Working Paper Series

No 149

Microstructure implications of ETF arbitrage with custom baskets

by

Berke Körükmez

ESRB
European Systemic Risk Board
European System of Financial Supervision

Abstract

Exchange-traded funds (ETFs) are typically considered to be passive investment vehicles designed to track a benchmark index. However, with the promulgation of the Securities and Exchange Commission's 2019 *ETF Rule*, funds are permitted the use of custom creation/redemption baskets. This change effectively enables a form of active basket management during the ETF's arbitrage process. In this paper, I show that the uptake of custom baskets has heterogeneous effects on the microstructure of corporate bond ETFs. While custom baskets enhance the liquidity transformation of bond ETFs, this comes at a cost, as they concurrently produce larger index tracking errors. To isolate these effects empirically, I exploit the 2019 *ETF Rule* as a quasi-natural experiment. My findings substantiate the presence of a trade-off between liquidity enhancement and tracking error minimization, and underscore the role of custom baskets as contributors to this trade-off.

JEL classifications: G12, G14, G18, D47

Keywords: exchange-traded funds, liquidity transformation, custom baskets

In this paper, I investigate whether an ETF's basket construction method is a key determinant of its (market) microstructure. Basket construction is a pivotal element of fund design. However, the economic implications of competing basket construction methods have received relatively little attention in the existing academic literature on ETFs. Until recently, all newly established basket-based funds used a standard (in-kind) pro-rata approach. Unless an exemptive order had been obtained by the fund sponsor prior to 2006, custom basket use was prohibited in the US.

The picture looks different today. With the promulgation of SEC Rule 06c-11, known informally as the 2019 *ETF Rule*, all funds get to choose between two competing basket-based replication methods: (i) pro-rata, where baskets replicate the full portfolio in exact proportion to its holdings; and (ii) custom, where the fund is allowed to select specific assets that may differ from the benchmark index for basket delivery. This means that, under custom basket representation, an ETF can hold index assets at weights that differ from the benchmark, and it can hold proxy assets that are not constituents of the benchmark it is designed to track. Custom basket replication is ubiquitous in bond ETFs due to the highly fragmented and illiquid nature of the underlying bond market, making precise replication more challenging. [1]

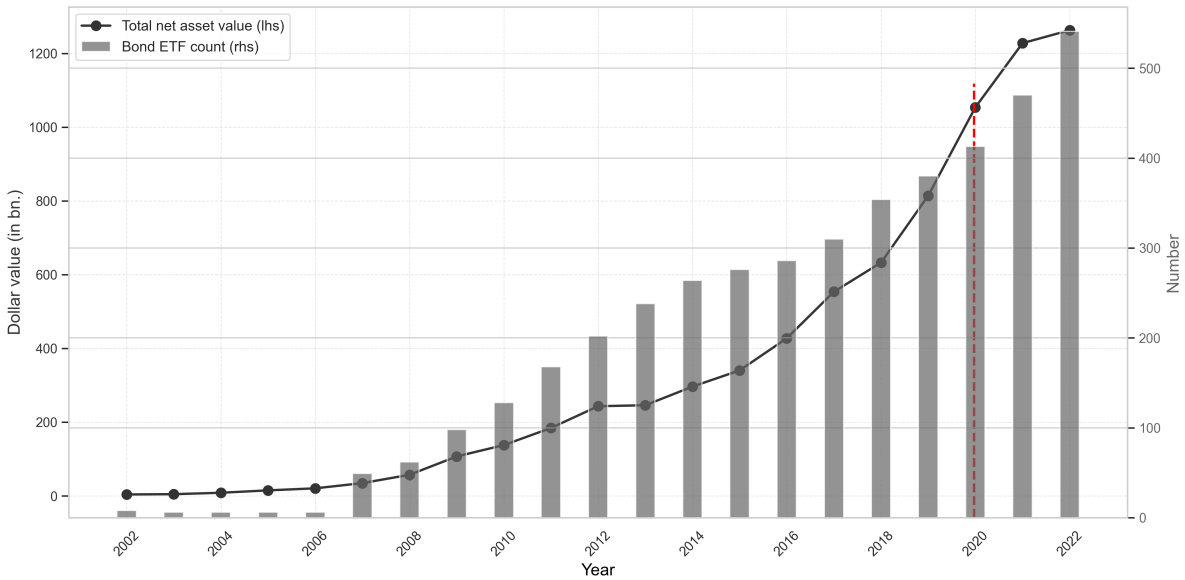
The contribution of this paper is to draw causal inferences on how basket replication methods impact ETF liquidity and tracking errors. This has important implications for ETF investors, fund managers and policymakers, as it informs on optimal ETF design and fund selection. Policy relevance emerges from the fact that custom basket replication relies on negotiations between the fund issuer and authorized participants (APs), who jointly agree on a list of accepted assets for delivery in creation and redemption baskets. Pan and Zeng (2023) find that this process may result in conflicts of interests when APs occupy a dual role as bond dealers and as ETF arbitrageurs. In such cases, APs may have incentives to leverage the flexibility of custom basket replication to optimize inventory management, potentially leading to increased ETF mispricing.

Custom basket replication is the predominant replication method in the rapidly growing bond ETF market. As illustrated in Figure 1, both net assets and fund counts for U.S. bond ETFs have grown significantly. Bond ETFs currently make up 19% of the overall U.S. ETF market, and they are capturing an increasing market share of the

[1] The prohibition on custom baskets primarily impacted new fund issuers, as many incumbent firms held exemptive orders and were therefore unaffected. Nonetheless, the introduction of the ETF Rule serves as a useful instrument that allows me to empirically assess the economic significance of custom basket replication.

market for ETFs. This growth is also fueled by the liquidity transformation service that bond ETFs provide. Unlike the underlying bond securities, which can be difficult to trade due to fragmentation and illiquidity, bond ETFs are traded on exchanges like equities, resulting in greater liquidity. This paper aims to examine the role of custom baskets in further facilitating liquidity transformation.

Figure 1.
Total net asset values and fund counts of US bond ETFs. This figure shows an annual time series plot of the the total net asset value and fund count of bond ETFs in the US from 2002 to 2022. The red line marks the effective date of the 2019 ETF Rule on December 23, 2019. Following a one-year transition period, previously granted exemptive orders were rescinded on December 22, 2020.



I show that custom baskets have a heterogeneous effect on the microstructure of corporate bond ETFs. While custom baskets enhance ETF liquidity, they concurrently widen index tracking errors. My findings corroborate recent work from Koont et al. (2024) and Brogaard et al. (2024), who also find evidence for the existence of a trade-off between liquidity transformation and tracking errors during ETF arbitrage. In my main regressions, custom baskets have a positive impact on two measures of liquidity: they tighten the effective half-spread by up to 62 basis points and the realized half-spread by up to 84 basis points, both at the 1% level. This result is economically significant because it amounts to approximately 10% and 13% of the mean effective and realized spreads, respectively. At the same time, index tracking errors widen by up to 1.38 percentage points.

My empirical identification strategy leverages a recent policy change in the US. I exploit the introduction of the 2019 *ETF Rule*, which lifts a prior ban on the use of custom baskets. It thereby induces variation in the adoption of custom baskets without feasibly impacting other determinants of ETF microstructure. Prior to the regulatory change, only funds that gained an exemptive order until 2006 were able to employ custom basket replication. These funds are unaffected by the legislation and serve as the control group. All other funds are affected and hence part of the treatment group. This setup is robust to potential endogeneity concerns because the treatment assignment induced by the policy change acts as a quasi-natural experiment and can reasonably be assumed to be random.

I conduct my empirical analysis in a two-step approach, following Foucault et al. (2011). The first part of my empirical strategy is to pursue a difference-in-differences setup. The identifying assumption in the above set-up rests on the condition that parallel trends are upheld. Provided that this condition holds, the interactive term represents the causal effect of the treatment on the dependent variable. However, if the empirical distributions of treated and control groups differ systematically on factors such as fund size, reference benchmark index, or average credit rating and duration of underlying holdings, and if the time series of ETF liquidity and tracking errors depends on these systematic differences, the estimate could be biased. In the subsequent analysis, I therefore test for systematic differences between treated and control funds, and perform quartile matching, percentage difference matching and propensity score matching to address this concern.

Subsequently, I conduct a number of further robustness checks. First, I check if my results depend on the length of the sample period around the treatment event. I perform all econometric regressions using a 12-month, a 24-month as well as a 36-month sample period, equally distributed around the effective date of the legislation. Second, I test for a potential anticipation effect using the announcement date and the starting day of the transition period as alternative treatment dates. Third, I winsorize the data at the 99% level to account for data quality concerns, while also presenting unwinsorized results for comparison. Fourth, I show a placebo test. Finally, I re-run the regressions excluding both the transition period and the COVID-19 period. This approach aligns with Haddad et al. (2021) and addresses the issue of stale NAV prices caused by the liquidity freeze in the bond market during the pandemic.

This paper is closely related to two recent studies on ETF microstructure. Koont et al. (2024) show that corporate bond ETFs actively steer their portfolios by selectively preferring liquid assets during physical replications. Similarly, Brogaard et al. (2024) find that ETF sampling increases return co-movement of liquid underlying securities. Consistent with both contributions, I also highlight the increasingly active decision-making of purportedly passive funds. While Koont et al. (2024) focus on the role of cash in the basket, Brogaard et al. (2024) investigate a phenomenon called index sampling, where fund managers create shares using a representative sample of an index's securities instead of including each index constituent. In contrast, my focus is on custom creation and redemption baskets, and the flexibility to deliver proxy assets, a related but distinct concept. Several other studies also highlight that ETFs may pursue active strategies. Cheng et al. (2019) find that some ETFs perform securities lending. Cong et al. (2024) find that some ETFs perform factor investing. In contrast, this paper focuses on active portfolio management achieved through the channel of authorized participants.

Custom baskets yield increased flexibility to APs, allowing them to substitute underlying assets with proxy assets that do not necessarily have to be part of the benchmark index. In practice, these decisions often involve negotiations between the AP and the fund sponsor. The AP selects proxy assets from a list of accepted assets published daily by the fund sponsor for the following trading day. With this paper, I add to the literature on APs (Pan and Zeng (2023), Gorbatikov and Sikorskaya (2022), Shim and Todorov (2023)). Dannhauser and Karmaziene (2023) find that bonds have higher inventory costs when they are included in creation baskets. Helmke (2023) finds that ETFs may be less liquid than mutual funds due to AP balance sheet costs. Reilly (2022) finds that APs deliver underperforming bonds in creation baskets.

This paper also builds on theoretical work on the asset pricing implications of exchange-traded funds. Malamud (2016) and Cespa and Foucault (2014) build a dynamic equilibrium model that presents share creations and redemptions as an information propagation process. Cong et al. (2024) build a model of optimal composite security design for liquidity trading with common risk factors. Pan and Zeng (2023) predict how APs optimally arbitrage in response to relative mispricings between the ETF and bond markets. Theoretical models often rely on an in-kind pro-rata representation, as this was customary in ETF markets until recently. This paper complements these models by showcasing the aggregate impact of a deviation to custom baskets. In doing so, it also informs the understanding of optimal ETF design.

In addition, this paper adds to the growing corpus of empirical studies on ETFs. In the context of equity markets, Ben-David et al. (2018) show that ETF ownership increases volatility of stocks, and Da and Shive (2018) show that it increases return co-movement. Israeli et al. (2017) determine that ETF trading amplifies the trading cost of stocks. This paper adds to the body of literature on bond ETFs. Dannhauser (2017) shows that ETF ownership lowers bond yields due to a migration of liquidity traders from the underlying to the ETF market. Holden and Nam (2017) observe its impact on corporate bond liquidity, while Agarwal et al. (2021) highlight its influence on liquidity co-movement.

The rest of the paper is structured as follows: section 2 describes details on an institutional framework on the creation and redemption mechanism with custom baskets, section 3 describes the data, section 4 provides motivating evidence, section 5 develops the hypotheses, section 6 outlines the identification strategy, section 7 discusses the main empirical results and robustness checks, and section 8 concludes.

2 THE CREATION AND REDEMPTION PROCESS

The purpose of this section is to provide the reader with a background on share creations and redemptions, the role of custom baskets within that process, and the institutional framework surrounding the ETF Rule.

ETF shares can be freely traded on the securities exchange but their value is designed to mimic a portfolio of underlying assets. Consequently, ETFs operate with a dual pricing structure: the market price of the ETF itself and the net asset value (NAV) of the underlying assets. To maintain alignment between these two prices, fund sponsors rely on designated market makers, known as authorized participants, to arbitrage away price deviations between ETF price and the NAV.

Authorized participants achieve this arbitrage by creating or redeeming ETF shares through one of two methods: cash or basket-based replication. In cash creations, the AP provides cash to the ETF fund, which the fund sponsor then uses to buy underlying securities, with the AP receiving new ETF shares in return. For cash redemptions, the AP returns ETF shares to the fund and receives cash. The fund sponsor maintains responsibility for buying and selling the underlying assets. This method is straightforward but less common, often used for specialty funds or those with specific regulatory constraints.

The basket-based replication method is more prevalent and involves the AP in directly managing the underlying securities. When the ETF trades at a premium to NAV, the AP purchases the underlying assets in the open market, bundles them into distinct unit sizes known as creation units, and exchanges them with the fund sponsor for new ETF shares. Conversely, when the ETF trades at a discount to NAV, the AP buys ETF shares on the open market and simultaneously sells the underlying assets in the ETF portfolio. They then redeem those shares with the fund sponsor in exchange for the basket of underlying assets. Figure 2 depicts this process.

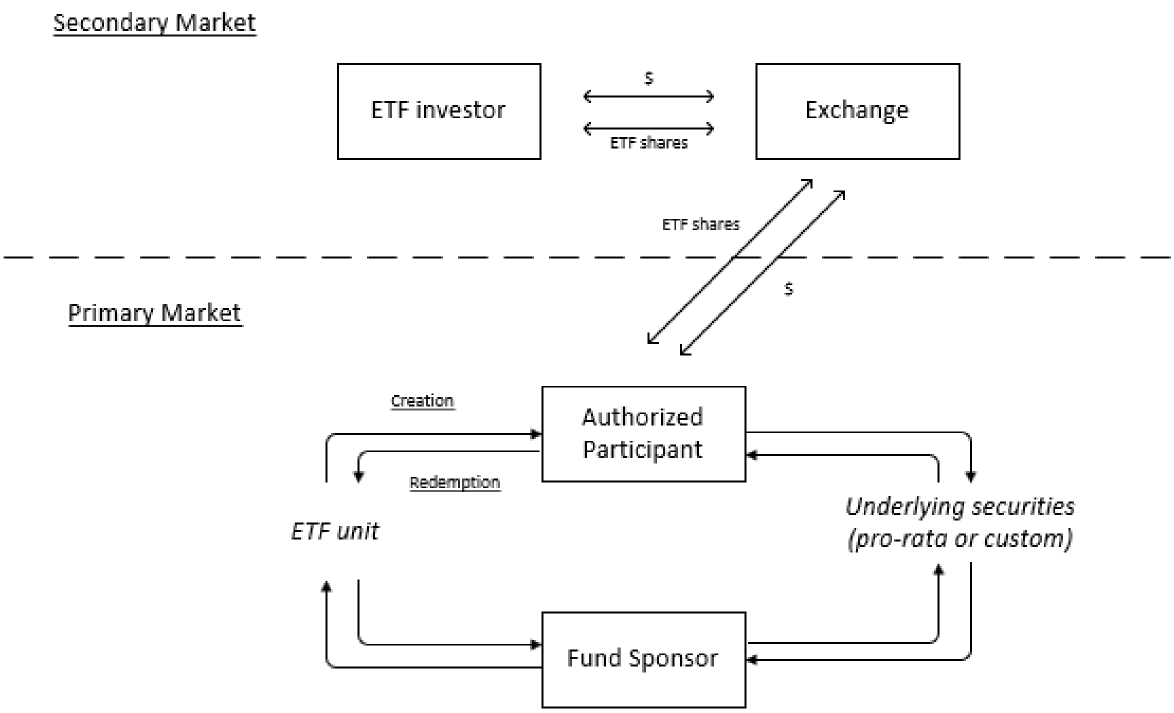
The standard basket-based approach is in-kind pro-rata, which means that the AP buys or sells the underlying asset in the exact weights as they appear in the benchmark. With the promulgation of the ETF Rule, the role of custom baskets has become increasingly important. In a custom basket, the AP can deliver or redeem a unique mix of securities that may not exactly reflect the ETF's benchmark, as long as it aligns with the fund's investment objectives and regulatory guidelines. This approach is preferred by funds with illiquid underlying securities that are difficult to trade, such as corporate bonds. In a custom basket, the AP can deliver or redeem a unique mix of securities that may not exactly reflect the ETF's benchmark, as long as it aligns with the fund's investment objectives and regulatory guidelines.

Under custom basket replication, the ETF sponsor and AP negotiate on the names and quantities of assets accepted in a basket. Every day, the ETF sponsor publishes the creation and redemption baskets for the next trading day. The objective function of the ETF sponsor is to minimize the index tracking error and transaction costs. The profit function of the AP depends additionally on the inventory risk on its balance sheet. Once a day, typically until 4 pm or sometimes earlier for bond ETFs, the AP can issue an order to create or redeem shares with the fund sponsor, who then approves the order.

Funds gained the ability to use custom baskets with the introduction of the ETF Rule on September 26, 2019. The rule became effective on December 23, 2019, with a one-year transition period, meaning that by December 22, 2020, all prior exemptive orders were rescinded, and funds were required to comply with the new regulations. Since the SEC stopped issuing broad exemptive orders for custom baskets around 2006, only newer funds were significantly impacted by the ETF Rule, while many larger, older funds were largely unaffected and serve as a control group in my econometric analysis.

To use custom baskets under the ETF Rule, funds must adopt specific internal procedures, including creating detailed policies on the construction and acceptance of custom baskets, updating disclosures on their websites, and maintaining updated records. This includes amending SEC Form N-1A, Form N-8B-2, and Form N-CEN. However, these filings do not provide a precise timeline for when funds began complying, as many funds were already in compliance with the first two forms, and Form N-CEN is only filed annually. As a result, for the purpose of analysis, I run an at-will treatment approach, using the rule's effective date as the first point at which funds could adopt custom baskets. This approach implies that my results represent a lower bound, as not all funds would have immediately transitioned from pro-rata to custom basket replication at that time.

Figure 2.
ETF Creation and Redemption Process. This figure illustrates the share creation and redemption process of an ETF. In a creation event, an authorized participant (AP) purchases the underlying assets of the ETF and delivers them to the fund in exchange for newly issued ETF shares.These new shares are then sold on the secondary market, where ETF investors buy and sell shares among themselves. When ETF investors buy shares on the secondary market, they are purchasing existing shares from other investors. If there is high demand for the ETF and a shortage of available shares, APs may initiate the creation process to increase supply (Lettau and Madhavan (2018)). This is depicted in the figure below by flow of cash from the ETF investor to the AP in exchange for ETF shares. The AP acquires the ETF shares via the outer circle flow, by exchanging underlying securities for ETF units with the fund sponsor. It can source the underlying with the received cash from the open market, or from assets sitting on its balance sheet. The reverse flow occurs during a redemption event. The AP acquires ETF shares from the secondary market and returns them to the fund in exchange for a basket of the underlying assets, which reduces the number of shares in circulation. This typically occurs when ETF investors are selling their shares and there is an excess supply. These transactions are conducted in the primary market, where new shares are created or redeemed in large blocks known as creation units, typically in lots of 50,000 shares and sometimes ranging from 25,000 to 100,000 shares, depending on the specific fund. The size of these creation units can contribute to temporary premiums or discounts between the ETF price and its NAV, as APs only engage in creation or redemption activities when the price discrepancy is sufficiently large to cover the transaction costs and operational expenses involved. Additionally, the liquidity of the underlying assets influences the premium or discount of an ETF relative to its NAV, with illiquid assets typically leading to larger premiums or discounts due to higher transaction costs and market impact, while liquid assets facilitate more efficient arbitrage and tighter alignment with NAV.



3 DATA

Data on authorized participants is sourced from SEC Form N-CEN filings, which registered investment companies are required to submit annually. This form includes detailed information on authorized participants, including entity identifiers and the dollar amounts of each AP's creation and redemption volume per fund. It also contains a dummy variable indicating whether an ETF relied on Rule 06c-11 in a given fiscal year. Although the requirement to file Form N-CEN only began in July 2018, making it a relatively novel dataset, it has already been used in research by Gorbatikov and Sikorskaya (2022), Zurowska (2022) Xiao (2022) and Du (2023), among others. I clean the filings data by removing duplicate filings as well as filings that do not cover the full fiscal reporting period.

The resulting dataset serves as the basis for the motivational evidence discussed in section 5. It covers 1,106 unique ETFs spanning from Q3 2018 to Q2 2024 and involves 67 unique APs. Each ETF typically has a higher number of registered APs compared to active APs. On average, an ETF has 35 registered APs, though 28 of them never participate in creation or redemption transactions. Therefore, I differentiate between registered APs and active APs, with the latter defined as those that have a non-zero creation or redemption value in a given year.

The main empirical analysis draws on ETF data from multiple sources. I define the ETF universe using tickers extracted from the CRSP Mutual Fund database, filtering specifically for exchange-traded funds and corporate bond funds based on share codes and CRSP object codes. Data on each ETF's issuing company and inception date are sourced from Bloomberg. This information is used to assign funds to the treatment and control groups. Funds that belong to an issuing company established before 2006 are categorized into the control group, as they had the opportunity to file exemptive orders. Newer funds fall into the treatment group. The control group includes fewer ETFs than the treatment group (8,936 in the control group versus 9,115 in the treatment group). However, because many of the largest funds belong to older issuing companies in the control group, its total NAV is significantly higher. The NAV of the control group is 4.9 trillion USD, while the NAV of the larger treatment group is 4.5 trillion USD. Additionally, I use Bloomberg to calculate ETF tracking errors. From the Trade and Quote (TAQ) database, I collect intra-day effective and realized spreads, which are high-frequency liquidity measures. All variables are observed at the fund-month level, and further details on the variables of interest are provided in Section 3.1.

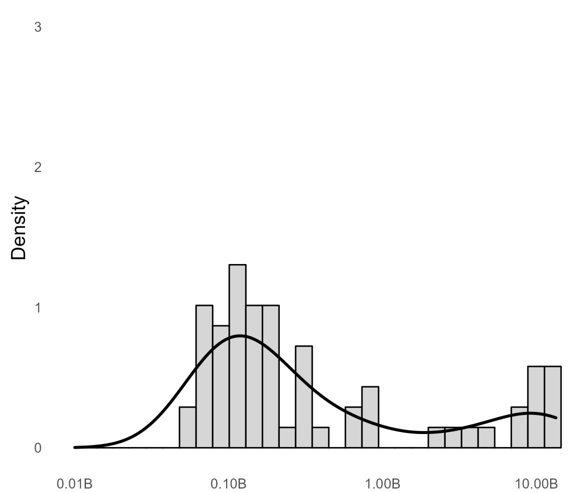
Table 1 shows summary statistics for treatment and control groups. I report the mean and standard deviation of all main variables of interest and of ancillary variables.

Table 1.
Summary Statistics for Treated and Control Groups. This table presents summary statistics for the data set used in this study. I report general and microstructure statistics, split between treatment and control groups. For each variable, I present mean, standard deviation and sample size.

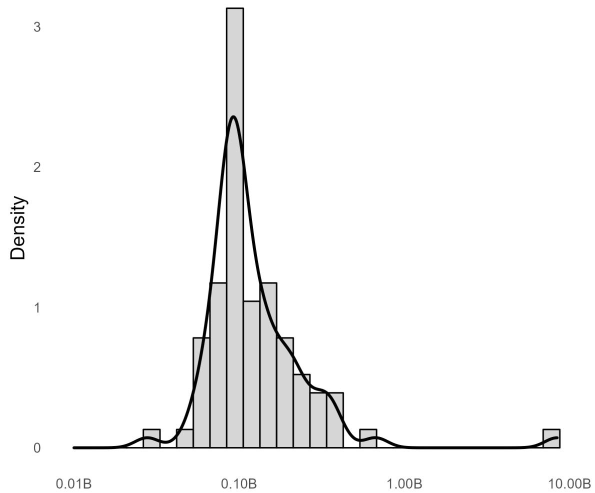
Variable	**Treated**			**Control**		
	Mean	**SD**	**N**	**Mean**	**SD**	**N**
Net Asset Value ($ bn)	4,545	1,021	9,115	4,931	666	8,936
Effective half-spread (%)	6.21	3.77	9,115	3.71	2.05	8,936
Realized half-spread (%)	6.25	53.67	9,115	2.08	9.64	8,936
Order imbalance ratio (%)	44.17	11.59	9,115	30.13	8.74	8,936
30-min variance ratio (%)	40.62	213.55	9,115	61.19	487.77	8,936
Tracking Error (%)	0.46	0.42	406	0.62	0.57	2,236

A potential concern is that the empirical distributions of treatment and control groups differ systematically. Figures 3 and 4 illustrate this point by showing the distributions of total net assets, benchmark indeces, as well as average credit ratings and duration of underlying bonds for each group at the beginning of my sample period. It can be seen that while empirical distributions differ, some overlap exists. This property is useful as it enables me to control for differences in these dimensions by using a matched sample approach. Section 5 explains the matching procedure in detail.

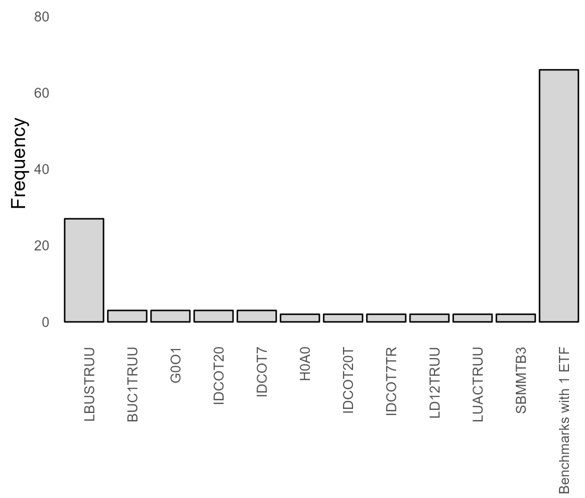
Figure 3.
Histograms and kernel densities for ETFs. This table presents histograms and kernel densities on the ETFs in the data set. Plot (a) shows the distribution of the log dollar value of NAV for the treatment group, plot (b) shows the equivalent distribution for the control group, plot (c) shows frequencies of the corresponding benchmark indices for the treatment group, and plot (d) shows the equivalent frequencies for the control group.



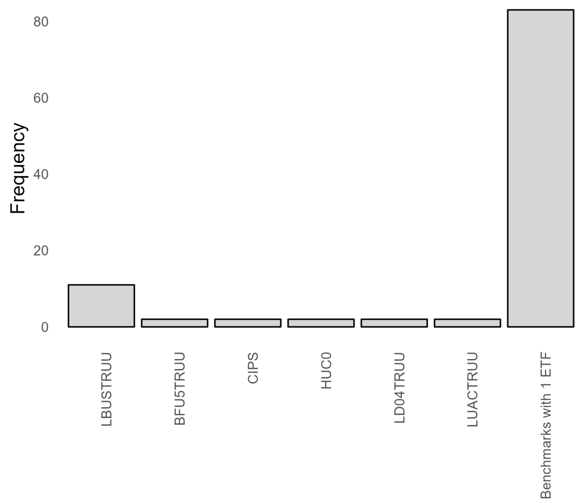
(a) Histogram and kernel density of log dollar value of NAV
(Treatment group)



(b) Histogram and kernel density of log dollar value of NAV
(Control group)

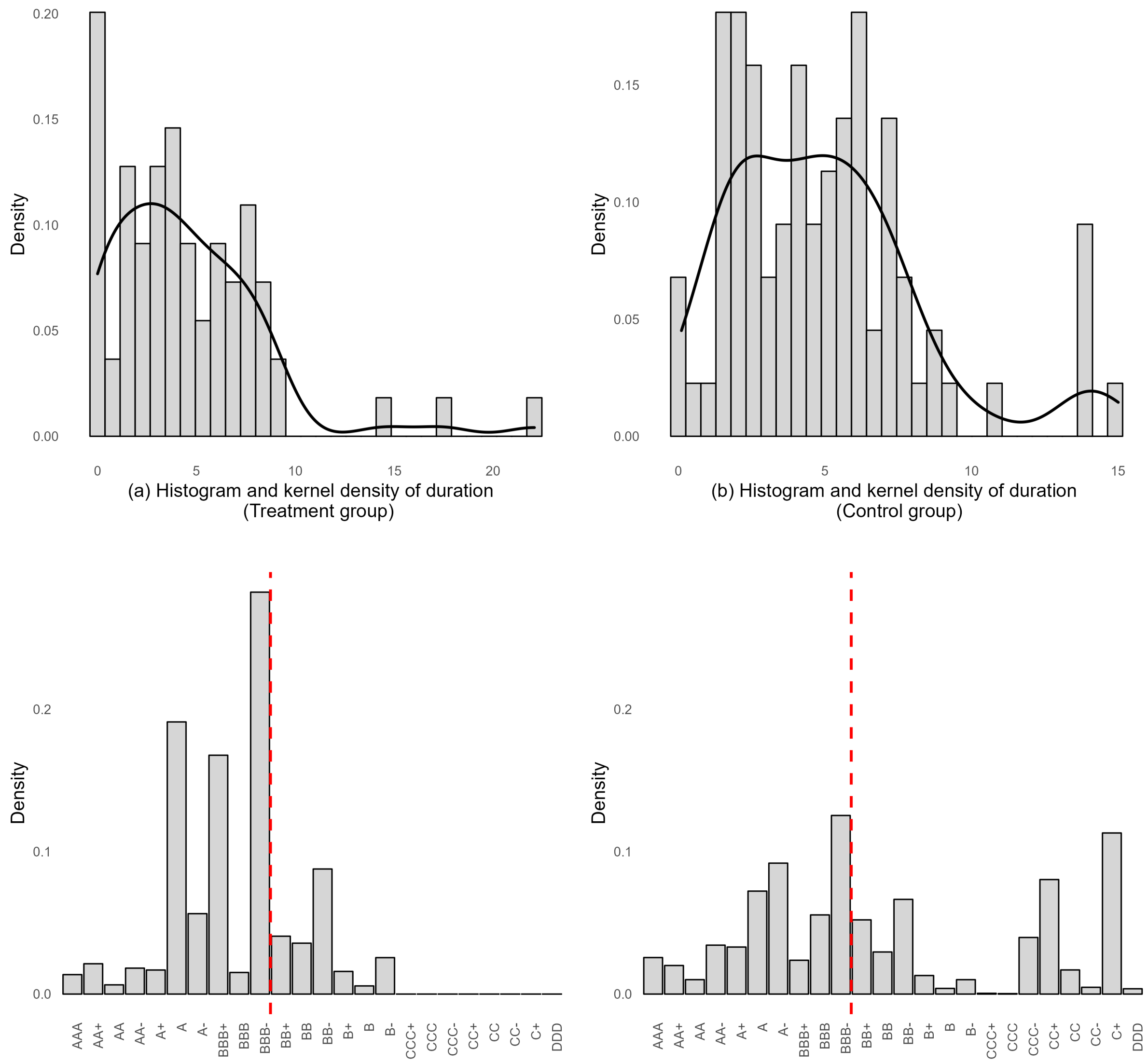


(c) Frequencies of benchmark indices
(Treatment group)



(d) Frequencies of benchmark indices
(Control group)

Figure 4.
Histograms and kernel densities for underlying holdings. This table presents histograms and kernel densities on the underlying assets of the ETFs in the data set. Plot (a) shows the distribution of the duration of the underlying bonds for the treatment group, plot (b) shows the equivalent distribution for the control group, plot (c) shows distribution of credit ratings of the underlying bonds within the treatment group, and plot (d) shows the equivalent distribution for the control group. The red line in plots (c) and (d) signifies the cutoff between investment grade and high yield ratings.



(a) Histogram and kernel density of duration (Treatment group)

(b) Histogram and kernel density of duration (Control group)

(c) Histogram of credit ratings (Treatment group)

(d) Histogram of credit ratings (Control group)

3.1 *Variables of interest*

The main variables of interest are the effective and realized half-spreads, and the tracking error.

The effective half-spread measures trading costs using prices actually obtained by investors. the transaction cost of executing a trade relative to prevailing bid and ask prices, reflecting the impact of liquidity on trade execution. In line with the microstructure literature (see Foucault et al. (2023)), it is given as:

$$S_e \equiv d \cdot \frac{(p - m)}{m}, \tag{1}$$

where d is the order direction indicator and equals 1 for buyer-initiated and -1 for seller-initated trades, and m is the midquote on the market prior to a transaction executed at price p. The effective spread can be interpreted as a measure of a transaction's impact on the price, because it measures the deviation of the actual execution price from the midprice prevailing just before the transaction. It's value is strictly positive.

In contrast, the realized half-spread adopts the viewpoint of liquidity providers. It measures the difference between the transaction price and the midprice at some time, Δ, after the transaction, where the interval Δ should be long enough to ensure that market quotes have adjusted to reflect the price impact of the transaction (Foucault et al. (2023)). Let p_t be the price of the transaction at time t, d_t the direction of the market order triggering it, and m_t the midprice at time t. The realized half-spread for this transaction is then given by:

$$S_r = d_t(p_t - m_{t+\Delta}) = d_t(p_t - m_t) - d_t(m_{t+\Delta} - m_t), \tag{2}$$

where the first term, $d_t(p_t - m_t)$, is the effective spread, and the second one, $d_t(m_{t+\Delta} - m_t)$, is the price impact of the transaction, defined as the change in the midprice that occurs after it. The overall expression can thus be seen as a measure of profit earned by the liquidity supplier on the transaction at time t if he unwinds his position at the midprice $t + \Delta$. Using the definition of the effective spread from equation 1 in equation 2, one can rewrite the average realized bid-ask spread as:

$$E(S_r) = E(S_e) - E(d_t(m_{t+\Delta} - m_t)) \tag{3}$$

This expression shows that the average realized spread is smaller than the average effective spread if $E(d_t(m_{t+\Delta} - m_t)) > 0$, that is, if on average transactions have a positive price impact. Interestingly, if the effective spread is on average smaller than the price impact, liquidity providers would lose money on average, as $E(S_e) < E(d_t(m_{t+\Delta} - m_t))$.

Finally, the ETF tracking error measures the deviation between ETF and benchmark index returns. It quantifies the volatility of the difference between the ETF's performance and that of the index it seeks to replicate, and is given as:

$$\text{Tracking Error} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(R_{ETF,t} - R_{Benchmark,t})^2} \tag{4}$$

where $R_{ETF,t}$ represents the return of the ETF at time t, $R_{Benchmark,t}$ denotes the return of the benchmark index at the same time, and N signifies the total number of periods considered for the calculation.

4 MOTIVATING EVIDENCE

The purpose of this section is to introduce stylized facts on the role of authorized participants during ETF arbitrage, and to illustrate how their involvement varies across different replication methods.

A natural starting point is to explore correlations between custom basket adoption and AP activity. Authorized participants are the market participants who actively drive creations and redemptions. Literature on Authorized participants. Comment on why custom basket use impacts APs.

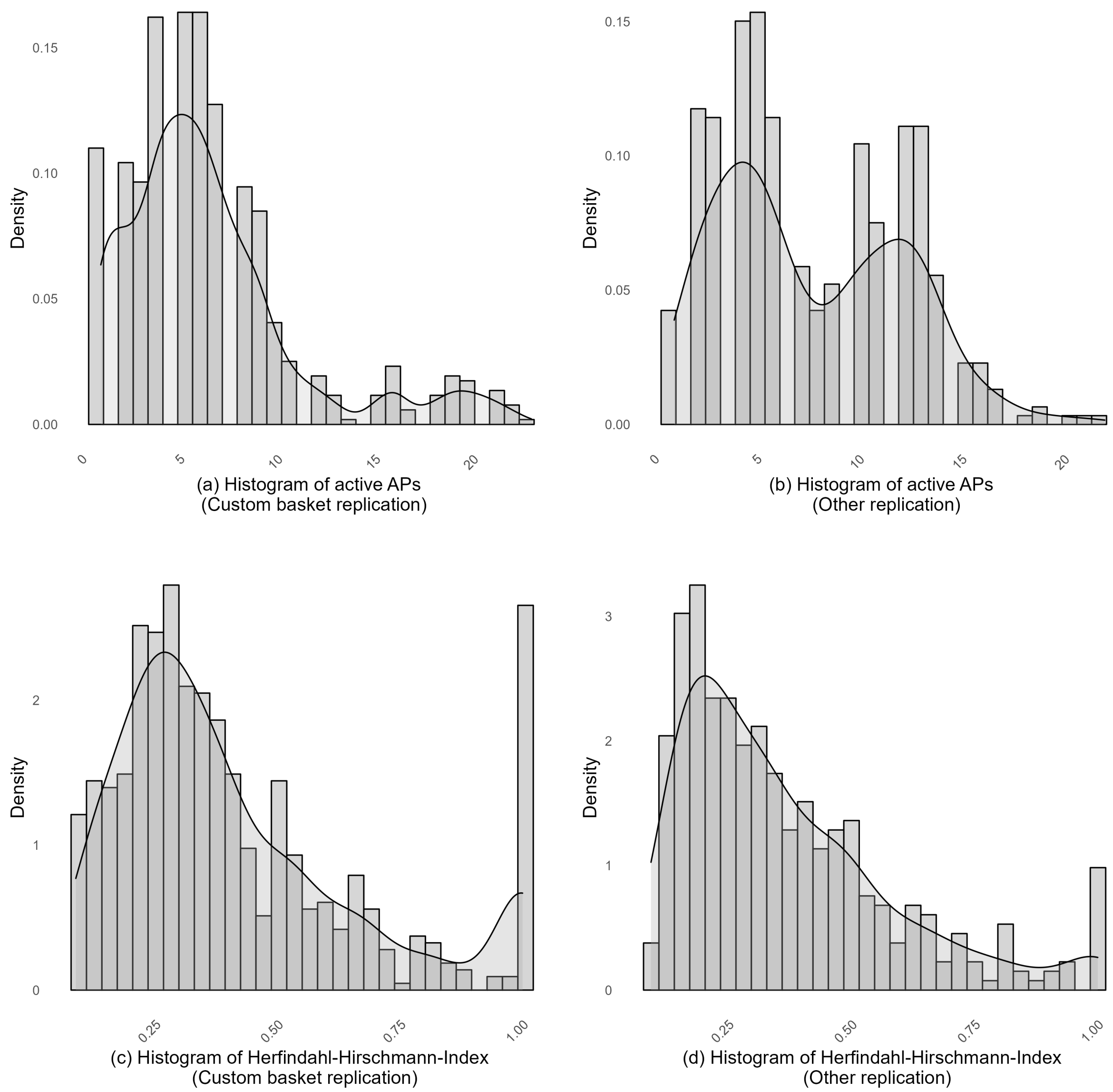
Using Form N-CEN filings data from the SEC, I compute the number of registered authorized participants and active authorized participants for each fund since 2018. I define an active AP as a registered authorized participant who has performed at least one creation or redemption in the filing period. Table A.2 in the appendix show a list of all unique APs in the dataset. To measure market concentration of APs, I calculate the Herfindahl—Hirschman index (HHI). This is given as the sum of squared market shares of each AP:

$$\text{HHI}_i = \sum_j \left(\frac{S_{ji}}{S_i}\right)^2$$

where S_{ji} denotes the j-th AP's aggregate transactions with the i-th ETF. Under this specification, HHI can take values between 0 and 1, with 1 corresponding to the case with only one active AP. I discard all fund-quarter observations where the number of active APs is 0, as the above ratio is then undefined. When the number of active APs is 0, the above ratio is undefined; however, in the empirical analysis, I assume it is equivalent to the monopoly by setting HHI to 1.

A preliminary examination of the data, as illustrated in Figure 5, reveals histograms that show the distribution of active APs and the HHI for ETFs utilizing custom basket replication in comparison to those employing conventional replication. The findings indicate that funds adopting custom basket replication generally exhibit a lower count of active APs. It is crucial to recognize that active and registered AP counts differ fundamentally; active APs are those who engage in transactions and contribute meaningfully to market activity, while registered APs may include entities that do not participate in creations or redemption at all. As such, relying solely on registered AP counts could lead to misleading conclusions regarding market dynamics. Furthermore, the histograms of the HHI substantiate the conclusion that custom basket replication promotes a more concentrated market structure. The HHI data further illustrate that while the number of active APs diminishes for custom basket ETFs, the concentration of creations and redemptions, as indicated by the HHI, increases, thereby underscoring the evolving dynamics within these funds.

Figure 5.
Histogram of active APs and HHI indeces This figure presents four histograms, each overlaid with a kernel density plot. Panels (a) and (b) show histograms for the number of active authorized participants (APs), while panels (c) and (d) show histograms for the Herfindahl-Hirschman Index (HHI). Panels (a) and (c) pertain to custom basket ETFs, whereas panels (b) and (d) correspond to non-custom basket ETFs.



I complement my observations more formally by examining whether there is a relationship between the usage of custom baskets in exchange-traded funds (ETFs) and various factors such as the count of registered authorized participants (APs), active APs, and the Herfindahl-Hirschman Index (HHI) values associated with creations, redemptions, and total flows. To study this question, I estimate the following binary logistic regression:

$$logit(CB_{it}) = \beta_0 + \beta_1 \mathrm{AP}_{it} + \epsilon_{it},$$

where CB_{it} is a binary indicator representing whether custom baskets are used for ETF i in quarter t, and AP_{it} represents registered and active AP counts.

Table 2.

Logit Regression Results for Custom Basket Use This table presents the results of logit regressions analyzing the determinants of custom basket usage among ETFs. The dependent variable CB_i is a binary dummy indicating whether a custom basket is used. The regression equation is specified as follows:

$$\mathrm{Logit}(CB_{it}) = \beta_0 + \beta_{\mathrm{AP}} \cdot \mathrm{AP}_{it} + \epsilon_{it}$$

where Y_i is the binary indicator for custom basket use, and AP_i represents the number of authorized participants. All models employ ETF-clustered robust standard errors.
In Model (1), the coefficient β_{AP} refers to the effect of the number of registered authorized participants on the log-odds of using a custom basket. In Model (2), β_{AP} reflects the effect of the number of active authorized participants on the log-odds of custom basket use. The intercepts and McFadden's pseudo-R^2 values are also reported.

	(1)	(2)
Number of registered APs	0.052^{***}	
	(0.056)	
Number of active APs		-0.055^{***}
		(-0.014)
Intercept	-1.299^{***}	0.869^{***}
	(0.21)	(0.116)
Observations	1,106	1,106
McFadden's Pseudo$-R^2$	0.061	0.011

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

My empirical results are shown in Table 2. The logit regressions reveal significant relationships between custom basket usage and the AP count. Custom basket use is associated with a higher registered AP count but a lower active AP count. The results from the logit regression presented in Table 2 indicate the relationship between the number of authorized participants and the likelihood of using a custom basket in ETFs. Odds ratios are given as $e^{0.052} \approx 1.0534$ for registered AP counts (model 1) and as $e^{-0.055} \approx 0.9460$ for active AP counts (model 2). This indicates that for each additional registered AP, the likelihood of custom basket use increases by approximately 5.34%, whereas for each additional active AP, the likelihood decreases by approximately 5.4%. This finding underscores that the likelihoods move in opposite directions, emphasizing that the active AP count may provide a more meaningful measure of market concentration when evaluating custom basket usage.

I complement the first analysis with further logistic regressions based on the Herfindahl-Hirschmann index for creations, redemptions and total flows. In Table 3, I estimate:

$$\text{Logit}(CB_{it}) = \beta_0 + \beta_{\text{HHI}} \cdot \text{HHI}_{it} + \epsilon_{it}$$

where CB_{it} is the binary indicator for custom basket use, HHI_{it} represents the Herfindahl-Hirschman Index (HHI) for creations, redemptions, or all flows, and ϵ_{it} is the error term. All models employ ETF-clustered robust standard errors.

In regression (1), the coefficient β_{HHI} for the Herfindahl-Hirschman Index (HHI) for creations is significant at the 5% level, indicating that a higher HHI for creations is associated with an increased likelihood of using a custom basket. Regression (2) shows a significant impact of the HHI for redemptions on the log-odds of custom basket use, also at the 5% level. Meanwhile, regression (3) demonstrates a strong positive association with custom basket usage, as the coefficient for the HHI for all flows is significant at the 1% level. The odds ratios indicate that a 0.1 increase in the HHI for creations leads to an approximate 6.21% increase in the odds of using a custom basket, calculated as $e^{0.621 \cdot 0.1} \approx 1.0621$. Similarly, for the HHI for redemptions, a 0.1 unit increase is associated with a 4.97% increase in odds, given by $e^{0.486 \cdot 0.1} \approx 1.0497$. Lastly, a 0.1 unit increase in the HHI for all flows results in an approximately 8.99% increase in the odds of custom basket use, calculated as $e^{0.861 \cdot 0.1} \approx 1.0899$.

Table 3.
Regression Results for Custom Basket Use This table presents the results of logit regressions analyzing the determinants of custom basket usage among ETFs. The dependent variable CB_i is a binary dummy indicating whether a custom basket is used. The regression equation is specified as follows:

$$\text{Logit}(CB_{it}) = \beta_0 + \beta_{\text{HHI}} \cdot \text{HHI}_{it} + \epsilon_{it}$$

where Y_i is the binary indicator for custom basket use, HHI_i represents the Herfindahl-Hirschman Index (HHI) for creations, redemptions, or all flows, and ϵ_i is the error term. All models employ ETF-clustered robust standard errors.
In Model (1), the coefficient β_{HHI} for HHI for creations indicates the effect of HHI on the log-odds of using a custom basket. In Model (2), the coefficient reflects the impact of HHI for redemptions on the log-odds of custom basket use. In Model (3), the coefficient pertains to the HHI for all flows, demonstrating its effect on the log-odds of custom basket use. The intercepts and McFadden's pseudo-R^2 values are also reported.

	(1)	(2)	(3)
HHI for creations	0.621**		
	(0.252)		
HHI for redemptions		0.486**	
		(0.218)	
HHI for all flows			0.861***
			(0.272)
Intercept	0.229*	0.279**	0.146
	(0.118)	(0.108)	(0.121)
Observations	1,106	1,106	1,106
McFadden's Pseudo$-R^2$	0.004	0.003	0.007

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

5 HYPOTHESIS DEVELOPMENT

In this section, I formulate testable hypotheses to guide my empirical analysis on the microstructure implications of custom basket replication.

My first hypothesis delineates the effects of custom basket replication on ETF liquidity. Since ETF sponsors use custom baskets to replace hard-to-source assets with more liquid proxy assets, I predict that the aggregate liquidity of underlying assets in custom basket replication is higher than in pro-rata replication, and that, as a result, the ETF itself will be more liquid. This should be reflected in ETF spreads.

Hypothesis 1. (ETF Liquidity: half-spreads).

An ETF that uses custom basket replication should trade with tighter effective and realized half-spreads than an equivalent ETF that uses pro-rata replication.

Hypothesis 1 relies on the premise that there exists a monotonically increasing function $f(\cdot)$ that maps the liquidity of the underlying assets, L_{UA}, to the liquidity of the ETF, L_{ETF}. Given the monotonic increase, I expect that $f'(L_{UA}) \geq 0$ for all L_{UA}, indicating that as L_{UA} increases, L_{ETF} also increases or remains constant, but does not decrease.

My second prediction is that custom baskets increase the ETF's index tracking error, $\sqrt{\text{Var}(R_{ETF} - R_{BM})}$. I motivate this hypothesis by considering the added flexibility that custom baskets provide in the composition of ETF portfolios. They allow deviations of the asset weights relative to the benchmark composition, and, in theory, they also allow the inclusion of assets that are entirely outside the benchmark. Tracking errors can be measured using various metrics such as the standard deviation of returns, mean absolute deviation, and root mean square deviation.

Hypothesis 2. (Index tracking errors).

Custom baskets increase the index tracking error of an ETF.

6 IDENTIFICATION STRATEGY

Isolating the effect of custom baskets on ETF dynamics is challenging because unobserved confounding variables may also affect ETF spreads and index tracking errors. For example, either could be driven by daily trading volumes, the number of authorized participants per ETF or other unobserved factors. To overcome this difficulty, I pursue a difference-in-differences approach and complement this with matching regressions.

6.1 *Diff-in-diff*

The promulgation of SEC Rule 06c-11 in December 2019 provides a robust quasi-natural experiment because its implementation is plausibly exogenous to ETF spreads and index tracking errors, and it is unlikely to have been influenced by the determinants of these variables. The legislation prescribes a clear separation between affected and unaffected funds, mitigating self-selection issues. In a diff-in-diff setup, I run the following regression:

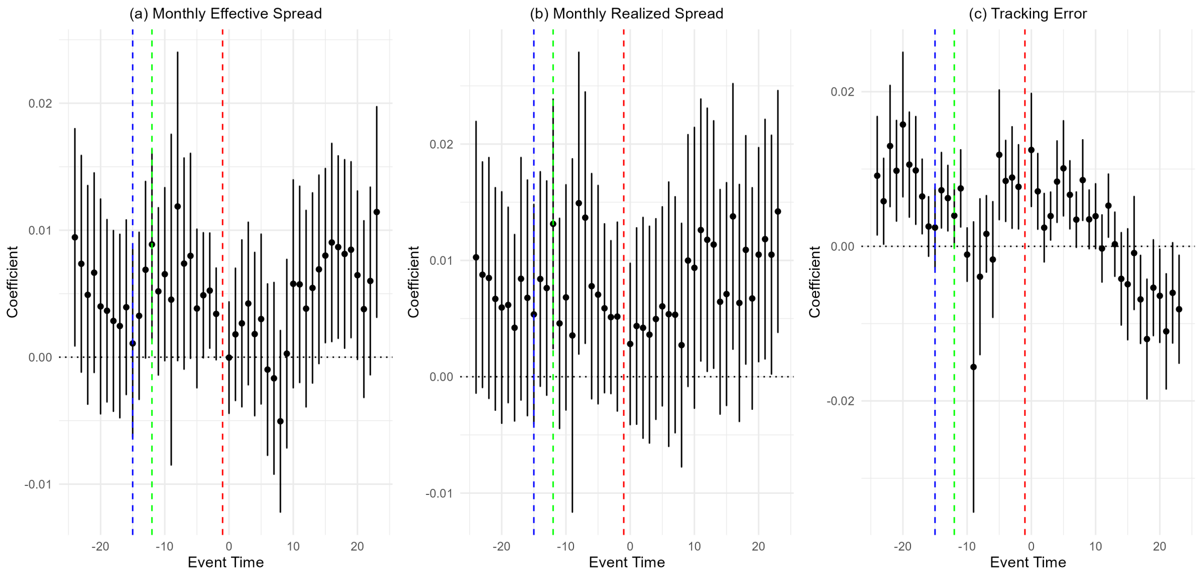
$$Y_{it} = \alpha + \beta_1 Treated_i + \beta_2 Post_t + \beta_3 Treated_i \times Post_t + \epsilon_{it}, \tag{5}$$

where Y_{it} is the measure of interest for fund i in month t. $Post_t$ is a dummy variable equal to one after 23 December 2019, which marks the post-transition date of the ETF Rule, and $Treated_i$ is equal to one if fund i has its inception date after 2006. In the event of serial correlation in the error terms for a given fund, the OLS standard deviations of the estimates in equation 5 may be biased. I therefore compute standard deviations for my estimates by computing double-clustered standard errors. In other words, I allow for correlation in residuals over time and across funds (see Thompson (2011)). In the above equation, $Treated_i$ controls for differences in outcomes between treated and control groups that are fixed over time and $Post_t$ controls for factors that may affect the development of the outcome variable over time across all funds.

6.2 *Parallel trends*

The identifying assumption in this empirical set-up rests on the condition that parallel trends are upheld. This implies that, absent the treatment, both the treated and control groups would have experienced similar changes in the outcome variable over time. The parallel trends assumption is inherently untestable. Tests of pre-treatment coefficients in event studies often lack sufficient power, making a failure to reject the null hypothesis inconclusive regarding the existence of pre-trends (see Roth (2022)). Instead, figure 6 presents event study plots on the variables of interest. The red line indicates the treatment date, corresponding to the post-transition date on 22 September 2019. The green line represents the pre-transition date, one year earlier, and the blue line denotes the announcement date in September 2018. The plots reveal that the confidence intervals for the pre-treatment coefficients predominantly include zero for effective and realized spreads but not for the tracking error. At the same time, the tracking error exhibits a more pronounced trend shift following the treatment.

Figure 6.
Event study plots. This figure presents event study plots on the difference-in-differences (DiD) analysis for (a) effective spread, (b) realized spread, and (c) tracking error. The y-axis represents the coefficient estimates. I use these plots to assess parallel trends by comparing the pre-event and post-event trajectories of the treated and control groups.



An issue with the diff-in-diff framework is that systematic differences between treated and control groups can undermine the parallel trends assumption, leading to biased estimates of treatment effects. In the next section, I perform various matching regressions as a further robustness check.

To address this, researchers can conduct pre-treatment tests, match covariates, or use alternative control groups to ensure that the parallel trends assumption holds.

6.3 *Matching regressions*

If treated and control groups differ systematically, these differences may confound the estimated treatment effect. For instance, variations in (i) net asset value, (ii) benchmark composition, or the (iii) average credit rating and (iv) average duration of underlying bonds between treated and control ETFs could lead to divergent pre-treatment trends, independent of the treatment itself. To address this issue, I match each fund in the treated group with a counterpart in the control group based on these four characteristics. I then estimate the following model:

$$Y_{it} - Y_{it}^{match} = \alpha + \delta_1 Post_t + \eta_{it} \tag{6}$$

where Y_{it} is the outcome variable for fund i in month t and Y_i^{match} is the outcome variable for the matched fund in the control group. δ_1 measures the average change in the difference $Y_{it} - Y_{it}^{match}$ associated with the introduction of the ETF Rule. Consistent with the diff-in-diff approach, I use double-clustered standard errors to account for both clustering at the fund level and within-period correlation in the data.

In line with Foucault et al. (2011), I use three different matching methods to ensure that the results are robust to the matching method: (1) quartile matching, (2) percentage difference matching, and (3) propensity score matching. For QM, I calculate average net asset values and duration of the ETFs in 2017, and group stocks in quartiles of NAV and duration. Thus, I obtain 16 groups of ETFs. The variable is then defined as the average value of over all control stocks that are in the same group as treated stock i. For PDM, I calculate for each treated fund the percentage differences between its NAV and the NAV of each control stock and its average duration and the average duration of each control stock in December 2017 (Foucault et al. (2011), Guo et al. (2011)). I match each treated fund i with a control fund that minimizes the maximum difference between the two computed differences for NAV and duration. I exclude treated funds from the sample where the distance between fund i and its nearest neighbour exceeds 10%. For PSM, I estimate propensity scores for each unit using a logistic regression model. The propensity score $\hat{e}_i$ is the predicted probability that fund i newly obtains the right to use custom baskets based on its observed characteristics X_i. The dependent variable of the logistic regression is the binary treatment indicator D_i, and the independent variables are the covariates X_i. Mathematically, $\hat{e}_i = \Pr(D_i = 1 \mid X_i)$ represents the conditional probability of unit i receiving the treatment given its observed covariates X_i. First, I estimate the following logistic regression:

$$Treated_i = \alpha + \beta ln(NAV_i) + \gamma BM_i + \theta D_i + \lambda CR_i + \nu_i \tag{7}$$

where *Treated* equals 1 if fund i has its inception after 2006, S_i is the average total net asset value of fund i over the year 2017, BM_i is a categorical variable for the corresponding benchmark, D_i is the average duration of underlying bonds, and CR_i is an indicator of the credit rating of underlying bonds with values 1 for investment grade ETFs and 0 for high-yield ETFs. With these estimates, I calculate the likelihood that a fund has an inception after 2006 given its total net assets and daily turnover. I then pair each treated fund with the control group fund that has the nearest score.

7 MAIN EMPIRICAL FINDINGS

This section presents empirical findings on custom basket replication in ETFs. Section 7.1 analyzes effects on ETF spreads, and Section 7.2 on index tracking errors. My results indicate that custom baskets tighten bid-ask spreads, enhancing liquidity, but at the cost of increased tracking errors. These findings are robust to various robustness checks.

7.1 *Spread size*

Following Hypothesis 5, custom basket replication should tighten spreads. The intuition behind this is straightforward: by replacing illiquid underlying securities with more liquid proxy assets, the overall liquidity of the ETF's holdings improves. Given the interconnectedness between the markets for ETFs and their underlying assets, this is likely to lead to liquidity spillover into the fund itself.

Table 4 presents the difference-in-differences results for effective and realized spreads in Panels A and B, respectively. The coefficient estimate β_2 for the interactive variable is consistently negative across both panels and all specifications. The effective spread values demonstrate economic significance across all models, ranging from -23 basis points within a 24-month window around the treatment (significant at the 10% level) to -62 basis points within a 12-month window around the event (significant at the 1% level). These estimates are economically significant, as they represent a tightening of up to 10% relative to the mean effective half-spread for ETFs in the treatment group. These results are based on double-clustered standard errors to allow for correlation in residuals over time and across funds.

In contrast, my estimates for β_2 related to the realized spread are also negative but are statistically significant only for the specification with a 36-month time window around the event. In this case, the magnitude of the tightening is 86 basis points (significant at the 1% level), which is substantial compared to the mean realized half-spread of 6.25% for treated funds in the sample. Although the specifications with 12- and 24-month windows are not statistically significant, it is important to note that the empirical design is based on an at-will treatment; thus, these estimates should be interpreted as a conservative lower bound.

The different time windows around the treatment date served as a first robustness check to my difference-in-differences results. As a further robustness check, I perform matching regressions. Table 6, Panels A.1 and A2. present regression results on effective and realized spreads with quartile matching, percentage difference matching and propensity score matching. The magnitude of the tightening more than doubles across all matching specifications relative to the diff-in-diff results. The effective spreads tighten by 140 bps (percentage difference matching) to 234 bps (propensity score matching). The realized spreads tighten by 115 bps (PDM) to 206 bps (PSM). All matching specifications are statistically significant at the 1% level.

The different time windows around the treatment date serve as an initial robustness check for the difference-in-differences results. To further validate these findings, I conduct matching regressions. Table 6, Panels A.1 and A.2, presents regression results for effective and realized spreads using quartile matching (QM), percentage difference matching (PDM), and propensity score matching (PSM). Across all matching specifications, the magnitude of tightening more than doubles relative to the difference-in-differences estimates. Effective spreads tighten by 140 bps (PDM) to 234 bps (PSM), and realized spreads tighten by 115 bps (PDM) to 206 bps (PSM), with all results statistically significant at the 1% level. The number of observations decreases in the matching analysis relative to the difference-in-differences specification due to the exclusion of funds missing data on fund size, benchmark, credit rating, or duration. Despite this reduction in sample size, the adjusted R-squared generally improves.

I conduct several further robustness checks. I test for anticipation effects surrounding the adoption of SEC Rule 06c-11, which was announced on September 26, 2019, and came into effect on December 23, 2019 with a one-year transition period for compliance. The main specifications above are based on the end of the transition period on 22 December 2020 as the treatment date, as funds were likely to have finalized the necessary administrative steps for custom basket adoption by then. Table A.4 presents robustness tests using the announcement date and the start of the transition period as alternative treatment dates. The evidence for anticipation effects is mixed: for effective spreads, the signs, magnitudes, and statistical significance of the coefficients vary across specifications, indicating an inconsistent effect. However, realized spreads display a more consistent anticipation effect, showing a tightening across all specifications, with statistical significance at the 1% level in most cases. Despite this, the observed magnitudes of spread tightening for the announcement and pre-transition dates are smaller than those for the post-transition date under the main specification. While some anticipation effect for realized spreads may be present, these results should be interpreted cautiously, as the relatively low adjusted R-squared suggests that the model captures only a small portion of the variation in realized spreads.

Due to the administrative steps required, funds likely adopted custom baskets at varying intervals within the transition period, though the exact adoption date for each fund is unobservable. However, it is reasonable to assume that most funds completed the adoption at some point during the one-year transition. To account for this variation, I conduct an additional robustness check by excluding data from the transition period. The results, shown in Table A.5, support the main findings.

The results in Table A.5 also serve as a robustness test for potential COVID-19 effects. During the COVID-19 period, the bond market experienced significant illiquidity, and many bond ETFs traded at a discount to NAV. By excluding data from this period, the robustness results in Table A.5 confirm that the main findings are not driven by the unusual bond market conditions during COVID-19. Notably, I exclude a relatively long time period, from December 2019 to December 2020. In comparison, Haddad et al. (2021) exclude the interval from 19 February 2020 to 16 April 2020.

Finally, I show a placebo test in Table A.6 and unwinsorized results in Table A.3. The placebo test is based on 01 June 2018, well in advance of the announcement date. That being said, the ETF Rule was first proposed by the SEC on 28 June 2018. A placebo test with a date will in advance of that date would have been preferable, but I use 01 June 2018 due to the limitations in my dataset. The results of the placebo test show mostly non-significant results for effective and realized spreads, thereby supporting the hypothesis. The unwinsorized results further confirm the main results.

7.2 *Index tracking error*

Hypothesis 5 predicts that index tracking errors should widen. Intuitively, substituting a benchmark asset with a proxy asset can contribute to an increase in tracking error when the proxy asset exhibits different return volatilities than the benchmark asset.

My results in Table 5 indicate that index tracking errors widen in the 24-month and 36-month windows surrounding the event, with increases ranging from 0.7 to 1.38 percentage points. In contrast, the 12-month window shows a decrease of 0.91 percentage points in the tracking error. That being said, the 12-month period coincides with the transition phase, which may influence the results. In the robustness check presented in Table A.4, I find that excluding the transition period data substantially increases tracking errors. This specification also improves the R-squared.

The matching regressions in Table 6 largely support my earlier findings. I observe statistically significant increases in tracking error under both the QM and PSM specifications, while the PDM specification shows no significant change. This pattern also persists in the robustness checks with alternative treatment dates shown in Table A.4. Specifically, when analyzing both the announcement date in September 2019 and the pre-transition date in December 2019, I find that tracking error increases under custom basket replication. In the placebo test outlined in Table A.6, the QM specification becomes statistically insignificant, while the difference-in-difference and PSM specifications remain significant at the 5% and 1% levels, though with reduced magnitudes.

8 CONCLUSION

In this paper, I demonstrate that the introduction of custom baskets significantly tightens both effective and realized spreads, indicating improved liquidity in the ETF market. This tightening effect is robust across various time windows, with realized spreads showing more pronounced tightening over longer horizons. However, while these improvements in liquidity and arbitrage efficiency are beneficial, they are accompanied by increased tracking errors over medium to longer time horizons. These findings suggest that regulatory interventions aimed at improving liquidity and arbitrage efficiency must also consider potential trade-offs in tracking accuracy.

Table 4.
Diff-in-Diff Results for ETF Spreads. In this table, I estimate the impact of the ETF Rule on ETF liquidity, using two alternate measures of ETF spreads. I estimate the following regression:

$$Y_{it} = \alpha + \beta_0 Treated_i + \beta_1 Post_t + \beta_2 Treated_i \times Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t, $Post_t$ is a dummy variable equal to one after 22 December 2020, and $Treated_i$ is equal to one if fund i was listed after 2006. In panel A, $Y_{it} = Se_{it}$ and in panel B, $Y_{it} = Sr_{it}$. The sample period starts in December 2017 and ends in December 2023. In brackets, I report double-clustered standard errors allowing for correlation in residuals over time and across funds.

	12-month window	24-month window	36-month window
Panel A: Dependent variable: Effective Spread (bps)			
Treated × Post (β_2)	-62.32***	-23.35*	-32.74***
	(21.26)	(13.11)	(10.56)
Treated	266.22***	230.94***	231.48***
	(61.07)	(46.00)	(42.86)
Post	-219.52***	-149.95***	-146.98***
	(12.78)	(7.28)	(5.56)
Constant	496.13***	458.64***	447.27***
	(35.88)	(28.78)	(26.13)
Observations	5,767	12,039	18,055
$Adj.R^2$	0.10	0.06	0.06
Panel B: Dependent variable: Realized Spread (bps)			
Treated × Post (β_2)	-45.33	-15.49	-84.367***
	(30.30)	(20.23)	(23.82)
Treated	270.92***	282.81***	348.34***
	(60.39)	(48.47)	(51.48)
Post	-165.36***	-128.67***	-132.35***
	(13.21)	(7.93)	(8.63)
Constant	384.06***	374.61***	370.85***
	(26.87)	(20.34)	(21.43)
Observations	5,767	12,039	18,055
$Adj.R^2$	0.03	0.02	0.02

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Table 5.
Diff-in-Diff Results for Tracking Errors. In this table, I estimate the impact of the ETF Rule on index tracking errors. I estimate the following regression:

$$TE_{it} = \alpha + \beta_0 Treated_i + \beta_1 Post_t + \beta_2 Treated_i \times Post_t + \epsilon_{it},$$

where TE_{it} is the index tracking error for fund i in month t, $Post_t$ is a dummy variable equal to one after 22 December 2020, and $Treated_i$ is equal to one if fund i was listed after 2006. The sample period starts in December 2017 and ends in December 2023. In brackets, I report double-clustered standard errors allowing for correlation in residuals over time and across funds.

	12-month window	24-month window	36-month window
Dependent variable: Tracking error (% pts.)			
Treated × Post (β_2)	-0.91***	0.70***	1.38***
	(0.22)	(0.16)	(0.13)
Treated	4.04***	2.50***	1.87**
	(1.10)	(0.92)	(0.82)
Post	-0.49***	-0.07*	-0.00
	(0.07)	(0.04)	(0.03)
Constant	1.15***	1.01*	0.85**
	(0.38)	(0.53)	(0.39)
Observations	3,545	7,474	11,368
$Adj.R^2$	0.03	0.01	0.03

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Table 6.
Matching Results. In this table, I estimate the impact of the ETF rule using various matching methods. I estimate the following regression:

$$Y_{it} - Y_{it}^{match} = \alpha_i + \delta_1 Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t and $Y_{matchit}$ is the value of this measure for the match of fund i in month t in the group of control stocks. I use three different methods to choose a match for fund i in month t: quartile matching, percentage difference matching, and propensity score matching. Estimates of the effect of the reform (δ_1) with each matching procedure are reported in columns 1, 2, and 3, respectively. In panel A.1, $Y_{it} = Se_{it}$; in panel A.2, $Y_{it} = Sr_{it}$; and in panel B, $Y_{it} = TE_{it}$. The sample period starts in December 2017 and ends in December 2023. In brackets, I report double-clustered standard errors allowing for correlation in residuals over time and across funds.

Variable	**Quartile Matching**	**Percentage Difference Matching**	**Propensity Score Matching**
Panel A.1: Dependent variable: Effective Spread (bps)			
Post (δ_1)	-187.39***	-140.41***	-233.94***
	(18.07)	(25.70)	(24.22)
Observations	777	336	777
$Adj.R^2$	0.10	0.06	0.09
Panel A.2: Dependent variable: Realized Spread (bps)			
Post (δ_1)	-152.96***	-115.34***	-206.45***
	(15.38)	(22.59)	(20.82)
Observations	777	336	777
$Adj.R^2$	0.09	0.05	0.09
Panel B: Dependent variable: Tracking Error (%)			
Post (δ_1)	0.18***	0.13	0.14*
	(0.06)	(0.11)	(0.08)
Observations	777	336	777
$Adj.R^2$	-0.01	-0.02	-0.02

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

REFERENCES

Agarwal, V., Hanouna, P., Moussawi, R., and Stahel, C. W. (2021). Do etfs increase the commonality in liquidity of underlying stocks? Technical report, CFR Working Paper.

Ben-David, I., Franzoni, F., and Moussawi, R. (2018). Do etfs increase volatility? *The Journal of Finance*, 73(6):2471–2535.

Brogaard, J., Heath, D., and Huang, D. (2024). Etf sampling and index arbitrage. *Working Paper*.

Cespa, G. and Foucault, T. (2014). Illiquidity contagion and liquidity crashes. *The Review of Financial Studies*, 27(6):1615–1660.

Cheng, S., Massa, M., and Zhang, H. (2019). The unexpected activeness of passive investors: a worldwide analysis of etfs. *The Review of Asset Pricing Studies*, 9(2):296–355.

Cong, L. W., Huang, S., and Xu, D. (2024). The rise of factor investing. *NBER Working Paper*.

Da, Z. and Shive, S. (2018). Exchange traded funds and asset return correlations. *European Financial Management*, 24(1):136–168.

Dannhauser, C. D. (2017). The impact of innovation: Evidence from corporate bond exchange-traded funds (etfs). *Journal of Financial Economics*, 125(3):537–560.

Dannhauser, C. D. and Karmaziene, E. (2023). The dealer warehouse–corporate bond etfs. *Working Paper*.

Du, L. (2023). Inventory in liquidity transformation: Evidence from corporate bond etf creations. *Working Paper*.

Foucault, T., Pagano, M., and Röell, A. (2023). *Market liquidity: theory, evidence, and policy*. Oxford University Press, USA.

Foucault, T., Sraer, D., and Thesmar, D. J. (2011). Individual investors and volatility. *The Journal of Finance*, 66(4):1369–1406.

Gorbatikov, E. and Sikorskaya, T. (2022). Two aps are better than one: Etf mispricing and primary market participation. *Working Paper*.

Guo, S., Hotchkiss, E. S., and Song, W. (2011). Do buyouts (still) create value? *The Journal of Finance*, 66(2):479–517.

Haddad, V., Moreira, A., and Muir, T. (2021). When selling becomes viral: Disruptions in debt markets in the covid-19 crisis and the fed's response. *The Review of Financial Studies*, 34(11):5309–5351.

Helmke, A. (2023). Will etfs drive mutual funds extinct? *Working Paper*.

Holden, C. W. and Nam, J. (2017). Market accessibility, corporate bond etfs, and liquidity. *Kelley School of Business Research Paper*, (18-1).

Israeli, D., Lee, C. M., and Sridharan, S. A. (2017). Is there a dark side to exchange traded funds? an information perspective. *Review of Accounting Studies*, 22:1048–1083.

Koont, N., Ma, Y., Pástor, L., and Zeng, Y. (2024). Steering a ship in illiquid waters: Active management of passive funds. *Working Paper*.

Lettau, M. and Madhavan, A. (2018). Exchange-traded funds 101 for economists. *Journal of Economic Perspectives*, 32(1):135–154.

Malamud, S. (2016). A dynamic equilibrium model of etfs. *Working Paper*.

Pan, K. and Zeng, Y. (2023). Etf arbitrage under liquidity mismatch. *Working Paper*.

Reilly, C. (2022). The hidden cost of corporate bond etfs. *Working Paper*.

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights*, 4(3):305–322.

Shim, J. J. and Todorov, K. (2023). Etfs, illiquid assets, and fire sales. *Working Paper*.

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of financial Economics*, 99(1):1–10.

Xiao, H. (2022). The economics of etf redemptions. *Working Paper*.

Zurowska, E. A. (2022). Etf primary market structure and its efficiency. *Working Paper*.

Appendix

Table A.1.

Descriptive Statistics on Authorized Participants. This table presents the measures of ETF primary market concentration. Number of registered APs represents the number of all APs listed by ETF issuer in N-CEN filing. Each individual AP entry made by an ETF is counted as a separate AP, even if it falls under the same ultimate holding company. Number of active APs represents the number of APs who created or redeemed shares at least once during the reporting period. HHI concentration index was computed for each ETF's primary market as a sum of squares of ratios of the individual AP's creation (redemption or primary market transactions volume) transaction relative to the total level of creations (redemptions or primary market transactions volume) for that particular ETF. The observation unit is N-CEN filing. Data span N-CEN filings filed July 2018 — June 2024. Reports that span the period of less than 12 months are excluded. In cases where the number of active APs is zero, the HHI index is set to 1.

	count	mean	sd	min	p10	p25	p50	p75	p90	max
Number of registered APs	1,106	35	12	1	22	28	38	43	45	50
Number of active APs	1,106	7	5	1	2	4	6	9	13	23
HHI for creations	1,106	0.41	0.25	0	0.15	0.22	0.34	0.53	0.80	1
HHI for redemptions	1,106	0.42	0.29	0	0.11	0.22	0.35	0.55	1	1
HHI for total	1,106	0.39	0.09	0.16	0.22	0.32	0.50	0.77	1	1

Table A.2.
List of Authorized Participants. This table presents the financial institution that appear in the N-CEN filings as authorized participants for filings submitted between July 2018 - June 2024. I filter for filings with full report length, filings from ETFs, and without missing data for reliance on Rule 6c-11. The cleaned database contained over 197 different AP names with 67 unique LEI identifiers. After grouping APs at the parent organization, I was able to identify 51 unique AP names listed in the table below.

Authorized Participants	
ABN Amro	Jane Street
Banca IMI	Jefferies
Bank of America	Knight Capital
Barclays	Macquarie Capital
BMO Capital Markets	Mirae Asset Securities
BNP Paribas	Mizuho
BNY Mellon	Morgan Stanley
Cantor Fitzgerald	MUFG
CIBC	National Bank of Canada
Citadel Securities	National Financial Services
CitiGroup	NATIXIS
Commerzbank	Natwest
Cowen and Company	Nomura
Credit Suisse	Pershing
Daiwa	RBC
Deutsche Bank	Scotiabank
EWT	SG Americas
Goldman Sachs	State Street
HSBC Holdings	Stifel Nicolaus
Hudson River Trading	TD
Industrial and Commercial Bank of China	Timer Hill
Interactive Brokers	U.S. Bancorp Investments
ITAU	UBS
ITG	Virtu Americas
J.P. Morgan	Wedbush
Wells Fargo	

Table A.3.
Unwinsorized Regression Results. In this table, I estimate the impact of the ETF Rule on ETF liquidity and tracking errors, using unwinsorized data. For the difference-in-differences regressions, I estimate the following:

$$Y_{it} = \alpha + \beta_0 Treated_i + \beta_1 Post_t + \beta_2 Treated_i \times Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t, $Post_t$ is a dummy variable equal to one after 22 December 2020, and $Treated_i$ is equal to one if fund i was listed after 2006. For the matching regressions, I estimate the following:

$$Y_{it} - Y_{it}^{match} = \alpha_i + \delta_1 Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t and $Y_{matchit}$ is the value of this measure for the match of fund i in month t in the group of control stocks. I use three different methods to choose a match for fund i in month t: quartile matching, percentage difference matching, and propensity score matching. Estimates of the effect of the reform (δ_1) with each matching procedure are reported in columns 2, 3, and 4, respectively. In panel A.1, $Y_{it} = Se_{it}$; in panel A.2, $Y_{it} = Sr_{it}$; and in panel B, $Y_{it} = TE_{it}$. The sample period starts in December 2017 and ends in December 2023. In brackets, I report double-clustered standard errors allowing for correlation in residuals over time and across funds.

Variable	**DD**	**Quartile Matching**	**Percentage Difference Matching**	**Propensity Score Matching**
Panel A.1: Dependent variable: Effective Spread (bps)				
Treated × Post (β_2)	-66.85***			
	(20.79)			
Treated	302.89***			
	(60.47)			
Post (δ_1)	-154.25***	-248.54***	-157.35***	-315.34***
	(8.15)	(39.53)	(33.01)	(50.24)
Constant	461.46***	226.39***	228.83***	309.94***
	(26.31)	(39.53)	(33.01)	(115.89)
Observations	12,039	777	336	777
$Adj.R^2$	0.03	0.05	0.06	0.05
Panel A.2: Dependent variable: Realized Spread (bps)				
Treated × Post (β_2)	-1048.22*			
	(602.24)			
Treated	909.82			
	(619.26)			
Post (δ_1)	160.57	-209.47***	-131.46***	-281.07***
	(205.99)	(35.70)	(28.93)	(45.06)
Constant	78.34	192.66**	200.78***	264.79***
	(236.37)	(88.10)	(28.93)	(100.26)
Observations	12,039	777	336	777
$Adj.R^2$	0.00	0.04	0.05	0.05
Panel B: Dependent variable: Tracking error				
Treated × Post (β_2)	0.5699*			
	(0.3079)			
Treated	2.3849			
	(2.0291)			
Post (δ_1)	-0.0469	0.26*	0.13	0.22
	(0.0791)	(0.15)	(0.11)	(0.16)
Constant	1.56	1.20	-0.10	1.07
	(1.90)	(0.98)	(0.20)	(0.94)
Observations	7,474	777	336	777
$Adj.R^2$	0.00	0.00	0.00	0.00

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Table A.4.
Regression Results with Announcement and Pre-transition Dates as Treatment Dates. In this table, I estimate the impact of the ETF Rule on ETF liquidity and tracking errors, using the announcement date and the pre-transition date as alternative treatment dates. For the difference-in-differences regressions, I estimate the following:

$$Y_{it} = \alpha + \beta_0 Treated_i + \beta_1 Post_t + \beta_2 Treated_i \times Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t, $Post_t$ is a dummy variable equal to one after 26 September 2018 in regressions (1), (3), (5) and (7), and equal to one after 23 December 2019 in regressions (2), (4), (6) and (8), and $Treated_i$ is equal to one if fund i was listed after 2006. For the matching regressions, I estimate the following:

$$Y_{it} - Y_{it}^{match} = \alpha_i + \delta_1 Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t and $Y_{matchit}$ is the value of this measure for the match of fund i in month t in the group of control stocks. I use three different methods to choose a match for fund i in month t: quartile matching, percentage difference matching, and propensity score matching. Estimates of the effect of the reform (δ_1) with each matching procedure are reported in columns 2, 3, and 4, respectively. In panel A.1, $Y_{it} = Se_{it}$; in panel A.2, $Y_{it} = Sr_{it}$; and in panel B, $Y_{it} = TE_{it}$. The sample period starts in December 2017 and ends in December 2023. In brackets, I report double-clustered standard errors allowing for correlation in residuals over time and across funds.

Variable	**DD**		**Quartile Matching**		**Percentage Difference Matching**		**Propensity Score Matching**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A.1: Dependent variable: Effective Spread (bps)								
Treated × Post (β_2)	7.23	10.60						
	(14.16)	(14.15)						
Treated	220.02***	187.43***						
	(50.70)	(44.23)						
Post (δ_1)	13.45	19.84***	-3.11	-5.30	-77.6**	-83.12***	58.41**	37.85
	(7.33)	(7.40)	(21.98)	(22.03)	(31.21)	(31.30)	(23.04)	(24.91)
Constant	381.25***	371.35***	117.83	110.06	244.65**	239.07**	194.78**	204.33**
	(26.74)	(24.72)	(80.13)	(78.68)	(121.79)	(114.54)	(84.38)	(93.62)
Observations	10,712	10,940	766	768	336	336	776	768
$Adj.R^2$	0.01	0.01	0.00	0.00	0.02	0.02	0.01	0.00
Panel A.2: Dependent variable: Realized Spread (bps)								
Treated × Post (β_2)	-72.91**	-75.18**						
	(35.99)	(33.85)						
Treated	300.67***	306.38***						
	(67.70)	(63.49)						
Post (δ_1)	-30.54	-14.28	-87.22***	-73.26***	-119.16***	-115.37***	-60.22*	-38.98
	(14.14)	(12.55)	(29.31)	(25.58)	(44.81)	(41.13)	(31.50)	(28.15)
Constant	337.48***	310.52***	133.65*	128.09*	264.50*	251.66*	268.21***	231.73**
	(29.65)	(24.79)	(81.23)	(76.79)	(143.73)	(132.30)	(85.29)	(90.40)
Observations	10,712	10,940	766	768	366	336	766	768
$Adj.R^2$	0.01	0.01	0.01	0.01	0.02	0.02	0.00	0.00
Panel B: Dependent variable: Tracking error (%)								
Treated × Post (β_2)	1.11***	1.05***						
	(0.14)	(0.14)						
Treated	2.36***	2.18***						
	(0.38)	(0.38)						
Post (δ_1)	2.39***	0.22***	0.34***	0.35***	-0.03	-0.01	0.32***	0.43***
	(0.90)	(0.82)	(0.06)	(0.05)	(0.07)	(0.07)	(0.01)	(0.07)
Constant	0.56*	0.64**	0.68	0.71	-0.05	-0.04	0.46	0.45
	(0.30)	(0.27)	(0.60)	(0.63)	(0.12)	(0.13)	(0.68)	(0.68)
Observations	6,057	6,302	766	768	336	336	766	768
$Adj.R^2$	0.04	0.03	0.04	0.06	-0.00	-0.00	0.02	0.05

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.5.
Regression Results without Transition Date. In this table, I estimate the impact of the ETF Rule on ETF liquidity and tracking errors, omitting data inside the transition period. For the difference-in-differences regressions, I estimate the following:

$$Y_{it} = \alpha + \beta_0 Treated_i + \beta_1 Post_t + \beta_2 Treated_i \times Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t, $Post_t$ is a dummy variable equal to one after 22 December 2020, and $Treated_i$ is equal to one if fund i was listed after 2006. For the matching regressions, I estimate the following:

$$Y_{it} - Y_{it}^{match} = \alpha_i + \delta_1 Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t and $Y_{matchit}$ is the value of this measure for the match of fund i in month t in the group of control stocks. I use three different methods to choose a match for fund i in month t: quartile matching, percentage difference matching, and propensity score matching. Estimates of the effect of the reform (δ_1) with each matching procedure are reported in columns 2, 3, and 4, respectively. In panel A.1, $Y_{it} = Se_{it}$; in panel A.2, $Y_{it} = Sr_{it}$; and in panel B, $Y_{it} = TE_{it}$. The sample period starts in December 2017 and ends in December 2023. In brackets, I report double-clustered standard errors allowing for correlation in residuals over time and across funds.

Variable	**DD**	**Quartile Matching**	**Percentage Difference Matching**	**Propensity Score Matching**
Panel A.1: Dependent variable: Effective Spread (bps)				
Treated × Post (β_2)	-20.20*			
	(10.41)			
Treated	216.62**			
	(40.87)			
Post (δ_1)	-91.65**	-156.76***	-151.70***	-149.39***
	(4.71)	(15.50)	(19.71)	(16.94)
Constant	406.06***	159.96**	223.18**	288.45***
	(23.24)	(66.42)	(97.47)	(55.79)
Observations	11,714	777	336	777
Adj. R^2	0.04	0.12	0.15	0.09
Panel A.2: Dependent variable: Realized Spread (bps)				
Treated × Post (β_2)	-126.65***			
	(37.03)			
Treated	396.61***			
	(61.97)			
Post (δ_1)	-98.18***	-126.95***	-155.53***	-119.27***
	(11.65)	(12.46)	(25.35)	(14.44)
Constant	341.47***	151.84**	224.79**	288.64***
	(22.98)	(71.73)	(104.07)	(62.07)
Observations	11,714	777	336	777
$Adj.R^2$	0.01	0.12	0.10	0.08
Panel B: Dependent variable: Tracking error (%)				
Treated × Post (β_2)	0.97***			
	(0.09)			
Treated	1.56***			
	(0.58)			
Post (δ_1)	0.16***	0.48***	-0.08	0.53***
	(0.02)	(0.05)	(0.01)	(0.06)
Constant	0.75**	0.09	-0.00	0.80
	(0.34)	(0.77)	(0.11)	(0.77)
Observations	7,123	781	336	781
$Adj.R^2$	0.07	0.10	-0.00	0.10

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Table A.6.
Regression Results with Placebo Treatment. In this table, I estimate the impact of the ETF Rule on ETF liquidity and tracking errors, based on a placebo treatment on 01 June 2018. For the difference-in-differences regressions, I estimate the following:

$$Y_{it} = \alpha + \beta_0 Treated_i + \beta_1 Post_t + \beta_2 Treated_i \times Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t, $Post_t$ is a dummy variable equal to one after 01 June 2018, and $Treated_i$ is equal to one if fund i was listed after 2006. For the matching regressions, I estimate the following:

$$Y_{it} - Y_{it}^{match} = \alpha_i + \delta_1 Post_t + \epsilon_{it},$$

where Y_{it} is one of the measures of ETF dynamics for fund i in month t and $Y_{matchit}$ is the value of this measure for the match of fund i in month t in the group of control stocks. I use three different methods to choose a match for fund i in month t: quartile matching, percentage difference matching, and propensity score matching. Estimates of the effect of the reform (δ_1) with each matching procedure are reported in columns 2, 3, and 4, respectively. In panel A.1, $Y_{it} = Se_{it}$; in panel A.2, $Y_{it} = Sr_{it}$; and in panel B, $Y_{it} = TE_{it}$. The sample period starts in December 2017 and ends in December 2023. In brackets, I report double-clustered standard errors allowing for correlation in residuals over time and across funds.

Variable	**DD**	**Quartile Matching**	**Percentage Difference Matching**	**Propensity Score Matching**
Panel A.1: Dependent variable: Effective Spread				
Treated × Post (β_2)	25.52			
	(19.80)			
Treated	305.67***			
	(81-67)			
Post (δ_1)	-90.88***	-59.26**	4.03	-83.56***
	(10.77)	(29.57)	(31.02)	(28.15)
Constant	564.47***	177.21**	249.90*	292.05***
	(47.90)	(76.70)	(129.00)	(105.68)
Observations	9,781	670	284	670
$Adj.R^2$	0.02	0.00	0.01	0.01
Panel A.2: Dependent variable: Realized Spread				
Treated × Post (β_2)	-98.15			
	(70.63)			
Treated	549.78***			
	(130.22)			
Post (δ_1)	-134.49***	-48.99	-19.82	-92.91*
	(29.85)	(47.84)	(63.96)	(48.90)
Constant	493.96***	152.50*	285.11***	352.65***
	(60.99)	(78.10)	(103.26)	(104.05)
Observations	9,781	670	284	670
$Adj.R^2$	0.02	0.00	0.01	0.00
Panel B: Dependent variable: Tracking error				
Treated × Post (β_2)	0.41**			
	(0.16)			
Treated	3.67***			
	(1.16)			
Post (δ_1)	0.15***	0.08	-0.02	0.32***
	(0.04)	(0.11)	(0.06)	(0.09)
Constant	0.71*	0.86	-0.09	0.40
	(0.37)	(0.66)	(0.12)	(0.07)
Observations	4,883	670	284	670
$Adj.R^2$	0.01	-0.00	-0.00	0.08

Standard errors in parentheses
* p < 0.1, ** p < 0.05, *** p < 0.01

Imprint and acknowlegements

I would like to express my gratitude to the participants of the ECB Early Ideas Seminar for their insightful feedback and valuable suggestions.
Parts of this paper were written during my research visit at the European Central Bank whose hospitality is gratefully acknowledged. All errors are my own.

Berke Körükmez
University of St Gallen, St Gallen, Switzerland; email: **berke.koeruekmez@unisg.ch**

Postal address	60640 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Website	**www.esrb.europa.eu**

For specific terminology please refer to the **ESRB glossary** (available in English only).

PDF ISBN 978-92-9472-393-2, ISSN 2467-0677, doi:10.2849/0666303, DT-01-25-001-EN-N