CFR Working Paper NO. 21-03

Do ETFs Increase Liquidity?

M. Saglam • T. Tuzun •
R. Wermers

centre for financial research
cologne

Do ETFs Increase Liquidity?*

Mehmet Sağlam
University of Cincinnati
email: mehmet.saglam@uc.edu

Tugkan Tuzun
Federal Reserve Board
email: tugkan.tuzun@frb.gov

Russ Wermers
University of Maryland
email: wermers@umd.edu

September 2020

Abstract

This paper investigates the impact of exchange-traded funds (ETFs) on the liquidity of their underlying stockholdings. Using a difference-in-differences methodology for large changes in the index weights of stocks in the S&P 500 and NASDAQ 100 indexes, we find that increases in ETF ownership are associated with increases in commonly used measures of liquidity. Stocks with high ETF ownership have higher price resilience and lower adverse selection costs. However, ETFs are linked to higher liquidation costs during the 2011 U.S. debt-ceiling crisis, suggesting that stocks with high ETF ownership may experience impaired liquidity during major market stress events.

*The views expressed herein are those of the authors and do not necessarily reflect the views of the Board of Governors, the Federal Reserve System or their respective staff. We thank Turan Bali, Jonathan Brogaard, Kent Daniel, Darrell Duffie, Larry Glosten, Brian Hatch, Benjamin Holcblat (Luxembourg discussant), Andriy Shkilko (NFA discussant), Michael Pagano, Anna Scherbina (EFA Discussant), Elvira Sojli (SAFE discussant), Valeri Sokolovski, Kumar Venkataraman and seminar and conference participants at University of Oregon, Baruch College, the 2nd SAFE Market Microstructure conference, 8th Luxembourg Asset Management conference, NFA 2018 Annual Meeting, and EFA 2019 Annual Meeting for helpful comments and suggestions. We are also grateful to Philip Murphy of the S&P Global for explaining in great detail the methodologies of the S&P 500 additions and deletions.

1. Introduction

Exchange-traded funds (ETFs) have become a major investment vehicle over the past 15 years. To wit, net assets of ETFs traded in the United States have grown by roughly 25 percent per year, on average, during this period—reaching over $4 trillion by the end of July 2019.[1] The remarkable growth of ETFs warrants a deeper examination of their impact on financial markets. Of great importance to investors in financial markets is how ETFs might affect the liquidity of the stocks that ETFs hold—i.e., do ETFs improve or impair market quality? In this paper, we address this question by empirically studying the impact of ETF ownership on the liquidity of their underlying stocks.

There are competing arguments about how the introduction of ETFs might affect the liquidity of their underlying stocks, generally based on theories developed to examine the introduction of stock futures contracts to cash index markets. On one hand, ETFs might divert uninformed traders from holding and trading individual stocks—if the intention of such investors is to place a (low-cost) diversified "bet" on a sector or strategy, rather than on the fortunes of an individual firm. Hence, ETFs might increase the concentration of *informed* traders in individual stocks and, correspondingly, decrease their liquidity (Jegadeesh and Subrahmanyam, 1993). On the other hand, ETFs may attract additional uninformed traders into the market—who find a low-cost exposure to be attractive for hedging or for speculating on a particular sector or style—resulting in an increase in the investor base for all stocks owned by an ETF. Hence, the liquidity of stocks might benefit from the arrival of these new uninformed traders (Holden, 1995). In fact, by theoretically modeling arbitrageurs with execution risk in the underlying securities, Malamud (2016) finds that introducing new ETFs may cause deterioration or improvement in liquidity, depending on the trading needs of the ETF investors.

There are also competing arguments about how arbitrage activity between ETFs and stocks might affect the liquidity of the underlying stocks. Cross-market arbitrageurs are often modeled as traders that have better access to investment opportunities than others. Gromb and Vayanos (2002) show that arbitrage activity is beneficial for liquidity because arbitrageurs can provide liquidity in segmented markets. In contrast, Foucault et al. (2017) suggest that arbitrage activity may

[1] Hill et al. (2015) and Lettau and Madhavan (2018) offer several reasons for the growth of ETFs based on the cost of ownership, ease of access, transparency, liquidity, and tax efficiency.

not always be beneficial for liquidity. Under this framework, low-latency arbitrageurs could take liquidity from underlying stocks when they notice information arrival through ETF market activity and quickly trade the underlying stocks. In aggregate, these various (and sometimes competing) arguments create an unclear picture of the net effect of the presence of ETFs on the liquidity of their underlying stockholdings. In this paper, we contribute to this literature through a carefully designed empirical analysis.

Quantifying the causal impact of the presence of ETFs on the liquidity of their underlying stocks is clearly empirically challenging, as they are jointly determined in equilibrium. In order to alleviate endogeneity concerns, we use two separate identification strategies. In the first one, we exploit the difference between two types of S&P 500 index additions that affect ETF ownership significantly but in opposite directions. Remarkably, some S&P 500 additions cause ETF ownership of such stocks to increase, whereas other additions cause ETF ownership of such stocks to decrease. In the second identification, we examine the days around which stocks undergo a rare substantial change in their weights in the NASDAQ 100 index. Our *daily* ETF holdings data allow us to verify the validity of these identification strategies and to design a difference-in-differences (DiD) methodology to quantify the impact of ETF ownership on stock liquidity.

The addition of stocks to the S&P 500 affects the ETF ownership of stocks, because a large majority of ETFs passively track stock indexes. On the day of its inclusion in the S&P 500, a stock typically experiences an increase in its ETF ownership. Specifically, stocks that are not part of any S&P index at the time of addition experience an increase in their ETF ownership as ETFs tracking the S&P indices start to hold the newly added stock. However, stocks that are added to the S&P 500 but deleted from the S&P MidCap 400 index experience a decrease in their ETF ownership. This drop in ownership occurs because the weight of the security in the new index is significantly smaller than its weight in the old index. Although the amount of ETF assets (in dollars) tracking the S&P 500 index is larger than ETFs that track the S&P MidCap 400, this difference between ETF assets is not enough to offset the large decrease in the weight of a security moving from the S&P MidCap 400 index to the S&P 500 index. This previously unexploited divergence of ETF ownership in response to addition into the S&P 500 index provides a unique identification opportunity to study the impact of ETFs.[2]

[2]Clearly, examining only one type of S&P 500 additions would not be fruitful as additions could be correlated with

One might wonder whether this identification strategy can truly isolate the impact of ETFs, as many other indexed products are also affected by changes in the composition of the S&P indexes. To address this concern, using daily index weights of stocks and monthly assets of index funds, we estimate daily open-ended (non-ETF) index mutual fund ownership changes for the stocks included into the S&P 500, but excluded from the S&P MidCap 400 index. We verify that index mutual funds increase their holdings of these stocks. This increase in index fund ownership occurs because, as opposed to ETF assets, the dollar amount of non-ETF index fund assets tracking the S&P 500 index relative to the S&P MidCap 400 index is dramatically larger. Consequently, the index fund ownership increases despite the significant drop in the index weight of the security after the index switch. Hence, the divergence of ETF ownership in response to the inclusion into the S&P 500, where ownership of non-ETF index mutual funds increases for all newly included stocks, allows us to isolate the effect of ETF ownership.

To alleviate the concern that our identification strategy might capture some unobservable difference between the stocks added to the S&P 500 from the outside of the S&P managed indexes and the stocks added to the S&P 500 from the S&P MidCap 400 index, we run a placebo test on our identification method using an earlier time period. If we are capturing some unobservable characteristic, during the time period when the latter stocks also experience increases in their ETF ownership, we should get results similar to those from our original time period. However, we find that in the earlier sample period, the stocks added to the S&P 500 from the S&P MidCap 400 index (now with higher ETF ownership rather than the lower ETF ownership in the baseline time period) are not different from the stocks added to the S&P 500 externally with regards to their change in liquidity after the addition.

Further, we employ a completely separate identification strategy involving index weight changes in the NASDAQ 100 index in May 2011. Hence, this identification allows us to study the impact of ETFs without relying on changes in index membership list. NASDAQ 100 index is constructed by a proprietary algorithm based on the market-capitalization of each security, so that each stock in the index preserves its respective position relative to the actual unadjusted order (e.g., from largest

higher analyst or media coverage. Russell index reconstitutions have also been used to identify the impact of ETFs. S&P 500 additions might be better suited for studying the impact of ETFs because for some S&P 500 additions, the ETF ownership decreases while index fund ownership increases. However, for Russell index reconstitution, Ben-David et al. (2018) show that index mutual fund ownership and ETF ownership go in the same direction.

to the smallest). However, through time, the weights of the securities can diverge from the implied weights based on market capitalizations.[3] To address this issue, NASDAQ 100 index weights were adjusted on May 2, 2011 so that the underlying weights of the index constituents reflect a weight closer to the actual market capitalization of the security. This rebalancing substantially affected the ETF ownership of the stocks in the index as a very large ETF, the Powershares QQQ, tracks the NASDAQ 100 index. Index mutual fund ownership is little affected by this change, as the assets under management of index mutual funds tracking the NASDAQ 100 are much lower than that of ETFs. Indeed, 82 out of the 100 securities in the NASDAQ 100 index experienced a reduction in their ETF ownership. This variation around the rebalancing date allows us to study the impact of ETFs on the liquidity of the underlying stocks. Consistent with our results from the S&P 500 additions, we find an increase in the liquidity of the stocks whose ETF ownership increases.

In our main analyses, we use four measures of liquidity: i) effective spread (ES) ii) quoted spread (QS) iii) Amihud illiquidity measure, and iv) implementation shortfall (IS). The first three measures are available from publicly available data sets, such as the Trade and Quote (TAQ) database, whereas IS is computed using a commonly used data set of institutional trading provided by Abel Noser. Controlling for common determinants of liquidity, we find that stocks with lower ETF ownership after their inclusion to S&P 500 index experience lower liquidity compared to stocks with higher ETF ownership after their inclusion to the index. The effect of ETF ownership on liquidity is also economically significant. For example, our DiD regressions imply that on average, a 1% increase in ETF ownership translates into 4%, 6% and 17% reduction in *ES*, *QS*, and *IS*, respectively.

Next, we investigate various channels for liquidity improvement. We start with a channel that is specific to ETFs. If intraday arbitrageurs are transmitting the liquidity of ETFs to their underlying stocks, then high ETF ownership stocks should exhibit higher price resilience. On the other hand, if ETFs are making it easier for market makers to hedge their positions, then realized spreads should be lower for these stocks. Consistent with the first channel, we find that high ETF ownership is positively correlated with price resilience, market depth and negatively correlated with adverse selection costs. We do not find significant evidence for the lower hedging costs for market makers

[3]For example, at the time of this rebalancing, Apple Inc. (NASD:AAPL) had a market capitalization of roughly $300 billion, which was approximately twice that of Google Inc. (NASD:GOOG), but the weight of Apple in the index was five times that of Google.

as the realized spreads do not drop with higher ETF ownership. Overall, these results also suggest that features specific to ETFs (e.g. intraday arbitrage activity) rather than their lower frequency passive investment style, which can also shared by index mutual funds, are driving improvements in liquidity.

Finally, we examine the relation between ETF holdings and liquidity during two stressful periods: the 2007-2009 financial crisis and the 2011 U.S. debt ceiling crisis. During the financial crisis period, the S&P 500 index dropped by nearly 40% and during the debt ceiling crisis period, ETFs experienced substantial outflows while realizing significant losses. The trading imbalance in ETFs, as evidenced by outflows, may put additional selling pressure on the underlying stocks during such stressful periods and may drain liquidity in the underlying stocks. Consistent with this hypothesis, we find that sell trades in stocks with high ETF ownership incurred higher transaction costs during the debt ceiling crisis period. However, transaction costs of buying stocks with high ETF ownership were lower during the same period. Overall, this asymmetric effect of the ETFs during market stress has important implications for asset managers. For example, short-term momentum-style investors can face additional transaction costs on stocks with high ETF ownership while competing with arbitrageurs trading in the same side. Similarly, value investors can trade these stocks with lower trading costs, as they may be trading against the arbitrageurs.

Our results are intriguing when interpreted in relation to Ben-David et al. (2018), who make an important contribution to our understanding of ETFs by showing that ETFs increase the volatility of underlying stocks. However, the focus of our paper is the impact of ETFs on liquidity. While liquidity and volatility are highly correlated, volatility is not the only driver of liquidity. As Kyle (1985) shows, the trading volume of uninformed traders is positively related to liquidity. Ben-David et al. (2018) argue that the arrival of short-term traders increases volatility. In addition to volatility, these traders can also increase the trading volume. Hence, the net effect of short-horizon investors on liquidity can be positive. In fact, we find that Amihud illiquidity measure, which is a function of trading volume, decreases when ETF ownership increases. Hence, this result could be interpreted as the volume effect of ETFs on liquidity being stronger.

2. Related Literature

Our paper is mainly related to the empirical literature studying the broad effects of ETFs on their underlying stockholdings. The literature has predominantly focused on equity ETFs[4] and their impact on price discovery,[5] informational efficiency,[6] volatility,[7] and asset correlations.[8] In this section, we specifically review the literature that examines the relation between ETF ownership and stock liquidity.

Researchers have found competing evidence on the impact of ETFs on the liquidity of their underlying stockholdings. Hamm (2014) finds that ETFs lead to increases in Kyle's lambda and argues that ETFs increase the adverse selection in individual stocks. Similarly, Israeli et al. (2017) find that ETFs are related to higher bid-ask spreads and higher Amihud illiquidity measures. Overall, these studies present correlational analysis between the ETFs and the liquidity of the underlying stocks. In contrast, Coles et al. (2017) design a regression discontinuity framework to address potential endogeneity problems and conclude that index funds do not affect the liquidity of stocks. Chinco and Fos (2017) study the liquidity implications of ETF rebalancing cascades. They argue that stocks that are susceptible to ETF rebalancing cascades experience higher liquidity, consistent with market participants interpreting these cascades as noise trading. Chang et al. (2014) focus on the price effects of indexing by utilizing the discontinuity around the Russell 1000 cut-off. They find that stocks added to the Russell 2000 index have higher returns than stocks that just missed making it to the index.

Our study adds to this growing literature by investigating the *causal* impact of ETF holdings on commonly used liquidity measures over a long horizon with granular daily data. Previous studies used either *quarterly* or *annual* ETF holdings data, which often makes identification difficult. In

[4]There is also a growing literature studying ETFs in fixed-income and commodities markets. Holden and Jayoung (2019) and Ye (2017) find that corporate bond ETFs improve the liquidity of the bonds they hold. Dannhauser (2017) examine the pricing implications of the corporate bond ETFs as a form of new financial innovation. Brogaard et al. (2017) study the emergence of indexing in the commodities markets and highlight the negative effect of index investing on real activity.

[5]See e.g., Hasbrouck (2003) and Madhavan and Aleksander (2016).

[6]Glosten et al. (2016) find that ETFs increase the informational efficiency of stocks by helping stock prices incorporate the systematic component of company earnings news. On the other hand, Bhattacharya and O'Hara (2017) find that ETFs could distort the pricing efficiency of underlying securities in a theoretical model.

[7]ETFs can attract short-horizon uninformed traders and the non-fundamental price volatility of the underlying stocks can increase as shown by Ben-David et al. (2018).

[8]For example, Da and Shive (2012) and Israeli et al. (2017) find that ETFs increase the correlations of their underlying security returns.

contrast, our high frequency analysis with daily dataset can be useful for identification because it allows for a direct investigation of the days around two types of index inclusions with opposite effects on ETF ownership.

In addition to the literature on ETFs, we also contribute to the literature on the liquidity implications of S&P 500 index changes. Previous studies on S&P 500 index additions (e.g. Beneish and Whaley (1996); Hegde and McDermott (2003)) find that additions to the S&P 500 index lead to increases in the liquidity of those stocks without the benefit of further examining how additions change the ownership by ETFs and open-end index mutual funds. We contribute to this literature by providing evidence that not all stocks added to the S&P 500 experience improvements in liquidity. In particular, only the stocks added to the S&P 500 outside of the S&P MidCap 400 index experience significant improvements in liquidity.

3. Theoretical Framework

In this section, we review the relevant theoretical models that guide our empirical analysis on the impact of ETFs on transaction costs.

There are mainly three channels through which ETFs may improve the liquidity of individual stocks. First, arbitrageurs could transfer the liquidity of ETFs to underlying stocks. If ETFs are enlarging the investor base by attracting uninformed traders, they could bring more liquidity to the market (Holden, 1995).[9] Arbitrageurs simply transmit uninformed liquidity trades from ETFs to individual stocks. If a trader's orders move the price of a stock, it may lead to an arbitrage opportunity between the ETF and the stock. As an arbitrageur enters into an arbitrage position, the arbitraguer's trades in the stock will be in the opposite direction of the trader. If the trader buys the stock and generates a discount on the ETF, the arbitrageur will buy the ETF and sell the stock. This mechanism could lead to a liquidity transfer from the ETF to its underlying stocks, reducing these stocks' transaction costs.

Arbitrage is more likely to happen with higher liquidity. Even if the price of an ETF and the prices of its underlying stocks signal an arbitrage opportunity, arbitrageurs require a certain level of liquidity in the ETF and in these stocks to take an arbitrage position. If a trader submits passive

[9] Relatedly, Chen et al. (2004) suggest that higher uninformed trading after an index addition, which can affect ETF ownership, is in part caused by increased investor awareness.

orders in a stock even without price improvements, then market liquidity improves for this stock. This improved liquidity may be just enough for arbitrageurs to take advantage of the arbitrage opportunity already prevailing in prices (Roll et al., 2007).

The second channel is related to holding costs of market makers. ETFs can be used as a hedging instrument by market makers and can help them reduce their holding costs. Risk-averse market makers demand compensation for holding risky inventory (Ho and Stoll, 1981; Grossman and Miller, 1988). If they can hedge their inventories easily, their holding costs will be lower, allowing them to provide liquidity at a lower cost. ETFs can serve as an additional instrument for market makers to hedge their individual stock holdings, lowering their liquidity provision costs.

While these two channels are very similar, the former conjectures that arbitrageurs transfer the liquidity benefits to the individual stocks wheras the latter suggests that ETF liquidity is transferred through market makers.

Third, stocks with high ETF ownership could have higher media or analyst coverage that can lead to higher informational efficiency in these stocks. Amid lower information asymmetry in these stocks, stock liquidity may increase (Glosten and Milgrom, 1985). If prices of stocks with high ETF ownership are more informationally efficient, then there is relatively less information available which is not incorporated into prices. This situation can make it less likely for traders to acquire private information and trade on it. On the other hand, if ETFs lower the transaction costs of stocks, acquiring information for these stocks could be more profitable.

There are two different theories predicting that underlying stock liquidity may deteriorate with the introduction of ETFs. First, motivated from the effect of futures trading, some early studies such as Jegadeesh and Subrahmanyam (1993) conjecture that basket securities such as ETFs attract uninformed traders from individual stocks and lead to a decrease in liquidity for these stocks. This idea suggests that stocks with high ETF ownership should have a high concentration of informed trading, leading to higher execution costs for traders.

Second, considering that arbitrageurs in ETFs are prominently high-frequency traders, market makers may shy away from liquidity provision in the underlying stocks if they run the risk of leaving stale quotes than can be sniped by the arbitrageurs. In this circumstance, arbitrage links would not be beneficial for liquidity in the underlying stock. For example, Foucault et al. (2017) argue that arbitrage can be harmful for liquidity when prices adjust asynchronously in response to the arrival of information. Under this framework, low-latency arbitrageurs could take liquidity from underlying stocks when they notice information arrival in ETFs and quickly trade underlying stocks.

4. Institutional details

4.1. Institutional details on ETFs

Similar to stocks, investors can buy and sell ETF shares from one another in various trading venues. ETFs hold publicly traded stocks, which allows for arbitrage opportunities. Deviations of the ETF's price from the prices of its underlying stocks can lead to arbitrage opportunities during the trading day. Arbitrageurs – who are generally broker-dealers, high-frequency traders, hedge funds and other institutional traders – can take offsetting positions in the ETF and its underlying stocks to take advantage of this arbitrage opportunity. To close their positions, these arbitrageurs can either wait for ETF and stock prices to converge or use the primary market, where authorized participants (APs) trade with ETF sponsors. While only APs, who are generally broker-dealers, can participate in the primary market, many APs also collect client orders for ETF creation and redemption activity. Hence, if an arbitrageur is not an AP (e.g. high-frequency traders), this arbitrageur could still submit ETF creation/redemption orders through an AP.

Throughout the day, APs accumulate creation and redemption orders from their clients, or they accumulate ETF shares and underlying securities through their own arbitrage activity. Later, APs trade with ETF agents to create or redeem ETF shares. These primary market transactions between APs and ETF agents occur once a day, typically at 4:00 p.m. through the National Securities Clearing Corporation (NSCC), which acts as a central counterparty.

In these primary market transactions, ETF shares are generally exchanged for pre-specified redemption or creation baskets of securities. These redemption and creation baskets are determined by ETF portfolio managers the day before. Soon after markets close at 4:00 p.m. (ET), ETF sponsors submit their creation and redemption baskets valid for the following trading day. The NSCC compiles the baskets of US ETFs into a large file and publishes it around 7:00 p.m. (ET). Markit receives this file daily from NSCC and complements it with the actual holdings from ETF sponsors at around 10:00 p.m. (ET). Markit then distributes these data with the following day's date stamp. Hence, ETF holdings are, as of the beginning of the trading day, reported in the Markit data. Furthermore, these holdings are reported for one creation unit size of ETF shares. We adjust these holdings accordingly and compute the stock holdings in ETF portfolios. We scale ETF stock holdings with daily shares outstanding of stocks to compute the daily ETF ownership at the stock level. From the Markit ETF data, we use only the U.S. domestic equity ETFs excluding Leveraged and Inverse ETFs.[10] The Markit ETF data set is from January 2011 to December 2018.[11]

4.2. Institutional details on S&P 500 additions

A large majority of ETFs passively follow stock indexes weighted by firms' market capitalizations. Because index mutual funds employ a similar passive strategy, we aim to provide an empirical design that can solely capture exogenous variation in ETF ownership. For this reason, simple applications of major stock index changes (e.g. Russell 1000 and Russell 2000 reconstitutions) cannot directly lead to a perfect empirical design because both index fund ownership and ETF ownership changes are generally affected similarly. Further, some of these changes may also have a dramatic impact on variables that may be correlated with stock liquidity. For example, S&P 500 additions can lead to higher media coverage that can ultimately increase stock liquidity. Considering these confounding issues, we propose an empirical design based on two different types of the S&P 500 additions. As expected, both types of additions lead to increases in index mutual fund ownership, but interestingly (and important to our identification approach), they affect ETF ownership in opposite directions.

We identify two different types of S&P 500 index inclusions. To the best of our knowledge, this is the first paper that documents this institutional detail. In the first case, which we refer to as a "Type 1" addition, the stock has not previously been in any S&P 1500 Composite index, as it does not satisfy the eligibility criteria for S&P 1500 Composite index inclusion.[12] In the second case,

[10] We cross check the quarter-end values in the *daily* Markit ETF data against the *quarterly* CRSP Mutual Fund database, which also has ETF coverage. The correlation of our ETF ownership variable, ratio of shares held by ETFs, calculated from these two datasets is 0.93 for our sample of stocks suggesting high overlap between the two datasets at the quarter-ends.

[11] Although Markit ETF dataset starts earlier than 2011, the coverage of this data set prior to 2011 is not sufficient to reliably use it.

[12] The current S&P U.S. Indices Methodology (2018) states that the issuing company must be based in the United States, have an investable weight factor of at least 50%, be trading on an exchange for at least 12 months, have a positive sum for its most recent four consecutive quarters' GAAP earnings, and have a ratio of annual dollar value traded to market capitalization of at least 1. The S&P also mentions a set of rule exceptions. If its Index Committee decides that an inclusion will enhance the representativeness of the index as a market benchmark, the stock may still be added despite failing the eligibility criteria.

which we refer to as a "Type 2" addition, a stock is already in an S&P managed index, the S&P MidCap 400 index, and it gets removed from this index while being added to the S&P 500 index. First, using our daily ETF holding data that start from 2011, we examine how ETF ownership varies with Type 1 and Type 2 additions.

Panel A of Table 1 provides a subsample of Type 1 stocks that are added to the S&P 500 index, including the announcement date, effective date of addition, and the ETF ownership ratios on the effective date (Post) and prior trading day (Pre). At the time of the announcement, these stocks do not belong to any other S&P managed component index such as the S&P MidCap 400 index. Hence, they enter the S&P index universe as a brand new addition. We consistently observe that there is substantial increase in ETF ownership when these stocks are added into the S&P 500 index. For this sample, the average nominal change in ETF ownership is 1.38 percentage points and the average percent change in ETF ownership is 167% when compared with existing ETF ownership. These statistics point to a sharp and significant increase in ETF ownership for this set of stocks. The left panel of Figure 1 illustrates this sharp increase for Accenture Inc. (NYSE:ACN) by plotting the ETF holdings of the stock around the effective date of addition. Because the *ETF Holding* variable is as of the beginning of the day, the plot suggests that ETFs are acquiring the stocks added to the S&P 500 index right before the effective day of the addition, likely during the trading hours of the previous day.

Panel B of Table 1 provides a subsample of 10 stocks that are added to the S&P 500 index and excluded from the S&P MidCap 400 index along with the announcement date and effective date of addition. ETF ownership ratios on the effective date of addition (Post) and the prior trading day (Pre) are also reported. All of these stocks experience a substantial decline in their ETF ownership around the effective date. On average, their nominal ETF ownership goes down 1.19 percentage points, or roughly 23%, on average, compared with their pre-event ETF ownership values. The right panel of Figure 1 illustrates the sharp decline for Chipotle Inc. (NYSE:CMG) by plotting the ETF holdings of the stock around the effective date of the S&P 500 addition and the S&P MidCap 400 deletion.

One may find the sharp decrease in ETF ownership for Type 2 stocks surprising. However, the rationale for this decrease is quite simple. As the S&P indexes are weighted by market capitalization, to-be-added stocks in the S&P MidCap 400 index have relatively higher weight in the index due to their relatively larger size. When they are added to the S&P 500 index, their weights in the index becomes substantially smaller as they are now relatively small companies compared to existing large-cap stocks. However, to compute the exact change in ETF ownership, we also need to consider the dollar amounts in ETFs tracking both indexes. The ETF ownership can actually decrease if ETF assets tracking the S&P MidCap 400 index are relatively large. Let ΔETF Holding$_i$ be the change in the ratio of ETF ownership of a Type 2 stock i as a result of the inclusion in the S&P 500 index, $w_{i,400}$ and $w_{i,500}$ be the weights of the stock in the S&P MidCap 400 and S&P 500, respectively. If E_{400} and E_{500} denote the dollar amount of ETF assets, and MV_i is the market capitalization of the stock at the time of the addition, then

$$\Delta \text{ETF Holding}_i = \frac{w_{i,500}E_{500} - w_{i,400}E_{400}}{MV_i}. \tag{1}$$

Thus, ΔETF Holding$_i$ will be less than zero if $\frac{w_{i,400}}{w_{i,500}} > \frac{E_{500}}{E_{400}}$. Using our daily ETF holding data that covers 2011-2018, we verify that this condition holds for all Type 2 additions (excluding merger-driven additions). To extend our sample as much as possible, we also examine this relation since the introduction of the first ETF in 1993.

Table 2 provides a summary of the Type 2 additions over years. We estimate change in ETF ownership for each Type 2 addition using equation (1). We download the monthly weights of the S&P MidCap 400 index and the S&P 500 index from the S&P Global. From Morningstar Direct, we obtain monthly net assets of the ETFs and index mutual funds tracking the indexes based on the S&P MidCap 400 and S&P 500. In the final column, we also report the *exact* change in ETF ownership using the daily Markit ETF data, which is available after 2011.

In order to use Type 2 stocks as an identification strategy to capture only the effect of ETFs, we need to rule out the effect of index mutual funds. If the index fund ownership of Type 2 stocks also decreases, then this identification strategy not only capture the effect of ETFs, but also the effect of index mutual funds. Thus, in Table 2, we also report the average implied change in index ownership around each Type 2 addition by undertaking a similar computation as in equation (1) but now for index fund ownership. We again use the monthly weights of the indexes from S&P Global, and monthly net assets of the index mutual funds from Morningstar Direct.

Table 2 provides interesting findings with regard to the effect of Type 2 additions on ETF and index fund ownership. In 1990s, the ratio of the size of ETFs that track the S&P 500 index to the ones that track the S&P MidCap 400 index was large enough to compensate for the decrease in the index weight. Consequently, during this period, ETF ownership of Type 2 additions increased. However, as ETFs that track the S&P MidCap 400 indexes grew at a faster rate, the ratio of their size started shrinking, and ETF ownership changes of Type 2 additions became negative in early 2000s. ETF ownership changes from the daily Markit ETF data also suggest the ETF ownership decreases for these stocks in the later period. On the other hand, the relative size of index funds tracking the S&P 500 indexes relative to the ones that track the S&P MidCap 400 indexes have always been more than enough to make up for the decrease in the index weights of Type 2 stocks. Hence, the imputed daily changes in index mutual fund ownership around a Type 2 addition are all positive, suggesting that index mutual funds increase their holdings of Type 2 stocks in both 1990s and 2000s. These increases in index fund ownership for Type 2 stocks build further confidence for our identification strategy, which aimed at capture only the impact of ETFs. While we only report the average effect across each year in Table 2, ETF ownership decreases for each individual Type 2 addition (excluding merger driven additions) starting in 2002.

Overall, this analysis informs us that we can design a DiD framework to quantify the causal impact of ETFs on liquidity by taking advantage of the difference between two types of S&P 500 index additions.[13] Further, our historical analysis illustrates that the magnitude of the estimated change in ETF ownership is substantial starting from 2002 and consequently we will construct our sample beginning from this date. In the next section, we will provide detailed information about the construction of the final sample.

5. Data

We compile our final data set from several sources: (i) S&P 500 additions from Compustat, (ii) daily ETF holding data from Markit (iii) liquidity measures from WRDS Intraday Indicators (iv) institutional trading costs from Abel Noser dataset (v) stock-day level information from CRSP, (vi) trade and quote statistics from TAQ, (vii) the quarterly index mutual fund ownership of U.S

[13]One could also think about utilizing the set of S&P 500 deletions in the form of two types: deletion to the outside of the S&P 1500 universe and relegation to the S&P MidCap 400 index. However, as Chen et al. (2004) and Chan et al. (2013) document, approximately 75% of all deletions are driven by corporate events (e.g. mergers and bankruptcies), which makes using the set of deletions as an identification strategy difficult.

stocks from CRSP, (viii) the daily ETF price, daily ETF net asset value, and daily ETF flows from Morningstar Direct, (ix) net assets of ETFs and index funds from the monthly Morningstar direct data,[14] (x) The S&P 500 and S&P MidCap 400 monthly index weights from S&P Global. Our identification strategy is expected to work starting from 2002, thus our main data set will cover the periods between 2002 and 2018.[15] In the upcoming section, we provide detailed information about the construction of the final data set.

5.1. S&P 500 additions and control stocks

We obtain all S&P 500 additions from Compustat. This dataset includes the ticker of the stock and the effective date of the addition. We match the ticker symbols to CRSP PERMNOs by utilizing the matching code provided by WRDS. We obtain 359 additions consisting of 152 Type 1 and 207 Type 2 additions for which we have price data from CRSP around the effective date. As corporate events could have major impact on the underlying liquidity of the asset, e.g., mergers and share splits, we eliminate 56 additions (28 Type 1 and 28 Type 2) in which the outstanding number of shares changes abnormally before and after the effective date.[16] In summary, we obtain 303 S&P additions occurring between 2002 and 2018.

To employ our DiD framework, we match each of these added stocks to a control stock that is a constituent of the S&P 500 index. We use market capitalization as the matching criteria. Using market capitalization data from 30 days prior to the effective date, the matching stock is chosen so that it has the smallest market capitalization difference to the to-be-added stock.[17]

We use pre- and post-event periods of 30 business days and obtain 35,992 stock-day level observations between 2002 and 2018. For these stock-day observations, we first populate CRSP data, e.g., price, return, volume, and shares outstanding.

[14] Morningstar Direct starts covering the largest S&P 500 ETF (SPY) and the largest S&P MidCap 400 ETF (MDY) a few years after their inception dates. For those years, we use the net assets of SPY and MDY from the CRSP Mutual Fund database.

[15] We will also include one placebo test that includes data from 1993 to 1998.

[16] Formally, we compute the median number of shares outstanding in the 30 days before and after the effective date of the addition and if the absolute value of the normalized change is greater than 3%, we classify this addition as triggered by a corporate event. Our findings are robust to using different levels of threshold criteria.

[17] Announcement dates of the additions are typically five business days before the effective addition date so the matching is done using data before the announcement date.

5.2. Liquidity measures

We use the following liquidity measures for our main analysis: effective and quoted spreads, Amihud illiquidity measure, and implementation shortfall from institutional trades.

We obtain our spread data from WRDS intraday indicators which is available between September 2003 and December 2018. This data utilize millisecond time-stamped trade and quote information.

Our first measure is the effective spread (ES). This illiquidity proxy measures how transaction prices compare to mid-quote prices. For stock i, and trade occurring at t, the instantaneous effective spread is given by

$$\text{Effective Spread}_{i,t} = \frac{2|P_{i,t} - M_{i,t}|}{M_{i,t}},$$

where $P_{i,t}$ is the transaction price and $M_{i,t}$ is the mid-quote price of best bid price, $B_{i,t}$ and best offer price $O_{i,t}$. This measure is then aggregated to a stock-day level, by weighting each trade with its corresponding dollar size. This measure is referred to as *ESpread_Percent_DW* in the WRDS dataset.

Our second measure is quoted spread (QS). This proxy is derived from the distance between the best bid and offer price. For stock i, and time at t, the instantaneous quoted spread is given by

$$\text{Quoted Spread}_{i,t} = \frac{O_{i,t} - B_{i,t}}{M_{i,t}}.$$

This measure is then aggregated to a stock-day level, by weighting each spread with its corresponding time interval. This measure is referred to as *QSpread_Percent_tw* in the WRDS dataset. These two spread measures can specifically be useful to quantify small-sized trades originating from *retail* traders. We also utilize the two other measures–Amihud and implementation shortfall measures– than can be helpful to quantify the liquidity available for large-order trades originating from *institutional* investors.

Amihud illiquidity measure is a low-frequency measure that can be directly computed from daily CRSP data. We compute the Amihud illiquidity measure as the ratio of the absolute value of the stock's return to its daily share volume. Formally, for stock i on day t,

$$\text{Amihud}_{i,t} = \frac{|r_{i,t}|}{V_{i,t}}.$$

We populate this measure in our dataset for the complete period of the S&P 500 additions–between 2002 and 2018.

Our fourth and the last liquidity measure is the implementation shortfall (IS). Using Abel Noser institutional trading data for the periods between January 2002 and September 2014, we compute average institutional trading cost measure at the stock-day level. There are several academic studies that use Abel Noser's data for institutional trading cost measure (see e.g., Anand et al., 2011, 2013). Hu et al. (2017) report that the coverage of the Abel Noser dataset is relatively large, accounting for more than 10% of CRSP dollar trading volume. This dataset provides information on tickets sent by an institution to a broker where each ticket typically results in more than one execution. Using the value-weighted average cost of each ticker, we provide an institutional trading cost measure at the stock-day level, which can be a gauge for liquidity conditions large trades are facing. Spread-based measures usually account for the cost of small trades, whereas such an institutional trading cost proxy could measure the cost of large order trades.

Following Anand et al. (2011), we measure the cost of each ticket k by computing its implementation shortfall (IS) as follows:

$$ISTicket_k = D_k \frac{P_k^{\text{avg}} - P_{k,0}}{P_{k,0}}, \tag{2}$$

where D_k is the direction of the trade, 1 (buy) or -1 (sell), P_k^{avg} is the volume-weighted execution price of the ticket and $P_{k,0}$ is the price of the security (arrival price) when the ticket is received by the broker. We then calculate the volume-weighted average of the *ISTicket* of all trading tickets for stock i on day t and denote it as IS_{it}. Since the data availability is through September 2014, the data period for this measure will be between January 2002 and September 2014.

We winsorize all of our liquidity measures at 0.1% and 99.9% to reduce the effect of any outliers or erroneous values.

5.3. Summary Statistics

Table 3 presents the summary statistics of the main variables used in the paper. The average *effective spread* and *quoted spread* are 6.51 and 5.82 basis points, but they go from about 1.4 basis points to over 77 basis points for the *effective spread*. The average value of *implementation shortfall* is 12.50, slightly higher than the averages of the *effective spread* and *quoted spread*. However, *implementation shortfall* is more volatile, ranging from -511 basis points to 555 basis points. Our data primarily cover the universe of large stocks. Hence, market capitalization of stocks goes from $1 billion to $93.56 billion, with an average value of $11 billion. Finally, using the daily ETF holding data between 2011 and 2018, the average (median) ETF ownership of a stock is 4.55% (4.21%).

6. Evidence from the Difference between Type 1 and Type 2 Additions

In order to investigate the liquidity differential between Type 1 and Type 2 additions arising from their opposite impact on ETF ownership, we employ a pooled regression in a DiD design. As measures of market liquidity, we use *effective spread*, *quoted spread*, *Amihud* and *implementation shortfall*. In the following section, we first visually verify that the parallel trend assumption is satisfied for each liquidity measure.

6.1. Parallel Trends

In this section, we visually inspect the parallel trends assumption required for DiD analysis. For each liquidity measure, we compute the median estimate of the measure on trading days before the announcement date of the addition, which typically occurs within 5 prior trading days to the effective date. The minimum d we consider is -42, which is approximately 2 calendar months before the effective date of the addition. The left (right) panel of Figure 3 plots the median *effective spread* for Type 1 (Type 2) stocks and their matched controls. The left (right) panel of Figure 4 plots the median *quoted spread* for Type 1 (Type 2) stocks and their matched controls. The left (right) panel of Figure 5 plots the median *Amihud* for Type 1 (Type 2) stocks and their matched controls. Finally, the left (right) panel of Figure 6 plots the median *implementation shortfall* for Type 1 (Type 2) stocks and their matched controls. In all liquidity measures, we observe that both added and control stocks have roughly constant medians during the prior days of the event implying that the parallel trend assumption holds. We find that *implementation shortfall* is the noisiest measure and fluctuates around a wider range of values, which is consistent with the summary statistics on this measure.

6.2. Multivariate analysis with DiD

For each addition j, we have pre- and post-event data corresponding to the added and control stock. Added stock can be either Type 1 or Type 2 addition. In order to capture both the time-series and the cross-sectional properties of our data, we define the variable $X_{j,d,a}$ where j denotes the addition number, d denotes the distance (in trading days) to the effective addition date, and a can be "T1" (Type 1 stock), "T2" (Type 2 stock) or "C" (control stock). We let d be in the interval $\{-30, -29, \ldots, -2, 1, 2, \ldots, 28, 29\}$ excluding the effective addition day and the prior day to minimize the effect of index trackers' mechanical trading.[18] Formally, we run the following pooled regression with a set of control variables:

$$
\begin{aligned}
(3) \qquad \mathit{Illiquidity}_{j,d,a} = \xi &+ \beta \mathit{Post}_{j,d,a} \times \mathit{Type2Add}_{j,d,a} + \gamma \mathit{Post}_{j,d,a} \times \mathit{SP500Add}_{j,d,a} \\
&+ \delta_1 \mathit{Post}_{j,d,a} + \delta_2 \mathit{SP500Add}_{j,d,a} + \sum_i b_i \mathit{ControlVariable}_{i,j,d,a} + \varepsilon_{j,d,a}.
\end{aligned}
$$

where *Illiquidity* is one of our four illiquidity measure, $\mathit{Post}_{j,d,a}$ is a binary variable that takes a value of 1 if $d > 0$, $\mathit{Type2Add}_{j,d,a}$ is a binary variable that takes a value of 1 if $a = T2$, and $\mathit{SP500Add}_{j,d,a}$ is a binary variable that takes a value of 1 if $a \neq C$. We run this regression for all of the 303 additions during the period between 2002 and 2018. As control variables, we use turnover–defined as the ratio of the daily share volume to the number of shares outstanding–, inverse price and market capitalization, all of which tend to strongly co-vary with liquidity measures.[19] We also account for the time-series trend in illiquidity by including month fixed effects. We are mainly interested in the β coefficient as it captures the relative increase in the illiquidity of Type 2 stocks (when compared to Type 1 stocks) arising directly from the decrease in ETF ownership.

[18] Typically, all additions are announced within 5 business days to the effective date. In Section A of the Online Appendix, we show that our findings are unchanged if we were to exclude data from $d = -5$ to $d = -1$ and from $d = 1$ and $d = 5$.

[19] In Section A of the Online Appendix, we show that our findings are unchanged if we were to include volatility of the underlying stock as an additional control variable.

Table 4 reports the results of this regression. Our coefficient of interest is the loading on the interaction variable, *Post* $\times$ *Type2Add*, which measures the average change in illiquidity between Type 2 and Type 1 additions. Recall that Type 2 additions result in a sharp decrease in ETF ownership in our sample period, whereas Type 1 additions lead to a sharp increase. We find that this coefficient, β, is positive and statistically significant for all of the illiquidity measures considered with or without month fixed effects. We observe that stocks with Type 2 additions realize an increase of 0.57 bp in their *effective spread*, 1.13 bp in their *quoted spread*, and 5.7 bps in their *implementation shortfall* when compared with Type 1 stocks. These are economically significant effects. For example, during the 2011-2012 period, average ETF ownership increase (decrease) is 1.38 (1.19) percentage point after a Type 1 (Type 2) addition. This result implies that the cost increases should be scaled by 2.57 $(1.38 + 1.19)$ to compute the approximate impact of 1% change in ETF ownership. This back-of-the-envelope calculation implies that for every 1 percentage point increase in ETF ownership, *effective spread* decreases by 4%, *quoted spread* decreases by 6%, and *implementation shortfall* decreases by 17% (relative to their corresponding mean values).

One may worry that instead of the effect of ETFs, our identification strategy might be capturing some unobservable difference between the Type 1 and Type 2 stocks. If we are capturing some unobservable characteristic, then we should get similar regression results when Type 2 stocks experience increases in their ETF ownership. In order to address this concern, we run the same sets of regressions for the March 1993 - December 1998 period during which Type 2 stocks experience increases in their ETF ownership. Abel Noser and daily TAQ data is not available during this period, hence, we compute our spread measures using the monthly TAQ database. Table 5 reports the results of these regressions. The coefficients on the *Post* $\times$ *Type2Add* are all statistically insignificant, suggesting that during this period, there are no distinction between the changes in liquidity between Type 1 and Type 2 stocks . The signs of the coefficients are also negative in five out of the six cases. We still observe that the coefficients on *Post* $\times$ *SP500Add* are significant in four of six cases implying the liquidity improvement with Type 1 additions. The evidence that in the 1990s, Type 2 stocks did not experience a decline in illiquidity relative to Type 1 stocks supports our identification strategy, which is aimed to capture the effect of ETFs.

7. Evidence from NASDAQ 100 Rebalancing

One may worry that the liquidity benefits we find would be purely a result of the S&P 500 index additions, rather than an increase ETF ownership. In order to address this, we also analyze the NASDAQ 100 index rebalancing on April 5, 2011, when index weights of member stocks changed while index membership did not change.

NASDAQ OMX manages NASDAQ 100 index using a modified market-capitalization weighted indexing methodology so that each stock in the index preserves its respective position in the index relative to the actual unadjusted order (e.g., from largest to smallest). However, through time, the weights of the securities diverge from the implied weights based on market capitalizations, and a security may end up having higher weight in the index than a security with smaller market capitalization.[20] To assuage this issue, on April 5, 2011, NASDAQ OMX announced that the NASDAQ 100 index weights would be adjusted on May 2, 2011 so that the underlying weights of the index constituents reflect a weight closer to the actual market capitalization of the security.[21] This is considered to be a "special" rebalancing, as the main methodology to calculate weights is not changing and the universe of the stocks remains the same.

The rebalancing affected each security in the NASDAQ 100 index. 82 out of the 100 securities experienced a reduction in their index weight. On average, the weights of the stocks changed by 0.38 percentage points. The highest absolute weight change was in Apple Inc. (NASD:AAPL), whose weight dropped from 20.49% to 12.33%. Table 6 reports the top 25 largest weight changes in the NASDAQ 100 index due to the special rebalancing. These changes have substantially affected the ETF ownership of the stocks in the index, as a very large ETF, the Powershares QQQ, tracks the NASDAQ 100 index. The average absolute change in ETF ownership was 0.22% which is substantial given the persistence in daily ETF ownership. The highest absolute weight change occurs in Intel Corp. (NASD:INTC) whose holdings increase from 2.85% to 3.51%. Compared to the S&P 500 index additions, the absolute change in ETF ownership is slightly smaller but the main advantage of this event is its scope, which affects many securities at the same time. Figure 2

[20] Part of the divergence stems from a rebalancing that occurred during the initial set-up of the ETF tracking the index in 1998. If the index were constructed truly based on the implied market capitalization weights, it would not be diversified enough to meet a rule set by IRS.

[21] The official press release can be found at `http://ir.nasdaq.com/news-releases/news-release-details/nasdaq-100-index-undergo-special-rebalance`.

illustrates the sharp decline for Starbucks Corp. (NASD:SBUX) by plotting the ETF holdings of the stock around the rebalancing date. We again observe that the sharp change in ETF ownership occurs on the effective date of the rebalancing.

More importantly, this rebalancing had little effect on index fund ownership as there are only small-sized funds that use the NASDAQ 100 as a performance benchmark. From the monthly Morningstar data, we compute the aggregate net assets of ETFs and index mutual funds that track the NASDAQ 100. As of April 31, 2011, index mutual funds assets tracking the NASDAQ 100 have \$1.34 billion under management, whereas ETFs have \$27.32 billion under management, with almost all belonging to the Powershares QQQ ETF. That is to say, the ratio between index fund ownership and ETF ownership of these 100 stocks is roughly 1 to 20. Thus, this rebalancing disproportionately affects ETF ownership and gives us another identification strategy for the causal impact of ETF ownership on liquidity.

In order to investigate the impact of the rebalancing, we again employ a pooled regression in a DiD design with symmetric windows for the pre- and post-event periods. As a set of control stocks, we include 100 other common stocks that are NASDAQ-listed. These stocks are the largest 100 stocks in terms of market capitalization but they are not part of the the NASDAQ 100 index.

Similar to earlier analysis, we define the variable $X_{d,a}$ where d denotes the distance (in trading days) to the effective rebalancing date, and a can be "Up" (weight increase), "Down" (weight decrease) or "C" (control stock). We again exclude the data from the previous day and the implementation day to minimize the impact of rebalancing motives, i.e., $d \in \{-30, -29, \ldots, -2, 1, 2, \ldots, 28, 29\}$. Formally, we run the following pooled regression with stock and day fixed effects and our standard set of control variables:

$$\begin{aligned} (4) \qquad \mathit{Illiquidity}_{d,a} = {} & \xi + \beta \mathit{Post}_{d,a} \times \mathit{ETFIncrease}_{d,a} + \gamma \mathit{Post}_{d,a} \times \mathit{NASDAQ100}_{d,a} \\ & + \delta_1 \mathit{Post}_{d,a} + \delta_2 \mathit{NASDAQ100}_{d,a} + \sum_i b_i \mathit{ControlVariable}_{i,d,a} + \varepsilon_{d,a}. \end{aligned}$$

where *Illiquidity* is one of our four illiquidity measure, $\mathit{Post}_{d,a}$ is a binary variable that takes a value of 1 if $d > 0$, $\mathit{ETFIncrease}_{d,a}$ is a binary variable that takes a value of 1 if $a = Up$, and $\mathit{NASDAQ100}_{d,a}$ is a binary variable that takes a value of 1 if $a \neq C$. We use turnover–defined as the ratio between daily share volume to number of shares outstanding–inverse price and market capitalization, as all of these variables co-vary strongly with liquidity measures. We also account for time-series trends in illiquidity by including day fixed effects. We are mainly interested in the β coefficient as it captures the relative change in the liquidity of stocks whose weight increased in the index compared to other stocks treated by the rebalancing. We expect β to be negative as these stocks' ETF ownership increases substantially compared to the other affected stocks.

Table 7 reports the results of this regression with different sets of control variables. The coefficients on the $Post \times ETFIncrease$ are all negative and significant in all specifications. On average, the stocks that experienced an increase in their index weights have an average ETF weight differential of approximately 0.8 percentage points over the stocks that realized a decrease in their index weights. A back-of-the-envelope calculation suggests that a 1 percentage point increase in ETF ownership approximately decreases the *effective spread* by 0.6 (0.49/0.8) basis point, *quoted spread* by 0.8 (0.63/0.8) basis points and *implementation shortfall* by 5.5 (4.7/0.8) basis points. Overall, these findings, produced by a completely different identification technique, provide further evidence that higher ETF ownership at the stock level can increase liquidity.

8. Underlying Mechanism

One may worry that liquidity benefits found in earlier analysis are a result of the liquidity benefits of ETFs are solely due to their passive investment structure, and their ETF-specific structure do result in any liquidity benefits. In this section, we aim to explore the ETF-specific channels through, which ETFs can lead to improvements in the liquidity of the stocks they own. The most important feature specific to ETFs is that there are intraday arbitrageurs who continuously trade the ETF and the basket of securities the ETF owns.

8.1. Arbitrageur Channel and Higher Resiliency

Arbitrageurs could transfer the liquidity of ETFs to underlying stocks. If ETFs are enlarging the investor base by attracting uninformed traders, then arbitrageurs simply transmit uninformed trades from ETFs to individual stocks. If a large investor's trades move the price of a stock, it may lead to an arbitrage opportunity between the ETF and the replicating basket that includes the stock. If an arbitrageur trades against this potential mispricing, those trades in the stock will be in the opposite direction of the investor's trading. If the investor buys the stock and generates a discount on the ETF, the arbitrageur will buy the ETF and sell the basket of stocks. This mechanism could lead to a liquidity transfer from the ETF to its underlying stocks, reducing the transaction costs of these stocks.

Overall, it is difficult to examine the impact of the arbitrage-based liquidity transfer theory. We analyze the implications of this potential channel by utilizing price trajectories in large order executions. We conjecture that if ETF arbitrageurs are trading against mispricings due to transitory price changes in the underlying stock, then their trading activity should impact the price dynamics observed during a large order execution.

We undertake this analysis using an institutional large-order trading dataset with information on child-order trades.[22] This data set was originally obtained from an investment bank to study the implications of investor heterogeneity on the estimation of the price impact (see Sağlam et al., 2019). This data set includes approximately 20,000 parent-orders consisting of more than 2.5 million child-order trades. This sample includes 15 months of data on S&P 500 stocks from January 1, 2011 – March 31, 2012. These large orders are submitted by 146 distinct investors comprised of primarily institutional portfolio managers. All of the orders in the dataset are executed to match the VWAP realized during the lifetime of the parent-order. This dataset provides various attributes at the parent-order level including information at the execution- and day-level. These statistics include order size, direction of the order (buy or sell), order start and end times, participation rate (the ratio of order size to the total volume during the trading interval), average execution price, proportional quoted spread and mid-quote volatility based on the duration of the execution. For each execution, we also have daily statistics on the stock including volume, turnover and number of trades. We complement these statistics with stock-level data from CRSP. The average order size is roughly $1 million and corresponds to an average participation rate of roughly 1.8%.

To test the presence of the arbitrageur channel, we compare the average price realized during the initial and final stages of the parent-order execution and compute the sensitivity of the difference in prices to ETF ownership.[23] First, we partition each large-order execution into two parts by

[22] Unfortunately, the Abel Noser data does not have information on individual child order trades: thus, we cannot use this data set for this analysis.

[23] Before we test the arbitraguer channel hypothesis, we verify that *implementation shortfall* computed from this dataset is also negatively correlated with ETF ownership.

assigning the child-order trades from the first 50% of the order size to the *FirstHalf* bin and the remaining 50% to the *SecondHalf* bin.[24] Finally, we compute the average trade prices in the *FirstHalf* and *SecondHalf* bins of the execution to define the price impact measures corresponding to the initial and final periods, *FirstIS* and *SecondIS*:

$$\text{(5)} \qquad \mathit{FirstIS}_i = \operatorname{sgn}(Q_i)\,\frac{\bar{P}_{i,1} - P_{i,0}}{P_{i,0}}, \qquad \mathit{SecondIS}_i = \operatorname{sgn}(Q_i)\,\frac{\bar{P}_{i,2} - \bar{P}_{i,1}}{\bar{P}_{i,1}}$$

where $\bar{P}_{i,1}$ is the average trade price for the *FirstHalf* bin and $\bar{P}_{i,2}$ is the average trade price for the remaining *SecondHalf* bin. We expect that if the arbitrage activity is the underlying mechanism, then the arbitrageur's trading will move the prices in the opposite direction of *FirstIS*, creating a price resilience for subsequent executions. Thus, if *FirstIS* is positive (negative), we expect negative (positive) correlation between *SecondIS* and ETF holdings. First, we check this claim. We condition on *FirstIS* being positive or negative and then test how *SecondIS* is correlated with ETF holdings. Let P and N be the set of executions for which *FirstIS* is positive and negative, respectively. Formally, we run the following regression at the parent-order level:

$$\mathit{SecondIS}_i = \xi_P + \beta_P \mathit{ETF\ Holding}_i + \sum_j d_{j,P}\,\mathit{ControlVariable}_j + \sum_{k=1}^{S} \theta_{k,P}\mathbb{I}_{\{m(i)=k\}} + \varepsilon_{i,P},\ i \in P,$$

$$\mathit{SecondIS}_i = \xi_N + \beta_N \mathit{ETF\ Holding}_i + \sum_j d_{j,N}\,\mathit{ControlVariable}_j + \sum_{k=1}^{S} \theta_{k,N}\mathbb{I}_{\{m(i)=k\}} + \varepsilon_{i,N},\ i \in N.$$

where we use stock fixed-effects using the mapping $i \xrightarrow{m} k$ to identify the executed stock k. In addition to these stock dummies, we also control for execution-level control variables including participation rate, mid-quote volatility, execution duration (expressed as a fraction of trading day), turnover, inverse price, and market capitalization of the executed stock.

Table 8 reports the results of these regressions. Consistent with our conjecture, the coefficient on *ETF Holding* has different signs in these two different regimes. We also verify that this difference is highly statistically significant using a Wald test. Overall, this finding suggests that the higher the ETF ownership of the stock, the larger the reversal of stock price in the second stage of the execution, suggesting a higher price resiliency. Given that *FirstIS* is positive, on average, the

[24] Not all of the executions can be exactly split into two equal portions. Formally, *FirstHalf* bin for the ith execution includes the maximum number of child orders from the start of the execution so that the cumulative sum of these orders is still less than or equal to 50% of the total order.

arbitrageur activity through ETF linkage can reduce the total execution costs.

This finding is also important for linking the oberseved liquidity benefits to a feature that is specific to ETFs. Relating liquidity improvements to intraday arbitrage activity, can refute the hypothesis that observed liquidity improvements are caused by features that are common across passive funds.

8.2. Higher liquidity due to lower inventory costs or information asymmetry

ETFs can be used as a hedging instrument by market makers and help them reduce their holding costs. Risk averse market makers demand compensation for holding risky inventory (Ho and Stoll, 1981; Grossman and Miller, 1988). For example, Hendershott and Menkveld (2014) estimate these holding costs for NYSE specialists and Biais et al. (2016) suggest that proprietary traders manage to keep their holding costs low. If market makers can hedge their inventories easily, their holding costs will be lower, allowing them to provide liquidity at a lower cost. ETFs can serve as an additional instrument for market makers to hedge their individual stock holdings, lowering their liquidity provision costs.

While this mechanism is very similar to the arbitrage activity considered in the earlier subsection, this channel specifically suggests that ETF liquidity is transferred through market makers. Relatedly, stocks with high ETF ownership may exhibit lower information asymmetry, as the cost of acquiring information about these stocks decreases with potentially more public information about them. If these theories were to drive our findings, we would expect to see lower inventory holding costs for stocks with high ETF ownership. We decompose the *effective spread* into *realized spread* and *adverse selection* (*price impact*) costs as proposed by Huang and Stoll (1997). *Realized spread* is the net percent profit margin of uninformed liquidity providers while *adverse selection* component can be interpreted as the uninformed liquidity suppliers loss to informed liquidity demanders. Both of these measures are obtained from WRDS Intraday Indicators database. We employ our main DiD regression design in equation 3 using *realized spread* and *adverse selection* as dependent variables.

Columns 3–6 in Table 9 report the regression results with and without monthly fixed effects. The coefficient on $Post \times Type2Add$ is positive and highly significant for *adverse selection.* The coefficients for the *realized spread* are statistically insignificant and negative. Overall, these results suggest that the improvement in *effective spread* with higher ETF ownership seems to be driven by lower adverse selection costs implying that ETFs decrease the liquidity costs due information asymmetry.

9. ETFs and Stressful Periods

In our main analysis, we studied the relation between ETF holdings and underlying stock liquidity during a period that covers both normal and stressful times. However, it has been argued that although ETFs can be beneficial in general, they could drain liquidity from stocks particularly during stressful periods. These periods occur when the liquidity is most needed, as liquidity can dry up in these periods partly due to funding constraints (Brunnermeier and Pedersen, 2008) or higher compensation required by liquidity providers (Nagel, 2012).

Specifically, regulators are concerned that ETFs may worsen already poor liquidity conditions as liquidation of ETFs in times of market stress may put additional selling pressure on the underlying stocks.[25] ETFs attract short-horizon traders who can quickly react to market events. Hence, these traders could generate large trading imbalances in ETFs during stressful market events, which can spillover to underlying stocks, affecting their liquidity. For example, Dannhauser and Hoseinzade (2017) argue that outflows from corporate bond ETFs had a significant impact on corporate bond spreads during the 2013 Taper Tantrum. Given these plausible hypotheses, it is crucial to study the sensitivity of our findings under stressful periods of trading. We will examine two stressful periods during our sample study to investigate the liquidity implications of ETFs in these turbulent periods.

First, we examine the financial crisis period between December 2007 and April 2009. During this episode, U.S. financial markets experienced significant selling pressure leading to substantial losses in main stock indexes. If ETFs are responsible for amplifying illiquidity during stress events, then we expect stocks with higher ETF ownership to have lower liquidity during this stressful period. In order to analyze the effect of ETFs during this crisis period, we run our main DiD regression (Eq refeq:Type2) for the S&P 500 additions during the crisis period. If the ETFs have a negative effect on liquidity, then we expect the β coefficient to be negative and significant.

Table 10 reports the results of the regressions for the S&P 500 additions during the crisis period.

[25] On the other hand, Karmaziene and Sokolovski (2019) argue that during the short-sales ban of 2008, ETFs helped improve the liquidity of the stocks that were affected by the ban.

We observe that the coefficients on the *Post*× *Type2Add* variable are all positive except for *effective spread*. The coefficients are all insignificant in the presence of monthly fixed effects. These findings suggest that a decrease in ETF ownership does not lead to any deterioration in liquidity during periods of market stress.

Second, using our daily ETF holding data set, we examine the institutional trading costs during the U.S. Debt Ceiling crisis period. U.S. financial markets experienced significant turbulence. It started with the Congress struggling to pass a resolution to lift the debt ceiling. An agreement to raise the debt ceiling was finally reached on July 31, 2011 and the bill was signed into law on August 2, 2011. Nevertheless, the S&P downgraded the U.S credit rating from AAA to AA+ after markets closed on August 5, 2011, citing concerns about the bill falling short of stabilizing the debt dynamics of U.S. government debt. Following this downgrade, the S&P 500 dropped 6.7 percent on August 8, 2011. These events likely weighed on economic activity and August nonfarm employment came in well below market expectations on September 2, 2011. Against the backdrop of these developments, on October 3, 2011, the budget proposal of the Greek government, which was already under notable financial stress, fell short of market expectations and put further pressure on global stock prices. We define the stressful period as July 22 through October 3 because, as shown in Figure A1 in the Online Appendix, this period includes the peak of the VIX and the trough of the S&P 500 index in 2011. Between July 22 and October 3, the S&P 500 declined 18.3 percent on net, and the VIX averaged 34.26, significantly higher than its average value of 19.15 in June 2011. Meanwhile, U.S. equity ETFs are estimated to have received about $ 7 billion (2% of their TNA) in outflows, suggesting that ETFs came under considerable stress during this period.[26]

If ETFs are responsible for additional pressure on liquidity during stress events, then we expect stocks with higher ETF ownership to have higher transaction costs during this stressful period. We expect this effect to be more pronounced on the sell-side because ETF redemptions suggest that arbitrageurs were buying ETFs and selling underlying stocks. Using our implementation shortfall measure at the stock-day level from Abel Noser data between July 22, 2011 and October 3, 2011, we run a panel regression to examine the correlation between ETF holding and the trading costs of large orders. We separately examine the impact on buy- and sell-side trades for potential asymmetric

[26] As Staer (2017) suggests ETF shares outstanding is often reported with a one-day lag. Hence, we use "T+1" accounting to calculate net ETF flows reported in Morningstar Direct.

effect.

$$(6) \qquad IS_{i,t} = \xi + \beta ETF\ Holding_{i,t} + \sum_j b_j ControlVariable_{j,i,t} + \alpha_i + \alpha_{\text{month}(t)} + \varepsilon_{i,t}.$$

where control variables include turnover, inverse price, market capitalization, active and index fund ownership along with stock and month fixed effects.

Table 11 reports the results of these regressions. We observe that sell-side trades have positive and significant coefficient on *ETF Holding* implying that liquidating the underlying asset in the direction of the aggregate ETF trading may be considerably more expensive. However, we find improved liquidity for the buy-side transactions, suggesting that sell-side pressure originating from ETFs may decrease the price impact of the large buy orders in stocks. Using a Wald test, we verify that this difference in the *ETF Holding* coefficient between sell- and buy-side orders is statistically significant at the 1% level. When both buy-side and sell-side executions are included, the coefficient for the *ETF Holding* variable is negative and statistically significant. Overall, there is evidence for a one-sided destabilizing effect of the ETFs, i.e., liquidity dries up on the sell-side but improves on the buy-side. This potential liquidity asymmetry between the sell-side and the buy-side due to ETFs is an interesting finding with important implications for asset managers. Short-term momentum-style investors could experience additional transaction costs whereas value investors can build positions with lower trading costs in times of market stress.

10. Conclusion

The effect of ETFs on their underlying stocks is of considerable interest to academics, regulators and practitioners. Using several measures of liquidity, we find that ETFs increase the liquidity of their underlying stocks.

We use a previously undiscovered S&P 500 addition pattern to explore the impact of ETFs. Stocks that are added to the S&P 500 but removed from the S&P MidCap 400 index experience a decrease in their ETF holdings whereas stocks added to the S&P 500 index outside of the S&P universe experience an increase in their ETF ownership. The divergence in ETF ownership provides us an identification strategy to explore the impact of ETF on the liquidity of the stocks they hold.

We find that illiquidity measures, *effective spreads*, *quoted spreads*, *Amihud* and *implementation shortfall* all increase when the ETF ownership of the stock goes down. This finding can not be explained by changes in index mutual fund ownership as index mutual fund ownership goes up when stocks are added to the S&P 500 index and removed from the S&P MidCap 400 index.

We also take steps to increase our confidence in this identification strategy, which is intended to isolate the impact of ETFs. First, we run a falsification test using a time period when the two types of S&P 500 additions experience increases in both ETF and index fund ownership. Unlike the time period when stocks that are added to the S&P 500 but removed from the S&P MidCap 400 index experience a decrease in their ETF holdings, we do not find that stocks that are added to the S&P 500 but removed from the S&P MidCap 400 index experience a deterioration in liquidity. Second, our findings from the special rebalancing of the NASDAQ 100 index are largely the same as the ones from our main identification method.

Examining the price dynamics over large order executions, we find that one primary channel of higher liquidity is due to arbitrageurs trading against potential mispricings between the ETF and the basket of underlying stocks. Further, we find that higher ETF ownership is positively correlated with lower adverse selection costs for market makers. These findings could have implications for designing markets and developing optimal execution algorithms.

The relation between ETFs and lower transaction costs come with important caveats. First, stock investors may benefit from these lower transactions costs only under normal conditions. During the 2011 U.S. debt ceiling crisis, sell trades in stocks with high ETF ownership incurred higher transaction costs. This result is consistent with ETFs draining liquidity from their underlying stocks and increasing their transaction costs during market stress.

Second, it is important to emphasize that improvement in liquidity results can only be related to plain-vanilla equity ETFs. It is not clear whether other exchange-traded products such as leveraged and inverse ETFs, synthetic ETFs, corporate bond ETFs, and volatility ETFs can have liquidity benefits for their underlying securities even under normal times. In fact, Bai et al. (2015) and Tuzun (2013) examine the impact of leveraged ETFs on stocks and find higher volatility due to predictable late-day rebalancing of these funds.

References

Anand, A., P. Irvine, A. Puckett, and K. Venkataraman (2011). Performance of institutional trading desks: An analysis of persistence in trading costs. *The Review of Financial Studies 25*(2), 557–598.

Anand, A., P. Irvine, A. Puckett, and K. Venkataraman (2013). Institutional trading and stock resiliency: Evidence from the 2007–2009 financial crisis. *Journal of Financial Economics 108*(3), 773–797.

Bai, Q., S. A. Bond, and B. Hatch (2015). The impact of leveraged and inverse etfs on underlying real estate returns. *Real Estate Economics 43*(1), 37–66.

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

Beneish, M. D. and R. E. Whaley (1996). An anatomy of the s&p game: The effects of changing the rules. *The Journal of Finance 51*(5), 1909–1930.

Bhattacharya, A. and M. O'Hara (2017). Can etfs increase market fragility? effect of information linkages in etf markets. *Available at SSRN*.

Biais, B., F. Declerck, and S. Moinas (2016). Who supplies liquidity, how and when?

Brogaard, J., M. Ringgenberg, and D. Sovich (2017). The economic impact of index investing. *Available at SSRN*.

Brunnermeier, M. K. and L. H. Pedersen (2008). Market liquidity and funding liquidity. *The Review of Financial Studies 22*(6), 2201–2238.

Chan, K., H. W. Kot, and G. Y. Tang (2013). A comprehensive long-term analysis of s&p 500 index additions and deletions. *Journal of Banking & Finance 37*(12), 4920–4930.

Chang, Y.-C., H. Hong, and I. Liskovich (2014). Regression discontinuity and the price effects of stock market indexing. *The Review of Financial Studies 28*(1), 212–246.

Chen, H., G. Noronha, and V. Singal (2004). The price response to s&p 500 index additions and deletions: Evidence of asymmetry and a new explanation. *The Journal of Finance 59*(4), 1901–1930.

Chinco, A. and V. Fos (2017). The sound of many funds rebalancing.

Coles, J. L., D. Heath, and M. Ringgenberg (2017). On index investing.

Da, Z. and S. Shive (2012). Exchange traded funds and asset return correlations. *European Financial Management*.

Dannhauser, C. and S. Hoseinzade (2017). The transformation of corporate bond investors and fragility: Evidence on mutual funds and etfs.

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

Foucault, T., R. Kozhan, and W. W. Tham (2017). Toxic arbitrage. *The Review of Financial Studies 30*(4), 1053–1094.

Glosten, L., S. Nallareddy, and Y. Zou (2016). Etf activity and informational efficiency of underlying securities.

Glosten, L. R. and P. R. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*(1), 71–100.

Gromb, D. and D. Vayanos (2002). Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics 66*(2-3), 361–407.

Grossman, S. J. and M. H. Miller (1988). Liquidity and market structure. *the Journal of Finance 43*(3), 617–633.

Hamm, S. J. (2014). The effect of etfs on stock liquidity.

Hasbrouck, J. (2003). Intraday price formation in us equity index markets. *The Journal of Finance 58*(6), 2375–2400.

Hegde, S. P. and J. B. McDermott (2003). The liquidity effects of revisions to the s&p 500 index: An empirical analysis. *Journal of Financial Markets 6*(3), 413–459.

Hendershott, T. and A. J. Menkveld (2014). Price pressures. *Journal of Financial Economics 114*(3), 405–423.

Hill, J. M., D. Nadig, and M. Hougan (2015). A comprehensive guide to exchange-traded funds (etfs). *Research Foundation Publications 2015*(3), 1–181.

Ho, T. and H. R. Stoll (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics 9*(1), 47–73.

Holden, C. W. (1995). Index arbitrage as cross-sectional market making. *Journal of Futures Markets 15*(4), 423–455.

Holden, C. W. and N. Jayoung (2019). Market accessibility, corporate bond etfs, and liquidity.

Hu, G., K. Jo, Y. A. Wang, and J. Xie (2017). Institutional trading and abel noser data. *Available at SSRN*.

Huang, R. D. and H. R. Stoll (1997). The components of the bid-ask spread: A general approach. *The Review of Financial Studies 10*(4), 995–1034.

Israeli, D., C. M. Lee, and S. A. Sridharan (2017). Is there a dark side to exchange traded funds? an information perspective. *Review of Accounting Studies 22*(3), 1048–1083.

Jegadeesh, N. and A. Subrahmanyam (1993). Liquidity effects of the introduction of the s&p 500 index futures contract on the underlying stocks. *Journal of Business*, 171–187.

Karmaziene, E. and V. Sokolovski (2019). Short selling equity exchange traded funds and its effect on stock market liquidity. *Swedish House of Finance Research Paper* (14-05).

Kyle, A. S. (1985, November). Continuous auctions and insider trading. *Econometrica 53*(6), 1315–1335.

Lettau, M. and A. Madhavan (2018). Exchange-traded funds 101 for economists. *Journal of Economic Perspectives 32*(1), 135–54.

Madhavan, A. and S. Aleksander (2016). Price dynamics and liquidity of exchange-traded funds. *Journal of Investment Management 14*(2), 1–17.

Malamud, S. (2016). A dynamic equilibrium model of etfs. *Available at SSRN*.

Nagel, S. (2012). Evaporating liquidity. *The Review of Financial Studies 25*(7), 2005–2039.

Roll, R., E. Schwartz, and A. Subrahmanyam (2007). Liquidity and the law of one price: the case of the futures-cash basis. *The Journal of Finance 62*(5), 2201–2234.

Sağlam, M., C. C. Moallemi, and M. G. Sotiropoulos (2019). Short-term trading skill: An analysis of investor heterogeneity and execution quality. *Journal of Financial Markets 42*, 1–28.

S&P U.S. Indices Methodology (2018).

Staer, A. (2017). Fund flows and underlying returns: The case of etfs. *International Journal of Business 22*(4).

Tuzun, T. (2013). Are leveraged and inverse etfs the new portfolio insurers? *Federal Reserve Board Finance and Economics Discussion Series* (48).

Ye, S. (2017). How do etfs affect the liquidity of the underlying corporate bonds?

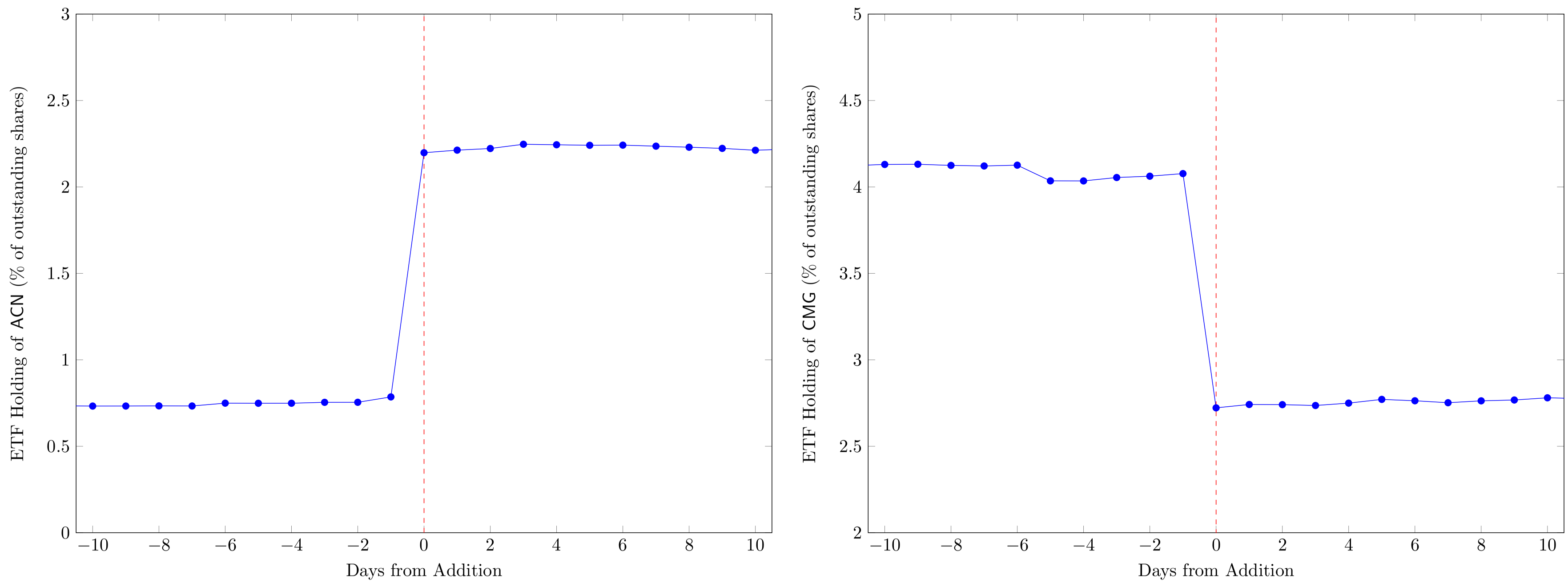

Figure 1: Change in ETF ownership for Accenture Inc. (NYSE:ACN, left) and Chipotle Inc. (NYSE:CMG, right) around the S&P 500 Type 1 and Type 2 additions, respectively.

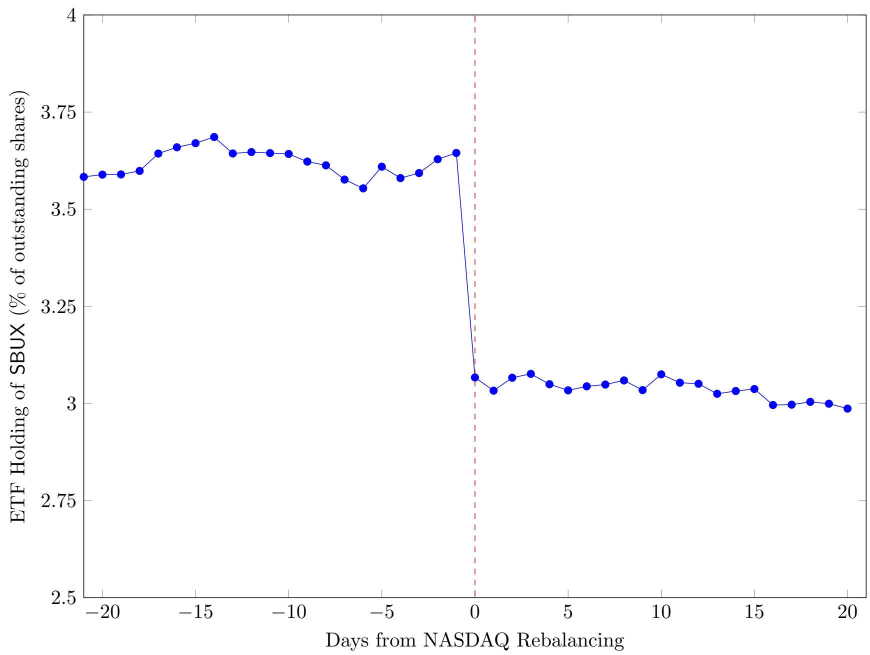

Figure 2: Change in ETF ownership for Starbucks Corp. (NASD:SBUX) around the NASDAQ 100 rebalancing.

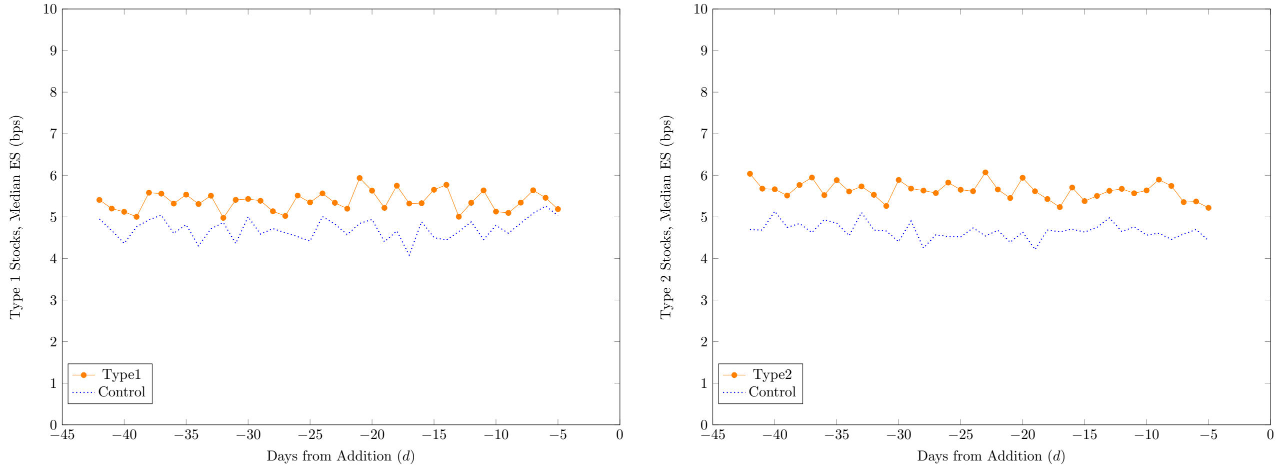

Figure 3: Median *effective spread* values for Type 1 (left panel) and Type 2 (right panel) stocks along with their matched control stocks before the typical announcement date of the addition.

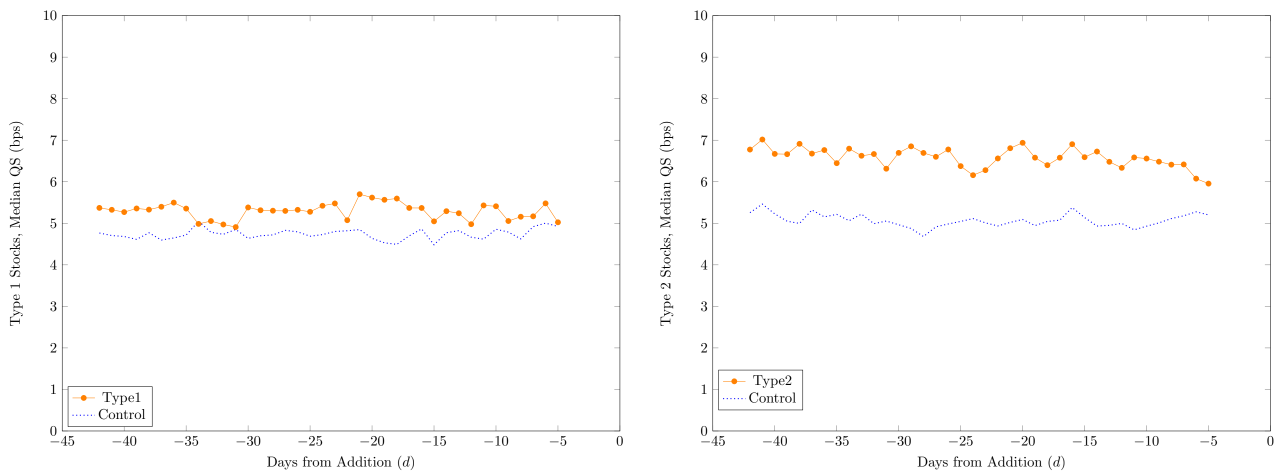

Figure 4: Median *quoted spread* values for Type 1 (left panel) and Type 2 (right panel) stocks along with their matched control stocks before the typical announcement date of the addition.

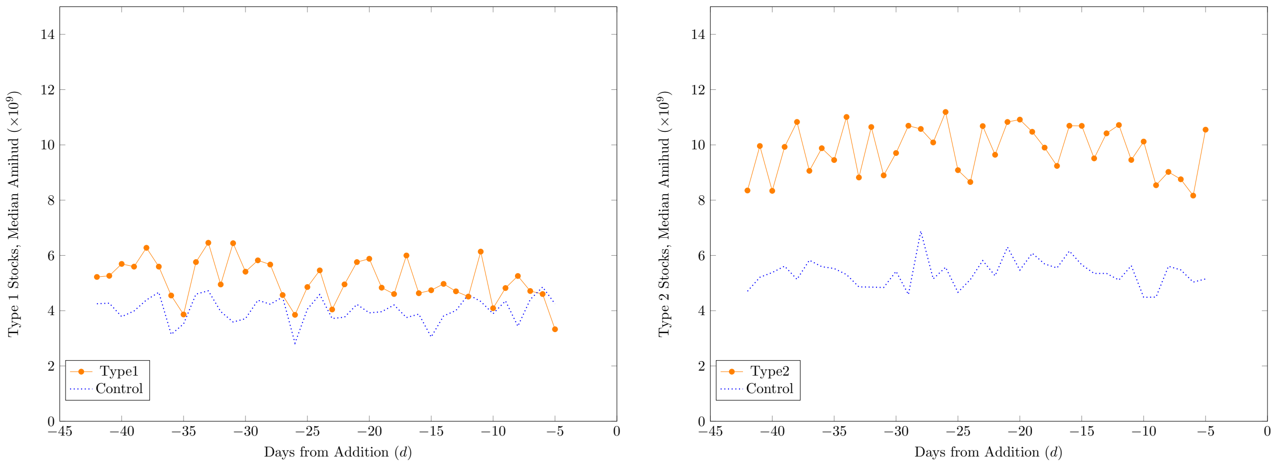

Figure 5: Median *Amihud* values for Type 1 (left panel) and Type 2 (right panel) stocks along with their matched control stocks before the typical announcement date of the addition.

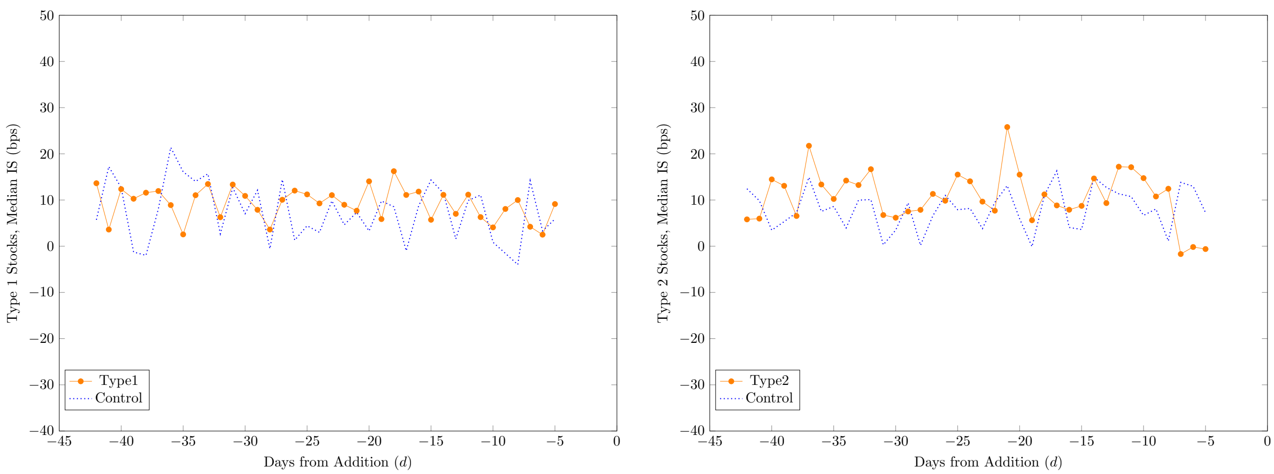

Figure 6: Median *implementation shortfall* values for Type 1 (left panel) and Type 2 (right panel) stocks along with their matched control stocks before the typical announcement date of the addition.

Table 1: Subsample of stocks added to the S&P 500 index. There are two types of additions. An addition is a Type 1 addition if it has not previously been in any S&P 1500 Composite index. In a Type 2 addition, the stock is removed from the S&P MidCap 400 index. We report the announcement date, effective addition date, and the ETF ownership ratios on the effective date of addition (Post) and the prior trading day (Pre).

Panel A: Type 1 Additions				
Announcement Date	Effective Addition Date	Trading Symbol	ETF Ownership (%) Pre	Post
1/11/2011	1/18/2011	NE	0.98	2.81
1/26/2011	3/1/2011	COV	0.81	2.16
3/29/2011	4/4/2011	BLK	0.5	1.59
6/27/2011	7/6/2011	ACN	0.78	2.19
9/21/2011	9/26/2011	MOS	1.06	2.67
10/11/2011	10/17/2011	TEL	0.48	1.92
11/18/2011	11/23/2011	CBE	1.23	2.62
3/7/2011	3/14/2012	CCI	1.4	2.77
5/17/2012	5/25/2012	KMI	0.51	2.41
5/21/2012	5/25/2012	ALXN	3.5	4.98
8/29/2012	9/5/2012	LYB	0.54	1.45
11/26/2012	12/3/2012	DG	0.88	2.14
12/05/2012	12/12/2012	GRMN	3.33	4.22
Panel B: Type 2 Additions				
Announcement Date	Effective Addition Date	Trading Symbol	ETF Ownership (%) Pre	Post
3/24/2011	4/1/2011	EW	3.94	2.78
4/20/2011	4/28/2011	CMG	4.07	2.72
12/8/2011	12/19/2011	DLTR	4.84	3.97
12/8/2011	12/19/2011	BWA	3.52	2.78
12/8/2011	12/19/2011	PRGO	5.43	4.74
3/28/2012	4/4/2012	FOSL	4.35	3.68
5/10/2012	6/5/2012	LRCX	5.14	3.94
6/21/2012	6/29/2012	MNST	4.59	4.07
9/24/2012	9/30/2012	PNR	8.86	5.09
10/1/2012	10/5/2012	PETM	4.11	3.22

Table 2: Type 2 Additions: Average changes in ETF and Index Fund Ownerships. We report averages for the weights of the Type 2 stocks in the S&P MidCap 400 (shown as S&P 400 for space constraints) and S&P 500 indexes as well as the averages for monthly net assets of the ETFs and index mutual funds tracking these indexes. ETF and index mutual fund ownership changes are computed from these variables. Average *Daily* ETF ownership change is computed from Markit ETF data.

	Weight (%)		ETF Net Assets ($ B)		Index Fund Net Assets ($ B)		Ownership Change (%)		
Year	S&P 400	S&P 500	S&P 400	S&P 500	S&P 400	S&P 500	ETFs	Index Funds	ETFs (Daily)
1993	1.37	0.21	0.00	0.37	0.10	16.61	0.01	0.55	n.a.
1994	1.03	0.14	0.00	0.45	0.13	19.31	0.01	0.56	n.a.
1995	0.84	0.11	0.02	0.60	0.18	27.79	0.01	0.67	n.a.
1996	0.93	0.12	0.08	1.22	0.74	48.15	0.02	1.08	n.a.
1997	0.73	0.08	0.23	2.87	0.30	86.74	0.01	1.26	n.a.
1998	0.80	0.10	0.76	7.73	0.39	138.80	0.02	1.74	n.a.
1999	0.98	0.09	1.83	13.13	0.46	209.54	-0.07	2.04	n.a.
2000	1.09	0.09	2.92	21.90	1.18	244.62	-0.12	1.97	n.a.
2001	0.70	0.06	4.68	30.35	2.34	215.46	-0.26	1.80	n.a.
2002	0.72	0.06	7.70	35.13	4.18	200.74	-0.58	1.60	n.a.
2003	0.63	0.07	7.15	42.70	5.29	186.81	-0.43	1.54	n.a.
2004	0.79	0.08	10.69	59.30	8.65	251.15	-0.48	1.50	n.a.
2005	0.70	0.07	14.41	70.48	11.33	277.09	-0.72	1.38	n.a.
2006	0.76	0.07	17.54	83.78	12.24	298.05	-0.77	1.34	n.a.
2007	0.67	0.07	19.01	98.99	13.65	336.00	-0.73	1.56	n.a.
2008	0.72	0.07	16.09	108.20	10.22	269.00	-0.69	1.64	n.a.
2009	0.61	0.05	14.43	98.80	8.40	229.06	-0.74	1.59	n.a.
2010	0.66	0.06	21.68	108.47	11.51	286.17	-1.12	1.36	n.a.
2011	0.72	0.07	26.29	131.33	13.84	327.87	-1.14	1.78	-0.45
2012	0.60	0.07	26.79	159.71	14.01	359.36	-0.89	1.89	-0.82
2013	0.86	0.08	39.74	220.01	18.13	440.67	-1.26	1.61	-0.58
2014	0.63	0.06	48.40	275.32	21.79	529.81	-1.32	1.79	-0.56
2015	0.70	0.07	54.68	331.86	22.98	604.40	-1.44	1.89	-0.78
2016	0.60	0.05	55.89	353.68	21.86	627.47	-1.68	2.21	-1.25
2017	0.60	0.05	76.45	474.91	25.80	768.53	-2.23	2.26	-2.17
2018	0.67	0.05	87.96	598.89	25.11	904.28	-2.13	2.52	-2.36

Table 3: Summary statistics. We report the data period, number of observations, mean, standard deviation, minimum, 25th percentile, median, 75th percentile and maximum values of the main variables we use in our analysis.

Statistic	Data Period	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Effective Spread (bps)	Sep2003–Dec2018	32,273	6.51	5.82	1.40	3.53	4.98	7.37	77.09
Quoted Spread (bps)	Sep2003–Dec2018	32,273	7.15	12.54	1.57	3.85	5.28	7.48	266.15
Amihud ($\times 10^9$)	Jan2002–Dec2018	35,990	9.92	17.27	0.00	2.08	5.48	12.32	354.86
IS (bps)	Jan2002–Sep2014	20,110	12.50	76.89	−510.55	−15.26	7.96	37.61	554.73
Inverse Price	Jan2002–Dec2018	35,992	0.03	0.03	0.0003	0.01	0.02	0.03	0.44
Mkt Cap ($B)	Jan2002–Dec2018	35,992	11.00	8.68	0.99	6.05	9.25	12.66	93.56
Turnover (%)	Jan2002–Dec2018	35,992	1.46	2.44	0.03	0.57	0.90	1.53	116.10
ETF Holding (%)	Jan2011–Dec2018	12,984	4.55	2.18	0.00	3.19	4.21	5.57	28.26

Table 4: Difference-in-differences analysis for Type 1 and Type 2 additions to the S&P 500 index with a control group. *Post* is a binary variable that takes a value of 1 after the effective date of the Type 1 or Type 2 addition, *Type2Add* is a binary variable that takes a value of 1 if the stock realizes a Type 2 addition that is removed from the S&P MidCap 400 index simultaneously, and *SP500Add* is a binary variable that takes a value of 1 if the stock is added to the S&P 500 index. Sample excludes the implementation date of the addition, and the prior day. Standard errors are given in parentheses and are adjusted by double clustering on the stock and calendar day level.

	Dependent variable:							
	ES (bps)		*QS (bps)*		*Amihud* ($\times 10^9$)		*IS (bps)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Type2Add	0.67**	0.57**	1.01**	1.13***	4.72***	5.33***	6.90***	5.70***
	(0.29)	(0.24)	(0.43)	(0.34)	(1.41)	(1.69)	(2.01)	(2.08)
Post × SP500Add	−1.01***	−0.97***	−1.64***	−1.73***	−4.83***	−5.22***	−7.98***	−7.25***
	(0.24)	(0.22)	(0.38)	(0.27)	(0.87)	(1.07)	(2.92)	(2.81)
Post	−0.003	−0.04	0.25	0.21	0.12	−0.48	1.22	1.71
	(0.12)	(0.13)	(0.24)	(0.17)	(0.25)	(0.52)	(2.59)	(2.58)
SP500Add	1.33***	1.20***	1.90***	1.78***	3.88***	3.51***	1.07	0.57
	(0.23)	(0.18)	(0.37)	(0.24)	(0.99)	(0.85)	(1.90)	(1.80)
Turnover	27.18***	23.43***	−22.58**	−31.62***	−86.62***	−94.91***	324.98***	312.29***
	(8.68)	(7.44)	(9.19)	(10.57)	(26.89)	(26.68)	(85.81)	(85.54)
Inverse Price	91.60***	80.86***	67.94***	56.97***	−77.85***	−104.22***	97.06***	84.66**
	(5.27)	(5.89)	(9.44)	(9.60)	(19.35)	(26.81)	(33.86)	(34.73)
Mkt Cap ($B)	−0.05***	−0.02**	−0.07***	−0.04***	−0.28***	−0.19***	−0.14**	−0.08
	(0.01)	(0.01)	(0.02)	(0.01)	(0.06)	(0.04)	(0.07)	(0.06)
Month FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	31,193	31,193	31,193	31,193	34,785	34,785	19,438	19,438
Adjusted R^2	0.22	0.29	0.03	0.07	0.06	0.14	0.01	0.01

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 5: Falsification test for the sample between 1993-03 and 1998-12. Difference-in-differences analysis for Type 1 and Type 2 additions to the S&P 500 index with a control group. *Post* is a binary variable that takes a value of 1 after the effective date of the Type 1 or Type 2 addition, *Type2Add* is a binary variable that takes a value of 1 if the stock realizes a Type 2 addition that is removed from the S&P MidCap 400 index simultaneously, and *SP500Add* is a binary variable that takes a value of 1 if the stock is added to the S&P 500 index. Standard errors are given in parentheses and are adjusted by double clustering on the stock and calendar day level.

	Dependent variable:					
	ES (bps)		*QS (bps)*		*Amihud* ($\times 10^9$)	
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Type2Add	−1.54	−0.16	−3.13	2.78	−2.13	−0.15
	(1.59)	(1.31)	(4.58)	(3.55)	(4.05)	(4.30)
Post × SP500Add	−1.96	-2.82^{***}	−2.41	-6.11^{***}	-16.58^{***}	-17.81^{***}
	(1.20)	(1.00)	(2.80)	(2.32)	(4.08)	(4.82)
Post	0.23	0.11	−1.05	−1.05	2.11	1.60
	(0.45)	(0.51)	(0.84)	(1.08)	(1.72)	(2.21)
SP500Add	3.03^{***}	3.24^{***}	5.97^{**}	6.52^{***}	9.21^{*}	9.23^{**}
	(1.00)	(0.86)	(2.85)	(2.29)	(5.02)	(4.62)
Turnover	301.64^{***}	269.85^{***}	−127.52	−187.74	$-1{,}756.79^{***}$	$-1{,}747.05^{***}$
	(64.20)	(63.24)	(171.80)	(139.19)	(220.13)	(198.93)
Inverse Price	834.18^{***}	822.01^{***}	$1{,}786.99^{***}$	$1{,}739.38^{***}$	-598.28^{**}	-639.30^{***}
	(61.08)	(52.83)	(161.87)	(123.38)	(246.71)	(244.93)
Mkt Cap ($B)	-0.52^{***}	-0.32^{***}	-1.97^{***}	-1.24^{***}	-3.26^{***}	-3.69^{***}
	(0.14)	(0.10)	(0.42)	(0.25)	(0.41)	(0.45)
Month FE	No	Yes	No	Yes	No	Yes
Observations	11,273	11,273	11,273	11,273	11,826	11,826
Adjusted R^2	0.34	0.37	0.51	0.65	0.12	0.14

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 6: This table reports the top 25 largest weight changes in the NASDAQ 100 index due to the special rebalancing.

Ticker	Old Weight	New Weight	Change
AAPL	20.49%	12.33%	-8.16%
MSFT	3.41%	8.32%	4.91%
ORCL	3.32%	6.68%	3.36%
INTC	1.75%	4.20%	2.45%
CSCO	1.56%	3.66%	2.10%
GOOG	4.18%	5.77%	1.59%
QCOM	5.00%	3.48%	-1.52%
AMGN	1.07%	1.92%	0.85%
SBUX	1.79%	1.08%	-0.71%
AMZN	2.50%	3.16%	0.66%
DELL	0.47%	1.08%	0.61%
COST	0.79%	1.26%	0.47%
YHOO	0.49%	0.86%	0.37%
ALTR	0.89%	0.53%	-0.36%
CMCSA	1.69%	2.03%	0.34%
BIDU	1.79%	1.46%	-0.33%
INTU	0.98%	0.65%	-0.33%
AMAT	0.49%	0.80%	0.31%
TEVA	1.68%	1.38%	-0.30%
BBBY	0.77%	0.48%	-0.29%
PCLN	1.24%	0.98%	-0.26%
PCAR	1.01%	0.76%	-0.25%
CTXS	0.78%	0.55%	-0.23%
GILD	1.55%	1.32%	-0.23%
RIMM	1.35%	1.13%	-0.22%

Table 7: Difference-in-differences analysis for the special rebalancing of the NASDAQ 100 index in May 2011. *Post* is a binary variable that takes a value of 1 for execution occurring after the effective date of the rebalancing, *ETFIncrease* is a binary variable that takes a value of 1 if the executed stock experienced increase in its weight in the NASDAQ 100 index and *NASDAQ100* is a binary variable that takes a value of 1 if the stock is affected by the rebalancing. As a set of controls, we include executions from NASDAQ-listed S&P 500 stocks within the same cutoff dates. We consider symmetric three-month windows for the pre- and post-event periods. Standard errors are given in parentheses and are adjusted by double clustering on the stock and calendar day level.

	Dependent variable:							
	Effective Spread (bps)		*Quoted Spread (bps)*		*Amihud* ($\times 10^9$)		*IS (bps)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × ETFIncrease	-0.49^{**}	-0.49^{**}	-0.63^{**}	-0.63^{**}	-2.18^{**}	-2.18^{**}	-4.39^{**}	-4.47^{**}
	(0.22)	(0.22)	(0.28)	(0.28)	(1.00)	(1.00)	(2.03)	(2.03)
Post × NASDAQ100	−0.12	−0.12	0.19	0.19	−0.72	−0.72	−3.17	−3.15
	(0.36)	(0.36)	(0.12)	(0.12)	(1.11)	(1.11)	(2.98)	(2.96)
Post	−0.14	−0.52	−0.16	-0.33^{**}	1.08	0.61	3.71	0.67
	(0.20)	(0.57)	(0.13)	(0.14)	(1.31)	(1.91)	(3.59)	(3.73)
NASDAQ100	-1.07^{***}	-1.07^{***}	-1.79^{***}	-1.79^{***}	-10.11^{***}	-10.11^{***}	−2.38	−2.39
	(0.41)	(0.41)	(0.34)	(0.34)	(2.01)	(2.00)	(1.90)	(1.90)
Turnover	4.66	4.75	1.73	1.77	-166.90^{***}	-166.78^{***}	126.63^{**}	126.82^{**}
	(5.34)	(5.39)	(7.19)	(7.21)	(48.88)	(48.90)	(54.46)	(54.09)
Inverse Price	69.96^{***}	69.98^{***}	76.65^{***}	76.67^{***}	-84.43^{**}	-84.42^{**}	15.08	16.64
	(5.02)	(5.01)	(11.06)	(11.06)	(39.36)	(39.34)	(12.78)	(12.69)
Mkt Cap ($B)	-0.01^{***}	-0.01^{***}	-0.01^{**}	-0.01^{**}	-0.04^{***}	-0.04^{***}	-0.03^{***}	-0.03^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.01)	(0.01)	(0.01)	(0.01)
Month FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	11,538	11,538	11,538	11,538	11,538	11,538	7,094	7,094
Adjusted R^2	0.34	0.34	0.65	0.65	0.10	0.10	0.01	0.01

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 8: Results of regressing *SecondIS* on ETF holdings and various control variables when executions are classified as positive and negative *FirstIS*. Standard errors are given in parentheses and are adjusted by double clustering on stock and the calendar day.

	Dependent variable: SecondIS	
	$FirstIS < 0$	$FirstIS > 0$
ETF Holding (%)	3.16	-3.64^{**}
	(2.32)	(1.68)
Participation Rate	45.93^{**}	38.09^{**}
	(23.27)	(16.76)
Volatility	−304.57	520.09^{**}
	(356.77)	(229.68)
Order Duration	−5.42	2.63
	(5.07)	(5.17)
Turnover (%)	-8.76^{**}	18.19^{***}
	(3.61)	(4.73)
Inverse Price	−62.12	−43.08
	(71.91)	(54.21)
Log Mkt Cap	−8.09	-26.78^{**}
	(9.83)	(10.78)
Stock FE	Yes	Yes
Observations	9,111	10,596
Adjusted R^2	0.03	0.07

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 9: Difference-in-differences analysis of *Realized Spread* and *Adverse Selection* for Type 1 and Type 2 additions to the S&P 500 index with a control group. *Post* is a binary variable that takes a value of 1 after the effective date of the Type 1 or Type 2 addition, *Type2Add* is a binary variable that takes a value of 1 if the stock realizes a Type 2 addition that is removed from the S&P MidCap 400 index simultaneously, and *SP500Add* is a binary variable that takes a value of 1 if the stock is added to the S&P 500 index. Standard errors are given in parentheses and are adjusted by double clustering on the stock and calendar day level.

	Dependent variable:			
	Realized Spread (bps)		*Adverse Selection (bps)*	
	(1)	(2)	(3)	(4)
Post × Type2Add	-0.17	-0.05	0.78^{***}	0.58^{***}
	(0.20)	(0.16)	(0.24)	(0.22)
Post × SP500Add	-0.10	-0.21	-0.88^{***}	-0.75^{***}
	(0.22)	(0.20)	(0.19)	(0.19)
Post	-0.06	-0.14	0.04	0.07
	(0.10)	(0.11)	(0.11)	(0.11)
SP500Add	0.62^{***}	0.48^{***}	0.68^{***}	0.69^{***}
	(0.18)	(0.15)	(0.19)	(0.16)
Turnover	21.35^{***}	22.75^{***}	5.16	-0.03
	(4.68)	(4.35)	(5.54)	(4.53)
Inverse Price	29.46^{***}	19.53^{***}	61.62^{***}	60.77^{***}
	(3.63)	(3.07)	(6.11)	(5.80)
Mkt Cap ($B)	-0.02^{*}	0.01	-0.03^{***}	-0.03^{***}
	(0.01)	(0.01)	(0.01)	(0.01)
Month FE	No	Yes	No	Yes
Observations	31,193	31,193	31,193	31,193
Adjusted R^2	0.03	0.06	0.09	0.13

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 10: Crisis period difference-in-differences analysis for Type 1 and Type 2 additions to the S&P 500 index with a control group. *Post* is a binary variable that takes a value of 1 after the effective date of the Type 1 or Type 2 addition, *Type2Add* is a binary variable that takes a value of 1 if the stock realizes a Type 2 addition that is removed from the S&P MidCap 400 index simultaneously, and *SP500Add* is a binary variable that takes a value of 1 if the stock is added to the S&P 500 index. Standard errors are given in parentheses and are adjusted by double clustering on the stock and calendar day level.

	Dependent variable:							
	Effective Spread (bps)		*Quoted Spread (bps)*		*Amihud* ($\times 10^9$)		*IS (bps)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Type2Add	0.84	0.28	2.46*	1.44	15.04	14.19	9.43	7.24
	(0.94)	(0.80)	(1.28)	(1.33)	(11.69)	(11.74)	(6.47)	(6.56)
Post × SP500Add	−0.57	−0.29	−2.32**	−1.97*	−9.29**	−9.03*	−8.53	−6.59
	(0.78)	(0.63)	(1.13)	(1.10)	(4.56)	(5.03)	(12.61)	(11.43)
Post	−0.63	−0.82*	0.08	0.13	0.20	−2.18	5.20	3.08
	(0.46)	(0.42)	(0.58)	(0.64)	(1.15)	(2.12)	(11.54)	(10.82)
SP500Add	1.28	1.19*	3.07**	2.63***	7.10	6.79	−1.05	−2.04
	(0.84)	(0.71)	(1.48)	(0.93)	(4.41)	(4.28)	(7.97)	(7.62)
Turnover	77.56***	72.18***	−60.54	−55.46*	−209.20	−157.24*	575.02***	601.98***
	(20.27)	(17.84)	(60.33)	(29.04)	(147.22)	(91.45)	(193.25)	(204.73)
Inverse Price	72.13***	69.90***	60.91***	45.81***	−147.40**	−159.11**	130.24	123.15
	(15.15)	(14.37)	(19.39)	(13.82)	(66.21)	(65.86)	(82.08)	(87.97)
Mkt Cap ($B)	−0.25***	−0.09**	−0.50**	−0.14	−1.23***	−1.21*	−1.20*	−0.90
	(0.06)	(0.05)	(0.22)	(0.10)	(0.32)	(0.66)	(0.64)	(0.56)
Month FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4,325	4,325	4,325	4,325	4,325	4,325	3,438	3,438
Adjusted R^2	0.23	0.28	0.03	0.07	0.08	0.13	0.01	0.01

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table 11: Results of regressing *IS* on ETF holdings and various control variables around the 2011 U.S. debt ceiling crisis. Standard errors are given in parentheses and are adjusted by double clustering on the stock and calendar day.

	Dependent variable: IS		
	All	*Sells*	*Buys*
ETF Holding (%)	-4.52^{***}	12.69^{***}	-21.73^{***}
	(0.60)	(2.65)	(3.33)
Turnover	146.55^{*}	122.69	169.86
	(75.08)	(156.49)	(192.61)
Inverse Price	46.03	394.50	−293.75
	(37.70)	(262.14)	(226.94)
Mkt Cap ($B)	0.03	-3.36^{*}	3.42^{*}
	(0.39)	(1.88)	(1.98)
Index Holding (%)	3.05	−7.58	13.66^{**}
	(4.66)		(5.36)
Active Holding (%)	−0.08	2.21^{**}	-2.34^{**}
	(0.50)	(0.93)	(0.95)
Stock FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Observations	41,575	20,778	20,797
Adjusted R^2	0.001	0.01	0.02

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Online Appendix to "Do ETFs Increase Liquidity?"

Mehmet Sağlam
University of Cincinnati
email: mehmet.saglam@uc.edu

Tugkan Tuzun
Federal Reserve Board
email: tugkan.tuzun@frb.gov

Russ Wermers
University of Maryland
email: wermers@umd.edu

September 2020

A. Robustness Tests

In this section, we provide two robustness tests for our main DiD analysis based on Type 1 and Type 2 additions. First, we include daily volatility as a control variable as reported by WRDS Intraday Indicators based on second-by-second quotes. This intraday volatility data are also available between September 2003 and December 2018. Table A1 illustrates that the loadings on $Post \times Type2Add$ remain very close to the ones reported in Table 4. Second, to account for any abnormal trading activity right before and after the effective addition date, we exclude all observations occurring within five business days to the S&P 500 effective addition date. Since the announcements of the additions typically occur in this interval, this robustness test also excludes any trading activity in response to the announcement itself. Table A2 again shows that our findings are robust to the exclusion of these trading days.

Table A1: Difference-in-differences analysis including daily volatility as a control variable for Type 1 and Type 2 additions to the S&P 500 index with a control group. *Post* is a binary variable that takes a value of 1 after the effective date of the Type 1 or Type 2 addition, *Type2Add* is a binary variable that takes a value of 1 if the stock realizes a Type 2 addition that is removed from the S&P MidCap 400 index simultaneously, and *SP500Add* is a binary variable that takes a value of 1 if the stock is added to the S&P 500 index. Sample excludes the implementation date of the addition, and the prior day. Standard errors are given in parentheses and are adjusted by double clustering on the stock and calendar day.

	Dependent variable:							
	Effective Spread (bps)		*Quoted Spread (bps)*		*Amihud* ($\times 10^9$)		*Implementation Shortfall (bps)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Type2Add	0.65^{**}	0.56^{**}	1.01^{**}	1.13^{***}	4.99^{***}	5.32^{***}	7.20^{***}	6.53^{***}
	(0.29)	(0.24)	(0.43)	(0.34)	(1.49)	(1.83)	(2.06)	(2.16)
Post × SP500Add	-1.01^{***}	-0.97^{***}	-1.64^{***}	-1.73^{***}	-4.67^{***}	-4.88^{***}	-7.09^{**}	-6.30^{**}
	(0.24)	(0.22)	(0.38)	(0.27)	(0.87)	(1.12)	(3.22)	(3.05)
Post	−0.003	−0.04	0.25	0.21	0.02	−0.57	0.14	0.37
	(0.12)	(0.13)	(0.24)	(0.17)	(0.22)	(0.54)	(2.84)	(2.74)
SP500Add	1.33^{***}	1.20^{***}	1.90^{***}	1.78^{***}	3.85^{***}	3.55^{***}	1.67	1.04
	(0.23)	(0.18)	(0.37)	(0.24)	(1.03)	(0.88)	(2.04)	(1.94)
Turnover	27.17^{***}	23.45^{***}	-22.58^{**}	-31.61^{***}	-63.37^{***}	-76.83^{***}	308.11^{***}	305.06^{***}
	(8.68)	(7.44)	(9.19)	(10.57)	(23.54)	(22.71)	(87.00)	(87.60)
Inverse Price	91.62^{***}	80.88^{***}	67.94^{***}	56.97^{***}	-90.82^{***}	-115.48^{***}	173.92^{***}	144.57^{***}
	(5.27)	(5.88)	(9.44)	(9.60)	(24.42)	(33.89)	(33.78)	(43.63)
Mkt Cap ($B)	-0.05^{***}	-0.02^{**}	-0.07^{***}	-0.04^{***}	-0.27^{***}	-0.19^{***}	-0.12^{*}	−0.08
	(0.01)	(0.01)	(0.02)	(0.01)	(0.06)	(0.03)	(0.07)	(0.07)
Volatility (bps)	0.09^{***}	0.09^{***}	0.002	0.004^{*}	-0.02^{***}	-0.02^{***}	0.003	−0.01
	(0.001)	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)	(0.01)
Month FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	31,193	31,193	31,193	31,193	31,192	31,192	16,803	16,803
Adjusted R^2	0.22	0.30	0.03	0.07	0.06	0.14	0.01	0.01

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

Table A2: Difference-in-differences analysis for Type 1 and Type 2 additions to the S&P 500 index with a control group. *Post* is a binary variable that takes a value of 1 after the effective date of the Type 1 or Type 2 addition, *Type2Add* is a binary variable that takes a value of 1 if the stock realizes a Type 2 addition that is removed from the S&P MidCap 400 index simultaneously, and *SP500Add* is a binary variable that takes a value of 1 if the stock is added to the S&P 500 index. Sample excludes the 5 trading days before and after the index addition. Standard errors are given in parentheses and are adjusted by double clustering on the stock and calendar day.

	Dependent variable:							
	Effective Spread (bps)		*Quoted Spread (bps)*		*Amihud* ($\times 10^9$)		*Implementation Shortfall (bps)*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post × Type2Add	0.68^{**}	0.58^{**}	0.92^{**}	1.12^{***}	4.87^{***}	5.57^{***}	7.26^{***}	6.09^{***}
	(0.31)	(0.26)	(0.44)	(0.37)	(1.44)	(1.75)	(2.07)	(2.16)
Post × SP500Add	-1.02^{***}	-0.98^{***}	-1.71^{***}	-1.84^{***}	-4.86^{***}	-5.31^{***}	-8.27^{**}	-7.53^{**}
	(0.26)	(0.23)	(0.40)	(0.30)	(0.92)	(1.13)	(3.26)	(3.15)
Post	−0.01	−0.07	0.30	0.20	0.08	−0.61	1.08	1.98
	(0.13)	(0.14)	(0.27)	(0.18)	(0.27)	(0.59)	(2.88)	(2.90)
SP500Add	1.39^{***}	1.25^{***}	2.02^{***}	1.92^{***}	4.28^{***}	3.91^{***}	2.63	2.14
	(0.24)	(0.18)	(0.37)	(0.25)	(1.01)	(0.87)	(2.11)	(1.99)
Turnover	36.06^{***}	30.57^{***}	-26.53^{**}	-39.15^{***}	-109.29^{***}	-120.97^{***}	377.30^{***}	352.77^{***}
	(8.96)	(7.14)	(10.37)	(11.63)	(25.18)	(23.52)	(81.67)	(81.24)
Inverse Price	91.38^{***}	80.49^{***}	63.55^{***}	54.22^{***}	-78.90^{***}	-105.49^{***}	88.70^{***}	69.97^{**}
	(5.35)	(5.90)	(8.81)	(10.03)	(19.66)	(27.35)	(32.48)	(32.42)
Mkt Cap ($B)	-0.05^{***}	-0.02^{**}	-0.07^{***}	-0.04^{***}	-0.29^{***}	-0.19^{***}	-0.18^{**}	−0.12
	(0.01)	(0.01)	(0.02)	(0.01)	(0.06)	(0.05)	(0.07)	(0.07)
Month FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	26,874	26,874	26,874	26,874	29,965	29,965	16,736	16,736
Adjusted R^2	0.21	0.28	0.02	0.07	0.06	0.14	0.01	0.01

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$

B. Benchmark Indexes of ETFs and Index Mutual Funds

Table A3: The list of the indexes based on the S&P 500 and S&P MidCap 400.

S&P 500 Indexes	S&P MidCap 400 Indexes
S&P 500 TR	S&P MidCap 400 TR
S&P 500 PR	S&P MidCap 400 PR
S&P 500 Value TR	S&P MidCap 400 Value TR
S&P 500 Growth TR	S&P MidCap 400 Growth TR
S&P 500 Pure Growth TR	S&P MidCap 400 Pure Growth TR
S&P 500 Pure Value TR	S&P MidCap 400 Pure Value TR
S&P 500 Equal Weighted TR	S&P MidCap 400 Equal Weighted TR
S&P 500 Buyback TR	S&P MidCap 400 Revenue-Weighted TR
S&P 500 Catholic Values NR	S&P MidCap 400 Low Volatility TR
S&P 500 Div and Free CF Yield TR	S&P MidCap 400 Dividend Arist TR
S&P 500 Dividend Aristocrats TR	S&P 400 Managed Risk 2.0 TR
S&P 500 Dynamic Gold Hedged TR	
S&P 500 Enhanced Value TR	
S&P 500 Ex-Financials & Real Estate	
S&P 500 Ex-Health Care	
S&P 500 Ex-Infor Tech & Telecom	
S&P 500 Fossil Fuel Free NR	
S&P 500 High Beta TR	
S&P 500 High Dividend	
S&P 500 High Momentum Value TR	
S&P 500 Low Volatility High Div TR	
S&P 500 Low Volatility Rate Rep TR	
S&P 500 Low Volatility TR	
S&P 500 Managed Risk 2.0 TR	
S&P 500 Minimum Volatility NR	
S&P 500 Momentum TR	
S&P 500 Quality TR	
S&P 500 Revenue-Weighted TR	
S&P 500 Top 50 TR	
S&P 500 Ex-Energy	
S&P 500 Volatility Response TR	

C. U.S. Debt Ceiling Crisis

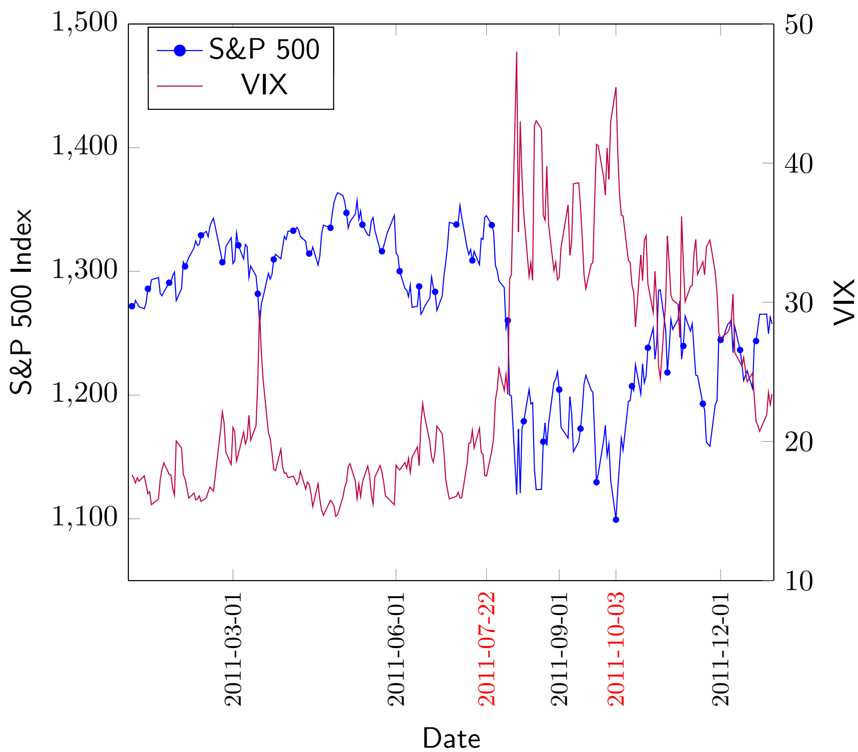

Figure A1: The S&P 500 index and VIX in 2011.

cfr working papers are available for download from www.cfr-cologne.de.

2021

No.	Author(s)	Title
21-05	G. Cici, P. Zhang	On the Valuation Skills of Corporate Bond Mutual Funds
21-04	V. Agarwal, P. Hanouna, R. Moussawi, C.W. Stahel	Do ETFs Increase the Commonality in Liquidity of Underlying Stocks?
21-03	M. Saglam, T. Tuzun, R. Wermers	Do ETFs Increase Liquidity?
21-02	C. Andres, D. Bazhutov, D. Cumming, P. Limbach	The M&A Rumor Productivity Dip
21-01	T. Bali, F. Weigert	Hedge Funds and the Positive Idiosyncratic Volatility Effect

2020

No.	Author(s)	Title
20-14	V. Agarwal, Y. Lu, S. Ray	Are hedge funds' charitable donations strategic?
20-13	M. Jagannathan, W. Jiao, R. Wermers	International Characteristic-Based Asset Pricing
20-12	E. Theissen, L. Zimmermann	Do Contented Customers Make Shareholders Wealthy? - Implications of Intangibles for Security Pricing
20-11	M. Hendriock	Implied Cost of Capital and Mutual Fund Performance
20-10	J. Fink, Stefan Palan, E. Theissen	Earnings Autocorrelation and the Post-Earnings-Announcement Drift – Experimental Evidence
20-09	E. Theissen, C. Yilanci	Momentum? What Momentum?
20-08	V. Agarwal, L. Jiang, Q. Wen	Why Do Mutual Funds Hold Lottery Stocks?
20-07	V. Agarwal, S. Ruenzi, F. Weigert	Unobserved Performance of Hedge Funds
20-06	M. Ammann, S. Fischer, F. Weigert	Factor Exposure Variation and Mutual Fund Performance
20-05	P. Limbach, P. R. Rau, H. Schürmann	The Death of Trust Across the U.S. Finance Industry
20-04	A. Y. Chen, T. Zimmermann	Open Source Cross-Sectional Asset Pricing

20-03	L. Ahnert, P. Vogt, V. Vonhoff, F. Weigert	Regulatory Stress Testing and Bank Performance
20-02	P. Schuster, E. Theissen, M. Uhrig-Homburg	Finanzwirtschaftliche Anwendungen der Blockchain-Technologie
20-01	S. Ruenzi, M. Ungeheuer, F. Weigert	Joint Extreme Events in Equity Returns and Liquidity and their Cross-Sectional Pricing Implications

2019

No.	Author(s)	Title
19-06	S. Greppmair, E. Theissen	Small Is Beautiful? How the Introduction of Mini Futures Contracts Affects the Regular Contract
19-05	G. Cici, M. Hendriock, A. Kempf	Finding your calling: Skill matching in the mutual fund industry
19-04	O. Korn, P. M. Möller, Ch. Schwehm	Drawdown Measures: Are They All the Same?
19-03	G. Cici, M. Hendriock, S. Jaspersen, A. Kempf	#MeToo Meets the Mutual Fund Industry: Productivity Effects of Sexual Harassment
19-02	T. Johann, S. Scharnowski, E. Theissen, C. Westheide, L. Zimmermann	Liquidity in the German Stock Market
19-01	A. Betzer, P. Limbach, P. R. Rau, H. Schürmann	Till death (or divorce) do us part: Early-life family disruption and investment behavior

2018

No.	Author(s)	Title
18-06	M. Goergen, P. Limbach, M. Scholz-Daneshgari	On the Choice of CEO Duality: Evidence from a Mandatory Disclosure Rule
18-05	F. Brochet, P. Limbach, D. Bazhutov, A. Betzer, M. Doumet	Where Does Investor Relations Matter the Most?
18-04	G. Cici, A. Kempf, C. Peitzmeier	Knowledge Spillovers in the Mutual Fund Industry through Labor Mobility
18-03	T. Rischen, E. Theissen	Underpricing in the Euro Area Corporate Bond Market: New Evidence from Post-Crisis Regulation and Quantitative Easing
18-02	S. Lesmeister, P. Limbach, M. Goergen	Trust and Shareholder Voting
18-01	G. Cici, M. Hendriock, A. Kempf	The Impact of Labor Mobility Restrictions on Managerial Actions: Evidence from the Mutual Fund Industry

This document only covers the most recent CFR working papers. A full list can be found at www.cfr-cologne.de.

centre for financial research
cfr/university of cologne
Albertus-Magnus-Platz
D-50923 cologne
fon +49(0)221-470-6995
fax +49(0)221-470-3992
kempf@cfr-cologne.de
www.cfr-cologne.de