbal	2.88***	1.45**	2.59***
	(0.51)	(0.69)	(0.63)
Bond IL	12.48***	11.78***	19.57***
	(0.19)	(0.19)	(0.35)
Avg Basket Size	-0.42	-1.40***	-1.81***
	(0.27)	(0.28)	(0.26)
Shared Index	0.42	1.62***	0.80*
	(0.54)	(0.51)	(0.48)
Observations	3255898	2832964	2902611
Adjusted R^2	0.23	0.13	0.44

A.5.6 Zero Trading Days

Table A.27: OLS Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 5 in the paper, with the additional Zero Trading Days control, which measures the bond's proportion of zero trading days during the past 20 week days.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-4.13***	-2.05***	-2.46***	-0.19***	-0.13***	-0.12***
	(0.25)	(0.27)	(0.24)	(0.02)	(0.03)	(0.02)
CR	-3.52***	-1.20***	-2.08***	-0.20***	-0.07***	-0.14***
	(0.14)	(0.14)	(0.13)	(0.02)	(0.02)	(0.02)
Bond IL	12.75***	11.79***	19.70***	11.30***	11.53***	19.02***
	(0.19)	(0.19)	(0.36)	(0.18)	(0.19)	(0.35)
Avg Basket Size	-0.80***	-1.45***	-2.19***	-1.53***	-1.72***	-2.63***
	(0.27)	(0.28)	(0.27)	(0.24)	(0.27)	(0.25)
Zero Trading Days	6.12***	-6.92***	-5.31***	-16.33***	-17.19***	-18.61***
	(1.07)	(1.04)	(1.35)	(0.99)	(1.12)	(1.29)
Amount Outstanding				-13.86***	-6.14***	-7.70***
				(0.31)	(0.37)	(0.33)
Observations	3254055	2831031	2901286	3254055	2831031	2901286
Adjusted R^2	0.23	0.13	0.44	0.23	0.13	0.44

Table A.28: First Stage Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 6 in the paper, with the additional Zero Trading Days control, which measures the bond's proportion of zero trading days during the past 20 week days.

(a) Number of Baskets

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	1.14***	1.16***	1.14***	-0.10*	-0.10*	-0.10*
	(0.10)	(0.10)	(0.10)	(0.05)	(0.05)	(0.05)
CR Instrument	-0.45***	-0.44***	-0.45***	-2.47***	-2.48***	-2.48***
	(0.05)	(0.05)	(0.05)	(0.13)	(0.13)	(0.13)
Bond IL	-0.95***	-0.53***	-0.76***	-1.97***	-0.81***	-1.67***
	(0.06)	(0.04)	(0.08)	(0.09)	(0.06)	(0.13)
Avg Basket Size	8.00***	7.97***	8.07***	9.80***	9.75***	9.91***
	(0.72)	(0.72)	(0.72)	(0.79)	(0.80)	(0.80)
Zero Trading Days	-18.40***	-18.44***	-18.78***	-38.14***	-38.72***	-39.09***
	(0.86)	(0.86)	(0.88)	(1.21)	(1.23)	(1.25)
Observations	3391931	3304918	3230624	3391931	3304918	3230624
Adjusted R^2	0.48	0.48	0.48	0.40	0.40	0.40

(b) Basket Shares

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	10.96***	11.08***	10.87***	-0.99**	-1.03**	-1.00**
	(1.10)	(1.10)	(1.10)	(0.41)	(0.41)	(0.41)
CR Instrument	-6.06***	-5.98***	-6.07***	-18.72***	-18.66***	-18.69***
	(0.56)	(0.56)	(0.56)	(0.88)	(0.88)	(0.88)
Bond IL	-5.08***	-3.57***	-4.16***	-6.61***	-3.59***	-6.34***
	(0.44)	(0.38)	(0.56)	(0.52)	(0.41)	(0.84)
Avg Basket Size	64.46***	64.24***	65.11***	71.95***	71.32***	72.47***
	(6.31)	(6.32)	(6.37)	(5.95)	(5.94)	(6.01)
Amount Outstanding	75.68***	76.44***	76.80***	117.61***	118.72***	118.81***
	(3.16)	(3.17)	(3.18)	(2.93)	(2.93)	(2.94)
Zero Trading Days	-87.65***	-86.82***	-87.47***	-151.66***	-154.36***	-152.60***
	(5.37)	(5.43)	(5.47)	(6.15)	(6.27)	(6.24)
Observations	3391931	3304918	3230624	3391931	3304918	3230624
Adjusted R^2	0.41	0.41	0.41	0.36	0.36	0.36

Table A.29: IV Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 7 in the paper, with the additional Zero Trading Days control, which measures the bond's proportion of zero trading days during the past 20 week days.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-15.16***	-15.49***	-7.77*	-1.07**	-1.42**	-0.56
	(4.96)	(5.53)	(4.21)	(0.52)	(0.58)	(0.45)
CR	-3.81*	-5.30**	-8.51***	-1.43***	-1.02***	-1.67***
	(2.26)	(2.58)	(1.90)	(0.33)	(0.38)	(0.28)
Bond IL	11.78***	11.14***	18.02***	10.31***	10.93***	17.46***
	(0.12)	(0.10)	(0.11)	(0.10)	(0.09)	(0.10)
Avg Basket Size	0.55	1.26**	-0.22	0.44	1.04**	-0.28
	(0.47)	(0.53)	(0.40)	(0.40)	(0.45)	(0.34)
Zero Trading Days	2.62**	-10.14***	-11.27***	-20.14***	-17.86***	-22.42***
	(1.33)	(1.56)	(1.20)	(0.79)	(0.90)	(0.70)
Amount Outstanding				-11.86***	-3.78***	-5.12***
				(0.42)	(0.50)	(0.37)
Observations	2020546	1753639	1803581	2020546	1753639	1803581
Adjusted R^2	0.01	0.01	0.02	0.01	0.01	0.02

Table A.30: Basket Inclusion on COVID Bond Liquidity

This table is the counterpart of Table 8 in the paper, with the additional Zero Trading Days control, which measures the bond's proportion of zero trading days during the past 20 week days.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-1.02	3.45***	0.80	0.10	0.33**	0.22
	(0.95)	(1.12)	(1.39)	(0.11)	(0.13)	(0.19)
CR	-3.03**	-1.08	-0.22	0.01	-0.14	0.19
	(1.17)	(0.83)	(1.62)	(0.13)	(0.10)	(0.18)
Bond IL	8.42***	7.94***	13.48***	8.15***	7.91***	13.46***
	(0.67)	(0.54)	(0.67)	(0.64)	(0.54)	(0.67)
Avg Basket Size	-0.31	-1.53	-1.23	-1.25	-1.30	-1.74
	(1.15)	(0.97)	(1.96)	(1.17)	(0.98)	(1.95)
Zero Trading Days	-56.99***	-17.67***	-121.50***	-69.18***	-12.78***	-120.57***
	(6.61)	(4.15)	(12.86)	(5.06)	(4.27)	(10.51)
Amount Outstanding				-8.32***	3.05**	-0.28
				(1.44)	(1.22)	(2.49)
Observations	111707	100263	101341	111707	100263	101341
Adjusted R^2	0.27	0.07	0.36	0.27	0.07	0.36

Table A.31: Basket Imbalance on Bond Liquidity

This table is the counterpart of Table 9 in the paper, with the additional Zero Trading Days control, which measures the bond's proportion of zero trading days during the past 20 week days.

	(1) IL1	(2) IL2	(3) IL3
Num Baskets	-10.78***	-5.03***	-7.29***
	(0.39)	(0.41)	(0.38)
Num Baskets × CR Imbal	1.84***	2.08***	1.85***
	(0.23)	(0.23)	(0.20)
Num Baskets × RD Imbal	2.92***	1.68**	2.80***
	(0.51)	(0.69)	(0.62)
Bond IL	12.52***	11.75***	19.55***
	(0.19)	(0.19)	(0.35)
Avg Basket Size	-0.44*	-1.29***	-1.91***
	(0.26)	(0.27)	(0.26)
Zero Trading Days	2.30**	-8.49***	-7.95***
	(1.04)	(1.05)	(1.32)
Observations	3254055	2831031	2901286
Adjusted R^2	0.23	0.13	0.44

A.5.7 All Controls

Table A.32: OLS Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 5 in the paper, with all additional controls.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-3.17***	-1.61***	-1.69***	-0.13***	-0.11***	-0.07***
	(0.25)	(0.27)	(0.24)	(0.02)	(0.03)	(0.02)
CR	-2.87***	-0.94***	-1.55***	-0.14***	-0.05***	-0.09***
	(0.13)	(0.14)	(0.13)	(0.02)	(0.02)	(0.02)
Bond IL	12.55***	11.75***	19.52***	11.16***	11.52***	18.89***
	(0.19)	(0.19)	(0.35)	(0.17)	(0.19)	(0.34)
Avg Basket Size	-0.12	-1.26***	-1.56***	-0.98***	-1.59***	-2.05***
	(0.27)	(0.28)	(0.27)	(0.24)	(0.27)	(0.26)
Shared Index	1.02*	2.03***	1.05**	0.05	0.19***	0.07
	(0.54)	(0.52)	(0.48)	(0.05)	(0.05)	(0.04)
ETF Ownership	-3.71***	-1.86***	-3.13***	-2.67***	-1.30***	-2.53***
	(0.23)	(0.21)	(0.22)	(0.20)	(0.20)	(0.20)
Zero Trading Days	2.85***	-8.59***	-8.08***	-18.35***	-18.14***	-20.45***
	(1.03)	(1.05)	(1.31)	(0.98)	(1.13)	(1.28)
Amount Outstanding				-13.75***	-6.05***	-7.52***
				(0.32)	(0.37)	(0.33)
Observations	3255850	2832929	2902571	3255850	2832929	2902571
Adjusted R^2	0.23	0.13	0.44	0.23	0.13	0.44

Table A.33: First Stage Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 6 in the paper, with all additional controls.

(a) Number of Baskets

	RD Basket			CR Basket		
	IL1	IL2	IL3	IL1	IL2	IL3
RD Instrument	1.03***	1.04***	1.02***	-0.29***	-0.30***	-0.30***
	(0.10)	(0.10)	(0.10)	(0.06)	(0.06)	(0.06)
CR Instrument	-0.48***	-0.47***	-0.48***	-2.51***	-2.52***	-2.52***
	(0.05)	(0.05)	(0.05)	(0.12)	(0.12)	(0.12)
Bond IL	-0.72***	-0.40***	-0.49***	-1.57***	-0.58***	-1.22***
	(0.05)	(0.04)	(0.07)	(0.08)	(0.05)	(0.11)
Avg Basket Size	6.44***	6.40***	6.50***	7.19***	7.13***	7.27***
	(0.69)	(0.69)	(0.70)	(0.74)	(0.74)	(0.75)
Shared Index Baskets	39.89***	39.95***	39.80***	65.02***	65.15***	65.07***
	(2.81)	(2.82)	(2.81)	(2.96)	(2.96)	(2.95)
ETF Ownership	6.04***	6.05***	6.12***	10.52***	10.62***	10.64***
	(0.29)	(0.29)	(0.29)	(0.33)	(0.33)	(0.34)
Zero Trading Days	-13.54***	-13.47***	-13.74***	-29.72***	-30.06***	-30.39***
	(0.73)	(0.74)	(0.75)	(1.07)	(1.08)	(1.09)
Observations	3392825	3305907	3231578	3392825	3305907	3231578
Adjusted R^2	0.50	0.50	0.50	0.42	0.42	0.42

(b) Basket Shares

	RD Basket			CR Basket		
	(1)	(2)	(3)	(4)	(5)	(6)
RD Instrument	9.91***	10.02***	9.81***	-2.29***	-2.34***	-2.31***
	(1.05)	(1.05)	(1.05)	(0.43)	(0.44)	(0.44)
CR Instrument	-6.11***	-6.03***	-6.13***	-18.86***	-18.79***	-18.83***
	(0.58)	(0.58)	(0.59)	(0.85)	(0.85)	(0.85)
Bond IL	-3.59***	-2.72***	-2.29***	-4.61***	-2.38***	-3.80***
	(0.40)	(0.36)	(0.50)	(0.48)	(0.37)	(0.76)
Avg Basket Size	51.31***	51.02***	51.89***	56.02***	55.40***	56.44***
	(6.12)	(6.13)	(6.19)	(5.58)	(5.57)	(5.64)
Amount Outstanding	66.26***	66.82***	67.30***	105.99***	106.85***	107.12***
	(2.92)	(2.93)	(2.94)	(2.66)	(2.66)	(2.67)
Shared Index Log Shares	34.31***	34.34***	34.18***	38.81***	38.76***	38.64***
	(2.41)	(2.42)	(2.41)	(1.84)	(1.84)	(1.83)
ETF Ownership	54.55***	54.49***	55.21***	74.45***	74.63***	74.94***
	(2.56)	(2.55)	(2.57)	(2.01)	(2.00)	(2.02)
Zero Trading Days	-57.29***	-55.88***	-55.93***	-109.70***	-111.54***	-109.50***
	(4.61)	(4.67)	(4.69)	(5.17)	(5.29)	(5.25)
Observations	3392825	3305907	3231578	3392825	3305907	3231578
Adjusted R^2	0.43	0.43	0.43	0.38	0.38	0.38

Table A.34: IV Basket Inclusion on Bond Liquidity

This table is the counterpart of Table 7 in the paper, with all additional controls.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-14.36***	-16.63***	-7.27	-1.13**	-1.58**	-0.58
	(5.36)	(5.98)	(4.57)	(0.55)	(0.62)	(0.48)
CR	-6.00***	-6.34**	-9.94***	-1.65***	-1.13***	-1.83***
	(2.20)	(2.50)	(1.85)	(0.32)	(0.37)	(0.28)
Bond IL	11.65***	11.12***	17.96***	10.26***	10.92***	17.44***
	(0.11)	(0.09)	(0.10)	(0.09)	(0.09)	(0.10)
Avg Basket Size	0.75*	1.25***	-0.21	0.56	1.03***	-0.28
	(0.42)	(0.47)	(0.36)	(0.36)	(0.40)	(0.31)
Shared Index	7.49***	10.91***	7.71***	0.97***	1.05***	0.80***
	(2.20)	(2.51)	(1.91)	(0.18)	(0.20)	(0.16)
ETF Ownership	-2.05***	-0.44	-1.46***	-0.86***	-0.03	-0.84***
	(0.37)	(0.43)	(0.32)	(0.33)	(0.38)	(0.29)
Zero Trading Days	0.49	-11.00***	-12.66***	-20.97***	-18.20***	-23.07***
	(1.10)	(1.26)	(0.99)	(0.67)	(0.76)	(0.60)
Amount Outstanding				-11.74***	-3.69***	-5.02***
				(0.40)	(0.47)	(0.35)
Observations	2020878	1754096	1803889	2020878	1754096	1803889
Adjusted R^2	0.01	0.01	0.02	0.01	0.01	0.01

Table A.35: Basket Inclusion on COVID Bond Liquidity

This table is the counterpart of Table 8 in the paper, with all additional controls.

	Number of Baskets			Basket Shares		
	IL1	IL2	IL3	IL1	IL2	IL3
RD	-0.13	3.38***	2.23	0.23*	0.29*	0.41*
	(1.03)	(1.18)	(1.55)	(0.12)	(0.14)	(0.21)
CR	-2.86**	-1.05	0.22	0.03	-0.14	0.23
	(1.16)	(0.79)	(1.61)	(0.13)	(0.10)	(0.18)
Bond IL	8.15***	7.89***	13.44***	7.84***	7.86***	13.41***
	(0.68)	(0.56)	(0.68)	(0.64)	(0.55)	(0.68)
Avg Basket Size	-0.23	-2.09**	-1.63	-1.18	-1.82*	-2.15
	(1.13)	(1.00)	(1.96)	(1.13)	(1.03)	(1.93)
Shared Index	-11.21***	-1.26	-18.02***	-1.05***	-0.08	-1.59***
	(2.83)	(2.77)	(3.65)	(0.24)	(0.23)	(0.35)
ETF Ownership	0.83	1.39*	1.21	0.76	1.34*	0.92
	(0.93)	(0.69)	(1.39)	(0.92)	(0.67)	(1.35)
Zero Trading Days	-57.49***	-16.11***	-121.99***	-70.62***	-11.56**	-122.02***
	(6.66)	(4.38)	(12.47)	(5.17)	(4.59)	(10.13)
Amount Outstanding				-8.85***	2.98**	-0.63
				(1.48)	(1.23)	(2.52)
Observations	111704	100311	101375	111704	100311	101375
Adjusted R^2	0.27	0.07	0.36	0.27	0.07	0.36

Table A.36: Basket Imbalance on Bond Liquidity

This table is the counterpart of Table 9 in the paper, with all additional controls.

	(1) IL1	(2) IL2	(3) IL3
Num Baskets	-8.82***	-3.95***	-5.44***
	(0.37)	(0.40)	(0.36)
Num Baskets × CR Imbal	1.35***	1.76***	1.33***
	(0.23)	(0.23)	(0.20)
Num Baskets × RD Imbal	2.64***	1.55**	2.58***
	(0.51)	(0.69)	(0.63)
Bond IL	12.41***	11.73***	19.45***
	(0.18)	(0.19)	(0.34)
Avg Basket Size	-0.06	-1.22***	-1.49***
	(0.26)	(0.27)	(0.26)
Shared Index Baskets	0.42	1.59***	0.77
	(0.54)	(0.51)	(0.47)
ETF Ownership	-2.60***	-1.46***	-2.46***
	(0.22)	(0.20)	(0.21)
Zero Trading Days	0.67	-9.38***	-9.44***
	(1.02)	(1.06)	(1.30)
Observations	3255850	2832929	2902571
Adjusted R^2	0.23	0.13	0.44

A.5.8 Basket Imbalance

Table A.37: Basket Imbalance: Interaction with a Simpler Measure

This table is the counterpart of Table 9 in the paper, except that it considers a simpler notion of basket imbalance that is defined in equation (20).

	(1) IL1	(2) IL2	(3) IL3
Num Baskets	-10.94***	-4.29***	-6.61***
	(0.41)	(0.41)	(0.41)
Num Baskets × Imbalance	2.00***	1.86***	1.79***
	(0.22)	(0.23)	(0.19)
Bond IL	12.52***	11.78***	19.59***
	(0.19)	(0.19)	(0.35)
Avg Basket Size	-0.43*	-1.28***	-1.89***
	(0.26)	(0.27)	(0.26)
Observations	3254055	2831031	2901286
Adjusted R^2	0.23	0.13	0.44