The Hidden Cost of Corporate Bond ETFs

Chris Reilly*

June 27, 2022

Click Here

for the latest version

Abstract

I document a hidden but substantial cost associated with the liquidity transformation that corporate bond exchange-traded funds (ETFs) provide. When creating new shares, authorized participants (APs) deliver a subset of the portfolio of bonds that underlie a corporate bond ETF. This subset contains bonds that realize low future returns, reducing ETF performance by 48 basis points per annum. This loss in performance cannot be attributed to forgone compensation for risk or illiquidity, but instead results from APs utilizing information regarding future changes in bond values to strategically deliver bonds when those bonds are expected to realize poor performance in the near future.

*Chris Reilly (christopher.reilly@utdallas.edu) is at the University of Texas at Dallas. I thank Caitlin Dannhauser, Jeff Pontiff, Edie Hotchkiss, Slava Fos, Max Clarke, John Griffin, Daniel Taylor, Ananth Madhavan, Jayoung Nam, Elisabeth Kashner, Adriana Robertson and seminar participants at the UT Austin PhD Student Symposium, Charles Schwab, Boston College, Southern Methodist University, Vanderbilt University, University of British Columbia, University of Toronto, Texas Christian University, Northeastern University, Rice, University of Texas at Dallas, Texas AM, University of Kentucky, Tulane, the Federal Reserve Board of Governors, Indiana University, HEC Paris, MIT, Arizona State University, the University of Chicago Booth School of Business, the Midwest Finance Association for helpful comments and discussion. I would additionally like to thank Caitlin Dannhauser for providing code that greatly assisted me in writing the paper.

1 Introduction

The ETF market is large and growing rapidly. Between 2010 and 2020, ETF assets under management increased from 1.3 to 7.7 trillion dollars. By 2021, ETFs surpassed passive index-tracking mutual funds in size. This growth was supported by the creation of numerous new funds that hold illiquid assets such as corporate bonds. Despite underlying asset illiquidity, ETFs remain relatively liquid, providing a valuable liquidity transformation service.

Under standard share-creation rules, ETFs utilize in-kind exchange of all underlying assets to allow Authorized Participants (APs) to conduct trades that arbitrage differences in ETF net asset values (NAVs) and secondary-market prices. Underlying asset illiquidity presents a challenge to ETF security design: ETF managers believe that standard creation rules coupled with portfolios that consist of numerous illiquid assets will prevent APs from conducting arbitrage trades that are necessary to improve ETF price efficiency. Corporate bond ETF managers therefore allow APs to deliver only a subset of the assets in portfolios that ETFs wish to track. These rule modifications embed a hidden cost that ETF investors incur in a manner not previously recognized by the literature. Share creations shift ETF holdings in a way that overweights delivered assets and underweights undelivered assets. I investigate whether bonds delivered to ETFs by APs subsequently underperform and to what extent this contributes to fund underperformance.

APs deliver corporate bonds that underperform bonds that are held by the same ETF that are not delivered during the creation process by 2.4 basis points. This causes ETFs on average to underperform their stated benchmarks following a creation event by 1.4 basis points. Given the frequent occurrence of creations within my sample, this results in a 48 basis-point-per-annum cost to the average corporate bond fund, which is higher than the 35 basis point average net expense ratio reported by the same funds. The valuable liquidity transformation provided by corporate bond ETFs is therefore not a "free lunch." I document a novel trade-off that ETF managers face. By allowing their APs greater flexibility, ETF managers encourage APs to more aggressively arbitrage tracking errors to the benefit of ETF investors while simultaneously allowing APs to interact strategically with ETF portfolios at the expense of ETF investors. My sample consists of funds that manage over 200 billion dollars in assets in 2019 and the mechanism described applies as well to other classes of ETFs that also utilize modified creation rules.

My empirical strategy leverages the fact that bonds delivered by APs to ETFs are accompanied by a natural control population: bonds held by the same ETFs but are not delivered during that specific creation event. Using this natural control group, bond fixed effects, fund-by-date fixed effects, and additional controls, I estimate the performance of delivered bonds relative to that of undelivered bonds in periods around deliveries versus the relative performance of those same bonds during periods that lack deliveries. My empirical method is similar to a differences-in-differences framework. This identification approach illustrates that delivered bonds do not underperform simply because ETFs hold low-risk or high-liquidity subsets of underlying benchmarks and receive lower compensation for their lower risk. Instead, ETF holdings shift in the time series and ETFs increase holdings in bonds when the short-term future return on those bonds is expected to be lower.

To estimate the size of such the hidden cost that is borne at the fund level, I estimate the performance of ETFs relative to that of their stated benchmarks conditional on experiencing a creation event relative to the performance of the same ETF when it does not experience a creation event. Specifically, I calculate spreads between ETF performance and underlying benchmark changes and regress them on a binary variable that indicates whether a creation event occurred. By investigating relative performance spreads with high-dimensional fixed effects, I am able to rule out many alternative explanations. Most importantly, I can reject the null hypothesis that such performance can be explained solely by the arbitrage motive for APs' trades and deliveries. The arbitrage trading that share creations are designed to facilitate implies that differences in long-run spreads should be mean zero. While arbitrage opportunities are still of the first order in APs' trade decisions and in explaining the time-series of spreads, I document a two-month underperformance that cannot be explained by arbitrage opportunities alone and instead represents the manifestation of a hidden cost.

I also investigate the underlying mechanism by which delivered bonds underperform. I consider two main sources of bond performance that embed the cost ETFs incur. First, asset values used during settlement may not accurately reflect underlying market-microstructure dynamics and may allow APs to settle with temporarily over-valued assets. Second, APs may utilize information, such as bond momentum or private information inferred from their own balance sheets in related market-making businesses, to predict future changes in asset values and offload securities that will be more costly to hold on their balance sheets. While both mechanisms make similar predictions of post-delivery underperformance by delivered bonds, they make contrasting predictions regarding pre-delivery performance. Bonds underperform prior to delivery, consistent with the utilization by APs of public and private information that is not reflected in asset values. Lastly, I consider one possible source of information: order flows from dealers related market making business. Consistent with the nexus of information being informed customers, I observe aggregate inventories raise in the week before a bond is delivered to an ETF but fall on the date of delivery.

I perform a number of robustness exercises that confirm the finding of a hidden cost. First, instead of relying on the idea that APs are incentivized to arbitrage cumulative spreads to mean zero over long horizons, I control for the main determinants of APs' arbitrage opportunities: premiums and discounts. In so doing, I isolate the spread dynamics that are attributable to the hidden cost and find them to be significant. Second, I consider spreads between the performance of ETFs' portfolios and that of underlying indexes. Since the hidden cost results from the fact that corporate bond ETFs hold portfolios that differ from target benchmarks, underperformance should be observable in NAVs, not only in the secondary-market prices of ETFs. Third, I show that ETF characteristics such as underlying liquidity and the size of a creation basket predict the size of the hidden cost, a finding that is consistent with the mechanism I posit. Last, I utilize ETFs that hold US equity in companies with large market capitalization (Large-Cap ETFs) that require all assets to be delivered during creation events as a sample for falsification tests. I find no empirical evidence of fund underperformance in this class of ETFs during creation events, a finding that is consistent with the idea that the hypothesized hidden cost drives corporate bond ETF underperformance.

My paper contributes to three major strands of literature. First, I contribute to the literature that investigates the costs of liquidity transformation in the investment management industry. In open-end mutual funds, entering-and-exiting investors demand for liquidity results in transaction costs that are borne by all investors (Edelen (1999)). These costs rise in the presence of stale NAVs. Investors can time their entry and exit decisions strategically at the expense of other investors (Chalmers et al. (2001)). The accuracy of NAVs attracted particular attention in the late 1990s and early 2000s, but has received renewed scrutiny with regard to fixed-income mutual funds (Zitzewitz (2006) and Choi et al. (2022)). It is commonly believed that the ETF mechanism design protects ETF buy-and-hold investors from the liquidity demands of other investors.[1] I document that many ETFs consistently pay a large cost for liquidity transformation in a manner that has not been recognized previously in the literature. This cost is not evident in tracking errors or premiums because such metrics describe the relative performance of ETF NAVs and secondary-market prices, while the cost instead arises as a result of wedges between ETF and benchmark portfolios. Additionally, because ETF creations are settled using the values that determine NAVs, this cost is theoretically larger in the presence of inaccurate bond pricing.

Prior literature has also considered implicit costs associated with the liquidity transformation that ETFs provide. In doing so, academics and policymakers have overwhelmingly focused on the potential for financial fragility in ETFs. Inspired by canonical models such as that proposed in Diamond and Dybvig (1983) and theoretical understanding of fragility

[1] For example, etf.com states that the "system is inherently more fair than the way mutual funds operate. In mutual funds, existing shareholders pay the price when new investors put money to work in a fund, because the fund bears the trading expense. In ETFs, those costs are borne by the AP (and later by the individual investor looking to enter or exit the fund)." (ETF.com (2021)).

in open-end mutual funds (e.g. Chen et al. (2010a)), many have considered whether ETF portfolios may be exposed to runs (see Pan and Zeng (2019); Dannhauser and Hoseinzade (2021); Haddad et al. (2020); and Ma et al. (2020)). Runs may result in persistent periods of mispricing, impacts on underlying asset markets, or spillovers to related asset markets. In its focus on financial fragility, the literature has overlooked another large and important cost: rule modifications necessary in face of underlying asset illiquidity result in performance drags on ETFs borne in good times. While they are theoretically plausible, the realization of permanent costs that are borne by buy-and-hold ETF investors due to financial fragility have been difficult to document empirically. The cost introduced in this paper has already been incurred and is shown empirically to be high. Koont et al. (2022) consider a similar trade-off faced by ETF managers when transforming liquidity. Theoretically, ETF managers accept deviations from the benchmark portfolio despite increased volatility of benchmarked performance in exchange for greater AP activity. I document that the deviations also on average lower the mean of benchmarked performance.

Second, since Jensen (1968), a strand of literature has assessed the relationship between flows in the investment-management industry and the subsequent performance of investment funds. Although such tests were originally interpreted as pertaining to the assessment of both the skills of investment managers and the rationality of investors in delegating capital to those managers, drawing inferences from such tests is difficult, for two reasons. First, if active management experiences decreasing returns to scale, in equilibrium flows will not predict future performance despite being motivated by managers possessing skills (Berk and Green (2004)). Second, open-end mutual funds must engage in transactions when they experience flows that incur transaction costs (Edelen (1999)). Despite such challenges, a number of papers have investigated the relationship between flows and future performance, often after accounting for transaction costs (for example, see Zheng (1999); Daniel et al. (1997); Wermers (2000); Edelen and Warner (2001); Frazzini and Lamont (2008); and Friesen and Sapp (2007)). I introduce a mechanism that relates ETF flows to short-run future performance in a manner that is consistent with Edelen (1999).[2] My paper introduces the need for work that assesses the market-timing ability of ETF investors to account for indirect costs of flows (akin to the mutual fund literature's accounting for transaction costs).

In related work, researchers have advocated utilizing investment-management flows as proxies for non-rational shifts in investor demand curves. Numerous studies have shown that investment-management flows are predictive of underlying asset returns (See for example Sirri and Tufano (1998); Coval and Stafford (2007); Frazzini and Lamont (2008); Ben-Rephael et al. (2012); Ben-David et al. (2018); Dannhauser (2017); and Doan (2020)). Specifically, ETF flows have been shown to negatively predict future fund and underlying asset returns (Brown et al. (2021)). Prior literature has failed to identify rational-agent-based explanations and instead has posited that flows act as a proxy for investor sentiment when explaining such return predictability.[3] I introduce for the first time a fully rational mechanism that relates ETF flows to short-term future returns. I also present new evidence pertaining to the relationship between ETF flows and future performance, as I measure performance as spreads between fund and benchmark returns rather than raw returns. Demand shocks suggested in the relevant prior literature would impact both ETF and underlying asset markets and thus would not be observable in spreads. The mechanisms I introduce also suggest an alternative explanation for a portion of the short-term return predictability that has been documented previously and helps to reconcile the return predictability documented in the sample utilized in Brown et al. (2021) and the lack of return predictability in the sample utilized by Dannhauser and Pontiff (2019): The former sample includes leveraged ETFs and exchange-traded notes that are more likely to use partial creation baskets or cash settlement.

[2]Edelen (1999) finds a relationship between flows and contemporaneous returns at a frequency of six months. I show that daily flows are associated with the following month's daily returns. The difference between focusing on contemporaneous and focusing on future returns reflects the different frequency that I utilize for estimation.

[3]Studies also investigate the relationship between past returns and future flows, (see e.g. Ippolito (1992) and Berk and Green (2004)). Clifford et al. (2014) and Dannhauser and Pontiff (2019) present evidence that flow-to-performance is stronger in ETFs than in active investment products, which is hard for many rational models to reconcile. My paper does not shed light on this empirical relationship to past returns and instead investigates the relationship between flows and contemporaneous or future returns.

In summary, I present a novel mechanism that helps to rationalize previous puzzles regarding the information content of ETF flows.

Third, my paper contributes to a very new and growing strand of literature that considers the incentives imposed by rules governing ETF creations. ETF managers face trade-offs when attempting to minimize tracking errors: flexibility encourages more creation and redemption activity and thus smaller premiums and discounts but shifts their portfolio away from target benchmarks and thus may increase the volatility of benchmarked performance (Koont et al. (2022)). On the other side of the transaction, APs may behave strategically to utilize the implicit liquidity provided by the share-creation process to offload inventory in periods of market turmoil after experiencing balance-sheet shocks. Such behavior could expose ETF portfolios to financial fragility (Pan and Zeng (2019)). Conversely, when markets are stressed, ETF managers may wish to avoid redemptions as their sale of received bonds may lower bond values and exacerbate balance sheet constraints. APs may thus help prevent fire sales in underlying bond markets (Shim and Todorov (2021)). While AP and ETF-manager incentives carry implications for financial fragility, I demonstrate that they also embed a consistent cost that is borne outside periods of market stress. This cost is motivated by the same strategic considerations that occupy APs as those discussed in Pan and Zeng (2019) but result in new and novel effects. The complimentary findings mirror those reported in the mutual fund literature. Pan and Zeng (2019) illustrate that ETFs face the same financial fragility theorized in Chen et al. (2010b) while I illustrate that ETF investors can pay consistent rents in a manner similar to those described by Chalmers et al. (2001). While stale values may benefit ETF managers when markets are stressed (Shim and Todorov (2021)), they impose a large and offsetting cost during normal times. My paper suggests that, even if financial fragility is not empirically likely, APs' incentives to interact strategically with ETF portfolios at the expense ETF investors must be considered when evaluating the liquidity transformation that ETFs provide.

In addition to these contributions to the academic literature, my paper also bears direct implications for policymakers and ETF managers. First, as discussed previously, my paper is the first to establish that ETF performance is affected by the accuracy of asset values. Recently, mutual funds have received renewed scrutiny over the practices they use to set values that determine NAVs. Many have argued that ETF-creation mechanisms render them immune to similar issues. My paper illustrates the flaw in such reasoning when applied to large subsets of ETFs and suggests that such ETFs should not be exempt from NAV-pricing regulations. Second, in 2019 the U.S. Securities and Exchange Commission (SEC) approved Rule 6c-11, which allows all ETFs to utilize custom creation baskets. When enacting such a rule, the SEC acknowledged the theoretical risk that custom creation baskets pose to ETF investors, but concluded that they are unlikely to occur in practice.[4] I present evidence that they do occur and the cost borne by investors is high. Nonetheless, investors receive valuable AP activity in exchange for incurring such costs and thus investors may knowingly and willingly accept the costs documented in the paper. If sophisticated ETF investors are aware of the dynamics documented in this paper, a simple revealed preferences argument allows one to conclude that the benefits of custom creation baskets must outweigh the costs for those investors.

Lastly, my paper helps resolve open questions regarding the organization of the investment management industry. ETFs provide many benefits, including empirically lower fees (Kostovetsky (2003)), lower tax obligations (Moussawi et al. (2022)), and the liquidity transformation on which this paper focuses. Researchers have identified very few costs associated with ETFs relative to those associated with mutual funds that would offset such benefits. While ETFs have grown massively, open-end mutual funds still represent the majority of funds in the investment-management industry. By documenting a novel and large cost, I help explain why ETFs do not represent the dominant security design, particularly when in-

[4]In rule 6c-11, the SEC defines "custom creation baskets" as distinct from "sampling baskets," the use of which was already permitted. The cost documented in this paper can result from both types of creation baskets: APs can interact strategically with any creation basket that does not exactly match the benchmark portfolio. Obviously, when APs are allowed to request custom creation baskets, the level of adverse selection and thus the cost is likely higher.

vestors wish to hold illiquid assets. My findings help extend the framework of Chordia (1996) to the ETF industry by showing that underlying asset liquidity is a first-order determinant of optimal security design. ETFs likely still present an attractive fund structure for many investors and while ETFs do underperform their benchmark by greater than their stated net expense ratios, ETFs have other offsetting sources of higher returns such as the participation in primary bond offerings or the ability to "reach for yield" and thus underperform by less than the 48 basis point hidden cost. The cost documented in this paper is relative to a counterfactual in which ETFs demanded all assets during creation and thus provided less liquidity transformation benefits for investors.

2 The Share-Creation Process

To provide a number of advantages, such as intraday liquidity, ETFs are designed to be traded on secondary markets. Insofar as ETFs will have secondary-market prices that do not necessarily equal the value of their underlying assets and because the vast majority of investors will purchase ETFs via this secondary market, ETFs require a mechanism that enables them to remove large discounts or premiums that differentiate net asset values from ETF share prices.[5] To facilitate the elimination of discounts and premiums, ETFs enter agreements with APs, who are allowed to create and redeem ETF shares by exchanging the assets that underlie an ETF share for a share of the ETF or vice versa. This process is referred to as "in-kind exchange." Consider a highly stylized and simplified example: suppose investors wish to hold an ETF that tracks the S&P 500 and purchase a sufficiently large quantity of the SPY ETF to positively impact SPY's secondary market price. This creates a premium in which the ETF trades at a higher price than the underlying assets. APs can seek to arbitrage this premium away by buying the underlying assets and selling short shares of the ETF. An AP notifies an ETF manager that it wishes to create a share

[5] This mechanism is the main security design innovation that distinguishes ETFs from close-end mutual funds.

as outlined in their agreement. At the end of the trading day, the AP delivers all of the underlying assets, receives newly created ETF shares and utilizes these shares to cover the short-sale position, thereby locking in the arbitrage profit. Because all relevant shares of the S&P 500 companies still underlie the SPY ETF, investors who hold the ETF throughout this entire process face no costs. Specifically, the ETF experiences a temporary tracking error induced by the premium but when the premium is arbitraged away the tracking error reverses and the buy-and-hold ETF investor receives a return that exactly tracks the performance of the S&P 500 gross of the ETF net expense ratio. Thus, when all underlying assets are exchanged in kind, ETF investors are not impacted by other investors' entry.

When ETFs hold fewer liquid assets or a very large number of assets, however, it is often difficult or overly cumbersome to exchange all underlying assets in kind when seeking to create shares. With corporate-bond ETFs in particular, a very small notional amount of a given bond may underlie each share, making it logistically unwieldy to locate and deliver such a small amount of a given bond. To overcome these barriers, facilitate creation activity, and thus remove premiums, ETF managers instead often allow creation to occur with a select subset of assets that they believe is representative of the full portfolio of underlying assets. For example, if a corporate bond ETF holds 1,000 separate bonds, they might at the start of the trading day identify 100 of those bonds as the "sampling basket,", also referred to as the "creation basket," which is utilized for share creation. In such a case, APs can deliver only these 100 bonds at the end of the day while delivering "cash-in-lieu" for the remaining underlying assets in exchange for a share of the bond ETF. Depending on the AP agreement established by an ETF manager, an AP can actually deliver cash in conjunction with the sampling basket. Much more commonly, APs can instead deliver the sampling basket alone and receive ETF shares on a pro-rata basis based on the NAV of the sampling basket versus the NAV of the ETF share.[6]

At the start of a trading day, ETF managers will disseminate to their APs which assets

[6]For example, if the sampling basket of 100 shares represented 10% of the NAV of the ETF share, they could deliver 10 shares worth of the sampling basket in exchange for one ETF share.

belong to that day's baskets via a clearing house such as the Depository Trust & Clearing Corporation. ETF managers have incentives to try to match the baskets to the overall portfolios that they wish to hold along key dimensions such as duration and credit risk. Specifically, to the extent that ETF managers believe there is a factor structure that characterizes bond returns, ETF managers have incentives for creation baskets to contain the same total factor exposure as their target portfolios. They also may consider asset liquidity and notional sizes to encourage more aggressive arbitrage activity, thus reducing premiums and discounts. Additionally, APs form relationships with ETF managers and may call them to request "custom creation baskets." They may request the omission of certain assets from baskets that they cannot readily locate or suggest alternative assets. ETF managers have the right to accept these custom creation baskets at their discretion and often do so in practice. Relatedly, ETF managers sometimes utilize custom creation baskets to facilitate "heartbeat trades," as documented in Moussawi et al. (2022). While ETF managers have incentives to ensure that creation baskets are fairly priced and match their target portfolios along key dimensions, many of them have longstanding relationships with APs and rely on APs to arbitrage away premiums and discounts that may deter future fund flows. Thus, it may often not be in ETF managers' interests to reject creation activity even if doing so embeds a cost that their customers bear.

When pro-rata settlement occurs, which is the norm for corporate bond ETFs, creations shift ETF portfolio weights at the point of creation. The portfolio weights of delivered assets naturally increase as they now hold more of these bonds. Portfolio weights as a percentage of an ETF's portfolio decrease for assets that are not delivered because the holdings of those assets are unchanged but the total assets under management of the fund increase. These shifts also imply that ETFs often hold portfolios that differ from the indices that they track. Naturally, one would expect ETF managers and APs to select larger and more liquid bonds for inclusion in the delivery basket as they are likely cheaper to secure. Thus, it is possible that many funds are systematically overweighted in liquid assets and thus ETFs fail to capture a liquidity premium but in turn face lower risk. While this may also be occurring, I produce the results of this paper while controlling for such persistent compositions of ETF portfolios and instead investigate whether the timing of portfolio shifts imposes a cost on ETF investors.

Bond-level estimates of performance control for time-varying bond liquidity and bond fixed effects and thus reveal whether bonds delivered to ETFs perform worse around delivery than when they are not delivered. Relatedly, fund-level results include fund fixed effects that would capture time-invariant portfolio differences between ETFs and their benchmark indices. Thus, the hidden cost described does not simply reflect differences in the risk/reward profiles of ETFs' actual portfolios as opposed to those of their benchmark portfolios. Instead, it reflects the nature of creations in which ETFs dynamically receive assets with lower expected future returns in such a way that creation events negatively impact the performance of ETFs relative to that of their benchmarks.

Insofar as pro-rata settlements are based on the values of underlying assets, it is important that ETFs set timely and correct values for underlying assets, as would be the case with an open-end mutual fund. If these values are not timely and accurate, APs can strategically time creations and redemptions in the same way as open-end mutual fund investors can strategically time their investments to profit from stale NAVs. Indeed, even if values that are used to price ETF assets are timely and accurate, APs may utilize information (public or private) that is not reflected in current prices when deciding which assets to deliver or when to deliver pre-specified assets. They may wish to offload inventory that will be costly to hold (based on lower expected future returns) or liquidate bonds in exchange for ETF shares that they can sell more easily. In so doing, ETFs will be left with adversely selected sets of assets, specifically assets that conditionally possess lower expected returns than the unconditional average expected returns on those assets. This will result in ETF underperformance relative to that of the underlying indexes that they state as their benchmarks.

When shares are created pro-rata, any market force that causes delivered assets to underperform non-delivered assets held by the same ETF will result in a performance drag on an ETF. Both incorrect values and the utilization of private information by APs will result in equivalent performance drags because both make the same predictions regarding post-delivery asset performance. Nonetheless, the two mechanisms may imply distinct remedies and hypothesize divergent asset performance prior to delivery. Values are often incorrect when assets that APs purchase to deliver temporarily impact prices in a way that is reflected in those bond values. Such temporary price impacts would on average revert, resulting in bond underperformance. Thus, bond values that fail to reflect market micro-structure forces predict that delivered bonds will outperform other bonds prior to delivery (with that outperformance attributable to the temporary price impact). If instead APs utilize information that is not reflected in bond values, there is no clear hypothesis to explain the performance of delivered assets. If APs utilize information indicating price reversions (not caused by their own trading), the predicted dynamics are similar to those that occur when values incorrectly reflect temporary price impacts. Instead, if APs utilize information embedded in past prices and bond values, such as bond momentum as in Jostova et al. (2013), bonds will underperform both prior to and following delivery. Lastly, if the information utilized by APs is not reflected in prices, i.e. it is private information acquired from their own balance sheets in related market-making businesses, delivered bonds may perform no differently prior to creation but underperform following creation. The results of my analysis reveal underperformance prior to delivery that is consistent with the utilization of information by APs, and is thus consistent with APs' taking advantage of bond-momentum information as well as private information.

ETF managers often recognize that the bond values utilized in these exchange processes fail to fully reflect the transaction costs that must be paid to re-balance an ETF portfolio, and thus charge a "fee on cash" for an actual cash settlement. This is particularly necessary because bond values are often priced using market bid prices rather than ask prices. Relatedly, perhaps because ETF managers know that creation events may impose costs on

ETF investors, ETF managers also charge flat creation or redemption fees to APs. Lastly, ETF managers have the right contractually to refuse creation or redemption if they deem it unfavorable to their funds, although this right is rarely exercised. It remains an empirical question whether fees are sufficiently large to offset or erase the hidden cost of share creation. My results suggest instead that fees on cash are not sufficiently high to fully offset this cost and thus ETF investors are affected by fund flows. ETF managers also face a difficult tradeoff between protecting investors from this hidden cost and encouraging APs to arbitrage discounts and premiums away. While it is possible that this hidden cost is understood by ETF investors who view it is as fair compensation for reduced tracking error, for many corporate bond ETF investors this cost may not be particularly salient as it does not show up in many key fund statistics such as tracking errors and expense ratios. Thus ETF managers who hope to attract fund flows may find it useful to pay a high but non-salient cost in exchange for the small but salient benefit of reduced tracking errors. The rationales that ETF investors follow are not observable, so it is impossible to distinguish between these two theories, but comparing this hidden cost with the explicit costs of administering a fund in the context of an established strand of literature on the behavioral biases of mutual fund investors suggests that some ETF investors may not be making informed decisions regarding these costs.

Unlike share creations, redemptions are much less likely to involve pro-rata settlements and thus do not embed the same cost as creations. It is much easier logistically for ETF managers to provide APs with portfolios that consist of all the assets that underlie an ETF rather just a subset as they do not have to enter the market and execute transactions to secure those assets. Additionally, APs may not need to sell those assets immediately, instead adding them to their balance sheets that they hold in their related market-making businesses. Thus, in practice, redemptions exert a far smaller distortionary effect on the portfolios of ETFs than creations and instead tend to deviate when doing so is optimal for ETF investors (such as when they wish to cycle out a bond that is coming to maturity to avoid transactions). As a result of this asymmetry between creations and redemptions, the tests conducted for this paper focus on creations and not simply flows of either kind. An empirical analysis of the dynamics surrounding redemptions can be found in the appendix.

3 Sample and Data

I construct a two samples with which to test for the existence of the hidden cost I have posited. I use the first sample to estimate the performance of bonds around delivery during creation events. I begin with the universe of holdings data of funds with "Fixed Income" as their asset class and "Corporate Bonds" as their category from ETF Global from the period of 2012 through 2019. ETF Global captures data reported by APs and ETF managers on creation activities and fund holdings. I exclude leveraged or inverse ETFs, as their buy-and-hold return characteristics are affected by other factors such as costs associated with leverage and realized benchmark volatility. Because I lack data that enable me to observe the stated creation baskets of ETFs or custom creation baskets, I infer what was delivered by measuring changes in ETF holdings. Thus, I assume that ETFs do not trade themselves and instead rely on creation and redemption events to manage their portfolios.[7] Such an assumption is in line with the motives ETFs possess to defer taxes (in-kind exchanges that occur during creations are not taxable events, whereas trades made by ETFs do create tax obligations for investors). Thus, in my sample of bonds, I rely solely on bonds with non-missing notional values to observe changes in holdings that are untainted by any measurement error in bond returns. Lastly, I exclude ETF dates in which no assets change value to remove the cases in which stale holdings are reported multiple times.

[7]Shim and Todorov (2021) and Koont et al. (2022) rely on the same methodology to identify the realized composition of creation baskets. Shim and Todorov (2021) provide evidence that ETF flows inferred from changes in holdings are highly correlated with observed ETF flows providing evidence that the use of holdings data to infer creations baskets is valid. If ETFs instead cash settle and in turn purchase assets, my tests may instead identify ETFs incurring transaction costs rather than costs due to AP's strategic behavior with similar consequences to fund performance. The occurence of transaction costs, such as temporary price impact, predict underlying bonds to outperform as they are purchased by ETFs which is not observable in the underlying bond data, further suggesting that changes in holdings act as a valid measure of realized creation and redemption baskets.

I augment these data with bond returns calculated using trades reported in TRACE and bond characteristics reported by FISD. I calculate bond returns as changes in dirty bond prices paid by customers who purchase bonds plus changes in accrued interest implied by bond coupon characteristics. Unlike in many other studies, in this study I calculate bond returns using prices that are implied by the most recent transactions inferred to hit bids rather than measuring the implied mid-point between bids and asks. Bond values are priced at dirty bid prices (with accrued interest accounted for during settlement); this return measure will therefore more accurately reflect the values used during settlements of ETF share creations.[8] If a bond increases its notional value while a fund increases in size, I classify the bond as "Delivered." I also calculate time-varying measures of bond liquidity utilizing TRACE data. I draw my liquidity measures from those utilized by Dannhauser (2017) in her construction of liquidity principle components. I calculate Amihud price impacts in a manner that is consistent with Amihud (2002), implied round-trip costs as described by Feldhütter (2012), and bid-ask spreads consistently with Hong and Warga (2000) and Chakravarty and Sarkar (2003). All liquidity values use the median measure from the calendar month prior to a creation event. Lastly, I calculate the change in aggregate bond dealer inventories as the net imbalances of customer to dealer trades reported over the prior 1 or 5 trading days.

As reported in Table 1, this creates a sample of 9,546,138 bond-by-fund-by-day observations across 9,819 bonds contained in 45,088 fund-by-day cross-sections over 1,258 trading days. Importantly, the same bond-day return may enter my sample multiple times if the bond is held by multiple funds but the fund characteristics assigned to that bond (such as delivered flags or fund fixed effects) will differ across these multiple observations.

Second, I create a sample of ETF performance measures, creation events, and ETF characteristics to investigate the hidden cost at the fund level. I begin with the universe of ETF Global, which reports daily creation and redemption activity via reporting the number of ETF shares outstanding. $Creation_{f,t}$ reports a simply binary indicating whether a fund

[8]In untabulated results I find that bond level dynamics are unchanged if mid prices are instead utilized to calculate returns.

experienced a net creation event on a given trading date (its shares outstanding increased). While it is possible that one AP creates shares while another redeems, this is unobservable and, theoretically, should be rare. While it is likely that on any given day factors other than the premium/discount (such as AP information sets or current balance sheets) would motivate one AP to create while another did not, it less likely that these factors would be so large as to make creation profitable for one AP while making redemption profitable for another. Unlike other data sources that are utilized in the literature to infer ETF creation and redemption activity, such as CRSP market data or CRSP mutual fund data, ETF Global captures data reported by APs and ETF managers on days of or following creation activity (depending on whether the particular fund reports on T or T+1) rather than a later share-settlement date. I again exclude leveraged or inverse ETFs. I utilize ETF Global to identify each fund's stated benchmark and restrict my sample to those with benchmarks, thus excluding the rare cases of actively managed ETFs. I obtain benchmark levels from Bloomberg and ETF return and dividend information from CRSP. As I report in Table 1, this yields a sample of 137 corporate bond ETFs over 1,887 trading days.

Table 1: Summary of Samples

Panel A of this table provides descriptive statistics for the sample of bonds analyzed in this study. $BondRet_{b,t}$ is the daily bond return implied by customer-to-dealer purchases reported in TRACE data. Bond returns cumulated over one month and two months are reported in $\prod_{n=-22}^{-1} BondRet_{b,f,t+n}$ and $\prod_{n=-22}^{20} BondRet_{b,f,t+n}$ respectively. $Delivered_{b,f,t}$ is a binary variable that takes the value of one if the notional value of a bond increases on a share creation day and zero otherwise. $Amihud_{b,t}$ is calculated in a manner that is consistent with Amihud (2002), $ImpliedRoundtripCost_{b,t}$ is calculated in a manner that is consistent with Feldhütter (2012), and $BidAsk_{b,t}$ is calculated in a manner that is consistent with Hong and Warga (2000) and Chakravarty and Sarkar (2003). All three liquidity values reflect the median value in the prior calendar month. Results for the same bond may be reported multiple times if the bond is held by multiple ETFs and fund-level variables such as $Delivered_{b,f,t}$ may vary across these observations. Panel B of this table provides descriptive statistics for the sample of analyzed ETFs. $Ret_{f,t}$ is the daily secondary-market return on an ETF. $NAVRet_{f,t}$ is the implied return on the underlying assets calculated as $NAVRet_{f,t} = \frac{NAV_{f,t}+Dividend_{f,t}}{NAV_{f,t-1}}$. $\Delta Index_{f,t}$ is the percentage change in the benchmark index. $RetSpread_{f,t,k}$ is the cumulative difference between an ETF return and the benchmark index cumulated through k days following the observation date: $RetSpread_{f,t,k} = \prod_{n=-22}^{k} ETFRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$. For example, $RetSpread_{f,t,-1}$ is the cumulative difference between the ETF and the index up to 1 day prior to observation (or share creation in the case of a creation event). $NAVSpread_{f,t,k}$ is the cumulative difference between the $NAVRet_{f,t}$ and the benchmark index cumulated through k days following the observation date. $Creation_{f,t}$ is a binary variable that takes the value of one if a fund's shares outstanding increased and zero otherwise. $Premium_{f,t}$ is the premium or discount as represented in percentage terms (%100 representing no discount or premium). All returns or spreads in both panels are reported in basis points.

Panel A: Underlying Bond Sample

			Percentile				
Variable	Mean	Std. Dev.	1st	25th	50th	75th	99th
$BondRet_{b,t}$	0.56	45.13	-179.80	-4.93	1.10	6.87	162.88
$\prod_{n=-22}^{-1} BondRet_{b,f,t+n}$	7.50	171.16	-580.24	-54.55	12.64	78.73	542.74
$\prod_{n=-22}^{20} BondRet_{b,f,t+n}$	18.56	246.31	-796.21	-88.26	20.68	129.41	769.37
$Delivered_{b,f,t}$	3.91%						
$Amihud_{b,t}$	0.04bp	0.05bp	0.00bp	0.00bp	0.02bp	0.04bp	0.27bp
$ImpliedRoundtripCost_{b,t}$	1.39bp	1.71bp	0.00bp	0.48bp	0.86bp	1.54bp	9.34bp
$BidAsk_{b,t}$	\$0.24	\$0.23	\$0.00	\$0.10	\$0.18	\$0.29	\$1.18
0.00	0.10	0.18	0.29	1.18			

Obs: 9,546,138; Trading Days: 1,258; ETF/Date Cross-sections: 45,088
Bonds: 9,819; ETF Bond Pairs: 84,781

Panel B: Fund Sample

			Percentile				
Variable	Mean	Std. Dev.	1st	25th	50th	75th	99th
$Ret_{f,t}$	1.5	29.2	-96.3	-10.7	1.0	14.7	94.7
$NAVRet_{f,t}$	1.5	23.0	-79.7	-6.4	1.1	10.4	77.1
$\Delta Index_{f,t}$	1.7	22.1	-76.7	-6.0	1.4	10.3	74.8
$Ret_{f,t} - \Delta Index_{f,t}$	-0.2	21.8	-72.7	-9.5	-0.2	9.1	73.6
$NAVRet_{f,t} - \Delta Index_{f,t}$	-0.1	16.1	-68.9	-1.9	-0.1	1.6	68.8
$RetSpread_{f,t,-1}$	-3.4	35.4	-126.2	-18.5	-2.8	11.6	117.3
$NAVSpread_{f,t,-1}$	-3.5	25.1	-112.7	-7.3	-1.9	2.0	89.9
$RetSpread_{f,t,20}$	-6.7	42.0	-157.4	-23.9	-5.5	10.8	136.2
$NAVSpread_{f,t,20}$	-6.8	31.6	-148.3	-12.4	-3.7	1.2	105.3
$Creation_{f,t}$	13.53%						
$Premium_{f,t}$	100.19%	0.42%	98.71%	99.99%	100.18%	100.37%	101.97%
$InvestmentGrade_f$	66.24%						
$AverageBasketSize_f$	64.9	91.7	1.0	16.9	40.3	75.6	600.9
$AverageAmihud_f$	0.04bp	0.02bp	0.00bp	0.03bp	0.04bp	0.04bp	0.09bp
$AverageIRC_f$	1.50bp	0.65bp	0.27bp	1.03bp	1.43bp	1.81bp	3.46bp
$AverageBidAsk_f$	\$0.30	\$0.15	\$0.05	\$0.21	\$0.26	\$0.36	\$0.80

Obs: 134,523; Trading Days: 1,887; Number of ETFs: 137

3.1 ETF Relative Performance

To estimate the underperformance of ETFs that is attributable to the hidden cost I describe, I calculate the following ETF spread metric:

$$RetSpread_{f,t,k} = \prod_{n=-22}^{k} ETFRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$$

This measure represents the net-of-fees buy-and-hold performance of an ETF share relative to its stated benchmark over a horizon of k days. This also represents the performance an ETF investor would actually achieve relative to the stated benchmark if that investor were to purchase and sell ETFs on the secondary market. This measurement is similar to tracking error measures used in the mutual fund or ETF literature but departs from them in a very significant way: it is signed. If an ETF experiences many mean zero daily tracking errors, this relative performance measure would be zero as the tracking errors would offset each other. Negative values of this measure represent negative drifts in tracking errors that harm investors by reducing returns rather than by increasing volatility. Because this is net-of-fee performance, one obvious driver of such underperformance relative to these untradeable benchmarks is the net expense ratio, including management fees, incurred by ETFs.

I calculate another spread in performance, the spread between NAVs and the underlying index, as follows:

$$\begin{aligned} NAVRet_{f,t} &= \frac{NAV_{f,t} + Dividend_{f,t}}{NAV_{f,t-1}} \\ &\approx \Delta NAV_{f,t} + DividendYeild_{f,t} \\ NAVSpread_{f,t,k} &= \prod_{n=-22}^{k} NAVRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n} \end{aligned}$$

This represents the net-of-fees buy-and-hold performance of the portfolio of assets that underlies an ETF share relative to that of its stated benchmark. Unlike the previous spread measure, this measure cannot be realized by a common investor (who is not an AP) who must purchase the ETF on the secondary market. Importantly, because creation and redemption activity seemingly seeks to arbitrage differences between ETF prices and NAVs, creation and redemption activity should have a very different effect on an ETF's relative performance and relative NAV performance. For example, creation activity should reduce ETF prices via the sale of ETF shares and thus lower ETF relative performance, but it should not affect differences between fund NAVs and the underlying index. If creations do contain information indicating differences between NAVs and underlying indexes, it would likely indicate that NAVs are too low and thus would outperform, not underperform, the underlying index. I report summary statistics for both performance measures in Table 1.

4 Main Findings

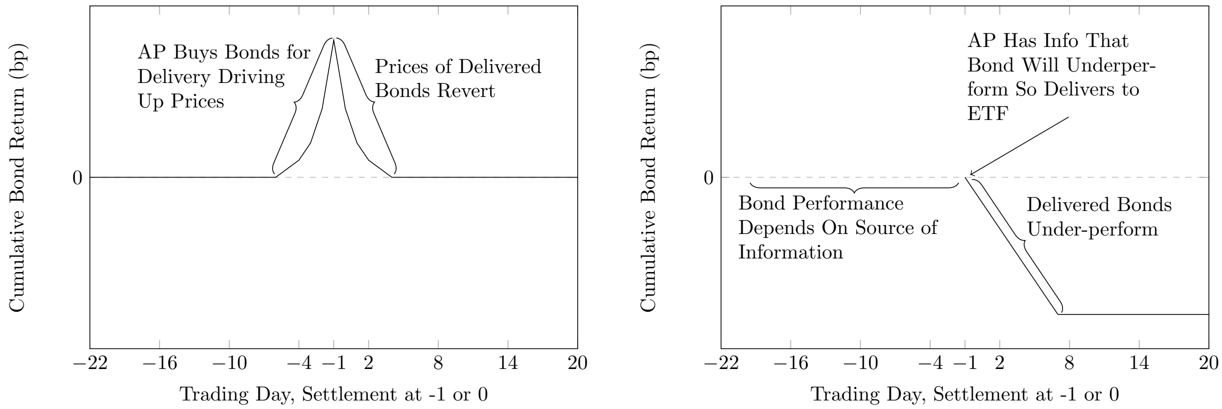
4.1 Corporate Bond Underperformance around Delivery

In this section I ask how corporate bonds perform before and after they are delivered by APs to an ETF during a creation event. As discussed in the above discussion of creation, any performance following t=-1 (because of T+1 reporting, some but not all creations are reported with a one-day lag), will embed a hidden cost that ETF investors incur. To determine whether this occurs, I estimate regressions of bond returns cumulated from 22 days prior to a creation event to various horizons on a binary variable that indicates whether a bond was delivered, time-varying bond liquidity controls, bond fixed effects, and fund-by-date fixed effects.

$$\prod_{n=-22}^{k} BondRet_{b,t+n} = \beta_k^D Delivered_{b,f,t} + \beta^L Liquidity_{b,t} + \gamma_{f,t} + \alpha_b + \varepsilon_{b,f,t}$$

By including fund-by-date fixed effects, I take advantage of the fact that delivered bonds have a natural control: bonds that are held by the same ETF but not delivered during that specific creation event. ETF managers have incentives to include more liquid and larger bonds in the creation basket. While this would lower ETF returns, ETFs would also bear lower risk in such a case. Bond fixed effects (which account for time-invariant bond sizes) and time-varying liquidity controls help to rule out this mechanism. Instead, by estimating this performance I can measure differences in performance between delivered bonds during a given creation event and undelivered bonds during a creation event and compare these performances with the relative performances of those same bonds at other times that do not surround creation days, much like what occurs with a differences in differences framework. Additionally, bond fixed effects will rule out alternative explanations in which delivered bonds are less risky in a time-invariant way. As discussed above in relation to creations, there are two main reasons why bonds may underperform following delivery: 1) bond prices (and/or assigned values) are temporarily high as a result of the temporary impact on prices of APs' securing of bonds to deliver or 2) The utilization of information about future changes in bond values by APs when deciding which bonds to deliver to the creation basket and when to deliver them. If prices are temporarily too high in a manner that is consistent with APs' temporarily impacting prices before settlement, one would expect to see a dynamic of the sort illustrated in Figure 1a. If the main mechanism is APs' utilization of information, bond behavior prior to delivery depends on the sources of that information. If APs utilize bond-price reversals, the dynamics should be similar to what can be seen in Figure 1a. If APs take advantage of bond momentum the bonds should underperform before delivery. If APs utilize private information it likely would not be observable in prices before delivery. A combination and thus the weighted average of such effects should be observable, as depicted in Figure 1b.

I estimate the effects of asset delivery on bond returns over horizons ranging from 1 to 42 trading days. Figure 2 displays the dynamics of the estimates of β_k^D across various k. The figure also shows 95 % confidence intervals using standard errors that account for two-way clusters by bond and by date. Full parameter estimates for select horizons can be found in Appendix Table A1. As can be seen, bonds that will be delivered begin underperforming



(a) Bond Dynamics if APs Temporarily Impact Bonds' Prices around Delivery

(b) Bond Dynamics if APs Utilize Information About Future Returns

Figure 1: Conjectured Bond-Return Dynamics
Hypothesized dynamics of β_k^D estimated using

$$\prod_{n=-22}^{k} BondRet_{b,t+n} = \beta_k^D Delivered_{b,f,t} + \beta^L Liquidity_{b,t} + \gamma_{f,t} + \alpha_b + \varepsilon_{b,f,t}$$

for various k are shown. $Delivered_{b,f,t}$ is a binary variable that takes the value of one if a bond increases in notional amount while a fund experiences a creation. $Liquidity_{b,t}$ includes the prior month median Amihud measure, implied round-trip costs, and bid-ask spreads. $\gamma_{f,t} + \alpha_b$ are fund-by-date and bond fixed effects. Subfigure (a) displays hypothesized bond-return dynamics if bond values are temporarily impacted by AP trading activity. Subfigure (b) displays hypothesized bond return dynamics if APs utilize information when deciding when to create shares and which bonds to deliver.

in the month prior to delivery and the cumulative return is statistically significant by t=-12. By t=-1, the earliest date at which creation could occur,[9] bonds that will be delivered have underperformed by 2.6 basis points. Thus, the pre-delivery performance is consistent with APs' utilizing information when deciding which bonds to deliver and when. On net, it appears that APs rely on some combination of bond momentum and private information. Following t-1, bonds underperform by an additional 2.4 basis points in the following month. Thus, following receipt by ETFs, precisely when the ETF portfolio weights of these bonds increase, the bonds underperform going forward. This underperformance is evidence that ETFs receive an adversely selected set of bonds from APs that will hurt ETF performance

[9]Creation occurs at either t=-1 or t=0, depending on whether the fund utilizes T or T+1 reporting; it is unobservable for a given fund which reporting standard managers utilized.

relative to their underlying benchmarks. Importantly, this is identified with time-series shifts in portfolio weights, not persistent differences in ETFs' portfolios and benchmark indices' portfolios, and thus represents a pure performance drag.[10] Thus, put simply, when ETFs experience creation events they receive inferior assets (assets with low expected future returns) from APs.

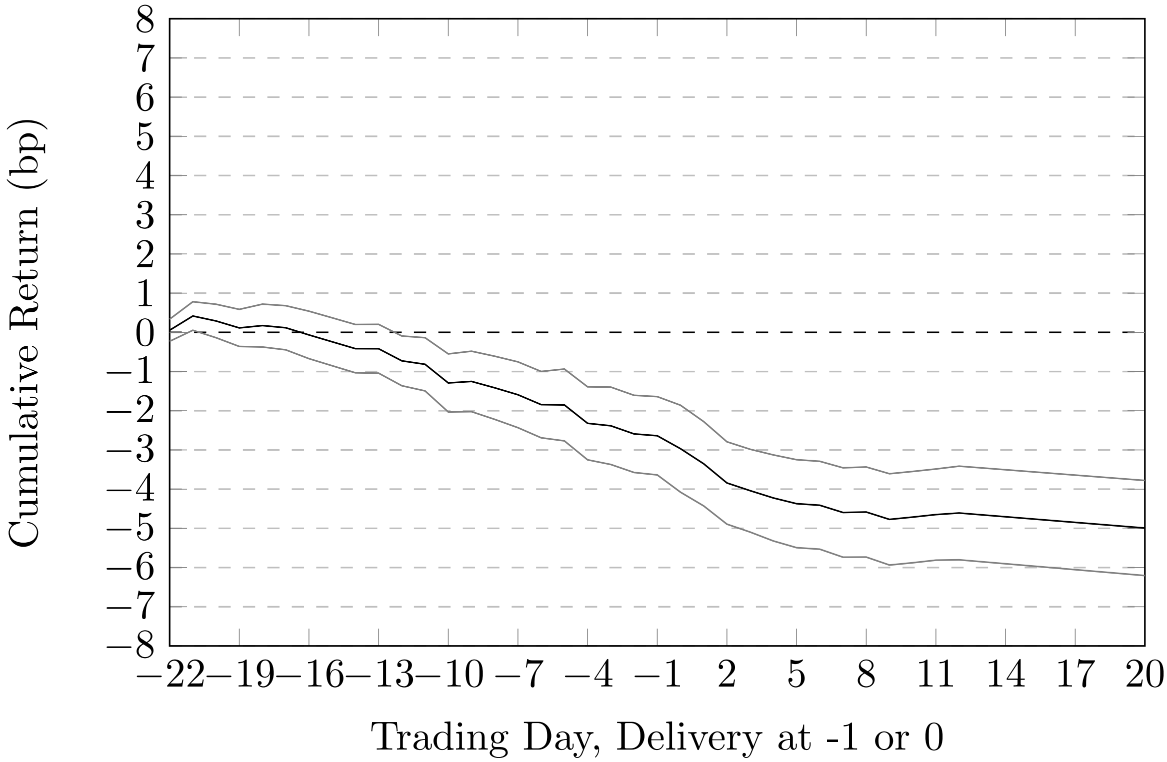


Figure 2: Bond Dynamics around Delivery to ETFs
This figure plots $\widehat{\beta_k^D}$ estimated using $\prod_{n=-22}^{k} BondRet_{b,t+n} = \beta_k^D Delivered_{b,f,t} + \beta^L Liquidity_{b,t} + \gamma_{f,t} + \alpha_b + \varepsilon_{b,f,t}$ for various k. The figure also displays 95% CIs using standard errors clustered by bond and trading date. $Delivered_{b,f,t}$ is a binary variable that takes the value of one if a bond increases in notional amount while a fund experiences a creation event. $Liquidity_{b,t}$ includes the prior month median Amihud measure, implied round-trip costs, and bid-ask spreads. $\gamma_{f,t} + \alpha_b$ are fund-by-date and bond fixed effects.

[10]Nevertheless, additional creations may also imply smaller premiums that may be desirable to ETF investors.

4.2 ETF Underperformance around Creation Events

As described above in the discussion of share creations, the bond underperformance documented in Figure 2 should translate to a hidden cost of lower relative fund performance when creations occur. To test this hypothesis at the fund level and obtain empirical estimates of the size of the hidden cost, I regress spreads over varying horizons between cumulative fund returns and changes in underlying indices on a binary variable that indicates whether a fund experienced a creation event on a given day, fund fixed effects, and time fixed effects:

$$RetSpread_{f,t,k} = \beta_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$$

Importantly, I estimate this hidden cost using spreads between ETF returns and changes in the underlying indices. Creation events still likely mean good news for ETF investors. Conditional on a creation event, ETF returns and underlying indices are both higher than they would be had no creation events occurred: this result likely reflects increased demand for a given asset. While flows raise price levels for both ETFs and assets in the underlying index, they also embed a hidden cost incurred for receiving assets that earn lower future returns. Thus, the hidden cost causes ETF investors to "miss a bit of the upside." By estimating the effects on spreads I am able to isolate these effects and increase the power of my tests significantly. Additionally, by including fund fixed effects I implicitly control for time-invariant characteristics, such as management fees and net expense ratios, that may have attracted investor flows during the sample period while explaining ETF performance. If more expensive ETFs were introduced during the sample period that were able to attract greater flows (for example by targeting a novel set of corporate bonds), flows might be correlated with, without being the source of, the poor performance. Fund fixed effects should control for this problem and instead allow the estimates to document the direct impact of share creations on performance.

The stated and desirable purpose of creation events is to arbitrage away premiums. Thus, conditional on observing a creation event at t=-1, one can infer that a premium was likely created in the days prior to the creation. The creation of this premium will likely result in positive cumulative spreads, in the periods prior to share creation. As the APs arbitrage away premiums, spreads should revert back to zero over some time horizon. Over long time horizons, such as the +/- one-month window that I have investigated, the premium arbitrage mechanism implies that cumulative spreads should be zero.[11] Thus, if no hidden cost incurred from creation events arise, the dynamics of return spreads should appear as in Figure 3a. If the receipt of assets with lower future returns indeed impacts fund performance, this effect would appear as an additional negative drift in spreads beginning at time zero, as illustrated in Figure 3b. I hypothesize that both of these mechanisms should occur and thus their joint effects should be observable, as in Figure 3c. Importantly, because the mechanism that drives the arbitrage motivation for creations implies a long-run zero impact on spreads (the null hypothesis that only the arbitrage mechanism is at play implies a long-run coefficient of zero), negative performance at the end of the +/- one-month window can instead be attributed to the hidden cost and acts as a point estimate of the size of this cost. Figure 9 displays the dynamics of the β_k coefficients, also showing 95 % confidence intervals using standard errors that account for two-way clusters by ETF and by date. Consistent with both the hypothesized hidden cost and the arbitrage mechanism that is associated with creations, ETFs outperform the underlying index prior to creations and revert following creations, but over the long horizon underperform the underlying index by 1.4 basis points. Full parameter estimates for select horizons can be found in Appendix Table A2. Thus, ETF investors suffer a 1.4 basis-point performance drag per creation event.

Redemptions also may theoretically impose hidden costs to ETF investors if APs are able to select bonds from ETF portfolios that will realized better future performance. In practice, redemptions are less likely to incur hidden costs for a number of reasons. First, ETF managers may have greater bargaining power when negotiationg the basket for re-

[11]If the spreads were not zero but the ETF successfully mirrored the underlying index, premiums and discounts would have been permanently drifting from the value of one.

demptions (as discouraging redemptions may be in ETF managers' interests). Secondly, empirically redemption baskets are much larger than creation baskets. Appendix Figure 9 reports the dynamics around redemptions. I find no evidence of the hidden cost occuring during redemptions.

4.3 ETF NAV Dynamics around Creation Events

I next estimate the performance of spreads between ETF NAVs and the underlying index. If ETF managers do well in maintaining portfolios that are representative of the underlying indices that they seek to track and they utilize the same pricing services as those used to set underlying indices, the NAV of an ETF and the underlying index should move closely together and deviations should be mean zero. Thus, the arbitrage mechanism predicts that creations will have no impact on spreads between NAVs and underlying indices. If the conjectured hidden cost exist, it occurs because an ETF portfolio fails to represent the underlying indices sufficiently over time. Specifically, this occurs when an ETF portfolio shifts to assets with lower expected future returns exactly at the point of creation. Thus, if the hidden cost exists, it should be observable in a dynamic that is very similar to that shown in Figure 3b. To test whether the hidden cost is observable in ETF NAVs, I estimate the following regression over various horizons. This estimation is the same as the estimation reported in Figure 9 but with an alternative spread as the dependent variable:

$$NAVSpread_{f,t,k} = \beta_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t,k}$$

Figure 5 reports the dynamics of β_k in the +/- one-month window around share creation. I am unable to reject the null hypothesis that the long-run spread between NAVs and underlying indices is zero. Thus, I am unable to reject the null hypothesis that the dynamics are fully explained by the arbitrage motives associated with a creation event. It is possible that I lack sufficient statistical power to identify such an effect given the long time horizon over which I cumulate returns. Additionally, the use of stale NAVs by ETFs could lengthen the time that passes until the hidden cost is observable in NAVs in a manner that would reduce the power of my test. Full parameter estimates for select horizons can be found in Appendix Table A3.

4.4 ETF Underperformance around Creation Events after Accounting for the Arbitrage Mechanism

In the prior section I explain why long-horizon spreads should be zero if the only force at play is the arbitrage mechanism that creations facilitate. To test the robustness of this result, I instead seek to alternatively estimate the impact of creations by excluding the arbitrage mechanism while isolating the dynamics of the hidden cost. To do so, I control for the premium at the time of share creation and describe the post-creation dynamics of spreads between cumulative ETF returns and underlying index changes. Insofar as creations occur during trades that seek to capture premiums, the sizes of such premiums are key determinants of the arbitrage mechanism, as described in the previous section, while controlling for this mechanism should isolate the hidden cost of creations. I therefore estimate the following regression over various time horizons:

$$RetSpread_{f,t,k} = \beta_k Creation_{f,t} + \beta^P Premium_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$$

This specification involves two key differences from those that underlie the results reported in Figure 9. First, the premium as of the end of day t=0 is included as a control. Second, cumulative returns are benchmarked to the end of day zero and only post-creation performance is accumulated. I present present the dynamics of β_k in Figure 6. Following a creation event and after controlling for the premium at the time of the creation, ETFs underperform their stated benchmarks by an additional 2.3 basis points, a finding that is consistent with the presence of a hidden cost of creations. Full parameter estimates for select horizons can be found in Appendix Table A4.

After controlling for premiums and discounts, the hidden cost my analysis reveals should also be observable in the implied performance of fund NAVs relative to changes in the underlying index. Thus, I estimate similar regressions but with spreads between NAVs and underlying indices:

$$NAVSpread_{f,t,k} = \beta_k Creation_{f,t} + \beta^P Premium_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$$

Figure 7 presents the dynamics of β_k. Full parameter estimates for select horizons can be found in Appendix Table A5. Following a creation event after controlling for premiums and discounts, ETF NAVs underperform on average by 0.6 basis points. Over the full 20-day horizon this 0.6 basis-point figure is significant at the 10% level but not at the 5% level. For 15 of the 19 horizons considered, the estimates are significant at the 5% level. The size estimates peak 18 days following creation at 0.8 basis points. These dynamics are hard to rationalize based on APs' arbitrage motives. If creations do not shift ETF portfolios away from a representative sample of the underlying index and the same prices are used to calculate both, NAVs and the underlying index should remain tightly linked even during arbitrage and no underperformance should be observable. Additionally, a reader may be concerned that controlling for premiums fails to rule out arbitrage and that pricing differences between the two may exist. Nonetheless, these results would still be hard to explain by reference to the arbitrage mechanism. APs engage in share creation to capture premiums that occur when ETF prices are too high relative to NAVs. Thus, one would expect creations to imply that ETF prices are too high, returns should be lower, and NAVs are too low and thus should outperform the index. Instead, Figure 7 documents NAV underperformance, contradicting even the hypothesis arising from arbitrage mechanisms in which NAVs are too low. NAV underperformance is not consistent with the objectives of the creation/redemption process and instead is best explained by the hidden cost embedded when share creation with a subset of assets with lower expected future returns is allowed.

5 Additional Findings

5.1 The Relationship between Fund Characteristics and the Size of the Hidden Cost

The return dynamics of delivered bonds both prior and subsequent to delivery are consistent with APs' utilizing information regarding future changes in bond values to decide strategically which assets to deliver into ETF portfolios. Such strategic interactions provide marginal incentives on top of APs' well-understood incentives to profitably arbitrage differences in ETF prices and underlying assets' value. Theory clearly predicts when strategic behavior should be more profitable,and thus more likely to occur, conditional on observing a creation event, and finally therefore where the hidden cost should be the highest.

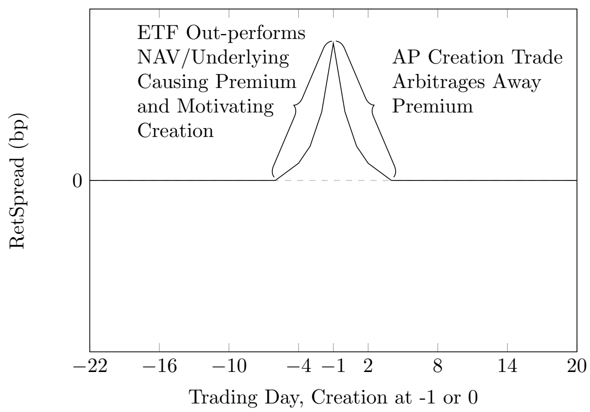
First, APs' strategic incentives will be stronger when their information regarding future bond values is more valuable. It is well understood that agents are likely to possess more valuable information in less liquid or less efficient markets (in fact, agents who possess valuable information can endogenously make those markets less liquid in many classical market-microstructure models). Additionally, ETF managers adopt less stringent share-creation rules in the presence of asset liquidity. Therefore, the hidden cost should be higher when underlying assets are less liquid.

Second, APs' ability to deliver bonds strategically is enhanced when they are required to deliver fewer bonds. If an AP has information that is relevant to the future performance of a given bond that will impact their expected profits from a creation trade, the impact will be much greater if that bond is one of ten bonds that needs to be delivered versus being one of a hundred bonds to be delivered. Moreover, it may be easier for APs to acquire information about the idiosyncratic performance of a small number of bonds. Thus, when creation baskets are large, it is more difficult for APs to acquire information regarding the average performance of a given creation basket and the strategic incentive will be relatively weaker, enabling the arbitrage incentive to dominate.

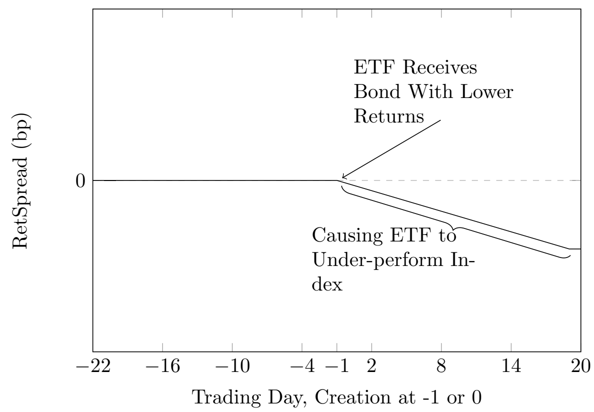
To assess theoretical predictions regarding fund differences and thus validate the described mechanism, I estimate how the cost per creation is related to fund characteristics. Specifically, I estimate the following regression:

$$RetSpread_{f,t,20} = \beta_1 Creation_{f_t} + \beta_2 Creation_{f,t} * FundCharacteristic_f + \alpha_t + \gamma_f + \varepsilon_{f,t}$$

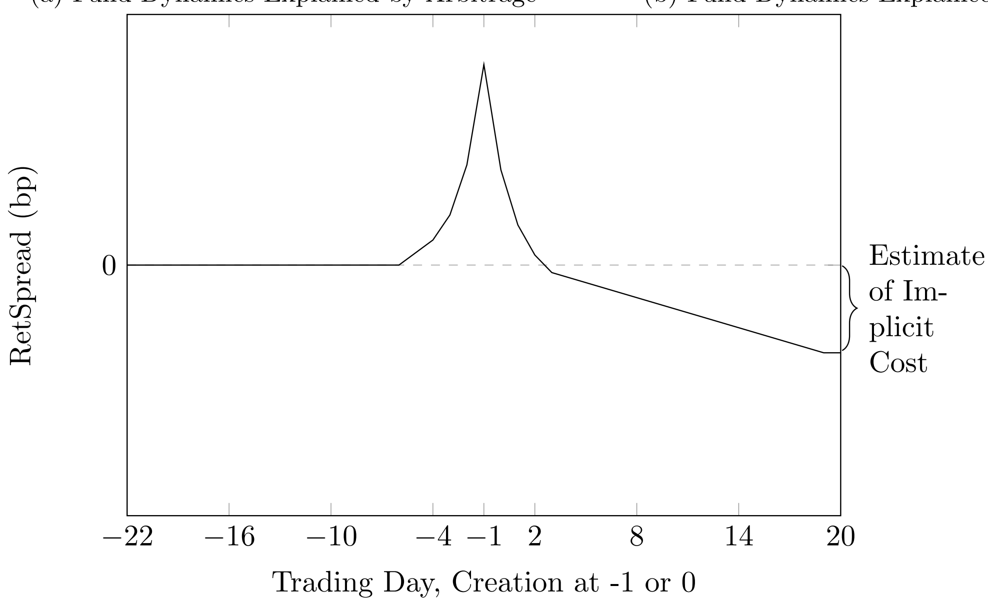
I consider five fund characteristics in particular. First, I examine the relationship between an ETF's asset focus and the cost per creation using $InvestmentGrade_f$. $InvestmentGrade_f$ takes the value of 1 if the focus as reported by ETF Global is on investment-grade corporate bonds. Most funds that do not focus on investment-grade bonds report focusing on high-yield bonds. Additionally, some funds focus explicitly on asset-backed securities, convertible bonds, and loans. In all such cases, the investment-grade bond market is likely relatively more efficient and liquid. In column 1 of Table 2 I report results indicating that investment-grade bonds typically incur a 2.41 basis-point lower cost per creation than other types of corporate bond funds, a finding that is consistent with theory. The T statistics associated with this estimate is 1.97 and the p-value is 5.1%.



(a) Fund Dynamics Explained by Arbitrage



(b) Fund Dynamics Explained by Hidden Cost



(c) Conjectured Fund Dynamics Where Both Occur

Figure 3: Conjectured ETF Return Spread Dynamics

This figure plots the hypothesized dynamics of β_k, estimated using $RetSpread_{f,t,k} = \beta_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for various k. $RetSpread_{f,t,k}$ is the cumulative difference between ETF returns and the benchmark index cumulated through k days following the observation date: $RetSpread_{f,t,k} = \prod_{n=-22}^{k} ETFRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$. For example, $RetSpread_{f,t,-1}$ is the cumulative difference between the ETF and the index up to 1 day prior to observation (or share creation in the case of a creation event). $Creation_{f,t}$ is a binary variable that takes the value of one if shares outstanding increased for a fund and zero otherwise. $\gamma_t + \alpha_f$ are date and fund fixed effects. Subfigure (a) displays the hypothesized dynamics if only the desired arbitrage mechanism is evident. Subfigure (b) displays the additional hidden cost for receiving bonds that earn lower returns. Subfigure (c) shows the hypothesized dynamics if a hidden cost exist.

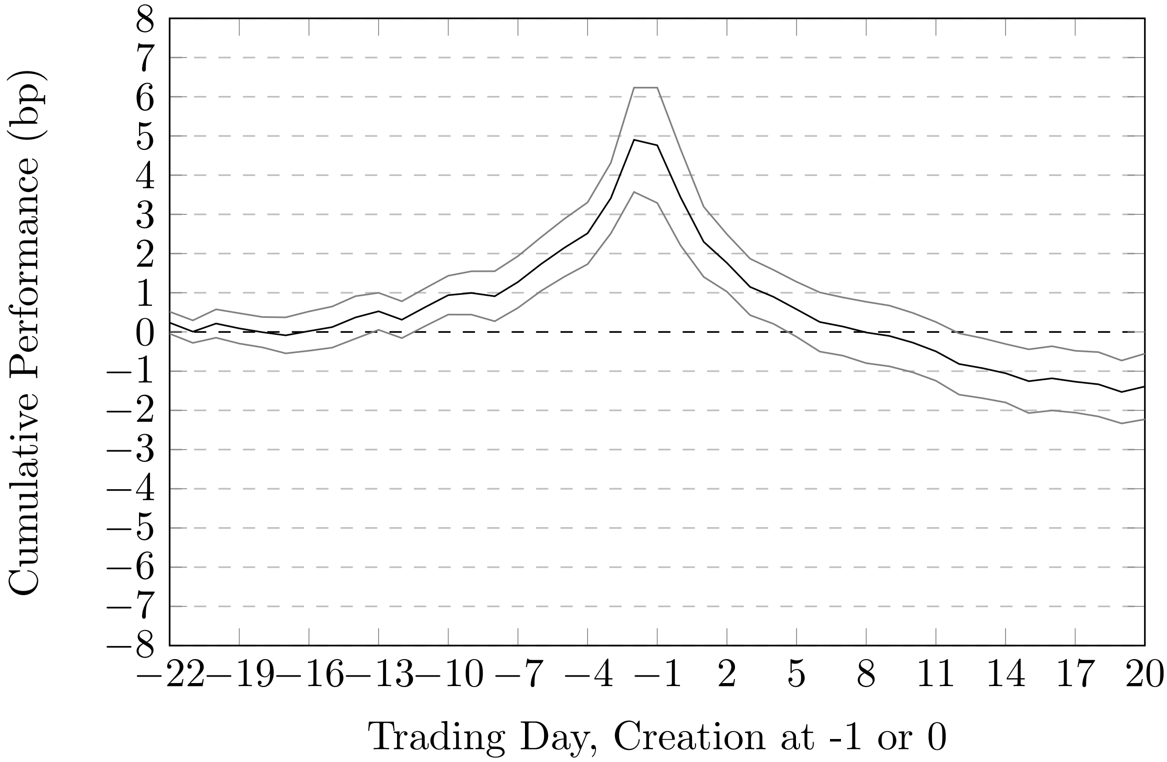


Figure 4: Return Spread Dynamics around Creation Events
This figure plots $\widehat{\beta_k}$ estimated using $RetSpread_{f,t,k} = \beta_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for various k. The figure also displays 95% CIs using standard errors clustered by fund and trading date. $RetSpread_{f,t,k}$ is the cumulative difference between ETF returns and the benchmark index cumulated through k days following the observation date: $RetSpread_{f,t,k} = \prod_{n=-22}^{k} ETFRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$. For example, $RetSpread_{f,t,-1}$ is the cumulative difference between the ETF and the index up to 1 day prior to observation (or share creation in the case of a creation event). $Creation_{f,t}$ is a binary variable that takes the value of one if a fund's shares outstanding increased and zero otherwise. $\gamma_t + \alpha_f$ are date and fund fixed effects.

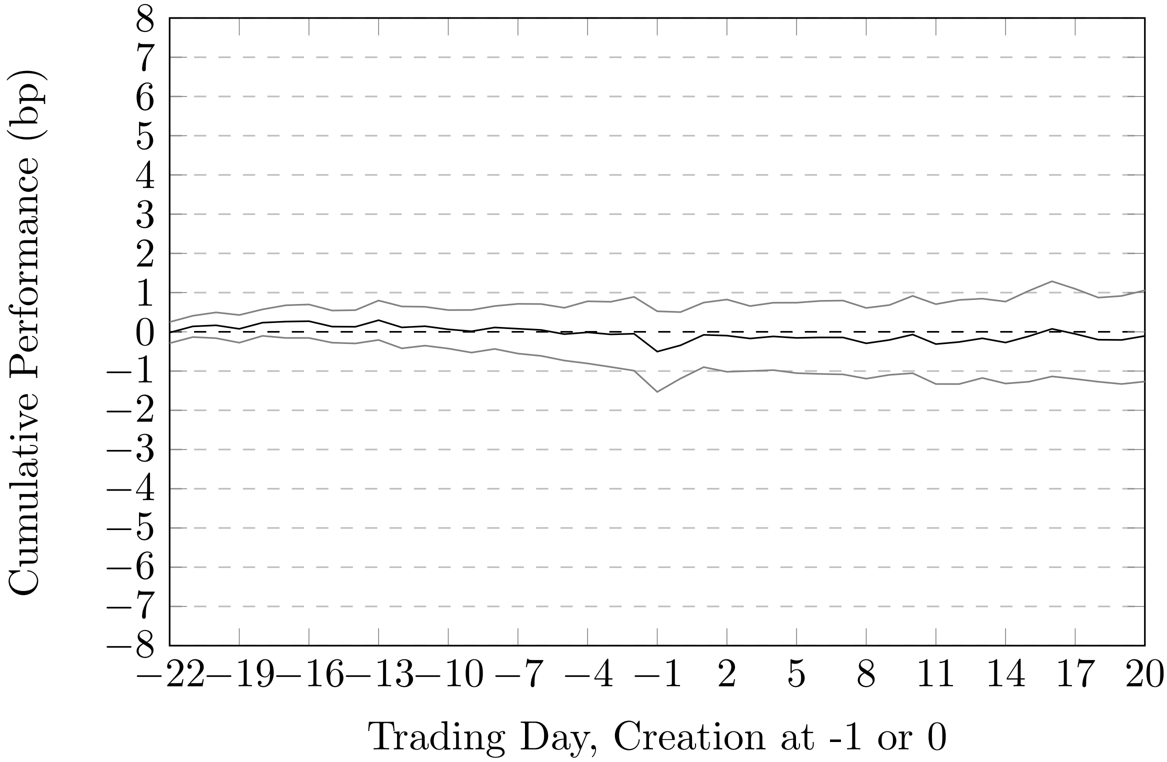


Figure 5: NAV Spread Dynamics around Creation Events
This figure plots $\widehat{\beta_k}$ estimated using $NAVSpread_{f,t,k} = \beta_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for various k. The figure also displays 95% CIs using standard errors clustered by fund and trading date. $NAVSpread_{f,t,k}$ is the cumulative difference between the return implied by fund NAVs and the benchmark index cumulated through k days following the observation date: $NAVSpread_{f,t,k} = \prod_{n=-22}^{k} NAVRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$. For example, $NAVSpread_{f,t,-1}$ is the cumulative difference between an ETF's assets and the index up to 1 day prior to observation (or share creation in the case of a creation event). $Creation_{f,t}$ is a binary variable that takes the value of one if a fund's shares outstanding increased and zero otherwise. $\gamma_t + \alpha_f$ are date and fund fixed effects.

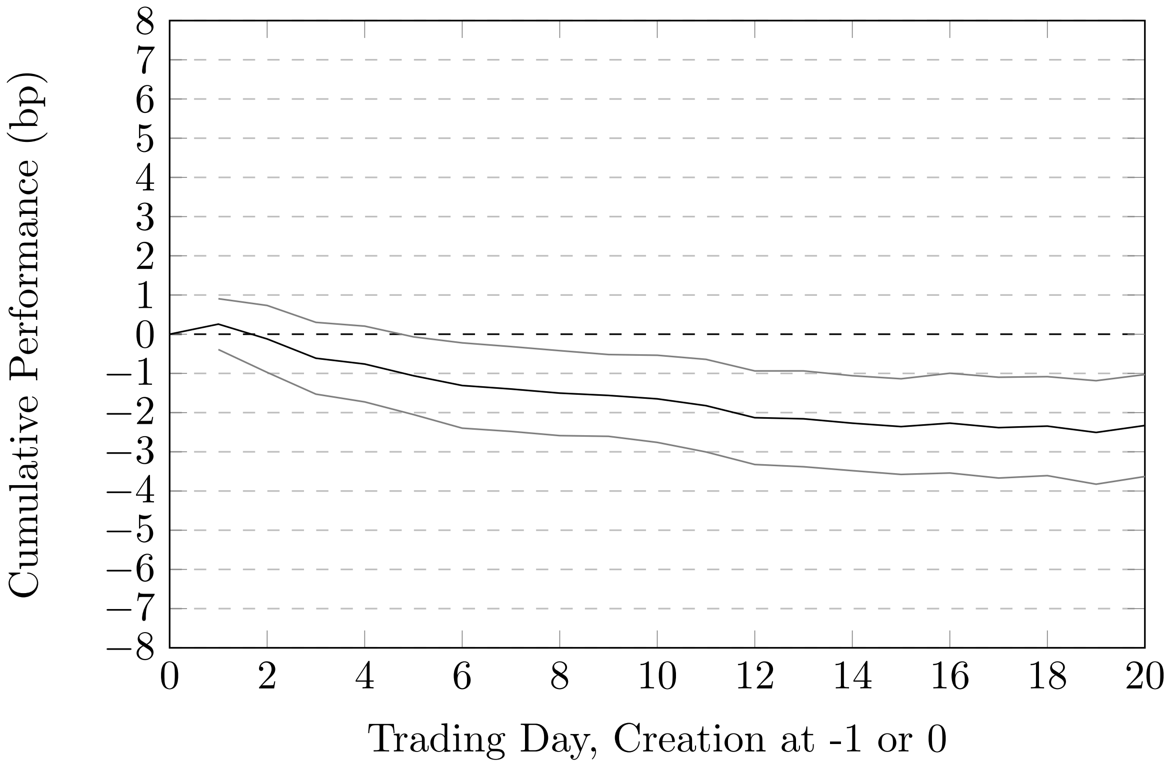


Figure 6: Return Spread Dynamics around Creation Events while Controlling for Premiums
This figure plots $\widehat{\beta_k}$ estimated using $RetSpread_{f,t,k} = \beta_k Creation_{f,t} + \beta^P Premium_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for various k. The figure also displays 95% CIs using standard errors clustered by fund and trading date. $RetSpread_{f,t,k}$ is the cumulative difference between ETF returns and the benchmark index cumulated through k days following the observation date: $RetSpread_{f,t,k} = \prod_{n=-22}^{k} ETFRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$. For example, $RetSpread_{f,t,-1}$ is the cumulative difference between the ETF and the index up to 1 day prior to observation (or share creation in the case of a creation event). $Creation_{f,t}$ is a binary variable that takes the value of one if a fund's shares outstanding increased and zero otherwise. $Premium_{f,t}$ is the premium or discount as represented in percentage terms (%100 representing no discount or premium). $\gamma_t + \alpha_f$ are date and fund fixed effects.

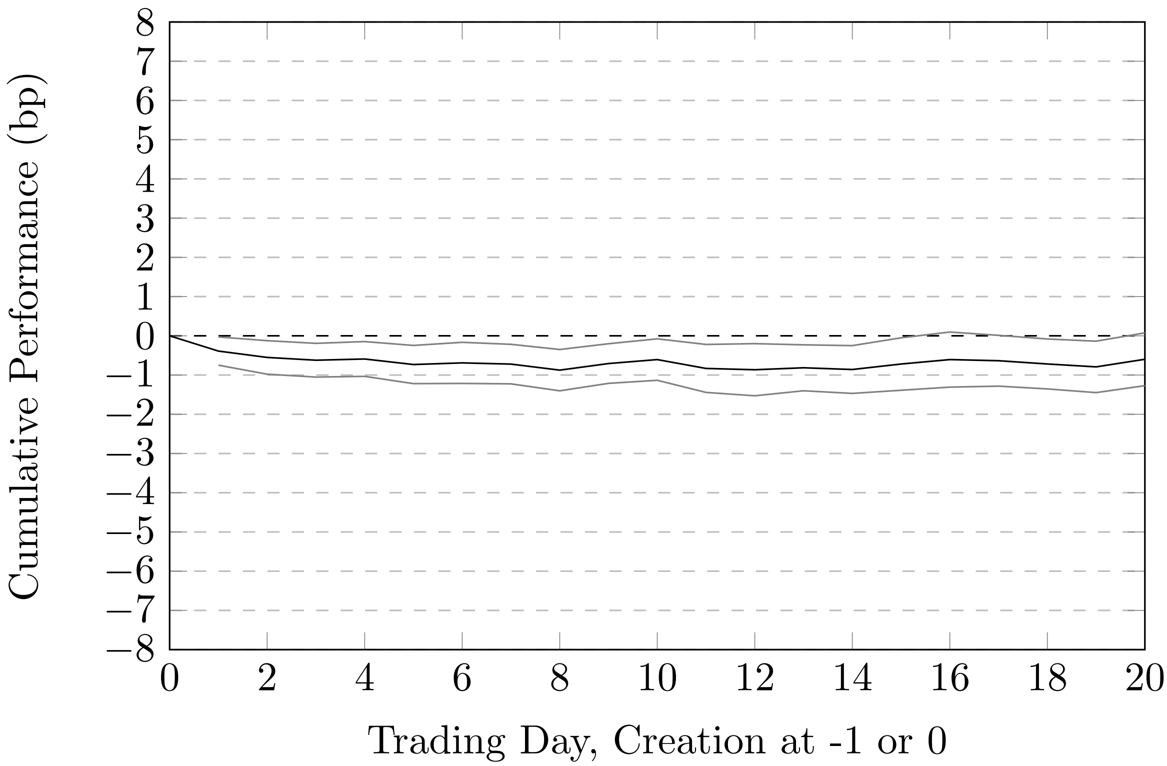


Figure 7: NAV Spread Dynamics around Creation Events while Controlling for Premiums
This figure plots $\widehat{\beta_k}$ estimated using $NAVSpread_{f,t,k} = \beta_k Creation_{f,t} + \beta^P Premium_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for various k. The figure also displays 95% CIs using standard errors clustered by fund and trading date. $NAVSpread_{f,t,k}$ is the cumulative difference between returns implied by fund NAVs and the benchmark index cumulated through k days following the observation date: $NAVSpread_{f,t,k} = \prod_{n=-22}^{k} NAVRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$. For example, $NAVSpread_{f,t,-1}$ is the cumulative difference between the ETF's assets and the index up to 1 day prior to observation (or share creation in the case of a creation event). $Creation_{f,t}$ is a binary variable that takes the value of one if a fund's shares outstanding increased and zero otherwise. $Premium_{f,t}$ is the premium or discount as represented in percentage terms (%100 representing no discount or premium). $\gamma_t + \alpha_f$ are date and fund fixed effects.

Table 2: Fund Characteristics' Relationship to Hidden Costs

This table reports the relationship between fund characteristics and estimates of the hidden cost per creation for corporate bond ETFs. Relationships are estimated based on the following specification:

$RetSpread_{f,t,20} = \beta_1 Creation_{f_t} + \beta_2 Creation_{f,t} * FundCharacteristic_f + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for five fund characteristics.

The first fund characteristic, $InvestmentGrade_f$, indicates whether the fund is a reported by ETF Global as focusing on Investment Grade corporate bonds. $\ln(\text{AverageBasketSize}_f)$ represents the natural log of the average number of securities delivered when a creation event occurs. $AverageAmihud_f$ represents the average prior month median Amihud measure for bonds held by the fund. $AverageIRC_f$ represents the average prior month median imputed round-trip cost measure for bonds held by the fund. $AverageBidAsk_f$ represents the average prior month median bid-ask spread measure for bonds held by the fund. $AverageAmihud_f$, $AverageIRC_f$, and $AverageBidAsk_f$ are winsorized at the 1% level. $RetSpread_{f,t,20}$ is the cumulative difference between the ETF return and the benchmark index cumulated through 20 days after the observation date: $RetSpread_{f,t,20} = \prod_{n=-22}^{20} ETFRet_{f,t+n} - \prod_{n=-22}^{20} \Delta UnderlyingIndex_{f,t+n}$. $Creation_{f,t}$ is a binary variable that takes the value of one if shares outstanding increased for a fund and zero otherwise. $\gamma_t + \alpha_f$ are date and fund fixed effects. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% levels respectively.

	Dependent Variable Spread Between Cumulative ETF and Benchmark Return Through 1 Month After Creation $\text{RetSpread}_{f,t,20}$				
$\text{Creation}_{f,t}$	-2.90***	-5.78***	0.10	-1.27	-1.16
	(-3.03)	(-3.11)	(-0.53)	(0.10)	(-1.22)
$\text{Creation}_{f_t} * InvestmentGrade_f$	2.41*				
	(1.97)				
$\text{Creation}_{f_t} * \ln(AverageBasketSize_f)$		1.03**			
		(2.38)			
$\text{Creation}_{f_t} * AverageAmihud_f$			-48.83*		
			(-1.75)		
$\text{Creation}_{f_t} * AverageIRC_f$				-0.27	
				(-0.40)	
$\text{Creation}_{f_t} * AverageBidAsk_f$					-1.71
					(-0.58)
Fund Fixed Effects	Y	Y	Y	Y	Y
Date Fixed Effects	Y	Y	Y	Y	Y
Standard Errors	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date
N	134,519	104,791	123,820	123,820	123,820
Adjusted R^2	8.99%	10.86%	10.01%	10.01%	10.01%
Within F.E. R^2	0.02%	0.03%	0.02%	0.02%	0.02%

Second, I examine the relationship between the size of a creation basket, as measured by $\ln(AverageBasketSize)$. To do so, I infer all assets that were delivered from ETF holdings data (including those without TRACE information) and calculate the average size of the basket conditional on observing any delivery. APs find it more difficult to utilize information regarding specific components of larger baskets. I report the corresponding results in column 2 of Table 2. The coefficient on $\ln(AverageBasketSize)$ is 1.03, implying that, when basket size doubles, the cost per creation decreases by 0.71 basis points, roughly half of the unconditional cost per creation of 1.39 basis points. The T statistic associated with this estimate is 2.38 and is statistically significant.

Last, I investigate three measures of portfolio liquidity: $AverageAmihud_f$, $AverageIRC_f$, and $AverageBidAsk_f$. $AverageAmihud_f$ represents the average prior month median Amihud measure for bonds held by a fund. $AverageIRC_f$ represents the average prior month median imputed round-trip cost measure for bonds held by a fund. $AverageBidAsk_f$ represents the average prior month median bid-ask spread measure for bonds held by a fund. All three values represent asset illiquidity and rise when underlying bonds are less liquid. I report the corresponding results in columns 3-5 of Table 2. I find a marginally statistically significant relationship only for $AverageAmihud_f$: the coefficient is -48.83 with a T statistic of -1.75. This coefficient, coupled with the previously discussed $InvestmentGrade_f$ estimate, provides suggestive evidence that the cost per creation is higher when assets are less liquid, a finding that is consistent with theory.

In summary, fund characteristics predicted by APs' strategic incentives are predictive of the estimated hidden cost per creation, further validating the mechanism described in this paper. When ETFs hold less liquid assets and when ETF managers allow smaller baskets to be delivered in the face of asset illiquidity, the cost is higher. Such results highlight that ETF investors pay an implicit cost, either knowingly or unknowingly, for the liquidity transformation services that corporate bond ETFs provide.

5.2 Changes in Dealer Inventory Prior to Delivery

While the bond level return dynamics surround deliveries is most consistent with APs utilizing information not reflected in bond settlement prices to reduce inventories of bonds they perceive will be costly to hold, returns dynamics alone reveal very little about the source of APs' information. If APs acquire information about the fundamental value of issuing firms, APs may seek to reduce their inventories by not only creating ETF shares but also by directly trading with customers. Thus, dealer inventories may fall prior to delivery to ETFs. Alternatively, a natural alternative source of information may arise from customer order flows in relatedly marking businesses. Thus, the ultimate nexus of information may lie in informed bond traders who trade against bond market makers. Dealer inventories may rise prior to delivery into ETF portfolios. In such a case, creations may act as a means of APs to avoid losses in the cases in which they are "run-over" by informed bond traders.

To assess possible sources of information, I thus estimate a linear probability model on the likelihood a bond is delivered during a creation:

$$Delivered_{b,t} = \beta_1 Daily\Delta DealerInv_{b_t} + \beta_2 Weekly\Delta DealerInv_{b_t} + \beta_3 LiquidityMeasures + \alpha_{f,t} + \varepsilon_{b,t}$$

Fund by time fixed effects are included to compare delivered bonds to other bonds held by the same ETF but not delivered. Importantly, in contrast to return specification, bond by fund fixed effects are not included in order to asses the relation between persistent liquidity measures and bond delivery choice.[12] I report the results in Table 3.

[12] Untabulated results with bond by fund fixed effects are consistent with the reported results.

Table 3: Bond Characteristics Relation to Inclusion in Realized Creation Baskets

This table reports the relationship between bond characteristics and the liklihood a bond is delivered to an ETF. The linear probability model is estimated based on the following specification:

$\text{Delivered}_{b,t} = \beta_1 Daily\Delta DealerInv_{b_t} + \beta_2 Weekly\Delta DealerInv_{b_t} + \beta_3 LiquidityMeasures + \alpha_{f,t} + \varepsilon_{b,t}$

$Delivered_{b,t}$ takes a value of $10,000$ if a bond was delivered and 0 otherwise (thus marginal probabilities are in basis points.) $Daily\Delta DealerInv_{b_t}$ and $Weekly\Delta DealerInv_{b_t}$ are the 1 and 5 day order imbalances of customer to dealer transactions reported in TRACE. $Amihud_{b,t}$ represents the prior month median Amihud measure for bonds held by the fund. $IRC_{b,t}$ represents the prior month median imputed round-trip cost measure for bonds held by the fund. $BidAsk_{b,t}$ represents the average prior month median bid-ask spread measure for bonds held by the fund. $Volume_{b,t}$ represents the prior month trading volume. All explanatory variables are winsorized at the 1% level. $\alpha_{f,t}$ are fund-by-date fixed effects. One, two, and three stars indicate statistical significance at the 10%, 5%, and 1% levels respectively.

	Probability Delivered (in bp) $\text{Delivered}_{b,f,t}$
Daily $\Delta DealerInv_{b_t}$	-2.67***
	(-12.56)
Weekly $\Delta DealerInv_{b_t}$	1.66***
	(14.51)
$\text{Amihud}_{b,t}$	0.00
	(0.70)
$\text{IRC}_{b,t}$	-0.00***
	(-3.28)
$\text{BidAsk}_{b,t}$	-14.48***
	(-3.50)
$\text{Volume}_{b,t}$	1.12***
	(30.01)
Fund by Date Fixed Effects	Y
Standard Errors	2 way Clustered by Bond and Date
N	14,453,118
Adjusted R^2	47.07%
Within F.E. R^2	0.18%

Consistent with APs learning from the order flow of their informed customers, dealer inventories increase in the 5 days before a bond is delivered into an ETF portfolio. On the day a bond is delivered to ETF portfolios, dealer inventories instead fall, consistent with APs simultaneously seeking to reduce their inventories through their market making activities and via ETF share creations. Such patterns are consistent with those documented by Pan and Zeng (2019). Additionally, consistent with Koont et al. (2022), more liquid bonds are more likely to be delivered into ETF portfolios.

APs need not possess superior bond picking skills nor act unscrupulously in order to deliver bonds with lower performance in a manner than embeds the hidden cost. APs may simply deliver the bonds that they can acquire easily because other informed investors wish to sell it to them. Nonetheless, since informed customer order flow is not fully incorporated into bond prices used to settle creations, ETF investors may still face the hidden cost in exchange for the valuable liquidity transformation services that APs in conjunction with the ETF security design delivers.

5.3 Large-Cap Equity ETFs: A Falsification Test

Unlike corporate bond ETFs, ETFs that hold large-cap US equities almost always require APs to deliver all underlying assets when they create new shares.[13] Therefore, large-cap equity ETFs lack the institutional features that are necessary to embed the hidden cost, making them an effective sample for a falsification test of my main specification. If large-cap equities underperform following a creation event, the associated underperformance of large-cap equity funds, and thus likely also of corporate bond funds, can be explained by an alternative mechanism. Therefore, I re-estimate the results I report in Figure 9, utilizing a sample of ETFs in the "Equity" asset class, with a "Large-Cap" and "North America" focus as reported by ETF Global. In Figure 8, I again report the dynamics of β_k coefficients

[13] These ETFs occasionally will utilize custom creation baskets to intentionally shift their portfolio. For example, ETFs will often utilize a heartbeart trade upon index insertion/deletions as described in Moussawi et al. (2022). These custom creation baskets are often designed by ETF managers at set dates and do not face the same adverse selection present in corporate bond ETFs.

and display 95% confidence intervals. I cannot reject the null hypothesis that large-cap US equity ETFs show no underperformance when they experience creation days, a finding that is consistent with theory. ETFs embed the hidden cost only when they modify the creation process to accommodate illiquid assets such as corporate bonds.

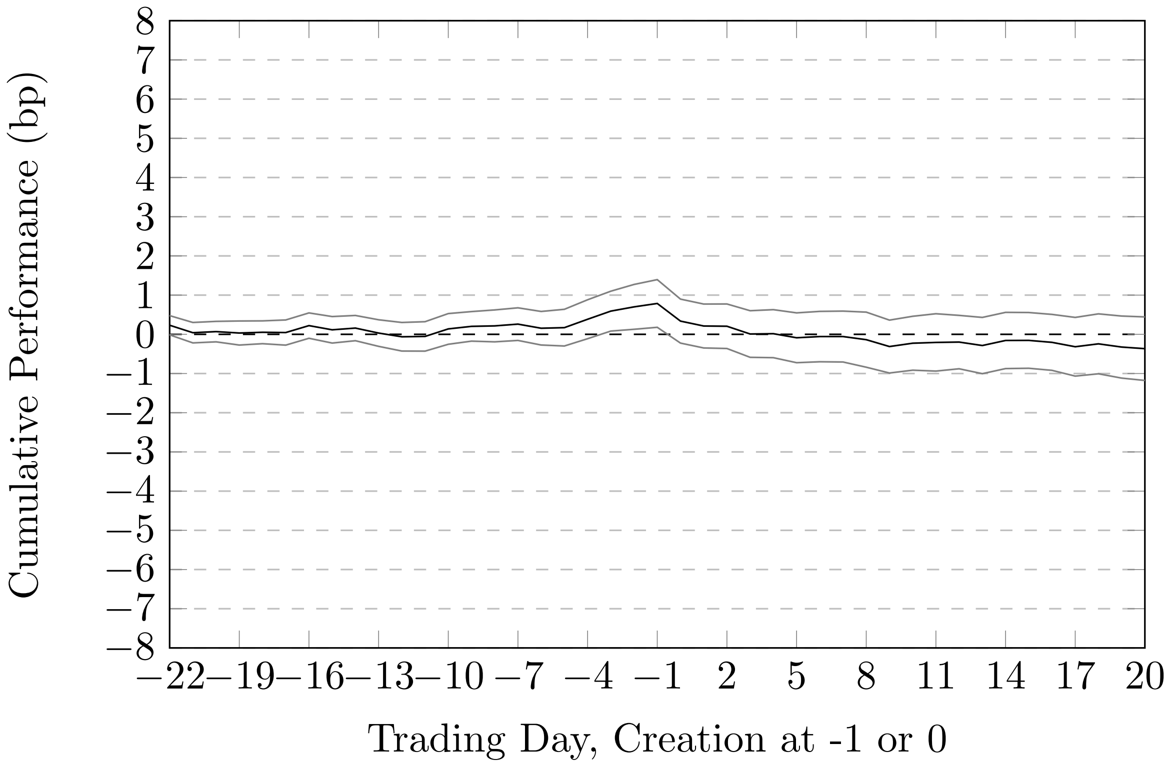


Figure 8: Falsification Test: Return Spread Dynamics around Large US Equity ETFs
This figure plots $\widehat{\beta_k}$ estimated using $RetSpread_{f,t,k} = \beta_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for various k. The figure also displays 95% CIs using standard errors clustered by fund and trading date. $RetSpread_{f,t,k}$ is the cumulative difference between ETF returns and the benchmark index cumulated through k days following the observation date: $RetSpread_{f,t,k} = \prod_{n=-22}^{k} ETFRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$. For example, $RetSpread_{f,t,-1}$ is the cumulative difference between the ETF and the index up to 1 day prior to observation (or share creation in the case of a creation event). $Creation_{f,t}$ is a binary variable that takes the value of one if a fund's shares outstanding increased and zero otherwise. $\gamma_t + \alpha_f$ are date and fund fixed effects.

6 Conclusion

In this paper I explain how the use of creation baskets that contain only subsets of assets that underlie corporate bond ETFs can embed a hidden cost that ETF investors pay. I document that APs deliver bonds that earn lower future returns to ETFs, with the result that ETFs underperform their stated benchmarks by an additional 1.4 basis points per creation event. Additionally, I isolate this cost by controlling for arbitrage motives and document the cost using ETF prices and ETF NAVs. While the arbitrage mechanism is certainly present around creation events, this set of facts cannot be fully explained by APs' arbitrage trades and instead illustrates the existence of a hidden cost of share creation. This hidden cost is high; investors in the average corporate bond ETF pay a 48 basis-points-per-annum hidden cost while the average corporate bond ETF reports only 35 basis points per annum in explicit costs. If ETF investors are fully aware of the hidden cost, the magnitude of the cost reveals a high willingness-to-pay for the liquidity transformation provided by corporate bond ETFs.

This hidden cost also has important policy implications. First, it highlights the need to scrutinize the prices that are used to calculate ETF NAVs and thus settle creations. Prior to this finding, regulators, academic researchers, and market participants may have reasonably believed that NAV accuracy is critical only to the performance of open-ended mutual funds, not ETF performance. Second, my finding highlights the cost attributable to partial creation baskets and in particular the cost attributable to custom creation baskets that are allowed under SEC rule 6c-11. While these custom creation baskets allow ETFs to engage in activities such as heartbeat trades to defer taxes that benefit ETF investors, they also can embed a high cost and thus represent an area of potential regulatory oversight or beneficial increases in transparency for ETF investors.

The existence of the hidden cost I have documented implies that investors, either knowingly or unknowingly, pay a high cost for the liquidity transformation that corporate bond ETFs provide. Specifically, the cost results from modifications to creation rules that are designed to induce APs to conduct arbitrage trades despite the illiquidity of underlying assets.

Many policymakers and academic researchers have expressed concerns over the fragility that might result from liquidity transformation. I demonstrate that, even if asset illiquidity poses no risk to financial stability, liquidity transformation is not a "free lunch." The concessions that ETF managers must make to induce arbitrage activity incur a high cost as a result of APs' ability to interact strategically with the rules. Because this cost results directly from asset illiquidity, the framework of Chordia (1996), which implies that the more illiquid the underlying assets are, the higher are the barriers to investor fund flows, likely extends from mutual funds to ETFs. This cost of liquidity transformation also helps to rationalize flow performance relationships in ETFs and to resolve open puzzles in the important strand of literature that examines flows to investment managers. Last, as the cost of such liquidity transformation is higher than the explicit fees reported by corporate bond ETFs, the findings I have reported in this paper are of the first order in describing the performance realized by ETF investors.

References

Amihud, Yakov (2002), "Illiquidity and stock returns: cross-section and time-series effects." *Journal of Financial Markets*, 5, 31–56, URL https://www.sciencedirect.com/science/article/pii/S1386418101000246.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi (2018), "Do etfs increase volatility?" *The Journal of Finance*, 73, 2471–2535.

Ben-Rephael, Azi, Shmuel Kandel, and Avi Wohl (2012), "Measuring investor sentiment with mutual fund flows." *Journal of financial Economics*, 104, 363–382.

Berk, Jonathan B and Richard C Green (2004), "Mutual fund flows and performance in rational markets." *Journal of political economy*, 112, 1269–1295.

Brown, David C, Shaun William Davies, and Matthew C Ringgenberg (2021), "Etf arbitrage, non-fundamental demand, and return predictability." *Review of Finance*, 25, 937–972.

Chakravarty, Sugato and Asani Sarkar (2003), "Trading costs in three us bond markets." *The Journal of Fixed Income*, 13, 39–48.

Chalmers, John MR, Roger M Edelen, and Gregory B Kadlec (2001), "On the perils of financial intermediaries setting security prices: the mutual fund wild card option." *The Journal of Finance*, 56, 2209–2236.

Chen, Qi, Itay Goldstein, and Wei Jiang (2010a), "Payoff complementarities and financial fragility: Evidence from mutual fund outflows." *Journal of Financial Economics*, 97, 239–262.

Chen, Qi, Itay Goldstein, and Wei Jiang (2010b), "Payoff complementarities and financial fragility: Evidence from mutual fund outflows." *Journal of Financial Economics*, 97, 239–262.

Choi, Jaewon, Mathias Kronlund, and Ji Yeol Jimmy Oh (2022), "Sitting bucks: Stale pricing in fixed income funds." *Journal of Financial Economics*, 145, 296–317, URL https://www.sciencedirect.com/science/article/pii/S0304405X21003706.

Chordia, Tarun (1996), "The structure of mutual fund charges." *Journal of Financial*

Economics, 41, 3–39, URL https://www.sciencedirect.com/science/article/pii/0304405X9500856A.

Clifford, Christopher P, Jon A Fulkerson, and Bradford D Jordan (2014), "What drives etf flows?" *Financial Review*, 49, 619–642.

Coval, Joshua and Erik Stafford (2007), "Asset fire sales (and purchases) in equity markets." *Journal of Financial Economics*, 86, 479–512.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers (1997), "Measuring mutual fund performance with characteristic-based benchmarks." *The Journal of finance*, 52, 1035–1058.

Dannhauser, Caitlin D (2017), "The impact of innovation: Evidence from corporate bond exchange-traded funds (etfs)." *Journal of Financial Economics*, 125, 537–560.

Dannhauser, Caitlin D and Saeid Hoseinzade (2021), "The unintended consequences of corporate bond etfs: Evidence from the taper tantrum." *The Review of Financial Studies*.

Dannhauser, Caitlin D and Jeffrey Pontiff (2019), "Flow." *Available at SSRN 3428702*.

Diamond, Douglas W and Philip H Dybvig (1983), "Bank runs, deposit insurance, and liquidity." *Journal of political economy*, 91, 401–419.

Doan, Viet-Dung (2020), "Exchange-traded funds and municipal bond market." *Available at SSRN 3703875*.

Edelen, Roger M (1999), "Investor flows and the assessed performance of open-end mutual funds." *Journal of Financial Economics*, 53, 439–466.

Edelen, Roger M and Jerold B Warner (2001), "Aggregate price effects of institutional trading: a study of mutual fund flow and market returns." *Journal of financial Economics*, 59, 195–220.

ETF.com (2021), "https://www.etf.com/etf-education-center/etf-basics/what-is-the-creationredemption-mechanism?nopaging=1."

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

Frazzini, Andrea and Owen A Lamont (2008), "Dumb money: Mutual fund flows and the cross-section of stock returns." *Journal of financial economics*, 88, 299–322.

Friesen, Geoffrey C and Travis RA Sapp (2007), "Mutual fund flows and investor returns: An empirical examination of fund investor timing ability." *Journal of Banking & Finance*, 31, 2796–2816.

Haddad, Valentin, Alan Moreira, and Tyler Muir (2020), "When selling becomes viral: Disruptions in debt markets in the covid-19 crisis and the fed's response." Technical report, National Bureau of Economic Research.

Hong, Gwangheon and Arthur Warga (2000), "An empirical study of bond market transactions." *Financial Analysts Journal*, 56, 32–46.

Ippolito, Richard A (1992), "Consumer reaction to measures of poor quality: Evidence from the mutual fund industry." *The Journal of Law and Economics*, 35, 45–70.

Jensen, Michael C. (1968), "The performance of mutual funds in the period 1945-1964." *The Journal of Finance*, 23, 389–416, URL http://www.jstor.org/stable/2325404.

Jostova, Gergana, Stanislava Nikolova, Alexander Philipov, and Christof W. Stahel (2013), "Momentum in Corporate Bond Returns." *The Review of Financial Studies*, 26, 1649–1693, URL https://doi.org/10.1093/rfs/hht022.

Koont, Naz, Yiming Ma, Lubos Pastor, and Yao Zeng (2022), "Steering a ship in illiquid waters: Active management of passive funds." Technical report, National Bureau of Economic Research.

Kostovetsky, Leonard (2003), "Index mutual funds and exchange-traded funds." *The Journal of Portfolio Management*, 29, 80–92.

Ma, Yiming, Kairong Xiao, and Yao Zeng (2020), "Mutual fund liquidity transformation and reverse flight to liquidity." *Available at SSRN 3640861*.

Moussawi, Rabih, Ke Shen, and Raisa Velthuis (2022), "Etf heartbeat trades, tax efficiencies, and clienteles: The role of taxes in the flow migration from active mutual funds to etfs." *Available at SSRN 3744519*.

Pan, Kevin and Yao Zeng (2019), "Etf arbitrage under liquidity mismatch." *Available at SSRN 2895478*.

Shim, John J and Karamfil Todorov (2021), "Etfs, illiquid assets, and fire sales." *Available at SSRN 3886881*.

Sirri, Erik R and Peter Tufano (1998), "Costly search and mutual fund flows." *The journal of finance*, 53, 1589–1622.

Wermers, Russ (2000), "Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses." *The Journal of Finance*, 55, 1655–1695.

Zheng, Lu (1999), "Is money smart? a study of mutual fund investors' fund selection ability." *the Journal of Finance*, 54, 901–933.

Zitzewitz, Eric (2006), "How widespread was late trading in mutual funds?" *American Economic Review*, 96, 284–289.

Table A1: Bond Dynamics Around Delivery to ETFs
$\prod_{n=-22}^{k} BondRet_{b,t+n} = \beta_k^D Delivered_{b,f,t} + \beta^L Liquidity_{b,t} + \gamma_{f,t} + \alpha_b + \varepsilon_{b,f,t}$ for select k

	Dependent Variable Cumulative Bond Returns		
	Up to Creation $\prod_{n=-22}^{-1} BondRet_{b,t+n}$	through 1 Week following Creation $\prod_{n=-22}^{5} BondRet_{b,t+n}$	through 1 Month following Creation $\prod_{n=-22}^{20} BondRet_{b,t+n}$
$\text{Delivered}_{b,f,t}$	-2.63***	-4.36***	-4.99***
	(-4.38)	(-6.45)	(-6.47)
$\text{Amihud}_{b,t}$	80.10***	104.97***	143.50***
	(4.87)	(5.29)	(5.15)
$\text{IRC}_{b,t}$	1.60***	2.34***	4.53***
	(4.30)	(5.38)	(8.13)
$\text{BidAsk}_{b,t}$	14.91***	18.17***	19.17***
	(5.93)	(6.32)	(5.56)
Bond Fixed Effects	Y	Y	Y
ETF*Date Fixed Effects	Y	Y	Y
Standard Errors	2-way Clustered by Bond and Date	2 way Clustered by Bond and Date	2 way Clustered by Bond and Date
N	9,544,076	9,544,076	9,544,076
Adjusted R^2	39.20%	43.05%	51.69%
Within F.E. R^2	0.14%	0.20%	0.32%

Table A2: Return Spread Dynamics Around Creation
$\widehat{\beta_k}$ estimated from $RetSpread_{f,t,k} = \beta_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for select k are shown.

	Dependent Variable		
	Spread Between Cumulative ETF and Benchmark Return		
	Up to Creation	Through 1 Week After Creation	Through 1 Month After Creation
	$\text{RetSpread}_{f,t,-1}$	$\text{RetSpread}_{f,t,5}$	$\text{RetSpread}_{f,t,20}$
$\text{Creation}_{f,t}$	4.76***	0.58	-1.39***
	(6.40)	(1.63)	(-3.28)
Fund Fixed Effects	Y	Y	Y
Date Fixed Effects	Y	Y	Y
Standard Errors	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date
N	134,519	134,519	134,519
Adjusted R^2	8.41%	8.53%	8.98%
Within F.E. R^2	0.20%	0.00%	0.01%

Table A3: NAV Spread Dynamics Around Creation Events
$\widehat{\beta}_k$ estimated from $NAVSpread_{f,t,k} = \beta_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for select k are shown.

	Dependent Variable Spread Between Cumulative ETF and Benchmark Return		
	Up to Creation $\text{NAVSpread}_{f,t,-1}$	through 1 Week following Creation $\text{NAVSpread}_{f,t,5}$	through 1 Month following Creation $\text{NAVSpread}_{f,t,20}$
$\text{Creation}_{f,t}$	-0.50	-0.16	-0.11
	(-0.97)	(-0.34)	(-0.18)
Fund Fixed Effects	Y	Y	Y
Date Fixed Effects	Y	Y	Y
Standard Errors	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date
N	134,519	134,519	134,519
Adjusted R^2	12.84%	12.95%	13.32%
Within F.E. R^2	0.00%	0.00%	0.00%

Table A4: Return Spread Dynamics Around Creation Events while Controlling for Premiums

$\widehat{\beta}_k$ estimated from $RetSpread_{f,t,k} = \beta_k Creation_{f,t} + \beta^P Premium_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for select k are plotted.

	Dependent Variable	
	Spread Between Cumulative ETF and Benchmark Returns	
	through 1 Week following Creation	through 1 Month following Creation
	RetSpread$_{f,t,5}$	RetSpread$_{f,t,20}$
Creation$_{f,t}$	-1.06**	-2.33***
	(-2.11)	(-3.54)
Premium$_{f,t}$	-2244.59***	-3157.40***
	(-9.45)	(-10.15)
Fund Fixed Effects	Y	Y
Date Fixed Effects	Y	Y
Standard Errors	2-way Clustered by Fund and Date	2 way Clustered by Fund and Date
N	139,080	139,080
Adjusted R^2	17.30%	19.10%
Within F.E. R^2	9.78%	11.98%

Table A5: NAV Spread Dynamics Around Creation Controlling For Premiums
$\widehat{\beta}_k$ estimated from $NAVSpread_{f,t,k} = \beta_k Creation_{f,t} + \beta^P Premium_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for select k are plotted.

	Dependent Variable	
	Spread Between Cumulative ETF and Benchmark Return	
	Through 1 Week After Creation	Through 1 Month After Creation
	NAVSpread$_{f,t,5}$	NAVSpread$_{f,t,20}$
Creation$_{f,t}$	-0.73***	-0.60*
	(-2.98)	(-1.76)
Premium$_{f,t}$	1083.44***	1211.17***
	(6.25)	(7.32)
Fund Fixed Effects	Y	Y
Date Fixed Effects	Y	Y
Standard Errors	2 way Clustered by Fund and Date	2 way Clustered by Fund and Date
N	139,080	139,080
Adjusted R^2	18.51%	15.99%
Within F.E. R^2	4.76%	3.67%

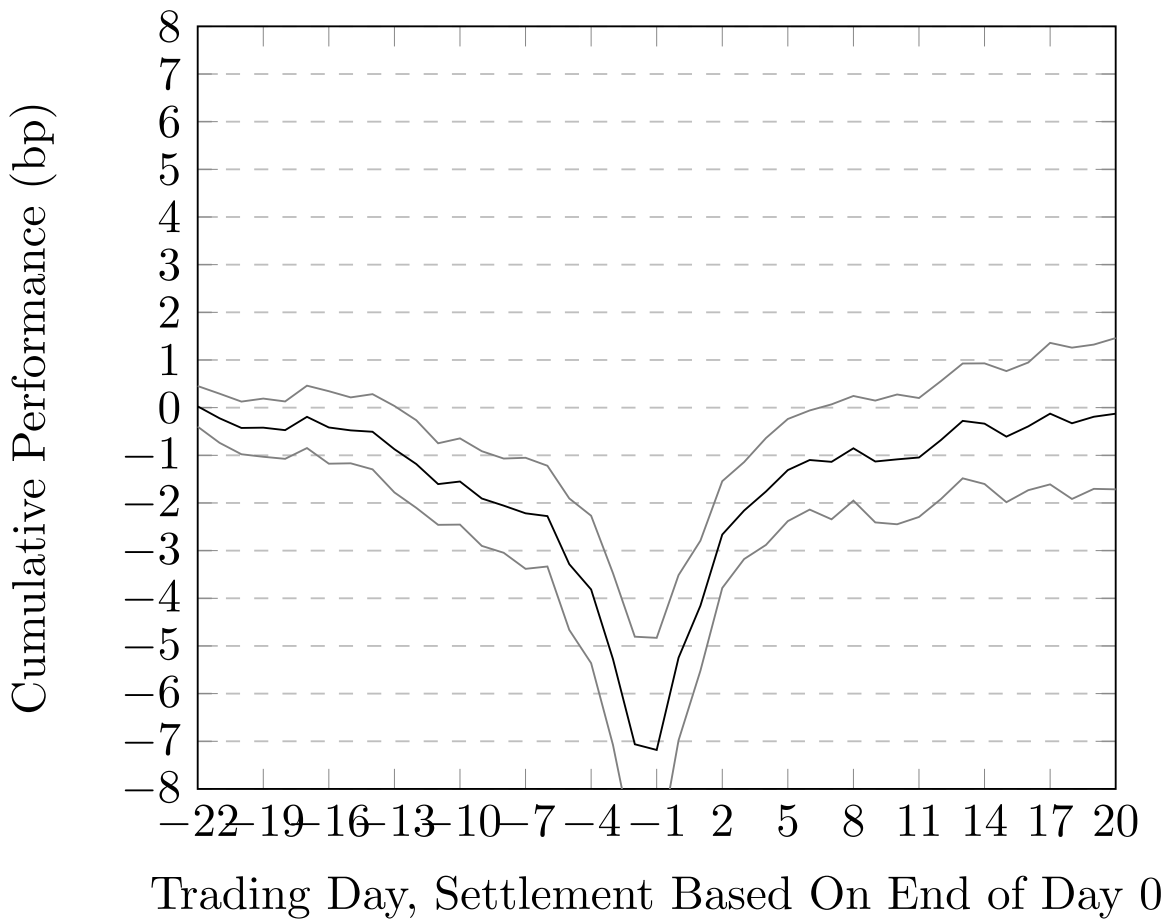


Figure 9: Return Spread Dynamics around Redemption Events
This figure plots $\widehat{\beta 1_k}$ estimated using $RetSpread_{f,t,k} = \beta 1_k Redemptionn_{f,t} + \beta 2_k Creation_{f,t} + \alpha_t + \gamma_f + \varepsilon_{f,t}$ for various k. The figure also displays 95% CIs using standard errors clustered by fund and trading date. $RetSpread_{f,t,k}$ is the cumulative difference between ETF returns and the benchmark index cumulated through k days following the observation date: $RetSpread_{f,t,k} = \prod_{n=-22}^{k} ETFRet_{f,t+n} - \prod_{n=-22}^{k} \Delta UnderlyingIndex_{f,t+n}$. For example, $RetSpread_{f,t,-1}$ is the cumulative difference between the ETF and the index up to 1 day prior to observation (or share creation in the case of a creation event). $Redemption_{f,t}$ is a binary variable that takes the value of one if a fund's shares outstanding decreased and zero otherwise. $Creation_{f,t}$ is a binary variable that takes the value of one if a fund's shares outstanding increased and zero otherwise. $\gamma_t + \alpha_f$ are date and fund fixed effects.