ETF Short Interest and Failures-to-Deliver: Naked Short-Selling or Operational Shorting?

Richard B. Evans
University of Virginia

Rabih Moussawi
Michael S. Pagano
John Sedunov
Villanova University

First Draft: March, 2016
This Draft: August, 2018

Abstract

ETFs constitute 10% of U.S. equity market capitalization but over 20% of short interest and 78% of failures-to-deliver. While this disproportionate share of short activity has raised concerns about excessive shorting/naked short-selling of ETFs, we identify an alternative source of ETF shorting related to creation/redemption activities. This source, "operational shorting", is associated with not only improved liquidity and greater price efficiency, but also increased counterparty risk and trading linkages between liquidity providers. In exploring possible mechanisms for this risk relationship, we document a commonality in operational shorting across ETFs that share the same authorized participant and the financial leverage of the authorized participant appears to amplify this commonality.

Keywords: Exchange-Traded Funds, Failure to Deliver, Financial Markets, Short Selling, Market Making, Security Settlement, Short Interest, Counterparty Risk, Authorized Participants

JEL Codes: G1, G12, G14, G23

We are thankful for helpful comments and feedback from Aleksandar Andonov, Jesse Blocher, Ethan Chiang, Francois Cocquemas, Shaun Davies, Darrell Duffie, Austin Gerig, Wes Gray, Jeff Harris, Pankaj Jain, Bryan Johanson, Dolly King, Stewart Mayhew, Tom McInish, Matt Ringgenberg, Vivek Sharma, James Simpson, and Jack Vogel, as well as participants at the Chicago Financial Institutions Conference, FINRA-Columbia Market Structure Conference, 5th Luxembourg Asset Management Summit, the Southern Finance Association, VSB Mid-Atlantic Research Conference in Finance, and seminar participants at the Federal Reserve Board, Federal Reserve Bank of Richmond (and its Charlotte branch), Loyola University of Maryland, University of North Carolina – Charlotte, Penn State – Harrisburg, University of Georgia, University of Memphis, University of Mississippi, and University of Virginia. We also greatly appreciate the capable research assistance of Alejandro Cuevas, Shreya Rajbhandary, and Austin Ryback.

1. Introduction

With over $4.7 trillion invested worldwide, and $3.5 trillion in the U.S.,[1] exchange traded funds (ETFs) are a financial innovation that has been embraced by investors. ETFs are a hybrid structure that combine features of both a stock and a mutual fund, and ETFs enable investors to purchase an equity security that tracks the performance of an index or asset class without ever trading in the underlying securities. This ETF mechanism works, in part, because of the trading activities of "Authorized Participants" (APs), which are institutional investors who arbitrage the difference in price between the ETF itself and the underlying securities. By trading in both the underlying securities and the ETF shares (i.e. purchasing the underlying basket and swapping it for ETF shares thereby creating shares when the arbitrage is profitable), APs are able to provide liquidity that improves market quality in both the ETF and underlying basket of securities held by the ETFs.

While the arbitrage activities of APs help to maintain the price efficiency of the underlying, transmitting investor demand and supply of ETF shares to the price of the underlying securities, there is a growing academic literature (e.g., Ben-David, Franzoni and Moussawi, 2017) that examines potentially negative effects of that transmission mechanism including excess volatility, commonality in liquidity, and increased trading costs. What has been overlooked in this literature, however, is that not all ETF trading results in trading of the underlying securities. In their *2018 Fact Book*, the Investment Company Institute, highlights this disconnect between the ETF and the underlying:

> "The vast majority of ETF trades occur between investors on the secondary market (i.e., the stock exchanges) and don't "touch" the stocks ETFs hold. Those stocks are affected only on the primary market—when there are net creations or redemptions of ETF shares."

Measuring the ratio of secondary ETF trading activity to the sum of both primary and secondary trading the ICI finds that on average, $1 worth of ETF trading translates into only $0.12 of trading in the underlying securities.

[1] 2018 Investment Company Fact Book, Investment Company Institute, pages 86-87.

The purpose of our paper is to examine the exemption granted to APs that allows them to sell new shares of the ETF, without purchasing the underlying basket of securities. The mechanism underlying this exemption is described as follows:

> "Market makers, often commercial banks or hedge funds, create ETFs for their issuers by buying the securities that the funds are supposed to represent. But they've discovered that they can make a predictable return by delaying the purchases and selling you nonexistent exchange-traded fund shares that they will create later. These transactions—a form of shorting—eventually may involve 50,000 shares—the amount typically in a "creation unit" authorized by the issuer..." [2]

Under this market maker exemption, an AP can sell new ETF shares to satisfy a bullish order imbalance, but opt to delay the physical share creation – by purchasing the basket of underlying securities and swapping that basket for the corresponding number of ETF shares – until a future date. There are a number of operational reasons why an AP might want to delay creation. First, ETF creation is done in discrete blocks of ETF shares called creation units (typically 50,000 ETF shares). If the order imbalance is smaller than the creation unit size, APs may wait until the the imbalance builds to a size equal to or greater than the creation unit. Second, if the underlying basket of securities is less liquid than the ETF itself and purchasing the securities to form the creation basket incurs price impact and liquidity costs, order flow might reverse during the time that creation is delayed. This reversal would enable the AP to earn the ETF bid-ask spread, without paying the trading costs associated with buying the basket of underlying securities. Both of these motivations become even more compelling if an inexpensive and liquid hedge is available through the futures or options markets.

The option to wait described above permits an AP to sell or buy ETF shares today and postpone the corresponding creation and redemption process. Exercising this option has a number of implications for trading in both the ETF and underlying shares. First, it reduces the connection between ETF trading

---

[2] Jim McTague, "Market Maker's Edge: T+6", *Barron's*, 12/24/2011, accessed online 10/4/16 at http://www.barrons.com/articles/SB50001424052748703679304577108520307148702. Emphasis added by the authors of this paper.

and trading of the underlying securities. On one hand, the "disconnect" between trading the ETF shares and creating / redeeming these shares can help to mitigate the transmission of volatility in the ETF price to volatility in the underlying. On the other hand, this disconnect can impact the information transmission mechanism between the ETF and the underlying, thus altering price efficiency in these markets.

Second, ETFs have been associated with a disproportionate share of short interest and failures-to-deliver. As Figure 1 shows at the end of 2016, the aggregate dollar value of ETF short interest was upwards of $150 billion, accounting for 20% of the overall dollar value of short interest in U.S. equity markets, even though ETFs constituted just under 10% of total U.S. equity market capitalization. Similarly, as Figure 2 and Table 2 show, since the introduction of SEC Regulation SHO Rules 203 and 204 addressing failure-to-delivers (hereafter FTDs), ETF FTDs have grown and now account for over 78% of the dollar volume of all equity-related FTDs. Recent enforcement actions[3] and the increasing role played by ETFs in short-selling activity has been interpreted by academics, regulators and practitioners as an indication of naked or abusive short-selling practices.[4] However, if this option to delay is the driving force behind the high short interest and FTD numbers, not only does it mitigate the concern about abusive short-selling, it may be an indication the option is being used as intended by regulators – to enhance liquidity.

Third, if market making activities by the AP contributes strongly to the high observed short interest and failures-to-deliver (FTD), this has important implications for the information content for such short

[3] In March of 2016, FINRA and Nasdaq fined Wedbush Securities, an ETF AP, for submitting "naked" ETF redemption orders on behalf of a broker/dealer client, Scout Trading, in a number of levered ETFs. If Scout Trading wanted to profit from the well-documented price decline/decay of these leveraged ETFs (i.e. Zhang and Judge, 2016), but was unable or unwilling to borrow shares due to short selling constraints, one way to access short exposure would be to redeem or sell shares they did not own ("naked" redemption/short-selling), and subsequently fail-to-deliver those shares to Wedbush.

[4] Thomas Gira, the FINRA Executive Vice President of Market Regulation and Transparency Services, explains, the regulatory concern of interest is "naked" short-selling of ETFs: "Timely delivery of securities is a critical component of sales activity in the markets, particularly in ETFs that rely on the creation and redemption process. Naked trading strategies that result in a pattern of systemic and recurring fails flout such principles and do not comply with Regulation SHO. Authorized Participants and their broker-dealer clients need to have adequate supervisory procedures and controls in place to ensure that they are properly redeeming and creating shares of ETFs." FINRA News Release, "FINRA and Nasdaq Fine Wedbush Securities Inc. $675,000 For Supervisory Violations Relating to Chronic Fails to Deliver by a Client in Multiple Exchange-Traded Funds", 3/21/2016, accessed 6/2/2017 at http://www.finra.org/newsroom/2016/finra-and-nasdaq-fine-wedbush-securities-inc-675000-supervisory-violations-relating.

selling indicators. When APs exercise the option to delay, they are selling ETF shares that have not been created and which they do not currently own and are, in effect, have "sold short" the shares. Unlike such activity which may be informationally motivated, 'operational shorting' – a term we coin to describe the AP's shorting activity – is motivated by liquidity provision. This finding has important implications for the extant short-selling literature because it underscores the need to account for the different motivations behind ETF short selling: directional/informational vs. operational/liquidity provision.

Fourth, the common incentives of APs to exercise this option has implications for counterparty risk and potential financial contagion. While the option to delay does improve underlying stock liquidity and price efficiency, there remains a concern that FTDs increase the possibility of financial contagion through a commonality in the liquidity provision activities of an inter-connected network of ETF APs and market makers. We show that APs exhibit both within-firm and across-firm commonality in exercising this option to delay and that APs facing greater financial leverage constraints are more likely to exercise the option. Because the option involves selling ETF shares to a counterparty and then failing-to-deliver those shares until a later date, this suggests correlated shocks to counterparty risk across APs.

We propose a simple and novel methodology to estimate the operational shorting of ETFs and show that the estimate is consistent with the economics behind the proposed mechanism. The description of the above activity suggests that operational shorting occurs when new ETF shares are purchased by investors but there is a delay in the creation of those shares by the AP. To measure operational shorting, we use: a) the buy-sell imbalance (measured using signed intra-daily trade data) of a given ETF to proxy for the purchase of new ETF shares by investors and b) changes in the daily shares outstanding of the ETF to proxy for the delayed, or non-contemporaneous, net share creation activity. If the buy-sell trade imbalance is positive at a given point in time but there is no contemporaneous creation of the ETF shares, the AP is operationally short those shares because they have yet to create and deliver them to investors. Figure 3 presents a daily timeline that depicts the evolution of an operational short position for an AP. This timeline demonstrates how the rules related to "bona fide market making" can extend the actual delivery of the ETF shares for several days past the traditional T+3 settlement.

With our measure of operational shorting, we first examine the relation between operational shorting and FTDs. Figure 4 plots the aggregate dollar value of operational shorting and FTDs across all ETFs. Comparing the two time series, we see that there is a strong positive correlation between the two, consistent with operational shorting playing an important role in ETF FTDs. Repeating the analysis at the ETF level and controlling for the other potential determinants, we confirm this statistically and economically significant relation. The result is especially striking given that our operational shorting measure only identifies cases where there is excess demand for ETF shares (i.e., there is a buy imbalance that is greater than the number of shares created). More operational shorting in an ETF's shares is found to be driven by: 1) a higher liquidity mismatch with the ETF's underlying basket of securities and 2) the presence of efficient hedges. Our results provide important support and a rationale for why APs have an incentive to wait and delay the assembly of the basket and creation of new ETF shares until a future date.

After establishing how operational shorting results from liquidity provision in the ETF share market, we turn our attention to the basket of underlying securities, and we examine the impact of such "operational shorting" on common stocks that are held by those ETFs. The exception to Rule 204 for market makers is granted only when the operational short is "attributable to bona fide market making activities." Given that caveat, examining the impact of operational shorting on ETF liquidity is an important verification that indeed these short-sales represent legitimate market making activities. Fotak, Raman, and Yadav (2014)[5], along with Merrick, Naik, and Yadav (2005), argue that FTDs can serve as an "important release valve" that removes any binding constraints on market participants' ability to supply liquidity and perform valuable arbitrage activities.[6] We run a similar analysis of the impact of operational

[5] Fotak, Raman, and Yadav (2014) examined stock FTDs during 2005-2008 and found that increased levels of stock-related FTDs led to improved market quality in terms of reduced pricing errors, as well as lower levels of intraday volatility, bid-ask spreads, and order imbalances. In addition, they find that FTDs during the 2008 financial crisis did not distort prices.

[6] This notion of a "release valve" is also supported in terms of short selling activity's impact on loosening institutional constraints and sharpening price discovery. For example, Chu, Hirshleifer, and Ma (2016) show that the introduction of Regulation SHO (which reduced short selling constraints) has led to a reduction in returns to asset pricing anomalies. The authors suggest that this increase in short selling ability has made arbitrage of asset pricing anomalies easier and thus has decreased the returns to these strategies. In effect, like FTDs, Regulation SHO acted as another form of release valve which can lead to increased market efficiency.

shorting on the liquidity of underlying stocks by examining the relation between ETF operational shorting activities, stock volatility, and best bid and offer spreads on an intraday basis.

Consistent with Ben-David, Franzoni and Moussawi (2015) and a growing literature on ETFs,[7] we find that ETF ownership is positively associated with higher volatility and intraday spreads of the ETF's underlying basket of securities. However, we also show that operational shorting is negatively related to intraday spreads and volatility, thus acting as a "release valve." As operational shorting increases due to a sudden surge in buying demand, the APs can provide liquidity in the ETF market without (or before) entering the market for the underlying stocks. Therefore, our evidence suggests that operational shorting serves as a buffer that reduces the transmission of large ETF liquidity shocks to underlying stocks, especially when higher frequency investors are increasingly attracted to ETFs due to their greater degree of liquidity (Ben-David, Franzoni and Moussawi, 2015).

As a separate test of the underlying economics behind operational shorting, we also examine its predictive power for future returns. There is a long literature documenting that short-selling activity is predictive of future underperformance, consistent with a "directional" motive for informed investors to short sell. In analyzing the relationship between operational shorting and future (next week's) return, we find that operational shorting is negatively related to the following week's ETF return, but there is no statistically significant relationship between operational shorting and the future return on the underlying securities. This, combined with the strong positive relationship between operational shorting and concurrent ETF return suggests that APs anticipate the short-term ETF price reversal, but their operational shorting is not informative about the value of the underlying securities. Further evidence that operational shorting is liquidity driven is given by a split of the sample. The statistically significant negative relationship between operational shorting and future returns is driven by non-equity ETFs and "high liquidity mismatch" equity ETFs, where the ETF is substantially more liquid than the underlying securities.

[7] For example, Da and Shive (2014), Hamm (2014), Sullivan and Xiong (2012), Chinco and Fos (2016), Bhattacharya and O'Hara (2016), Dannhauser (2017), and Israeli, Lee, and Sridharan (2017). See Ben-David, Franzoni, and Moussawi (2017) for a survey of ETF literature.

These results have important implications for the extant short-selling literature because it underscores the need to account for the different motivations behind ETF short selling: directional/informational vs. operational/liquidity provision. In addition, while previous research has shown that common stock short interest is an important predictor of aggregate stock returns consistent with a primarily directional motivation for short-selling (i.e. Rapach, Ringgenberg, and Zhou, 2016), we document that operational shorting is one of the most significant drivers of an ETF's short interest.

While operational shorting does improve underlying stock liquidity, there remains a concern that FTDs impact financial stability through a commonality in the liquidity provision activities of an inter-connected network of ETF APs and market makers. This mechanism might be important, because ETFs, as hybrid investment vehicles, form an essential nexus between several areas of the financial system. In a 2011 report, the Financial Stability Board (FSB) raised concerns about ETFs and their potential impact on financial markets because the size and complexity of the ETF market could increase both counterparty risk and systemic risk.[8] The FSB report noted "the expectation of on-demand liquidity may create the conditions for acute redemption pressures on certain types of ETFs in situations of market stress." The unique redemption / creation process of ETFs, as well as the risks of trading, clearing, and settling these securities, are different than those present in the equity markets.[9]

While there are a number of potential channels through which ETF operational shorting and FTDs could relate to financial stress, Malamud (2015) models ETF liquidity provision and proposes one such channel. In that model, Malamud (2015) notes that ETF liquidity providers are fundamentally different because they typically play a dual role as ETF market makers and arbitrageurs between the market for ETF

[8] As financial crises involving U.S. and European financial institutions in recent years have shown, problems in one market can quickly create negative "spillover" or "contagion" effects to financial institutions that were not thought to be closely related. These spillover effects can lead to sudden, sharp spikes in a financial system's overall risk, commonly referred to as systemic risk. To the extent that ETFs can also employ financial leverage and derivatives, one can see that ETFs are at the nexus of the markets for cash equities, options, futures, credit, and securities lending. Thus, shocks to any of these markets can affect many other areas of the financial system via their linkages to ETFs and the institutions that serve as ETF market makers.

[9] For research on the FTDs of U.S. equities, see Boni (2006); Stratmann and Welborn (2013); Fotak, Raman, and Yadav (2014); Autore, Boulton, and Braga-Alves (2015); and Jain and Jain (2015). For FTDs in option markets and linkages to common stocks, see Evans, Geczy, Musto, and Reed (2009); Battalio and Schultz (2011); and Stratmann and Welborn (2013).

shares and the market for the ETF's basket of underlying securities. As a consequence of this dual role, he shows that the creation and redemption mechanism in the ETF markets can serve as a "shock propagation channel through which temporary demand shocks may have long-lasting impacts on future prices."

To assess whether or not this channel plays a role in the observed relationship between ETF operational shorting/FTDs shorting and systemic risk, we first identify all of the different ETFs served by a given lead market maker.[10] We then explore whether or not operational shorting/FTDs in a given ETF could lead to decreased liquidity and increased operational shorting/FTDs in other ETFs for which the same participant serves as the lead market maker. In addition, most ETFs have more than one AP and thus a sudden spike in operational shorting/FTDs (coinciding with a drop in liquidity in the ETF) by an AP in one ETF could spill over to other APs if they make markets in a common set of ETFs. This, in turn, could create a ripple effect throughout the entire ETF market and consequently increase counterparty risk and system-wide stress not only with ETFs but also with ETF-related common stocks and derivatives. Indeed, we find that increases in operational shorting and FTDs for one ETF are related to the operational shorting/FTDs of other ETFs that are traded by the same lead market maker. Thus, there appears to be commonality in operational shorting/FTDs across ETF market makers which suggests that such a contagion phenomenon may play a role in explaining the observed relation between operational shorting and systemic risk. Further, we examine a channel by which FTDs and operational shorting are associated with financial instability: leverage. We find that the financial leverage of lead market makers is positively related to both of our key dependent variables. This finding is consistent with our earlier observations and an AP's business strategy of increasing its return on equity by economizing not only on creation fees and trading costs but also capital requirements. Although this strategy might be profitable at the level of an individual AP, it can also lead to a more highly levered and inter-connected ETF market that is vulnerable to financial stress on an aggregate basis.

[10] Because the authorized participants for a given ETF are not reported in public sources, we use the lead market maker as our proxy for the authorized participant. Antoniewicz and Heinrichs (2015) report similar numbers of active APs and APs registered as market makers, suggestive that lead market maker would be a viable proxy for an active AP.

The remainder of the paper is organized as follows. Section two motivates and defines the empirical models used in our analysis. Section three describes the data, while section four presents our results, and section five concludes.

2. ETF Market Making and Fails-to-Deliver

2.1 The Mechanics of ETF Trading and Market Making

Madhavan (2014) describes ETFs as more than "exchange-traded versions of index mutual funds," as they have a mixture of elements related to both open-end and closed-end mutual funds, as well as the ability to be traded intraday and engage in "in-kind" securities transfers that have tax advantages for investors.[11] Similar to stocks and closed end funds, ETF shares trade on exchanges, and such secondary market trading constitutes the majority of ETF trading activity. ETF market makers ensure the liquidity of ETF trading in secondary markets by assuming obligations to provide continuous bid and ask quotes on ETFs. In instances of buy/sell imbalances in the ETF secondary markets or when trading cannot be met with existing shares, ETF market makers can also improve liquidity by either working with affiliated APs or serving as APs themselves to create (or redeem) blocks of ETF shares called creation units.

APs are institutions, typically broker-dealers or banks, that have contractual agreements with the ETF sponsor allowing them to trade directly with the sponsor to create and redeem ETF shares in the primary market.[12] For U.S. equity ETFs, such transactions are typically in kind, and a creation basket of securities is exchanged for a creation unit of ETF shares.[13] APs do not receive compensation from the ETF sponsor, but rather pay a creation fee for the transaction, and have no legal obligation to participate in ETF

---

[11] Antoniewicz and Heinrichs (2014) note that ETFs can use in-kind redemptions by redeeming "low basis" securities for purchases of new securities to reduce unrealized capital gains. In effect, ETF investors can defer most of their capital gains until they sell their shares.

[12] An AP is typically a market maker or large institutional investor that has a legal agreement with the ETF to create and redeem shares of the fund. APs do not receive any compensation from the ETF and have no obligation to create or redeem shares of the ETF. Instead, APs earn commissions and fees from customer orders as well as potential profits from ETF-common stock arbitrage. APs must also pay a flat fee for any creation or redemption orders.

[13] According to Ben-David, Franzoni and Moussawi (2017), 70% of ETFs traded in the U.S. have creation units with blocks of 50,000 ETF shares, but creation unit sizes can range from 25,000 to 200,000 shares.

primary markets. However, they do have strong financial incentives to participate, as the price discrepancy between the ETF share (market price) and the underlying basket (net asset value or NAV) represents a potentially profitable arbitrage. Through these incentives, APs help keep the ETF prices in the secondary market aligned with their intrinsic values. ETF shares are redeemed (in effect, taken out of circulation and thus lowering the supply of shares outstanding) when this process is reversed: the AP delivers a block of ETF shares equivalent to one or more creation units to the ETF investment manager in exchange for the specific basket of cash securities. Note that this redemption process is the result of selling pressures on ETF shares in the open market that can cause a discount in ETF prices relative to NAV, which creates an arbitrage opportunity for APs to redeem ETF shares with the fund for the constituent basket that are worth more in such instances.

2.2 Failures-to-Deliver

Stratmann and Welborn (2013) describe failures-to-deliver (FTDs) as "electronic IOUs" where a market participant who has engaged in a short sale does not deliver the underlying security at the time of settlement, which was typically 3 days after the sale in the U.S., and referred to as "T+3" in the parlance of securities trading and settlement.[14] Failure-to-deliver can occur with any type of security, and Table 2 shows FTD summary statistics overall and broken out by security type in terms of the aggregate market value of fails (Panel A) and fails as a percentage of aggregate shares outstanding (Panel B) from 2004 to 2016. Comparing the aggregate value of all FTDs in 2016 to the aggregate value of FTDs in different security types, we see that ETFs accounted for over 78% of all FTDs.

Existing research on FTDs in the U.S. equity market provides evidence of both positive and negative effects related to "limits to arbitrage" and "search and bargaining frictions" models. This literature includes Merrick et al. (2005) and Fotak et al. (2014) who argue that a more permissive policy towards FTDs can improve market quality. Additionally, Battalio and Schultz (2011) and Stratmann and Welborn (2013) find evidence supportive of Fotak et al.'s (2014) "release valve" view that FTDs can have

[14] While a shortened T+2 settlement cycle was implemented for most securities on September 5, 2017, T+3 was the settlement cycle during most of our sample period.

positive benefits for the overall market by encouraging traders to supply more liquidity and engage in useful arbitrage activities. Autore, Boulton, and Braga-Alves (2015) explore the issue from the perspective of valuation. They show that stocks with high levels of failures are more likely to be over-valued but this apparent trading opportunity is difficult to arbitrage due to the high costs of short selling in these relatively illiquid securities. Thus, less-liquid stocks can remain over-valued even in the presence of high levels of FTDs. In contrast, Jain and Jain (2015) report not only a significant decline in the level of equity FTDs but also a weakening in the relation between short selling activity and FTDs after the implementation of SEC Rules 203 and 204 in 2008-2009.

Additionally, Boni (2006) shows that FTDs were pervasive and persistent in U.S. equities during three settlement dates: September 2003, November 2003, and January 2004. This finding is consistent with market makers' incentive to "strategically fail" when borrowing costs are high. Boni's result suggests that one market participant's FTDs can spill over to other parts of the market and cause increased stress on the broader market. Using detailed data from a large options market maker, Evans et al. (2009) finds similar strategic failure behavior in U.S. equity options markets during 1998-1999. The authors observe that the use of FTDs is due to the relatively low cost of failing. They compute an FTD's cost as "the cost of a zero-rebate equity loan plus the expected incidence of buy-in costs" and find that it amounts to only 0.1 basis points in their sample.[15] Accordingly, Evans et al. (2009) conclude that failing to deliver securities can be profitable for market makers and that this activity can affect options prices.

2.3 ETF Failures-to-Deliver

While the high level of ETF short interest and FTDs combined with evidence in the literature about strategic failing in equities may raise concerns about abusive ETF short-selling, the unique liquidity provision mechanism underlying ETFs provides a potential alternative explanation. One important objective of APs in the primary ETF market is to harvest the difference between ETF market price and its

[15] "Buy-in costs" refer to the expenses incurred by a market participant who is forced to close out its FTD via the clearinghouse, the National Securities Clearing Corp. (NSCC). For an excellent description of the process of short selling, rebates, FTDs, and buy-ins, see Appendix A of Evans et al. (2009).

NAV, or the price of the underlying securities that comprise the ETF basket. As demand for the ETF grows from investors in the secondary market, the ETF's market price should increase, creating a potentially more attractive arbitrage bewteen the market price and NAV. One might assume the AP then immediately purchases the basket of underlying at the NAV, swaps it for ETF shares, which are then sold in the secondary market. However, these two different legs of the trade (i.e. selling ETF shares and buying the underlying basket/creating the ETF shares) are not necessarily instantaneous.[16] Under prevailing market making rules, the AP sells the new ETF shares to satisfy a bullish order imbalance but can opt to delay the physical share creation until a future date. By selling ETF shares that have not yet been created, the AP incurs a short position for operational reasons (as opposed to informational advantages) that we hereafter call an "operational short" position.

This option to delay the purchase of the underlying basket and the subsequent ETF share creation of shares that have already been sold is valuable in the presence of transient, mean-reverting ETF order flows, especially when an inexpensive hedge is available through the futures and options markets. For underlying basket securities that are less liquid than ETF securities, and might incur steeper price impact and liquidity costs for amassing creation baskets, such an option to delay creation becomes even more valuable.

As an example of these incentives, consider the following scenario. APs must create ETF shares in creation units, which are typically blocks of 50,000 ETF shares. While an AP with an open operational short position of 75,000 shares would ideally create 1.5 creation units to close out this position, share creation can only occur in blocks of 50,000. In this example, the AP would be forced to either create 1

[16] Index Universe explains below using "Bob", a hypothetical market maker, they can actually sell the ETF shares before they enact the ETF creation, effectively generating an uncovered short position:
"Market makers are given more time to settle their accounts than everyone else: While most investors' trades must settle in T+3, market makers have up to T+6. Market makers often have reason to delay settlement for as long as they can, particularly for ETFs. If Bob is a market maker trading ETFs, it might deliberately sell more and more shares of SPY short until it's sold enough to warrant creating a basket with the ETF issuer, thus making good on its sales. The longer Bob delays basket creation, the longer it can avoid paying the creation fee (often $500 or $1,000) and related execution costs. Moreover, it can delay the time it takes before taking on responsibility for a full creation basket of ETF shares (often 50,000 shares)."
"ETF.com Briefing Book", *Index Universe*, 10/18/2011, pg. 14.

block of 50,000 shares or 2 blocks of 100,000 shares, both of which deviate from the AP's desired quantity of 75,000 shares. Due to the indivisibility of creation units, the AP might defer the creation of the second unit if he/she thinks the ETF's order flow is persistent and mean-reverting over time. By creating one unit of 50,000 shares today and then waiting for the next day's order flow to (hopefully) mean-revert to a negative 25,000 share order imbalance, the AP can cover the full 75,000 share short position because the -25,000 share imbalance can be offset by the AP buying 25,000 shares. Thus, by "partially cleaning up" the position with 1 creation unit and then waiting a day (or longer) with an open short position of 25,000 shares, the AP might be able to create a zero net position without having to incur the extra transaction costs and capital outlay for a second block of 50,000 shares. While this is a cursory description of the value of the option to wait, a more explicit numerical example of trade-off between the costs and benefits of an AP covering a short position either immediately or waiting for up to 6 days given in Appendix A. As with our simple example above, the more in-depth example in Appendix A suggests that the option to delay has the potential to generate large, predictable profits for APs (e.g., by avoiding creation fees and delaying the outlay of capital to accumulate the full creation basket of underlying securities).

If an AP or other market participant sells ETF shares that it does not already own and subsequently does not deliver to the NSCC within T+3 days, a failure to deliver (FTD) occurs. This can happen due to operational shorting, as part of bona fide market making activity, as well as directional shorting, or naked short selling with the purpose of obtaining a negative exposure in the ETF shares in anticipation of a future decline in ETF price. Our tests aim to distinguish between these two distinct motivations. The NSCC can then force a "buy-in" of an outstanding FTD by typically contacting the market participant with the oldest FTD and requiring them to purchase or borrow the shares in the open market. As Evans et al. (2009) reports, buy-ins are a relatively rare occurrence and the expected cost of failures is relatively low. Thus, there are economic incentives to failing, especially in the ETF market because of the difficulty in distinguishing between FTDs that are due to abusive short-selling and those FTDs that are due to liquidity provision and market-making.

While the literature on equity FTDs described above is much richer and more established, there are a handful of studies focusing on ETFs and their results suggest a greater potential for these hybrid investment vehicles to perturb financial markets. For example, as noted earlier, Madhavan (2012) and Ben-David et al. (2015) demonstrate that ETFs may have consequences for the volatility of financial markets. Furthermore, in contrast to earlier findings, Stratmann and Wellborn (2016) find that ETF-related FTDs Granger-cause higher stock market volatility and lower future returns which can ultimately lead to increased market instability.

Overall, it remains an empirical question as to: 1) how FTDs affect the risks, returns, and costs of trading ETFs, 2) whether or not the underlying rationale behind FTDs is the same for stocks and ETFs, as well as 3) what effects ETF FTDs have on the underlying securities.

2.4 Operational Shorting

We propose a simple measure to estimate operational shorting or the short-selling that arises from ETF liquidity provision. The motivation and empirical predictions behind operational shorting are distinct from those of *directional shorting,* or naked short-selling, that can also result in FTDs. To estimate operational shorting, we compare the buy-sell imbalance for trading in the ETF (our proxy for excess demand or supply of the ETF) to changes in the share creation. The formula for our measure of operational shorting is:

$$Operational\ \ Shorting = \frac{\max[0,\big(0,\text{Cumulative Buy/Sell Imbalance}(t-3,t-1) - \Delta\text{Shares Outstanding}(t-1,t)\big)]}{\text{Shares Outstanding(t - 3)}} \quad (1)$$

To calculate the buy-sell imbalance, we classify intraday trades in the ETF as buys or sells by comparing the execution price of the trade with the national best bid and offer (NBBO).[17] We then aggregate the buy-sell imbalance from time t-3 to t-1 because 3 days is the typical time between a short sale and its delivery for trades other than bona fide market making by an AP. We take the maximum of the buy-sell imbalance

[17] NBBO stands for the national best bid and offer, which is obtained from the NYSE TAQ database.

and 0 to ensure our measure captures only buy imbalances. We then subtract from this the daily net create/redeem activity, which is computed as the changes in ETF shares outstanding from t-1 to t because it is at time-t when prior short sales are expected to be covered. Finally, we use the maximum function to focus on those cases where excess buys (as measured by a large, positive buy-sell imbalance) exceed the actual creation of shares, as measured by the changes in shares outstanding. We normalize this result by dividing by the number of ETF shares outstanding. To ensure that our measure of operational shorting is solely capturing excess buys beyond contemporaneous creation activity, and not driven by excess redemptions relative to a sell-imbalance (i.e., ΔShares Outstanding(t-1,t) < Cumulative Buy/Sell Imbalance(t-3,t-1) < 0), we set operational shorting to 0 whenever there is a sell imbalance.

To understand the timing of these measures, consider the AP's decision of whether or not to submit a create order on date t. Observing excess demand for the ETF shares on date t (e.g., Cumulative Buy/Sell Imbalance (t-3,t-1) > 0), APs "acting as market makers or agents to market makers" might submit a create order on that date and have 3 trading days, until t+3, to deliver the basket of underlying to complete the creation.[18] If they deliver the underlying basket by the cutoff time on t+3, the ETF shares are created and the shares outstanding at t+4 would reflect the increased number of shares outstanding. However, if they fail-to-deliver, the ETF shares outstanding will not change.

Figure 5 contains an illustrative example of how the cumulative buy-sell imbalance, change in shares outstanding, and fails-to-deliver might relate, further motivating our measure. The figure shows these cumulative quantities for the iShares Core S&P Total U.S. Stock Market ETF (ticker: ITOT) over the year 2012. Early on, there are sharp increases in the cumulative *buy/sell imbalance* (black line) indicative of excess demand for the ETF. The cumulative *change in shares outstanding* (dark grey line) responds to this imbalance consistent with APs submitting orders to create new ETF units. However, the response of

[18] Antoniewicz and Heinrichs (2014) explain how failing-to-deliver in the primary market can generate fails in the secondary market: "Market makers, which can include APs acting as market makers or agents to market makers, have up to three additional days to settle trades (a total of T+6) if their failure to deliver is the result of bona fide market making. This mismatch in timing can create delays in the settlement of both primary market ETF redemptions and secondary market ETF trades, as market makers often use ETFs to hedge their inventories."

the cumulative *change in shares outstanding* lags behind the excess demand, possibly due to the reasons described above. Precisely when demand for the ETF increases sharply and the increase in the supply of ETF shares lags is when a spike in the percentage of *fails-to-deliver* (light grey line) occurs in ITOT shares. It would appear that APs and market makers are accommodating the demand, but the delay in creating them generates the FTDs observed. The operational shorting measure we propose above compares the cumulative *buy-sell imbalance* to the cumulative *change in shares outstanding* as an estimate of the potential short positions and failures-to-deliver that result due to the lagged response of APs/market makers to the excess demand.

3. Data

Because ETFs sit at the intersection of many different markets, our empirical analysis requires data from nine different sources. A complete listing of variables, definitions, and sources is provided in Appendix B. The FTD data[19] are from the SEC's website and are made available to the SEC by National Securities Clearing Corporation's (NSCC).[20] The FTD database contains CUSIP numbers, issuer names, prices, and the total number of fails-to-deliver shares recorded in the NSCC's Continuous Net Settlement (CNS) system on a daily basis. The total number of fails-to-deliver represents the total outstanding balance of shares failed, that are aggregated over all NSCC members, regardless of when the original fail position was initiated.[21] We collect these data from March 22, 2004, which is the beginning of the dataset, through December 31, 2016.[22]

We supplement the SEC data with additional variables from other sources. We merge the data with

[19] The FTD data can be downloaded from the following SEC page: http://www.sec.gov/foia/docs/failsdata.htm.
[20] The National Securities Clearing Corporation (NSCC) is regulated by the SEC, and is a subsidiary of the Depository Trust and Clearing Corporation (DTCC). See http://www.dtcc.com/about/businesses-and-subsidiaries/nscc and http://www.dtcc.com/~/media/Files/Downloads/legal/rules/nscc_rules.pdf for more info.
[21] The total number of fails reported on day (t) reflect the fails originating at day (t) as well as the remaining outstanding fails that were not closed out from previous days. FINRA and the SEC do not distribute the actual timing of the share settlement fails, and instead disseminate the outstanding balance of fails at a given day.
[22] Prior to September 16, 2008, only securities with aggregated fails of 10,000 shares or more were reported in the data. After that date, however, all fails regardless of the outstanding fail amounts are included in the fail to deliver data that the SEC disseminates.

Compustat, CRSP, and Mergent FISD to determine the asset class of each of those securities, as well as the total shares outstanding or issue size. Stock price and volume data come from the CRSP database, and are used to calculate variables such as *market capitalization, stock turnover, illiquidity,* and *idiosyncratic volatlity*. ETF characteristics are extracted from CRSP Mutual Fund database, and we use the ETF Global database for additional ETF-specific information, such as the ETF lead market maker and the historical creation unit size and fee amounts. The ETF holdings of underlying stocks is drawn from Thomson-Reuters and CRSP Mutual Fund Database. Buy and sell trade volume information, intraday spread, and return volatility are calculated from the NYSE TAQ database. Short interest information is extracted from Compustat on a biweekly basis, and represents the level of consolidated short interest in shares as reported by exchanges and compiled by FINRA. We supplement these short interest data with daily information on securities lending supply, utilization, and lending fees using Markit Securities Finance database (formerly Data Xplorers).

To compute our measure of operational shorting, we need both the ETF net creation/redemption activity in the primary market and the daily buy-sell trade imbalances of ETF shares in the secondary market. For the daily ETF creation and redemption activity, we rely on the daily changes in the ETF total shares outstanding. We follow Ben-David, Franzoni, and Moussawi (2015) and extract the ETF shares oustsanding data from Bloomberg because they are not not reported accurately in CRSP and Compustat.[23]

In order to compute the daily buy-sell imbalances in ETF shares, we need first to appropriately sign the ETF trades into buys and sells. To do that, we use the TAQ millisecond database to classify every trade between 2004 and 2016 into a buy or sell trade using a modified algorithm that combines the methods of Lee and Ready (1991) and Ellis, Michaely, and O'Hara (2000). First, for each trade, we compute the national best bid and offer (NBBO) quote at the end of the previous millisecond. Then, we compare the

[23] Bloomberg sources the ETF shares outstanding data directly from ETF administrators and custodians, which provide the new shares outstanding information that reflects accepted new create and redeem orders after market hours on the transaction date. While Bloomberg reports this information on same the day the create/redeem orders are submitted and accepted, it might take several days for other data vendors and exchanges to reflect this information.

trade price to the best bid and best offer. The midpoint reference inherent to the Lee and Ready (1991) algorithm does not take into consideration the "outside trades" which are not permitted under the Reg NMS rules, and therefore are less likely to occur in recent periods. For this reason, we use a modified quote test based on Ellis, Michaely, and O'Hara (2000) who propose a clever methodology that acknowledges the clustering of buys on the offer price, and sales on the bid prices.[24]

Table 1 presents summary statistics for the key variables in our analysis. These data are computed on a daily basis for the entire ETF sample in the top portion of the table while the bottom portion reports statistics based on a sub-sample comprised solely of ETFs that invest in U.S. equities. Strikingly, the short interest ratio for the full sample, measured as a percentage of shares outstanding has a standard deviation of 11.84%, and the 99th percentile of its distribution is equal to 83.76%. This may be a product of the operational shorting mechanism that we described above. Moreover, we find that 0.42% of the average ETF's shares are considered failures (FTDs) at any given time. Lastly, the average value of our operational shorting measure is 1.01%, with a standard deviation of 2.89%.

4. Empirical Methodology and Results

4.1 FTD Summary Statistics

Table 2 presents the average daily FTDs in dollar volume (Panel A) and as a percentage of shares outstanding (Panel B) by year and by asset class. Over the course of our sample, the total volume of FTDs across all asset classes is concentrated in stocks and ETFs, and Figure 2 provides a graphical representation of the FTD volume for these two security types. The total dollar volume of FTDs increased until it reached

[24] According to Ellis, Michaely, and O'Hara (2000), the quote test is less accurate when the trades are not executed at the ask or the bid. Most importantly, the authors' argument is especially valid when the Lee and Ready algorithm fails to take into consideration trades executed outside the quotation. Additionally, once an executed trade price crosses the prevailing NBBO within a millisecond, we stop using the quote test for the rest of the millisecond. Instead, and for the rest of the trades during this millisecond, we rely on the tick test as it is likely that the quote test is not accurate, especially when there is intense high frequency algorithmic trading that is faster than the refresh rate of the quotes within a millisecond period. So, our modified Ellis, Michaely, and O'Hara method takes into consideration that buys are more likely to be executed at the ask, and sales on the bid price, and whenever an outside trade is observed during that millisecond, then the algorithm relies instead on the tick test until the end of the millisecond. After signing all trades during market hours, we sum all the buys and sales at 4:00 pm to construct our buy and sale volume for the day.

over $7 billion in 2007 and over $6 billion in 2008, but exhibits a dramatic decline in 2009, coincident with the passage of SEC rules 203 and 204. From this point forward, it appears as the SEC rule change was effective in curbing common stock FTDs, which remain relatively low at around $500 million, but ETF FTDs begin to increase again, peaking at just under $2.6 billion in 2016. In fact, Table 2 shows the average dollar value of ETF-related FTDs now represents 78.5% of all FTDs (up from 29.5% in 2008).

4.2 Operational Shorting and its Impact on Short Interest and Failures-to-Deliver (FTDs)

While the literature has used short interest and FTDs as measures of informed, directional short-selling, the motivation behind operational shorting, as described in Section 2, has very different implications. As a first test of whether or not operational shorting plays a role in ETF shorting activity, we examine the determinants of ETF biweekly short interest (specifications 1 to 3) and daily FTDs (specifications 4 to 6) both scaled by shares outstanding, and we include our measure of operational shorting. The results are given in Table 3.

To capture alternative motivations for short-selling, we include the lagged short interest and the cost of borrowing/short-selling ETF shares to control for the intensity of prior shorting activity. Including the lagged *Short Interest Ratio* captures any prior short-selling motivations (directional or operational) that persist, so that the coefficient on operational shorting will only be statistically significant if the innovations in short interest coincide with operational shorting activity. We also control for the cost of borrowing the ETF shares by including the *Daily Cost of Borrow Score*, which measures the cost of borrowing the ETF shares based on a decile rank score (10 corresponds to ETFs with the highest borrowing costs) of the securities lending fees. Similar to including lagged short interest, the inclusion of the *Daily Cost of Borrow Score* is another important control variable, as it captures both directional and operational motivations for short-selling. As it becomes more expensive to borrow and short ETF shares, directional short-sellers are more likely to fail-to-deliver. Likewise, as ETF borrow costs increase, APs are less likely to borrow shares to hedge their position, and more likely to fail-to-deliver. In both cases, including these variables biases our results towards the null hypothesis of operational short-selling not playing a role, because these variables proxy for both directional and operational motivations.

Beyond the *Short Interest Ratio*, *Operational Shorting*, and *Daily Cost of Borrow Score*, our regressions also include control variables based on the findings in Fotak et al. (2014) and Stratmann and Welborn (2016) related to the effects of ETF liquidity and options listing. We control for the ETF's liquidity by including size (*log of Market Cap*) and trading volume (*Share Turnover*). We expect that the ETF's asset size should be negatively related to FTDs because larger funds are typically more liquid and thus it is easier to locate shares prior to the T+6 deadline. We expect ETF trading volume to be positively related to FTDs because greater share turnover increases the likelihood that some shares might not be delivered in a timely fashion. We also include an options listing dummy (*Available Options Dummy*) equal to 1 if options are traded on the ETF. Those ETF's with listed options are likely to be larger thus leading to a lower amount of FTDs. All regressions used in our analysis include ETF and date fixed effects, and standard errors are clustered by ETF and date.

Table 3 shows that all coefficients have the expected sign and are statistically significant at the 1% level. Regressions (2), (3), (5) and (6) confirm that greater trading volume, short-selling activity, and securities borrowing costs are related to increased ETF FTDs and short interest levels. Of particular interest is the positive and statistically significant coefficient on *Operational Shorting* in regressions (3) and (6). Operational shorting appears to be a very strong determinant of short interest, suggesting that it represents a significant component of the short interest ratio. Additionally, the operational shorting coefficient is a stronger determinant of FTDs and appears to have a higher economic and statistical significance than short interest ratio. Thus, when operational shorting is high, short interest and FTDs both increase, even after controlling for prior short-selling activity, securities lending costs, and an ETF's liquidity-related variables such as the fund's asset size and trading volume. This finding underscores the need to decompose the effects of short-selling that might be directional or informational in nature from shorting activity that is due to liquidity provision (as measured by our *Operational Shorting* variable).

4.3 The Persistence of ETF net creation activity and order imbalances

When faced with a large buying imbalance, APs have two primary trading strategies: 1) locate or create a sufficient number of shares to satisfy this buyer-initiated demand, or 2) sell the ETF shares now without locating or creating them and then wait up to T+6 days to obtain and deliver the shares. As described in section 2.3 and Appendix A, if order flows are persistent and alternate between positive and negative imbalances over time, the AP typically has a strong incentive to follow the second strategy. However, if there are no clear patterns associated with net creations and order flow, then APs would have less incentive to engage in operational shorting of ETF shares.

In this section, we examine daily patterns of ETF creations (net of any redemptions), ETF order flows, their persistence, and potential reversal patterns to assess whether or not these patterns support the suggested underlying economics. We examine these dynamics and inter-relations between *Net Creation Activity* and *ETF Order Imbalance* using lagged values (days t-8 to t-1) of the dependent variables along with the liquidity-related *Controls* (fund size and trading volume), as follows:

$$Net\ \ Creation\ \ Activity_t \quad or\ \ ETF\ \ Order\ \ Imbalance_t = \alpha_0 + \ \ \alpha_1 Controls + \sum_{n=0}^{8} \beta_n ETF\ \ Order\ \ Imbalance_{t-n} + \sum_{n=0}^{8} \gamma_n Net\ \ Creation\ \ Activity_{t-n} + \ \ \epsilon_t \qquad (2)$$

Equation (2) provides a parsimonious way to identify any autoregressive patterns in the dependent variables as well as possible inter-relations between order imbalances and past creation activity, and vice versa.

The results from estimating equation (2) are contained in Table 4. Models (1)-(3) use contemporaneous and lagged values of order imbalances (days t-8 to t), as well as lagged values of net creation activity (days t-8 to t-1) to estimate their effects on the current level of *Net Create/Redeem Activity*.[25] *Net Create/Redeem Activity* is constructed on a daily basis as the percentage change in the overall ETF shares outstanding. Since this variable is a percentage change that, like a stock's return, is bounded below at -100%, we construct our flows variable, *Net Create/Redeem Activity,* as the *log*(1 + % change in shares outstanding). This variable is likely to be more symmetrical for AP creation as well as redemption activities. After controlling for the two ETF liquidity variables, regressions (1) – (3) show that net creation activity is highly persistent with all of the net creation and order imbalance variables yielding positive and significant parameters at the 1% level. Thus, the prior sequence of net creation activity and order

[25] We use lags up to 8 days to control for possible effects from prior short selling and FTD activity. To compute the operational shorting and order imbalance measures, we focus on buyer- and seller-initiated trades during U.S. market hours (9:30 am – 4:00 pm Eastern time) and do not include after-hours trading activity.

imbalances support the idea that past behavior plays an important role in the subsequent creation and redemption of ETF shares.

Model (4)-(6) repeat this analysis using *ETF Order Imbalance* as the dependent variable. The persistent, autoregressive pattern is also apparent in these regressions although there are some important differences when compared to *Net Create/Redeem Activity*. For example, a comparison of the parameter estimates for the first autoregressive variable shows that the lagged 1-day *ETF Order Imbalance* parameter is much higher in model (6), 0.105, than its corresponding lagged *Net Create/Redeem Activity* parameter in model (3), 0.0358. This result indicates that order imbalances are much more persistent than net creations, consistent with the discrete nature of net creation activity.

In contrast to the discrete nature of net creation activity, order imbalances are continuous in nature and can respond quickly to changes in the buying and selling demand of ETF investors. Thus, it is not surprising that we find in models (5) and (6) of Table 4 that today's ETF order imbalances are more positively autocorrelated with yesterday's order imbalances than the *Net Create/Redeem Activity* regressions reported in models (1)-(3). In addition, when lagged values of both net creations and order imbalances are included in model (6), there is evidence of an inverse relation between today's *ETF Order Imbalance* and lagged *Net Create/Redeem Activity* variables, as can be seen by the negative parameters for lagged values of net creations/redemptions from day t-6 to t-2. For example, the *Net Redeem/Create Activity* parameter at t-3 is the most significant and most negative (-0.00797) while the contemporaneous time-t parameter for this variable is 0.0404, thus suggesting that order imbalances are highest when APs' net creations are currently positive while prior net creations were negative over the past 2-7 trading days (i.e., the APs were experiencing net redemptions in the past, especially at time t-3). Taken together, the results reported in Table 4 for order imbalances and net creation activity show that order imbalances are more persistent than net creations, suggesting both the potential value of the option to delay and the exercise of that option, as seen by the discrete and discretionary behavior of APs when creating blocks of ETF shares.

4.4 The effects of ETF net creation activity and order imbalances on FTDs

Given the potential autoregressive and dynamic patterns outlined in the above discussion, it is also useful to examine the effect of order imbalances and net creations on ETF-related FTDs. We then regress FTDs and short interest level as a percentage of shares outstanding on lagged values of *ETF Order Imbalance* and *Net Create/Redeem Activity*, as well as the controls for ETF liquidity, including lagged FTDs. This can also help confirm our proposed AP trade motivations and Operational Shorting measure timing.

Panel A of Table 5 presents regression results for FTDs and Panel B contains the results of Short Interest Ratio regressions. By focusing on the full specification of model (6) in Table 5, Panels A and B, we can see that the lagged value of *ETF Order Imbalance* at t-3 has the largest and most signifcant positive coefficient when compared to all other variables in both the FTD regression (i.e., 0.121 with a t-statistic of 13.82) and the short interest regression (i.e. 0.0126 with a t-statistic of 6.18). Given that FTDs occur after time t+3, it is not that suprising that order imbalances from 3 days prior can have such a large impact on today's FTD metric. This result shows that large positive order imbalances (symptomatic of strong excess buying demand by ETF investors) can lead to higher operational shorting, which consequently shows up in higher short interest, and eventually higher FTDs. The finding is consistent with the idea that APs can provide liquidity in an excess buying situation by engaging in operational shorting activity. However, some of these operational short positions might not be covered within 3 days and thus can result in a surge in FTDs. This pattern is confirmed by the relatively large positive coefficient on the t-3 *ETF Order Imbalance* variable.

Model (6) of Panels A and B in Table 5 also shows an alternating pattern between lagged values of *Net Create/Redeem Activity* at days t-4 to t-1 and the current level of short interest and FTDs (at day t). For the shortest lag, net creations are positively related to short interest and FTDs at t-1 (0.0976) and could be driven by the "partial clean-up" of past operational short positions. In contrast, net creations at t-3 are negatively related to FTDs (-0.0715) and short interest (-0.0103). It is also noteworthy that the higher and more positive the *Net Create/Redeem* activity before t-3, the lower the the ETF short interest level in Panel B, consistent with the closing of operational short positions. Keep in mind that the short interest data are disseminated on a biweekly basis and are refreshed once every two weeks in our sample. The large variation in coefficients in the FTD regression for net creations over a few days is similar to the relation observed between net creations and order imbalances reported earlier in Table 4. Thus, the discretion that APs exhibit when making creation/redemption decisions in the recent past appears to correspond to not only current order imbalances but also the current level of FTDs. Further, Table 5 shows that the time period between t-3 and t-1 is the most important in terms of economic and statistical signficance. Consequently, we have formulated our definition of *Operational Shorting* in Equation (1) over this critical t-3 to t-1 period and then use this variable in the following section to analyze the key factors that explain variations in this type of shorting activity across ETFs.

4.5 Incentives for AP's Operational Shorting Activity

We next examine the determinants of operational shorting activity to confirm the incentives described in section 2 for APs to participipte in making markets for ETFs in Table 6. The variables included in this analysis relate to the liquidity of the ETF itself (*log(Market Cap), Average Share Turnover*), fixed costs associated with ETF share creation activity (*Creation Unit Dollar Size, Creation Unit Fee*), proxies for the ease of hedging the ETF and underlying (*Maximum Rolling R-Squared with Available Futures Contracts*, *Available Options Dummy*), the potential arbitrage profits available to the AP from share creation and redemption (*Mispricing*, *Premium*, *Discount*), and a proxy for the liquidity differential between the ETF and the underlying securities in the basket (*Proxy for Liquidity Mismatch*). Fixed effects for ETF and date are included and standard errors are clustered at the ETF and date levels. The number of observations decreases across the different specifications due to data availability issues[26].

In this analysis, we examine the hypothesis that operational shorting activity is driven by AP's incentives to capture arbitrage profits resulting from buy pressure on ETF shares and the subsequent mispricing of the ETF relative to the NAV or the fundamental value of the underlying basket. Arbitrage

[26] The *Creation Unit* variables are only available for a subset of ETFs. Similarly, the *Maximum Rolling R-Squared* can only be computed for those ETFs which state the underlying index they are tracking. The *Proxy for Liquidity Mismatch* variable can only be calculated for equity ETFs where both the underlying holdings data and the associated spread variable are available.

profits represent a primary incentive for APs to engage in providing liquidity in the ETF market. The *ETF Mispricing* variable is included because operational shorting activity could also be motivated by APs' and other investors' incentives to arbitrage differences between the ETF's market price and its NAV. While this measure captures situations where the ETF market price is at a premium (positive *Mispricing*) and a discount (negative *Mispricing*) relative to the NAV, operational shorting exists only when APs are statisfying buying demand pressure, suggesting an asymmetry in how deviations of ETF prices from the NAV are related to operational shorting incentives. To address this issue, we decompose ETF mispricing into two separate variables (i.e., a "premium" vs. a "discount") which are equal to the *absolute* value of the mispricing variable only when the mispricing is positive vs. negative, and zero otherwise. Because operational shorting is a strategic response to excess ETF buying pressure, this specification accounts for the asymmetry and we expect only premium mispricing (ETF price greater than the NAV) to relate positively.

Additionally, we expect that the availability of options to hedge the underlying is an important determinant of operational shorting. These hedges would shield APs who operationally short from unanticipated market swings. Market makers are likely to be more inclined to delay creation when the underlying basket stocks are less liquid until they are have a better gauge of the permanent component of the ETF order flow before committing to a basket create order and therefore incurring related trading costs. Our model includes multiple proxies for the availability of efficient hedges. The *Maximum Rolling R-squred with Available Futures Contracts* variable measures the how well available futures contracts correlate with the returns of the ETF's stated benchmark index. For each date, the previous 252 days of ETF NAV returns are regressed on the futures return from S&P 500-mini, the S&P MidCap 400-mini, and the Russell 2000-mini contracts.[27] The maximum R^2 across these three regressions is the value assigned to the *Maximum*

[27] The futures data are taken from the Quandl web site and the roll assumption used in constructing the daily futures returns is the 'last-trading-day' or 'end-to-end roll' method. This assumption "…allows you to use the front contract for as long as possible; however, the danger is that activity may have switched to the back contract prior to your roll. A trading strategy based upon this rule runs the risk of unwanted delivery and/or close-out of your positions, if you do not roll in time (the margin for error is very limited)."

Futures R-squared variable. If an AP or other investor wanted to hedge their exposure to an ETF, this R-squred variable serves as a proxy for the suitability of using futures on one of the three equity indexes as a reliable hedging vehicle.[28] Options listed on the ETF would also facilitate the hedging of ETF-specific risk. One would expect a positive relation between these hedging-related variables and *Operational Shorting* because the presence of such hedging instruments can allow an AP to provide more liquidity when they can use the futures and/or option markets to hedge this short position (e.g., via a long futures position or long call option).

We also include a proxy for *Liquidity Mismatch* between the ETF and its underlying basket of securities, to capture another incentive to delay creation. We follow Pang and Zeng (2016) and measure liquidity mismatch as the difference between the trade-weighted average intraday bid-ask spread of the ETF's underlying securities and the trade-weighted average bid-ask spread for the ETF. We expect that the option to delay creation by APs is more valuable when there is a greater mismatch between the liquidity of the basket of securities relative to the ETF, as APs would prefer to observe the ETF order flows in subsequent days before committing to gathering the less-liquid underlying basket of securities and incurring related transaction costs.

Our model also includes controls for ETF-specific transaction costs / frictions: *the natural log of Creation Unit Dollar Size* and *Creation Unit Fee*. As we discuss in Section 2.3, higher creation unit sizes and fees are expected to encourage APs to engage in more operational shorting in order to minimize these costs. Regression (2) in Table 6 includes four main independent variables: the ETF liquidity-related control variables as well as two proxies for ETF-specific transaction costs or frictions: *the natural log of Creation Unit Dollar Size*, and *Creation Unit Fee* (for a single creation unit transaction). We find in model (2) that

---

[28] As the example in Appendix A demonstrates, the use of a long position in a futures contract to hedge an AP's short position can be an effective way to lock in an arbitrage profit while providing time for any order imbalance to reverse so that the AP's costs to deliver the ETF shares are reduced. Thus, a strategy of operationally shorting first, then hedging in the futures market, and ultimately covering the short position at a later date, can be more profitable than immediately covering any short position with the creation of new ETF shares. This approach can also be accomplished using options on the ETF but would entail greater upfront costs to purchase a long call position (but also provide potentially greater profit potential).

the coefficients on *ln(Creation Unit Dollar Size)* and *Creation Unit Fee* are positive and statistically different from zero at the 1% level. In both cases, we find that the more costly it is to create or maintain ETF shares, the more likely it is that APs will turn to operational shorting, perhaps to wait for excess buying demand to subside and order flows to reverse. Alternatively, the APs could simply be buying time until they need to pay a relatively higher creation fee and save on the capital outlay required to accumulate the requisite shares in the underlying securities.

Regression (3) adds proxies for the ability to use futures and options markets to hedge a long or short exposure to an ETF (*Maximum Rolling R-Squared with Available Futures Contract* and *Available Options Dummy*). This model confirms the positive relation between operational short positions and the hedging proxies.

As expected, arbitrage profits are an important driver of AP's market making and related operational shorting activity. Positive values for *ETF Mispricing* mean that the ETF's market price is too high relative to the NAV and thus one would expect *more* operational shorting to bring these two values in line (and vice versa when this variable is negative). Model (7) confirms our expectation that greater premium coincides with increased operational shorting due to the highly significant positive parameter estimate on *ETF Mispricing* (0.370 with a t-statistic of 10.23). In addition, consistent with our prior that greater liquidity mismatches should be positively related to operational shorting, models (5), (6) and (7) show that *Proxy for Liquidity Mismatch* is positively and significantly related to the *Operational Shorting* dependent variable (although the relation is not as strong statistically as the *ETF Mispricing* variable).

Overall, the signs and significance of the variables in model (7) suggest that smaller, actively traded ETFs that have greater potential profits from capturing the ETF premium, with decent available hedging alternatives, and with larger liquidity mismatches engage in more operational shorting activity. Since arbitrage activity by APs is an important role in the proper functioning of the ETF market, it is important to examine how operational shorting influences the mispricing between ETF prices and NAVs. In addition, the liquidity mismatch between the underlying basket and the ETF could also lead APs to wait longer before covering their operational short positions. This could alleviate demands for liquidity in the underlying securities market while also increasing FTDs of ETF shares. Table 6's results are consistent with the numerical example in Appendix A, which formulates the trade-offs an AP faces when it decides to hedge its short position in order to wait for excess buying imbalances to reverse. Table 6's findings also lead us to explore the effects of operational shorting activity on ETF mispricing and the liquidity of the underlying securities which are held by ETFs.

4.6 ETF Operational Shorting, Creation and Short-Term ETF Returns

The academic literature has documented a strong negative relation between short-selling constraints and future stock returns. Whether the measure of short-selling constraints is: a) short interest (e.g. Figlewski (1981); Asquith and Meulbroek (1996); Desai, Ramesh, Thiagarajan, and Balachandran (2002)), b) short interest relative to institutional ownership (e.g. Asquith, Pathak, and Ritter (2005); Nagel (2005)), c) rebate rates (e.g. Jones and Lamont (2002)), d) rebate rates combined with the lendable supply of shares (e.g. Cohen, Diether, and Malloy (2007)), e) trade-level indications of a short-sale (e.g. Boehmer, Jones and Zhang (2008); Diether, Lee and Werner (2009)), or f) FTDs (Autore, Boulton, and Braga-Alves (2015)), the result is the same: constrained short-selling is associated with over-valuation. One common interpretation of this strong predictive relation is that short-sellers are informed, but constraints prevent them from fully incorporating their information in market prices.

While operational shorting is an important component of ETF short interest and FTDs, as we have shown, our earlier results suggest that the operational shorting activity of APs is centered on liquidity provision and not informed, speculative short-selling. To test whether operational shorting is directional/informed vs. operational/liquidity-providing, we compute the average cumulative ETF weekly returns (both total returns and Fama-French 4-factor risk-adjusted returns) and regress them on a number of ETF characteristics. The key independent variable in these regressions is *Operational Shorting %*. Table 7 reports the results of this test.

In regression (1) of Table 7, we see that ETF *Operational Shorting* is related contemporaneously to higher ETF returns but in regressions (2) and (3), however, we see that *Operational Shorting* is negatively related to future ETF total returns (marginally) and 4-factor alphas (significantly). While such evidence is consistent with an informed and directional motivation for operational shorting, in regressions (4) and (5) we repeat the analysis with the total return and 4-factor alpha of the underlying (i.e. calculated from NAV returns). While the operational shorting activity of APs has some predictive power for a return reversal in the ETF, is has no predictive power for the return on the underlying. This result demonstrates that APs are operationally short when reversals in the price of the ETF are forthcoming – an indication that operational shorting is a cost-effective way for APs to handle liquidity mismatches. In unreported results, we find that the predictive power of operational shorting on ETF returns does not persist for returns two weeks later, further suggesting that the effect is due to liquidity provision rather than "directional" short selling.

To better understand what drives the ETF predictability, we run a sub-sample analysis in regressions (6) through (9). In regressions (6) and (7), we separate out non-equity and equity ETFs respectively and see that the predictive power is much greater for non-equity ETFs where we would expect the ETF to be much more liquid than the underlying securities. Regressions (8) and (9), where we have sub-divided the equity ETF sample into 'Low' and 'High' liquidity mismatch ETFs, further emphasizes this point. We measure liquidity mismatch as the difference between the ETF's intraday spread and the average intraday spread of the basket of securities held by the ETF. In the 'Low' liquidity mismatch sample, where the ETF and underlying exhibit similar liquidity, we find no predictive power. In the 'High' liquidity mismatch subsample, where the ETF is more liquid than the underlying, however, we find a statistically significant, negative relationship between operation shorting and future returns.

Taken together, these results indicate that the relation between AP behavior and future returns is not rooted in changes to the values of the underlying securities, but rather to the liquidity and supply of the shares of the ETF itself. These results are consistent with Kyle's concept of strategic market making where APs actively create new ETF units and these liquidity providers could be observing the order flow and using these patterns to identify when informed traders are most likely to be active in the market. The market makers could then be taking action in both the ETF and the underlying securities that ultimately reveals this private information in current prices, thus providing not only profitable trades for the APs but also making the market for these securities more informationally efficient. In addition, when APs are not actively creating new ETF units, it appears that they are providing liquidity rather than speculating on the future direction of ETF prices. We infer this last point because we observe that the NAV returns, which are determined by trades in the underlying basket of securities rather than by APs' ETF trades, do not exhibit any significant pattern relationship with operational shorting activity. To sum up, APs and market makers are more likely to operationally short in anticipation of reversals while creations are more likely to happen when there is a positive trend in returns. This effect is more pronounced when the liquidity mismatch is higher (for non-equity ETFs or for less-liquid equity ETFs). And by providing liquidity, the APs / MMs are acting as a buffer in the underlying cash market (and thus there is no effect on the NAV returns).

4.7 The effects of Operational Shorting on an ETF's Underlying Securities

4.7.1 The effects of Operational Shorting on ETF Mispricing

While Reg SHO enables failing-to-deliver by APs/market makers because it extends the settlement date until T+6, this exception is only allowed for "bona fide" market making activity, suggestive that the FTDs from operational shorting might be an acceptable outcome if ETF liquidity and pricing efficiency is improved. In this section, we examine the impact of operational shorting on measures of ETF mispricing as well as the informational efficiency and liquidity of the underlying basket as a test of whether or not the purpose behind this exception is being achieved.

In Table 8, we examine the impact on ETF mispricing, a measure of both the potential arbitrage profits for APs and other investors. In order to gauge the effect of operational shorting on mispricing, we use the change in ETF mispricing as the dependent variable in specifications 1 through 4, measured by the difference between the ETF market price and NAV as a percentage of the ETF price. In specifications 5 through 8, we also use *Absolute Mispricing Change*, which allows us to test whether operational shorting activities that are aimed to harvest the mispricing arbitrage opportunities are associated with a decrease in the magnitude of mispricing after such arbitrage takes place. We also include controls for ETF liquidity and hedging alternatives.

After controlling for the significant negative coefficients on ETF fund size/market capitalization, model (2) of Table 8 shows that the contemporaneous level of *Operational Shorting* has a strong negative relation with the (signed) ETF mispricing variable while model (3) also confirms that the lagged *Operational Shorting* variable leads to a reduction in ETF mispricing. Such a result is in line with our prior finding that operational shorting are incentivized by arbitrage opportunities in the form of ETF mispricing, and are now found to eventually take advantage and reduce these mispricing oportunities (consistent with recent papers on ETF arbitrage such as Brown, Davies and Ringgenberg (2017)). Models (5) and (7) repeat these tests with the *Absolute Mispricing Change* and confirm the significant negative relation between operational shorting and ETF mispricing, suggesting that operational shorting is a manifestation of the arbitrage that takes advantage of and eliminates mispricing opportunities.

4.7.2 The effects of Operational Shorting on an ETF's Underlying Securities

The decreased mispricing associated with operational shorting is consistent with APs using this exception to engage in "bona fide" market making activities that is incentivized by mispricing arbitrage opportunities. To directly explore such "bona fide" market making activities, we examine the impact of operational shorting on the liquidity of the underlying securities held by ETFs.[29] To do this, we follow Fotak et al. (2014), and use average spreads and intraday volatility as measures of the liquidity and market quality of trading in individual securities. Our first measure, the *Average Intraday Spread* of the underlying stocks is computed in two steps: first, for each stock and in each day, we compute the intraday spread by weighting each intraday NBBO spread by the size of the trade immediately following the NBBO quote. Then, we aggregate this measure across all stocks held by the ETF using the ETF's portfolio weights. Our second measure, the *Average Intraday Volatility* represents the second-by-second return volatility that is calculated from the last traded price recorded during each second of the trading day, and they are aggregated at the ETF level similar to the spread measure. The results are shown in Table 9.[30]

[29] As noted in Fotak et al. (2014), Stratmann and Welborn (2013), and others, FTDs for common stocks can have an impact on the liquidity of these equity securities. Here we explore if operational shorting, through a similar mechanism, can have the same effect.

[30] In Table 9, our sample is restricted to U.S. equity-only ETFs because these are the only funds that we can reliably

The basic structure of the empirical model is similar to the one used in Table 8 and includes contemporaneous and lagged measures of *Operational Shorting*, which are our main variables of interest. Since our measure of operational shorting activity is computed at the ETF level, we run all of our analysis of liquidity and volatility effects using these ETF-level liquidity measures, as well as the underlying stock measures aggregated at the stock level. Additionally, we control for the "liquidity level" effect of ETF ownership that is documented in earlier literature. In particular, Ben-David, Franzoni, and Moussawi (2015) document causal evidence that links ETF ownership with increased volatility of underlying securities. To control for this effect, we construct and include a measure of average ownership by ETF in the underlying basket stocks, using all the stocks in the ETF basket at the end of the previous month. We also explicitly control for the lagged dependent variable and for the ETF-based liquidity measures by including up to three lags of these ETF liquidity measures to control for the persistence in volatility and spread measures, as well as address potential reverse causality concerns.

Panel A of Table 9 reports the results of regressions (1)-(7) with underlying stocks' average bid-ask spread as the dependent variable. We also control for the contemporaneous and lagged forms of the ETF's intraday spread, as well as the lagged intraday spread of the underlying stocks held by the ETF. We find that operational shorting is negatively related to the underlying stocks' average spread, thus coinciding with an improvement in liquidity for these stocks. Consistent with prior literature, we find that increased ETF ownership in basket stocks is associated with higher spreads in these underlying stocks, as prior evidence suggests there can be a migration by liquidity-demanding investors from the underlying securities to the more-liquid ETF securities (e.g., see Dannhauser, 2017; Hamm, 2014; Israeli, Lee, and Sridharan, 2017).[31]

---

identify the holdings in each of the underlying stocks from Thomson Reuters. This sub-sample also facilitates our estimates of the national best bid and offer (NBBO) bid-ask spreads for both the ETFs as well as their underlying holdings. This reduces our sample sizes to around 800,000 ETF-day observations (compared to over 2.5 million in earlier tables).

[31] We limit the regression results to the lagged operational shorting levels for cleaner and more rigorous identification, despite the fact that the improvement to underlying stock liquidity is the strongest on the day the operational shorting is initiated.

In Panel B of Table 9, a similar set of regressions are performed on the intraday volatility of the ETF underlying stocks. We find that operational shorting is also negatively related to intraday return volatility after controlling for the ETF market capitalization and trading volume. Consistent with prior literature, we also find that the level of ETF ownership is positively associated with the average underlying stock volatilities. This effect could be due to increased exposure to high frequency traders and other liquidity demanders that transmit their liquidity shocks to the ETF and ultimately to the ETF's underlying basket (Ben-David et al, 2015). We interpret our findings that when APs do not engage in operational shorting and decide to physically create new units of ETF shares immediately, then these APs will buy shares of the underlying stocks and transmit liquidity shocks to the underlying securities. These shocks, related to the creation activity, in turn, can worsen the liquidity in the underlying stocks.

The negative relation between operational shorting and intraday spread and volatility of underlying stocks, confirms that liquidity in the underlying stocks held by ETFs improves as operational shorting increases. An AP's operational shorting activity can thus have an overall beneficial effect on the market for the underlying stocks. This is consistent with operational shorting acting as a "release valve" that improves liquidity (Fotak et al. (2014)). As operational shorting increases due to a sudden surge in ETF demand, the AP provides liquidity in the ETF market *without* entering the market for the underlying stocks. Through operational shorting, an AP in the ETF acts as a buffer that does not immediately transmit the liquidity shocks that hit ETFs to the underlying basket, thus cushioning the underlying stocks from higher volatility or widening spreads. If, on the other hand, the AP does not engage in operational shorting and decides to create new units of ETF shares immediately, then the AP will have to buy shares of the underlying stocks and transmit those liquidity shocks directly to the underlying securities. This can perturb the market for these underlying stocks, especially if this market is less liquid than the market for ETF shares. Thus, operational shorting at the ETF level can improve liquidity for the underlying stocks by enabling APs to delay transient trades in the fund's basket of less-liquid securities until future ETF order flows are observed.

Panle C of Table 9 examines another aspect of an ETF's impact on a fund's underlying securities: the informational efficiency of securities prices. As Huang, O'Hara, and Zhong (2018), Xu, Yin, and Zhao

(2018), and others have noted, informed traders in the ETF and underlying securities can influence prices and order flow which, in turn, can be used by market makers to learn more about the value of these securities. By providing more liquidity to the market in response to this informed order flow, APs can help establish more efficient prices for both the ETF and underlying securities.

One common way to measure price efficiency is by computing the variance ratio of the intraday returns of the underlying securities over short time intervals such as 5 and 15 seconds. Ideally, the 15-second return variance should be 3 times the 5-second return variance, which would result in variance ratio of 1.0 if the trading in these securities was perfectly efficient. Similar to Mao and O'Hara (2011), we compute deviations from this perfect value of 1.0 by taking the absolute difference between 1 and the observed variance ratio. Thus, values that deviate greatly above or below 1.0 will denote larger deviations from perfectly efficient markets. Panel C of Table 9 reports the results of regressing the operational shorting measure lagged one day on our modified version of the 15 second-to-5 second variance ratio (calculated over all 15-second intervals on a daily basis). Additional control variables are also included such as the current ETF ownership of the underlying stock, the ETF's market capitalization and trading volume, as well as lagged values of the ETF's second-by-second return volatility.

We find that operational shorting has a consistently negative effect on our variance ratio measure. This shows that greater operational shorting activity coincides with decreases in deviations of the variance ratio from its ideal level of 1.0. The results shown in this panel confirm that price efficiency in the udnerlying securities held by ETFs improves as markets engage in higher levels of operational shorting activity. Consistent with the results in Table 7 related to when APs actively create new ETF units, it appears that ETF liquidity providers could be observing the order flow and using these patterns to identify when informed traders are most likely to be active in the market. In turn, the market makers could be taking action in both the ETF and the underlying securities that reveals this private information in current prices, thus making the market for these securities more informationally efficient.

Overall, our evidence suggests that operational shorting can dampen the potentially adverse effects of ETFs on the volatility and liquidity of underlying stocks in their baskets as well as enhance these stocks' price efficiency, which is consistent with our earlier results, as it represents additional incentive for APs to delay covering these short positions. Operational shorting thus appears to act as a buffer that reduces the effects of liquidity shocks that ETFs are receiving from their clients' orders, which is consistent with the notion that operational shorting is a potentially beneficial by-product of liquidity provision by these market makers / APs.

4.8 Is it All that Beneficial? Operational Shorting and Financial Contagion

While operational shorting does seem to decrease ETF mispricing and positively affect the liquidity of the underlying securities, it also results in increased FTDs. One concern regarding FTDs is that they have the potential to increase the systemic risk of financial markets. In testimony before the Senate Committee on Banking, Housing, and Urban Affairs, Harold Bradley and Robert E. Litan, the CIO and Vice-President for Research and Policy of the Kauffman Foundation, respectively, discussed these risks:

> "Market-makers enjoy significant and historically arcane exemptions from rules applying to trading and settlement that extend to all other market participants—we worry these special privileges may lead to high levels of trading "fails" and greater systemic risks to the overall market. Such trading "fails" in ETFs during times of market stress could domino into a greater systemic risk issue for our markets."[32]

In this section, we explore the relation between ETF FTDs, operational shorting, financial contagion, and leverage.

4.8.1 Potential Spillover effects on FTDs and Operational Shorting activity

One possible effect of operational shorting is an increase in counterparty risk and financial contagion. Given that FTDs represent a delay in payment from one party to another, increases in the overall volume of FTDs might lead to the impairment of some market participants' ability to meet their other obligations in a timely way, thus leading to greater counterparty risk. More specifically, it may be

[32] Bradley, Harold and Robert E. Litan, 2009, ETFs and the Present Danger to Capital Formation, Prepared Testimony Before the United States Senate Committee on Banking, Housing, and Urban Affairs, Kauffman Foundation.

the case that when an AP delays share creation in one ETF due to a lack of liquidity, for example, the AP also delays share creation in other ETFs for which the AP is also a market maker. Similarly, problems at one AP may spill over to other APs that are trading in the same or similar ETFs, thus creating a market-wide contagion effect across several APs and other ETFs they trade. In such a scenario, widespread delays in ETF share creation or, equivalently, operational shorting surges, could spread across the market of APs and ETFs.

We use as dependent variables the FTDs at the individual ETF level as well as our *Operational Shorting* variable. Our key independent variables are the *FTD* and *Operational Shorting* variables aggregated at the lead market maker level. For each of these variables, we sum the lead market maker's activity (for both FTDs and operational shorting) on a specific day over all ETFs in which it trades, but excluding the FTDs or operational shorts in the ETF of interest. These varibles are denoted *Lead Market Maker FTDs* and *Lead Market Maker Operational Shorts*, respectively. In order to test for broader commonality effects associated with FTDs and operational shorting, we also compute these FTD and operational shorting variables for the entire market of ETFs. We do this by summing all FTDs and operational shorts on a given day for all ETFs while excluding the individual ETF's FTDs and operational shorts on that day. These are denoted *Overall Market FTDs* and *Overall Market Operational Shorts*, respectively. For completeness, we also include the liqudity-related control variables and the proxies for futures and options hedging vehicles. If there are contagion effects between APs and other ETF market makers, then we expect the dependent variables to be positively related to the Lead Market Maker's FTDs and Operational Shorting activity.[33]

In Table 10, we find that the coefficients on FTDs generated in other ETFs by the same lead market maker in regressions (1) and (2) are positive and statistically different from zero. Further, we find a positive and significant association with aggregate, market-wide FTDs in other ETFs from all other APs

[33] In these tests, we only include fixed effects for each ETF because the market-wide measures of FTDs and Operational Shorting are the aggregated values across all ETFs on a given day, after excluding the individual ETF values. Thus, we omit date fixed effects in order to exploit the time varying element of both the market-wide levels and the lead market maker levels in our regression.

(net of the FTDs by the lead market maker affiliated with each ETF). This positive relation between individual ETF FTDs and overall market-wide FTDs suggests that the impact of other non-affiliated market maker's FTD activity is similar to the contagion effects of a lead market maker's FTDs. These regressions also show that the liquidity-related control variables are consistent with earlier tests and the hedging-related variables are negatively related to FTDs. This latter finding is consistent with the notion that investors are more likely to use futures and/or options to engage in directional short-selling, thus causing an ETF's FTDs to be lower because these FTDs have been shown in our analysis to be more highly correlated with liquidity provision activities (which are non-directional in nature).

Regression (3) and (4) repeat these tests using individual ETF-level *Operational Shorting* as the dependent variable. Similar to the results for FTDs, we find that a lead market maker's other operational shorts are positively related to an individual ETF's operational shorting activity and the *Overall Market-wide Operational Shorts* also exhibit a positive relation with an individual ETF's operational shorting. These findings suggest there is a commonality in AP liquidity provision across the AP's portfolio of ETFs and therefore FTDs and operational shorting may serve as a contagion mechanism in ETF markets. Additionally, the activity by non-affiliated APs other than the lead market maker also appear to play a reinforcing role in exacerbating some of these spillover effects. As in regressions (1) and (2), the liquidity control variables are consistent with our earlier findings. Regressions (3) and (4) also indicate that the hedging-related variables are positively related to operational shorting, as the presence of hedging vehicles can encourage APs to engage in more operational shorting. This latter result is also consistent with our earlier findings.

4.8.2 Market Maker Leverage as a Potential Channel for Spillover effects

As a final exploration into the possible mechanisms for the relationship between ETF FTDs/operational shorting and financial contagion effects, we examine the financial leverage of APs as one of the channels by which FTDs and operational shorting can affect financial system stress. The leverage of financial institutions is a primary concern for regulators because excessively high leverage can jeopardize stability in the financial system. We gather capital positions for many of the lead market makers in our sample from the CFTC's Futures Commission Merchants Financial Reports.[34] We measure a broker's leverage by examining the firm's capital constraints. This is done by computing the following ratio:

$$Capital\ \ Constraints = \frac{Net\ \ Capital\ \ Required}{Adjusted\ \ Net\ \ Capital} \tag{3}$$

This ratio is closely monitored by the CFTC because it measures how much capital a market maker holds (the denominator) relative to the CFTC's net capital requirements (the numerator). Also, this ratio is bounded between 0 and 1, where the ratio for more capital-constrained (and more highly leveraged) market makers will approach 1. Thus, the variable is increasing in financial leverage. We then use this variable as an independent variable in a regression setting to estimate the effect of financial leverage on FTDs and operational shorting.

Table 11 displays the results of this test. Regresions (1)-(3) examine FTDs and Regressions (4)-(6) examine operational shorting. Across all specifications, we find that the lead market maker's leverage is positively related to both FTDs and operational shorting at the individual ETF level. This relation is statistically significant at the 1% level. We do not find a similar effect for market-wide leverage. The positive relation between market maker leverage and our key variables of interest, ETF FTDs and operational shorting, is evidence of one of the potential channels by which FTDs and operational shorting at the ETF level can increase counterparty risk and financial contagion. Thus, individual APs that follow a business strategy of economizing on both trading costs (via operational shorting) and capital investment (via higher leverage) might collectively impose a significant negative externality in terms of increased system-wide financial stress through the inter-connected nature of AP-led liquidity provision.

5. Conclusion

This study is the first comprehensive analysis of short interest and failures-to-deliver (FTDs) of Exchange-Traded Funds (ETFs) and how the liquidity provision activities of APs for ETFs can have important effects on short-selling, FTDs, market quality, ETF returns, as well as counterparty risk and

[34] http://www.cftc.gov/MarketReports/financialfcmdata/HistoricalFCMReports/index.htm

potential market-wide contagion. We propose, test, and find evidence that ETF short interest and FTDs are related to "operational shorting," a manifestation of market making efforts when APs and market makers are faced with unexpectedly large excess buying demand from investors. This behavior can be explained by the dynamics of daily order imbalances and frictions associated with the creation of new ETF units.

Consistent with the economics underlying the proposed explanation, we find that operational shorting is associated with increases in ETF short interest and FTDs, and that operational shortng is negatively predictive of future 1-week ETF returns, specifically when the ETF has a high liquidity mismatch with its underlying basket of securities. Our novel measure of operational shorting can help disentangle the effects of liquidity-induced short-selling from informed / directional short-selling and thus has important implications for extant theories and research on short-selling. Additionally, we document that share creation is observed in the changes of ETF shares outstanding several days after the true flows appear in the ETF trade imbalance. This finding is consistent with the existence of a valuable "option to wait," where the AP has an incentive to delay the delivery of ETF shares in order to maximize its profit from liquidity provision and arbitrage activities (while also increasing the level of FTDs in the financial system).

Although these liquidity provision activities can be profitable to APs and other market makers, they might create a negative externality in terms of greater levels of FTDs that spill over from one ETF to another with the same AP, or from one AP to another, because they make markets for ETFs with similar underlying securities. We find that increased levels of FTDs and operational shorting at the lead market maker's other ETFs can have a contagion-like effect. In addition, the FTDs and operational shorting activity aggregated across the overall market are positively related to an individual fund's FTDs and operational short-selling activity.

These results suggest ETF trading relies on an inter-connected network of liquidity providers which, at times, pursue positively correlated trading strategies that can be detrimental to the overall market. We also identify a potential channel by which market makers' financial leverage can reinforce the effects of operational shorting on other market participants. Thus, the unique structure of ETFs and the AP's ability to sell shares on an intraday basis that have not been created (or at least the underlying basket of securities has not been delivered) can lead to improvements to trading in individual ETFs but, at an aggregate level, it can create greater counterparty risk that has potentially destabilizing effects for not only ETFs but also the underlying securities held by these ETFs.

Given that the above results have focused on operational shorting, one possible avenue for future research pertains to examining the asymmetry of AP behavior when there is excess selling pressure from ETF investors. In this alternative situation, the AP could provide liquidity by engaging in "operational buying" of the (relatively) cheap ETF shares and potentially minimizing the cost of this activity by redeeming ETF shares quickly in order to receive the underlying basket of (more-valuable) securities. However, it is unclear how quickly operational buy positions are covered relative to operational short positions. Thus, additional research into this asymmetry between and operational buying vs. operational shorting is warranted.

References

Antoniewicz, R., and J. Heinrichs, 2014, Understanding Exchange-Traded Funds: How ETFs work, *ICI Research Perspective* 20(5), 1-39.

Antoniewicz, R., and J. Heinrichs, 2015, The Role and Activities of Authorized Participants of Exchange-Traded Funds, *ICI Research Report* March, 1-13.

Asquith, P., and L. Meulbroek, 1996, An empirical investigation of short interest, Working paper.

Asquith, P., P. Pathak, and J. Ritter, 2005, Short interest, institutional ownership, and stock returns, *Journal of Financial Economics* 78, 243–276.

Autore, D., T. Boulton, and M. Braga-Alves, 2015, Failures to deliver, short sale constraints, and stock overvaluation, *Financial Review* 50, 143-172.

Bhattacharya, A., and M. O'Hara, 2016, Can ETFs Increase Market Fragility? Effect of Information Linkages in ETF Markets, Working paper.

Battalio, R., and P. Schultz, 2011, Regulatory uncertainty and market liquidity: The 2008 short sale ban's impact on equity option markets, *Journal of Finance* 66, 2013-2053.

Ben-David, I., F. Franzoni, and R. Moussawi, 2015, Do ETFs Increase Volatility? *Journal of Finance*, forthcoming, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1967599.

Ben-David, I., F. Franzoni, and R. Moussawi, 2017, Exchange Traded Funds (ETFs), *The Annual Review of Financial Economics*, 9 (6), November 2017.

Boehme, R., B. Danielsen, and S. Sorescu, 2006, Short sale constraints, differences of opinion, and overvaluation, *Journal of Financial and Quantitative Studies* 41, 455–487.

Boehmer, E., C. Jones, and X. Zhang, 2008, Which Shorts Are Informed? *The Journal of Finance* 63, 491–527.

Boni, L., 2006, Strategic delivery failures in U.S. equity markets, *Journal of Financial Markets* 13, 397-421.

Brown, D., S. Davies, and M. Ringgenberg, 2017, ETF Arbitrage and Return Predictability, Working paper, University of Arizona.

Chu, Y., D. Hirshleifer, L. Ma, 2016, The causal effect of limits to arbitrage on asset pricing anomalies, Working paper.

Chinco, A., and V. Fos, 2016, The Sound of Many Funds Rebalancing, Working paper, University of Illinois at Urbana-Champagne.

Cohen, L., K. Diether, and C. Malloy, 2007, Supply and demand shifts in the shorting market, *Journal of Finance* 62, 2061–2096.

Da, Z., and S. Shive, 2014, Exchange Traded Funds and Asset Return Correlations, Working paper, Notre

Dame University.

Dannhauser, C., 2017, The Impact of Innovation: Evidence from Corporate Bond ETFs, *Journal of Financial Economics* 125, 537-560.

Desai, H., K. Ramesh, S. Thiagarajan, and B. Balachandran, 2002, An Investigation of the Informational Role of Short Interest in the Nasdaq Market, *Journal of Finance* 57, 2263-2287.

Diether, K., K. Lee, I. Werner, 2009, Short-sale strategies and return predictability, *Review of Financial Studies* 22, 576-607.

Ellis, K., Michaely, R., O'Hara, M., 2001, The accuracy of trade classification rules: Evidence from Nasdaq, *Journal of Financial and Quantitative Analysis* 35, 529-551.

Evans, R., C. Geczy, D. Musto, Reed, A., 2009, Failure is an option: Impediments to short selling and option prices, *Review of Financial Studies* 22, 1955–1980.

Figlewski, S., 1981, The informational effects of restrictions of short sales: Some empirical evidence, *Journal of Financial and Quantitative Analysis* 16, 463-476.

Fotak, V., V. Raman, and P. Yadav, 2014, Fails-to-deliver, short selling, and market quality, *Journal of Financial Economics* 114, 493–516.

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

Israeli, D., C. Lee, and S. Sridharan, 2017, Is there a Dark Side to Exchange Traded Funds (ETFs)? An Information Perspective, *Review of Accounting Studies* 22, 1048–1083.

Jain, A., and C. Jain, 2015, Fails-to-Deliver before and after the implementation of Rule 203 and Rule 204, *Financial Review* 50, 611-636.

Kyle, A., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315-1335.

Madhavan, A., 2012, Exchange-Traded Funds, market structure, and the flash crash, *Financial Analysts Journal* 68(4), 20-35.

Madhavan, A., 2014, Exchange-Traded Funds: An overview of institutions, trading, and impacts, *Annual Review of Financial Economics* 6, 311-341.

O'Hara, M., & Ye, M., 2011, Is market fragmentation harming market quality? *Journal of Financial Economics* 100, 459-474.

Malamud, S., 2015, A Dynamic Equilibrium Model of ETFs, Working paper.

Nagel, S., 2005, Short sales, institutional investors, and the cross-section of stock returns, *Journal of Financial Economics* 78, 277–309.

Nutz, M., and J. Scheinkman, 2017, Supply and shorting in speculative markets, Working paper, Princeton University.

Ofek E., and M. Richardson, 2003, DotCom mania: The rise and fall of Internet stock prices, *Journal of*

Finance 58, 1113-1138.

Pan, K., and Y. Zeng, 2016, ETF arbitrage under liquidity mismatch, Working paper, University of Washington.

Rapach, D., M. Ringgenberg, and G. Zhou, 2016, Short interest and aggregate stock returns, *Journal of Financial Economics* 121, 46-65.

Stratmann, T., and J. Welborn, 2013, The option market maker exception to SEC Regulation SHO, *Journal of Financial Markets* 16, 195–226.

Stratmann, T., and J. Welborn, 2016, Informed short selling, fails-to-deliver, and abnormal returns, *Journal of Empirical Finance* 38, 81–102.

Sullivan, R., and J. Xiong, 2012, How Index Trading Increases Market Vulnerability, *Financial Analysts Journal* 68(2), 70–84.

Zhang, T., and T. Judge, 2016, Investment Analysis of Leveraged ETFs, Working paper.

Table 1 – Summary Statistics: This table presents summary statistics for key variables used in our analysis. The sample period is March, 22 2004 – December 31, 2016. We provide summary statistics for both the entire ETF sample and the subsample of US equity ETFs. Listed below are the mean, standard deviation, and the 1st (p1), 25th (p25), 50th (p50), 75th (p75) and 99th (p99) percentiles of the distribution of each varable. A complete list of variable names, sources, and definitions is provided in Appendix B.

	Variable	Obs	Mean	Std.Dev.	p1	p25	p50	p75	p99
Entire ETF Sample	Fail-to-Deliver Shares / Shares Outstanding	3,007,239	0.42%	1.53%	0.00%	0.00%	0.00%	0.11%	11.45%
	Operational Shorting, as % of Shares Outstanding	3,006,555	1.01%	2.89%	0.00%	0.00%	0.00%	0.65%	20.83%
	Net Create/Redeem Activity: log (1 + % change in Shares Outstanding)	3,006,045	0.11%	1.37%	-5.72%	0.00%	0.00%	0.00%	8.82%
	ETF Order Imbalance: (Buys - Sells) / Average Shares Outstanding	2,772,648	0.15%	1.81%	-7.15%	-0.15%	0.03%	0.29%	10.63%
	Market Capitalization, $ million	3,007,054	$867.19	$2,600.87	$1.38	$16.81	$86.20	$427.69	$18,523.09
	Daily Share Turnover, % of Shares Outstanding	2,950,760	4.0%	8.6%	0.1%	0.6%	1.2%	2.8%	55.5%
	Amihud Illiquidity Measure	2,756,643	0.11	0.37	0.00	0.00	0.00	0.04	2.59
	% Mispricing: % difference between ETF price and NAV	2,912,330	0.029%	0.572%	-2.332%	-0.118%	0.016%	0.184%	2.115%
	Maximum Rolling R-Squared with Available Futures Contracts	2,673,729	53%	29%	0%	30%	59%	77%	96%
	Available Options Dummy	3,007,239	0.31	0.46	0.00	0.00	0.00	1.00	1.00
	Creation Unit Size	931,999	69,602	35,005	25,000	50,000	50,000	100,000	250,000
	Creation Unit Fee	931,999	$1,577.56	$2,664.75	$100.00	$500.00	$500.00	$1,400.00	$15,000.00
	Bid-Ask Spread, at Close	2,956,434	0.330%	0.542%	0.011%	0.067%	0.147%	0.339%	3.544%
	Intraday NBBO Bid-Ask Spread, Trade Size Weighted	2,772,053	0.269%	0.395%	0.012%	0.064%	0.135%	0.288%	2.470%
	Intraday Volatility, using second-by-second intraday returns	2,703,755	0.0083%	0.0083%	0.0000%	0.0037%	0.0061%	0.0100%	0.0511%
	Daily Cost of Borrow Score	1,768,565	3.19	1.47	1.00	2.00	3.00	4.00	7.00
	Indicative Fee	1,588,220	4.37%	3.44%	0.38%	1.75%	3.50%	6.00%	18.00%
	Short Interest Ratio	2,946,535	4.66%	11.84%	0.00%	0.28%	0.90%	3.20%	83.76%
US Equity ETF Sample	Intraday NBBO Bid-Ask Spread, Trade Size Weighted	856,148	0.1785%	0.2895%	0.0117%	0.0485%	0.0963%	0.1843%	2.4695%
	Intraday Volatility, using second-by-second intraday returns	847,880	0.0086%	0.0076%	0.0000%	0.0045%	0.0065%	0.0099%	0.0511%
	Daily Cost of Borrow Score	571,069	2.63	1.26	1.00	2.00	3.00	3.00	7.00
	Indicative Fee	525,851	3.134%	2.596%	0.375%	1.125%	2.500%	4.000%	18.000%
	Short Interest Ratio	863,150	6.40%	15.24%	0.00%	0.30%	0.97%	3.98%	83.76%
	Average Intraday NBBO Bid-Ask Spread for Underlying Basket Stocks	866,921	0.1036%	0.1176%	0.0263%	0.0454%	0.0683%	0.1122%	0.8835%
	Average Intraday Volatility of Underlying Basket Stocks	866,897	0.0195%	0.0102%	0.0084%	0.0130%	0.0164%	0.0223%	0.0668%
	Underlying Basket Stocks, Average Daily Cost of Borrow Score	813,294	1.12	0.24	1.00	1.00	1.03	1.12	2.564050
	Underlying Basket Stocks, Average Indicative Fee	866,921	0.572%	0.475%	0.281%	0.386%	0.430%	0.544%	3.688%
	Underlying Basket Stocks, Average Short Interest Ratio	866,921	4.15%	2.40%	0.81%	2.34%	3.51%	5.41%	12.63%

Table 2 – Failures-to-Deliver (FTDs) Summary Statistics: This table presents summary statistics specifically for Failures-to-Deliver (FTDs). Panel A reports the average daily dollar volume of FTDs, and Panel B reports the average daily FTDs as a percentage of shares outstanding. Both panels report figures by asset class. Panel B reports the statistics only for securities that we were able to identify in CRSP, Compustat, and Mergent FISD databases. The sample period is March, 22 2004 – December 31, 2016

A. Average Daily Fail-To-Deliver Dollar Volume, by Asset Classes, $ million

Year	Total Dollar FTD	ETF	Common Stock	OTC Stocks	Corporate Bond	ADR	Structured Products	Units and Trusts	Other Securities	# of Securities with Positive FTD
2004	$3,439.9	$936.0	$2,103.8	$36.7	$35.9	$212.7	$21.2	$102.6	$2.8	2,739
2005	$3,011.3	$974.4	$1,691.4	$43.2	$25.5	$201.1	$14.6	$65.4	$0.3	2,488
2006	$3,443.6	$994.1	$2,040.2	$42.6	$88.7	$211.1	$19.7	$50.7	$1.2	2,639
2007	$7,129.6	$2,540.9	$3,520.4	$50.5	$451.3	$359.4	$40.9	$57.5	$117.1	2,937
2008	$6,401.6	$1,887.7	$3,931.2	$47.2	$45.8	$342.6	$66.1	$46.7	$44.2	4,545
2009	$1,430.0	$866.4	$402.0	$10.3	$15.9	$91.7	$25.4	$13.0	$10.6	6,465
2010	$1,953.3	$1,272.4	$495.0	$14.9	$13.9	$114.1	$20.2	$15.7	$12.4	6,265
2011	$2,479.4	$1,705.2	$543.1	$16.9	$15.5	$142.3	$30.8	$15.5	$19.2	6,109
2012	$1,877.0	$1,183.7	$509.0	$11.3	$20.5	$99.3	$23.8	$20.8	$18.3	5,731
2013	$2,065.3	$1,313.6	$552.4	$10.4	$20.1	$106.7	$29.2	$24.4	$17.6	5,588
2014	$2,704.9	$1,734.0	$746.4	$11.8	$20.0	$137.3	$36.3	$14.7	$12.0	6,074
2015	$3,460.1	$2,506.3	$734.2	$9.1	$15.1	$137.6	$37.4	$11.2	$15.9	6,190
2016	$3,304.1	$2,592.5	$522.1	$8.2	$10.3	$122.0	$32.1	$14.5	$7.0	5,951

B. Average Daily Fail-To-Deliver % of Shares Outstanding, As Percent of Security Shares Outstanding, by Asset Classes

Year	Total FTD, % of Shares Outstanding	ETF	Common Stock	OTC Stock	Corporate Bond	ADR	Structured Products	Units and Trusts	Other Securities	# of Securities with Positive FTD
2004	0.83%	3.94%	0.63%	1.12%	1.29%	1.01%	1.49%	0.47%	1.57%	1,943
2005	0.57%	2.40%	0.39%	1.02%	0.78%	0.63%	0.65%	0.27%	0.58%	1,756
2006	0.73%	3.35%	0.33%	1.72%	1.05%	0.49%	0.48%	0.20%	1.42%	1,834
2007	0.99%	5.24%	0.37%	2.01%	1.01%	0.46%	0.55%	0.22%	0.82%	2,124
2008	0.82%	4.05%	0.31%	1.66%	0.32%	0.23%	0.97%	0.14%	0.45%	3,507
2009	0.22%	0.85%	0.03%	1.20%	0.05%	0.03%	0.21%	0.02%	0.03%	5,400
2010	0.18%	1.02%	0.03%	0.61%	0.09%	0.02%	0.17%	0.02%	0.00%	5,373
2011	0.23%	1.15%	0.04%	0.53%	0.07%	0.04%	0.33%	0.02%	0.00%	5,216
2012	0.17%	0.87%	0.03%	0.28%	0.07%	0.03%	0.24%	0.02%	0.00%	5,185
2013	0.23%	1.10%	0.03%	0.14%	0.05%	0.11%	0.27%	0.02%	0.00%	5,061
2014	0.17%	0.80%	0.03%	0.18%	0.04%	0.06%	0.31%	0.01%	0.00%	5,553
2015	0.17%	0.68%	0.02%	0.34%	0.03%	0.08%	0.31%	0.01%	0.00%	5,664
2016	0.18%	0.83%	0.02%	0.31%	0.02%	0.02%	0.14%	0.01%	0.00%	5,504

Table 3 – The Deterimants of ETF Short Interest and Failures-to-Deliver: This table displays Ordinary Least Squares (OLS) regression results. The dependent variables are *Short Interest* and *Fail-to-Deliver Shares*. Both dependent variables are normalized by total *Shares Outstanding*. The independent variables are the ETF's *Short Interest Ratio*; the ETF's *Share Turnover*; the *Daily Cost of Borrow Score*; an *Available Options Dummy*; and *Operational Shorting*, which measures the propensity for the operational shorting of ETF shares. A complete list of variable names, sources, and definitions is provided in Appendix B. All independent variables are lagged. The sample period is March, 22 2004 – December 31, 2016. All specifications include ETF and date fixed effects, and the standard errors are clustered at the ETF and date levels. The t-statistics are based on standard errors clustered at the ETF and date level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Short Interest / Shares Outstanding (t)			Fail-to-Deliver Shares / Shares Outstanding (t)		
	(1)	(2)	(3)	(4)	(5)	(6)
log (Market Cap) at (t-1)	-0.00203***	-0.00183***	-0.000878**	-0.00352***	-0.00326***	-0.00258***
	(-5.708)	(-4.766)	(-2.213)	(-13.78)	(-8.799)	(-6.923)
Share Turnover, as % of Shares Outstanding at (t-1)	0.0394***	0.0310***	0.0284***	0.0721***	0.0755***	0.0737***
	(5.427)	(4.577)	(4.228)	(7.650)	(6.381)	(6.185)
Short Interest Ratio, as % of Shares Outstanding at (t-1)	0.697***	0.767***	0.767***	0.0469***	0.0332***	0.0322***
	(37.34)	(43.58)	(43.52)	(8.947)	(6.807)	(6.655)
Daily Cost of Borrow Score at (t-1)		0.000558***	0.000536***		0.000421***	0.000408***
		(2.803)	(2.683)		(3.038)	(2.933)
Available Options Dummy at (t-1)		0.00258***	0.00230***		-0.00282***	-0.00300***
		(3.244)	(2.873)		(-4.645)	(-4.891)
Operational Shorting, as % of Shares Outstanding at (t-1)			0.105***			0.0753***
			(7.493)			(9.613)
Observations	260,352	163,454	163,454	2,925,879	1,755,400	1,755,400
R-squared	0.787	0.848	0.849	0.100	0.125	0.129

Table 4 – The Dynamics of Net Creation Units and Order Imbalances: This table displays Ordinary Least Squares (OLS) regression results. The dependent variables are *Net Creation Units (Flows) and ETF Order Imbalance*. These measures estimate the inter-relationships between the net demands on creating new ETF units and buying ETF shares. Independent variables include the ETF's 15-day lagged *log(Market Cap)*; 15-day lagged *Share Turnover;* zero- through eight-day lagged *ETF Order Imbalance*; and the zero- through eight-day lagged *Net Create/Redeem Activity*. A complete list of variable names, sources, and definitions is provided in Appendix B. The sample period is March, 22 2004 – December 31, 2016. All specifications include ETF and date fixed effects, and the standard errors are clustered at the ETF and date levels. The t-statistics are based on standard errors clustered at the ETF and date level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Net Create/Redeem Activity at day t: log (1 + % change in Shares Outstanding)			ETF Order Imbalance at day t: (Buys - Sells) / Average Shares Outstanding		
	(1)	(2)	(3)	(4)	(5)	(6)
log (Market Cap), at (t-15)	-0.000509***	-0.000268***	-0.000208***	-0.00146***	-0.00217***	-0.00145***
	(-16.78)	(-5.648)	(-5.104)	(-16.24)	(-13.27)	(-15.89)
Share Turnover, as % of Shares Outstanding, at (t-15)	0.00530***	0.00525***	0.00422***	0.00350***	0.00512***	0.00343***
	(12.79)	(10.73)	(9.840)	(7.124)	(6.666)	(7.050)
ETF Order Imbalance at (t)		0.0246***	0.0249***			
		(8.368)	(8.763)			
ETF Order Imbalance at (t-1)		0.0668***	0.0660***	0.108***		0.105***
		(13.07)	(13.20)	(23.26)		(22.22)
ETF Order Imbalance at (t-2)		0.0466***	0.0433***	0.0643***		0.0626***
		(14.15)	(13.94)	(17.95)		(16.65)
ETF Order Imbalance at (t-3)		0.0288***	0.0240***	0.0455***		0.0450***
		(11.65)	(10.43)	(15.04)		(14.09)
ETF Order Imbalance at (t-4)		0.0193***	0.0144***	0.0419***		0.0421***
		(9.251)	(7.560)	(14.06)		(13.33)
ETF Order Imbalance at (t-5)		0.0151***	0.0103***	0.0375***		0.0379***
		(7.628)	(6.021)	(12.75)		(12.11)
ETF Order Imbalance at (t-6)		0.0126***	0.00766***	0.0321***		0.0326***
		(7.166)	(5.121)	(12.37)		(11.90)
ETF Order Imbalance at (t-7)		0.00968***	0.00513***	0.0331***		0.0337***
		(6.286)	(3.722)	(11.92)		(11.50)
ETF Order Imbalance at (t-8)		0.00695***	0.00208	0.0359***		0.0362***
		(5.039)	(1.568)	(15.47)		(14.68)
Net Create/Redeem Activity at (t)					0.0699***	0.0404***
					(19.18)	(10.91)
Net Create/Redeem Activity at (t-1)	0.0507***		0.0358***		0.0249***	-0.00232
	(11.93)		(7.821)		(11.78)	(-1.045)
Net Create/Redeem Activity at (t-2)	0.0463***		0.0362***		0.0169***	-0.00526***
	(20.54)		(14.87)		(10.55)	(-2.928)
Net Create/Redeem Activity at (t-3)	0.0318***		0.0235***		0.0109***	-0.00797***
	(11.16)		(7.887)		(6.906)	(-4.210)
Net Create/Redeem Activity at (t-4)	0.0223***		0.0148***		0.0110***	-0.00519***
	(8.134)		(4.801)		(6.444)	(-2.719)
Net Create/Redeem Activity at (t-5)	0.0299***		0.0246***		0.00947***	-0.00437**
	(8.221)		(6.078)		(5.724)	(-2.356)
Net Create/Redeem Activity at (t-6)	0.0125***		0.00784***		0.00809***	-0.00397***
	(5.298)		(2.966)		(6.030)	(-2.863)
Net Create/Redeem Activity at (t-7)	0.0195***		0.0160***		0.00813***	-0.000557
	(10.13)		(7.323)		(5.514)	(-0.372)
Net Create/Redeem Activity at (t-8)	0.0172***		0.0150***		0.00806***	0.00334**
	(10.20)		(7.763)		(5.376)	(2.223)
Observations	2,950,589	2,136,427	2,136,427	2,136,427	2,364,099	2,136,427
R-squared	0.024	0.038	0.043	0.091	0.055	0.092

Table 5 – The Effects of Net Creation Activity and Order Imbalances on Failures-to-Deliver and Short Interest: This table displays Ordinary Least Squares (OLS) regression results. The dependent variable in Panel A is *Failure-to-Deliver*, and the dependent variable in Panel B is *Short Interest* (both are scaled by Total ETF Shares outstanding). Independent variables include the ETF's 15-day lagged *log(Market Cap)*; 15-day lagged *Share Turnover;* zero- through eight-day lagged *ETF Order Imbalance*; and the zero- through eight-day lagged *Net Create/Redeem Activity*. Regressions in Panel A include lagged *Fail-to-Deliver* and regressions in Panel B include lagged *Short Interest.* A complete list of variable names, sources, and definitions is provided in Appendix B. The sample period is March, 22 2004 – December 31, 2016. All specifications include ETF and date fixed effects, and the standard errors are clustered at the ETF and date levels. The t-statistics are based on standard errors clustered at the ETF and date level are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Failures to Deliver Regressions

	Fail-to-Deliver Shares / Shares Outstanding (t)					
	(1)	(2)	(3)	(4)	(5)	(6)
log (Market Cap), at (t-15)	-0.00384***	-0.00112***	-0.00377***	-0.00120***	-0.00375***	-0.00113***
	(-11.07)	(-10.92)	(-13.99)	(-14.65)	(-10.85)	(-10.80)
Share Turnover, as % of Shares	0.0820***	0.0237***	0.0847***	0.0254***	0.0803***	0.0239***
Outstanding, at (t-15)	(8.638)	(8.098)	(8.768)	(8.658)	(8.490)	(8.206)
ETF Order Imbalance at (t-1)	0.0219***	0.0103***			0.0156***	0.00760***
	(4.477)	(4.382)			(3.395)	(3.301)
ETF Order Imbalance at (t-2)	0.0368***	0.0215***			0.0174***	0.0167***
	(6.911)	(7.447)			(3.605)	(5.889)
ETF Order Imbalance at (t-3)	0.145***	0.119***			0.127***	0.121***
	(13.46)	(13.55)			(12.59)	(13.82)
ETF Order Imbalance at (t-4)	0.111***	0.00799**			0.100***	0.0134***
	(12.28)	(2.403)			(11.69)	(4.245)
ETF Order Imbalance at (t-5)	0.0867***	0.00726***			0.0804***	0.0113***
	(11.45)	(2.719)			(10.84)	(3.911)
ETF Order Imbalance at (t-6)	0.0675***	0.00513*			0.0634***	0.00736**
	(9.793)	(1.797)			(9.337)	(2.565)
ETF Order Imbalance at (t-7)	0.0523***	0.00246			0.0489***	0.00365
	(8.810)	(1.085)			(8.252)	(1.542)
ETF Order Imbalance at (t-8)	0.0469***	0.00641**			0.0436***	0.00707***
	(7.960)	(2.426)			(7.500)	(2.631)
Net Create/Redeem Activity at (t-1)			0.265***	0.101***	0.251***	0.0976***
			(27.01)	(17.25)	(23.27)	(15.46)
Net Create/Redeem Activity at (t-2)			0.137***	-0.0398***	0.102***	-0.0684***
			(13.76)	(-3.933)	(9.605)	(-5.872)
Net Create/Redeem Activity at (t-3)			0.0361***	-0.0531***	-0.00472	-0.0715***
			(4.986)	(-12.17)	(-0.584)	(-13.91)
Net Create/Redeem Activity at (t-4)			0.0177***	-0.00280	-0.0166**	-0.0103***
			(2.855)	(-0.916)	(-2.304)	(-2.928)
Net Create/Redeem Activity at (t-5)			0.0152***	0.00580**	-0.0102	0.00242
			(2.860)	(2.319)	(-1.584)	(0.839)
Net Create/Redeem Activity at (t-6)			0.0118**	0.00561**	-0.00790	0.00144
			(2.364)	(2.267)	(-1.312)	(0.503)
Net Create/Redeem Activity at (t-7)			0.0161***	0.00941***	0.00240	0.00667**
			(3.778)	(4.024)	(0.476)	(2.512)
Net Create/Redeem Activity at (t-8)			0.0130***	0.00471**	0.00491	0.00245
			(3.599)	(2.479)	(1.183)	(1.147)
Fail-to-Deliver Shares / Shares		0.697***		0.699***		0.698***
Outstanding (t-1)		(84.72)		(88.66)		(82.59)
Observations	2,151,271	2,151,271	2,950,592	2,950,592	2,151,271	2,151,271
R-squared	0.128	0.557	0.104	0.541	0.137	0.559

Panel B: Short Interest Regressions

	Short Interest / Shares Outstanding (t)					
	(1)	(2)	(3)	(4)	(5)	(6)
log (Market Cap), at (t-15)	-0.00912***	-2.23e-05	-0.00886***	-0.000166***	-0.00885***	-6.22e-05*
	(-5.291)	(-0.636)	(-5.933)	(-4.729)	(-5.146)	(-1.732)
Share Turnover, as % of Shares Outstanding, at (t-15)	0.300***	0.00262***	0.292***	0.00369***	0.295***	0.00323***
	(8.295)	(3.894)	(8.430)	(5.370)	(8.183)	(4.772)
ETF Order Imbalance at (t-1)	0.0478***	0.00404***			0.0442***	0.00355**
	(4.402)	(2.964)			(4.147)	(2.573)
ETF Order Imbalance at (t-2)	0.0514***	0.00534***			0.0349***	0.00436***
	(5.082)	(3.302)			(3.456)	(2.750)
ETF Order Imbalance at (t-3)	0.0623***	0.0127***			0.0370***	0.0126***
	(6.070)	(6.327)			(3.551)	(6.180)
ETF Order Imbalance at (t-4)	0.0732***	0.00977***			0.0429***	0.0108***
	(7.081)	(5.017)			(4.036)	(5.452)
ETF Order Imbalance at (t-5)	0.0807***	0.00622***			0.0488***	0.00857***
	(7.508)	(3.573)			(4.351)	(4.837)
ETF Order Imbalance at (t-6)	0.0840***	0.00139			0.0522***	0.00481***
	(7.467)	(0.752)			(4.413)	(2.614)
ETF Order Imbalance at (t-7)	0.0890***	0.000680			0.0584***	0.00513***
	(7.488)	(0.395)			(4.686)	(2.886)
ETF Order Imbalance at (t-8)	0.0948***	-0.00112			0.0653***	0.00447***
	(7.302)	(-0.675)			(4.834)	(2.684)
Net Create/Redeem Activity at (t-1)			0.213***	0.0214***	0.185***	0.0170***
			(17.27)	(8.786)	(12.83)	(7.024)
Net Create/Redeem Activity at (t-2)			0.200***	-5.51e-05	0.171***	-0.00475**
			(17.38)	(-0.0272)	(12.58)	(-2.304)
Net Create/Redeem Activity at (t-3)			0.181***	-0.00657***	0.153***	-0.0103***
			(16.41)	(-3.109)	(11.67)	(-4.544)
Net Create/Redeem Activity at (t-4)			0.157***	-0.0149***	0.129***	-0.0182***
			(14.92)	(-6.295)	(10.29)	(-7.511)
Net Create/Redeem Activity at (t-5)			0.135***	-0.0160***	0.109***	-0.0174***
			(13.10)	(-7.028)	(9.000)	(-7.496)
Net Create/Redeem Activity at (t-6)			0.108***	-0.0201***	0.0866***	-0.0214***
			(10.50)	(-8.533)	(7.332)	(-8.788)
Net Create/Redeem Activity at (t-7)			0.0862***	-0.0203***	0.0686***	-0.0214***
			(8.001)	(-8.813)	(5.768)	(-8.942)
Net Create/Redeem Activity at (t-8)			0.0540***	-0.0286***	0.0429***	-0.0290***
			(4.908)	(-11.34)	(3.628)	(-11.26)
Short Interest / Shares Outstanding (t-1)		0.980***		0.979***		0.981***
		(448.1)		(440.0)		(452.7)
Observations	2,476,342	2,475,921	2,926,486	2,925,790	2,476,342	2,475,921
R-squared	0.663	0.988	0.652	0.986	0.664	0.988

Table 6 – The Determinants of Operational Shorting: This table displays Ordinary Least Squares (OLS) regression results. The dependent variable is *Operational Shorting* normalized by total shares outstanding. This measure estimates the propensity for operational shorting of ETF shares. Independent variables include the ETF's 15-day lagged *log(Market Cap)*; 15-day lagged *Average Share Turnover*; lagged *Creation Unit Dollar Size* and *Creation Unit Fee (per share)*; lagged *Maximum Rolling R-Squared with Available Futures Contracts*; lagged *Available Options Dummy*; lagged *Mispricing*; lagged *Discount*; and a lagged *Proxy for Liquidity Mismatch.* A complete list of variable names, sources, and definitions is provided in Appendix B. The sample period is March, 22 2004 – December 31, 2016, and t-statistics based on standard errors clustered at the stock and date level are in parentheses. All specifications include ETF and date fixed effects, and the standard errors are clustered at the ETF and date levels. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Operational Shorting, as % of Shares Oustanding at day (t)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log (Market Cap), at (t-15)	-0.00730***	-0.00778***	-0.00806***	-0.00802***	-0.00899***	-0.00687***	-0.00962***
	(-14.19)	(-11.02)	(-14.16)	(-13.99)	(-6.777)	(-7.636)	(-4.590)
Average Share Turnover, as % of Shares	0.0320***	0.0246***	0.0319***	0.0316***	0.0376***	0.0330***	0.0409***
Outstanding, at (t-15)	(11.81)	(6.848)	(11.44)	(11.39)	(4.597)	(5.108)	(3.307)
Creation Unit Dollar Size, log, at (t-1)		0.00536***					0.00228
		(7.596)					(1.433)
Creation Unit Fee, per share, at (t-1)		0.0455***					0.00167
		(3.165)					(0.0784)
Maximum Rolling R-Squared with Available			0.0112***	0.0111***	0.0129***	0.00912***	0.0149**
Futures Contracts at (t-1)			(7.100)	(7.004)	(3.190)	(3.034)	(2.551)
Available Options Dummy at (t-1)			0.00243***	0.00253***	0.00387***	0.00261***	0.00276*
			(3.745)	(3.881)	(3.034)	(2.894)	(1.690)
Mispricing at (t-1): % difference between ETF				0.286***	0.243***		0.370***
price and NAV at the close of the previous day				(22.69)	(5.974)		(10.23)
Premium at (t-1), if mispricing>0, and zero						0.253***	
						(5.208)	
Discount at (t-1), in absolute value, if						-0.238***	
mispricing<0, and zero otherwise						(-3.230)	
Proxy for Liquidity Mismatch, at (t-1): Average					0.219***	0.157***	0.250**
Intraday Basket Spread - Intraday ETF Spread					(3.674)	(3.670)	(2.302)
Observations	2,950,667	1,988,950	2,633,071	2,624,669	787,099	820,652	499,849
R-squared	0.164	0.201	0.166	0.168	0.199	0.184	0.262

Table 7 – Operational Shorting and Contemporaneous/Future Returns: This table displays Ordinary Least Squares (OLS) regression results. The dependent variable are 1-week contemporaneous (t) or forward-looking (t+1) total returns (Ret) or Fama-French 4-factor risk-adjusted alphas (FF3 α). This measure is based on the ETF or NAV price as indicated in the header. Independent variables are measured at time *t* and include the *Operational Shorting - Weekly %* (cumulative buy-sell imbalance in week (t) minus create orders in week (t) as a percentage of ETF shares outstanding), *Create Orders – Weekly %* (as a percentage of ETF shares outstanding), *log(Market Cap)* of the ETF, the ETF's *Average Share Turnover %* (as a percentage of ETF shares outstanding), and the ETF's *Amihud Illiquidity*. A complete list of variable names, sources, and definitions is provided in Appendix B. Specifications 1-3 include all ETFs, 4 non-equity ETFs, 5 through 9 equity ETFs. Specifications 8 and 9 are further split based on the liquidity mismatch with the Low Liquidity Mismatch indicating similar liquidity for the ETF and underlying. High Liquidity Mismatch indicates the ETF is more liquid than the underlying. The sample period is March, 22 2004 – December 31, 2016, and t-statistics based on standard errors clustered at the stock and date level are in parentheses. All specifications include ETF and date fixed effects, and the standard errors are clustered at the ETF and date levels. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Weekly Return								
	ETF Ret (t) (1)	ETF Ret (t+1) (2)	ETF FF4 α (t+1) (3)	NAV Ret (t+1) (4)	NAV FF4 α (t+1) (5)	ETF Ret (t+1) (6)	ETF Ret (t+1) (7)	ETF Ret (t+1) (8)	ETF Ret (t+1) (9)
Operational Shorting - Weekly % (t)	2.368***	-0.252	-0.0370***	-0.107	-0.00255	-0.378***	-0.232**	0.0652	-0.232**
	(8.93)	(-1.46)	(-2.93)	(-0.63)	(-0.20)	(-3.68)	(-2.11)	(0.21)	(-2.13)
Create Orders - Weekly % (t)	-0.00256	-0.00432	-0.00044*	-0.00215	0.00005	0.00286**	0.00313*	0.311	0.00464***
	(-0.73)	(-1.39)	(-1.92)	(-0.83)	(0.21)	(2.07)	(1.96)	(1.03)	(3.96)
log (Market Cap), at (t-1)	-0.0267	-0.0532***	-0.00740***	-0.0513**	-0.00675***	-0.0567***	-0.0503***	-0.0484***	-0.0564***
	(-1.31)	(-2.65)	(-3.91)	(-2.56)	(-3.82)	(-4.66)	(-4.44)	(-2.91)	(-4.19)
Average Share Turnover (t-1)	-1.047**	-0.829**	-0.0376**	-0.833**	-0.0270	-0.153	-0.0947	-0.436	-0.0550
	(-2.47)	(-2.24)	(-1.98)	(-2.36)	(-1.25)	(-0.86)	(-0.26)	(-0.83)	(-0.15)
Amihud Illiquidity Measure (t-1)	0.0979**	0.0482	0.00986	0.0384	0.00414	0.0475	0.0367**	0.0375	0.0357
	(2.31)	(1.42)	(1.41)	(1.45)	(1.08)	(1.46)	(2.09)	(0.64)	(1.51)
Observations	577,080	574,997	561,283	574,997	561,283	269,156	225,458	60,962	162,207
R-squared	0.292	0.291	0.084	0.281	0.095	0.420	0.695	0.747	0.679
Stock & Date Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock & Date Clustering	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ETF Sample	All	All	All	All	All	Non-Equity	Equity	Equity	Equity
Liquidity Mismatch (ETF vs Underlying)								Low	High

Table 8 – ETF Mis-Pricing and Arbitrage Activity: This table displays Ordinary Least Squares (OLS) regression results. The dependent variables are *Mispricing Change* and *Absolute Mispricing Change*. Independent variables include contemporaneous and lagged *Operational Shorting;* 15-day lagged *log(Market Cap)*; 15-day lagged *Average Share Turnover*, normalized by shares outstanding; lagged *Maximum Rolling R-Squared with Available Futures Contracts*; lagged *Available Options Dummy*; lagged *Mispricing Change*; and lagged *Absolute Mispricing Change*. A complete list of variable names, sources, and definitions is provided in Appendix B. The sample period is March, 22 2004 – December 31, 2016, and t-statistics based on standard errors clustered at the stock and date level are in parentheses. All specifications include ETF and date fixed effects, and the standard errors are clustered at the ETF and date levels. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Mispricing Change at (t)				Absolute Mispricing Change (at t)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Operational Shorting, as % of Shares	-0.00233***	-0.00225***			-0.00128***	-0.000877***		
Outstanding at (t)	(-12.80)	(-12.60)			(-6.058)	(-5.927)		
Operational Shorting, as % of Shares			-0.000243**	-0.00133***			-0.00115***	-0.000716***
Outstanding at (t-1)			(-2.225)	(-9.912)			(-5.491)	(-4.810)
log (Market Cap), at (t-15)	-2.12e-05*	-2.37e-05**	-7.31e-06	-1.85e-05**	-0.000243***	-0.000164***	-0.000242***	-0.000162***
	(-1.772)	(-2.400)	(-0.743)	(-2.130)	(-9.731)	(-9.629)	(-9.692)	(-9.553)
Average Share Turnover, as % of Shares	4.90e-05	5.32e-05	-1.10e-05	1.63e-05	-0.000182	-0.000122	-0.000185	-0.000126
Outstanding, at (t-15)	(0.688)	(0.765)	(-0.159)	(0.257)	(-0.883)	(-0.853)	(-0.901)	(-0.887)
Maximum Rolling R-Squared with		6.18e-05	3.97e-05	6.84e-05	-0.00150***	-0.000969***	-0.00150***	-0.000971***
Available Futures Contracts at (t-1)		(0.280)	(0.180)	(0.367)	(-7.076)	(-6.404)	(-7.081)	(-6.414)
Available Options Dummy at (t-1)		5.99e-06	1.00e-06	6.26e-06	-9.36e-05*	-6.37e-05*	-9.39e-05*	-6.41e-05*
		(0.285)	(0.0477)	(0.337)	(-1.754)	(-1.746)	(-1.759)	(-1.756)
Mispricing Change at (t-1)				-0.485***				
				(-59.10)				
Absolute Mispricing Change at (t-1)						0.330***		0.330***
						(41.17)		(41.17)
Observations	2,864,290	2,624,038	2,624,039	2,623,622	2,624,038	2,623,621	2,624,039	2,623,622
R-squared	0.039	0.040	0.039	0.266	0.369	0.438	0.369	0.438

Table 9 – Effects of ETF Operational Shorting on the Liquidity of the Underlying Securities: This table displays Ordinary Least Squares (OLS) regression results. The dependent variable in Panel A is the *Intraday NBBO Spread* of underlying stocks held by U.S. equity-only ETFs. The dependent variable in Panel B is the *Intraday Second-by-Second Return Volatility* of underlying stocks. The key independent variable is *Operational Shorting*. Additional independent variables include the lagged *Averate ETF Ownership* in underlying stocks; 15-day lagged *log(Market Cap)*; 15-day *Average Share Turnover* normalized by shares outstanding; 0-, 1-, 2-, and 3-day lagged *Intraday NBBO Spread* of the ETF; and the lagged *Average Intraday NBBO Spread* of underlying stocks in the ETF basket. A complete list of variable names, sources, and definitions is provided in Appendix B. The sample period is March, 22 2004 – December 31, 2016, and t-statistics based on standard errors clustered at the stock and date level are in parentheses. All specifications include ETF and date fixed effects, and the standard errors are clustered at the ETF and date levels. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Effect of Operational Shorting on Intraday NBBO Spread of Underlying Stocks

	Average Intraday NBBO Spread of Underlying Stocks in ETF Basket (t)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average ETF Ownership in Underlying Stocks in ETF Basket (t-1)	0.00216210***	0.00066539***			0.00216031***	0.00066489***	0.00066526***
	(3.14)	(3.03)			(3.14)	(3.03)	(3.03)
Operational Shorting, as % of Shares Outstanding at (t-1)			-0.00026142**	-0.00008075*	-0.00023448**	-0.00007327**	
			(-2.16)	(-1.93)	(-2.40)	(-2.03)	
Operational Shorting, as % of Shares Outstanding at (t)							-0.00009766***
							(-2.67)
log (Market Cap), at (t-15)	-0.00002512**	-0.00000740**	-0.00001784*	-0.00000493	-0.00002668***	-0.00000789**	-0.00000805**
	(-2.55)	(-2.30)	(-1.86)	(-1.59)	(-2.65)	(-2.40)	(-2.44)
Average Share Turnover, as % of Shares Outstanding, at (t-15)	-0.00027368**	-0.00008531**	-0.00029175*	-0.00008913*	-0.00026616**	-0.00008297*	-0.00008221*
	(-2.07)	(-2.03)	(-1.77)	(-1.74)	(-2.01)	(-1.96)	(-1.94)
Intraday NBBO Spread of ETF, at (t)	0.00493170**	0.00239582***	0.00935935**	0.00342491**	0.00490543**	0.00238771***	0.00238477***
	(2.41)	(2.60)	(2.24)	(2.39)	(2.40)	(2.60)	(2.59)
Intraday NBBO Spread of ETF, at (t-1)	0.00416669**	0.00089497			0.00413961**	0.00088665	0.00088090
	(2.20)	(1.08)			(2.19)	(1.07)	(1.06)
Intraday NBBO Spread of ETF, at (t-2)	0.00404646**	0.00105572			0.00401106**	0.00104477	0.00104227
	(2.19)	(1.62)			(2.18)	(1.61)	(1.60)
Intraday NBBO Spread of ETF, at (t-3)	0.00442017**	0.00131459*			0.00438403**	0.00130343*	0.00130040*
	(2.30)	(1.76)			(2.29)	(1.75)	(1.74)
Average Intraday NBBO Spread of Underlying Stocks in ETF Basket (t-1)		0.68773597***		0.69196466***		0.68770651***	0.68770000***
		(28.93)		(29.12)		(28.93)	(28.93)
Observations	837,347	837,333	853,554	852,955	837,347	837,333	837,333
R-squared	0.755	0.869	0.753	0.870	0.755	0.869	0.869

Table 9 – Effects of ETF Operational Shorting on the Liquidity of the Underlying Securities: *(continued)*

Panel B: Effect of Operational Shorting on Intraday Second-by-Second Return Volatility of Underlying Stocks

	Average Intraday Second-by-Second Return Volatility of Underlying Stocks in ETF Basket (t)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average ETF Ownership in Underlying Stocks in ETF Basket (t-1)	0.00018230***	0.00006649***			0.00018216***	0.00006645***	0.00006651***
	(2.79)	(2.78)			(2.78)	(2.77)	(2.78)
Operational Shorting, as % of Shares Outstanding at (t-1)			-0.00002750**	-0.00001036**	-0.00002604***	-0.00000998***	
			(-2.56)	(-2.51)	(-2.71)	(-2.73)	
Operational Shorting, as % of Shares Outstanding at (t)							-0.00000974***
							(-2.71)
log (Market Cap), at (t-15)	-0.00000288***	-0.00000103***	-0.00000207**	-0.00000073**	-0.00000306***	-0.00000110***	-0.00000110***
	(-3.08)	(-2.99)	(-2.08)	(-1.99)	(-3.26)	(-3.18)	(-3.18)
Average Share Turnover, as % of Shares Outstanding, at (t-15)	0.00004437**	0.00001635**	0.00004410*	0.00001601*	0.00004506**	0.00001662**	0.00001661**
	(2.09)	(2.09)	(1.85)	(1.85)	(2.11)	(2.12)	(2.12)
Average Intraday Second-by-Second Return Volatility of ETF, at (t)	0.11402739***	0.06758951***	0.11371293***	0.06704016***	0.11427900***	0.06769228***	0.06769561***
	(12.14)	(12.21)	(11.98)	(12.03)	(12.15)	(12.21)	(12.21)
Average Intraday Second-by-Second Return Volatility of ETF, at (t-1)	0.06845335***	-0.00042104	0.06819799***	-0.00115933	0.06872139***	-0.00030888	-0.00026267
	(10.17)	(-0.15)	(9.99)	(-0.39)	(10.18)	(-0.11)	(-0.09)
Average Intraday Second-by-Second Return Volatility of ETF, at (t-2)	0.06534263***	0.02290503***	0.06509124***	0.02246237***	0.06575231***	0.02306781***	0.02304690***
	(10.18)	(8.63)	(10.03)	(8.44)	(10.20)	(8.68)	(8.67)
Average Intraday Second-by-Second Return Volatility of ETF, at (t-3)	0.06200923***	0.01963887***	0.06205356***	0.01921328***	0.06239564***	0.01979273***	0.01975866***
	(9.47)	(7.55)	(9.34)	(7.18)	(9.49)	(7.58)	(7.57)
Average Intraday Volatility of Underlying Stocks in ETF Basket, at (t-1)		0.63641562***		0.64258288***		0.63632873***	0.63632892***
		(38.53)		(37.29)		(38.53)	(38.52)
Observations	822,739	822,712	823,270	822,712	822,739	822,712	822,712
R-squared	0.844	0.907	0.841	0.907	0.844	0.907	0.907

Table 9 – Effects of ETF Operational Shorting on the Liquidity of the Underlying Securities: *(continued)*

Panel C: Effect of Operational Shorting on Intraday Variance Ratios of Underlying Stocks

	Average Intraday Variance Ratio of Underlying Stocks in ETF Basket (t)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average ETF Ownership in Underlying Stocks	0.06956235**	0.06615123**	0.03782044**	0.06597038**	0.03772495**	0.06573597**	0.03770011**	0.03777137**
in ETF Basket (t-1)	(2.17)	(2.09)	(2.08)	(2.08)	(2.07)	(2.07)	(2.06)	(2.07)
Operational Shorting, as % of Shares				-0.02433442***	-0.01332624**	-0.02614266***	-0.01461224***	
Outstanding at (t-1)				(-2.68)	(-2.43)	(-2.88)	(-2.67)	
Operational Shorting, as % of Shares								-0.01509371***
Outstanding at (t)								(-2.61)
log (Market Cap), at (t-15)	0.00004206	-0.00027432	-0.00015225	-0.00033633	-0.00018622	-0.00038304	-0.00021312	-0.00021447
	(0.09)	(-0.56)	(-0.54)	(-0.69)	(-0.66)	(-0.78)	(-0.76)	(-0.76)
Average Share Turnover, as % of Shares	-0.00686759	-0.00944429	-0.00532467	-0.00913807	-0.00515750	-0.00951224	-0.00536589	-0.00536099
Outstanding, at (t-15)	(-1.05)	(-1.41)	(-1.37)	(-1.37)	(-1.33)	(-1.42)	(-1.38)	(-1.38)
Average Intraday Second-by-Second Return		0.02226453***	0.01727076***	0.02240687***	0.01734934***	0.02110853***	0.01664839***	0.01666383***
Volatility of ETF, at (t)		(5.90)	(6.21)	(5.96)	(6.25)	(6.17)	(6.33)	(6.34)
Average Intraday Second-by-Second Return		0.01257684***	0.00320915**	0.01273483***	0.00329685**	0.01133288***	0.00248557*	0.00251579*
Volatility of ETF, at (t-1)		(4.65)	(1.98)	(4.73)	(2.05)	(4.97)	(1.74)	(1.76)
Average Intraday Second-by-Second Return		0.00865806***	0.00338032**	0.00885789***	0.00349042**	0.00745230***	0.00269770*	0.00270143*
Volatility of ETF, at (t-2)		(3.17)	(2.00)	(3.26)	(2.08)	(3.27)	(1.85)	(1.86)
Average Intraday Second-by-Second Return		0.00941953***	0.00574568***	0.00961969***	0.00585574***	0.00792488***	0.00489855***	0.00489545***
Volatility of ETF, at (t-3)		(3.42)	(3.49)	(3.51)	(3.57)	(3.55)	(3.53)	(3.53)
Average Intraday Second-by-Second Return						0.00731230***	0.00398655***	0.00394140***
Volatility of ETF, at (t-4)						(3.11)	(2.69)	(2.66)
Average Intraday Second-by-Second Return						0.00724083***	0.00415932***	0.00415986***
Volatility of ETF, at (t-5)						(3.15)	(3.01)	(3.01)
Average Intraday Variance Ratio of			0.42906157***		0.42900779***		0.42889233***	0.42888924***
Underlying Stocks in ETF Basket (t-1)			(34.49)		(34.48)		(34.39)	(34.39)
Observations	744,187	720,069	720,056	720,069	720,056	713,693	713,680	713,680
R-squared	0.799	0.799	0.836	0.799	0.836	0.799	0.836	0.836

Table 10 – Market Makers' Spillover Effects on FTDs and Operational Shorting: This table displays Ordinary Least Squares (OLS) regression results. The dependent variables are ETF-related *Failures-to-Deliver (FTD)* and *Operational Shorting*. Independent variables include the 15-day lagged *log(Market Cap)*; 15-day lagged *Average Share Turnover*, normalized by shares outstanding; the lagged *Maximum Rolling R-Squared with Available Futures Contracts*; and the lagged *Available Options Dummy*. We also include independent variables at both the lead market maker level and the market-wide level. Lead market maker variables exclude individual ETF FTDs and Volume, while market-wide measures exclude affiliated lead market maker ETF FTDs and Total Volume. These variables include *Fail-to-Deliver* as a percentage of either total volume or ETF market cap; and *Operational Shorting* as a percentage of either total volume or ETF market cap. A complete list of variable names, sources, and definitions is provided in Appendix B. All specifications include ETF fixed effects, and the standard errors are clustered at the stock and date levels. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	Fail-to-Deliver Shares / Shares Outstanding, at day (t)		Operational Shorting / Shares Oustanding, at day (t)	
	(1)	(2)	(3)	(4)
log (Market Cap), at (t-15)	-0.00233***	-0.00209***	-0.00644***	-0.00625***
	(-19.79)	(-18.50)	(-19.76)	(-19.29)
Average Share Turnover, as % of Shares Outstanding, at (t-15)	0.0302***	0.0291***	0.0263***	0.0250***
	(13.85)	(13.06)	(11.39)	(10.84)
Affiliated Lead Market Maker Fail-to-Deliver, % of LMM Total Volume, *excluding individual ETF FtDs and Volume*	0.00950***			
	(10.34)			
Market-Wide Fail-to-Deliver, % of Overall Trading Volume, *excluding Affiliated Lead Market Marker ETF FtDs and Total Volume*	0.0225***			
	(10.29)			
Affiliated Lead Market Maker Fail-to-Deliver, % of All Affiliated ETF Market Cap, *excluding individual ETF FtDs and Market Cap*		0.354***		
		(10.59)		
Market-Wide Fail-to-Deliver, % of ETF Market Cap, *excluding Affiliated Lead Market Marker ETF FtDs and Market Cap*		0.737***		
		(13.52)		
Affiliated Lead Market Maker Operational Shorts, % of LMM Total Volume, *excluding individual ETF Operational Shorts and Volume*			0.00119**	
			(2.460)	
Market-Wide Operational Shorts, % of Overall Trading Volume, *excluding Affiliated Lead Market Marker ETF Operational Shorts and Market Cap*			0.0110***	
			(5.589)	
Affiliated Lead Market Maker Operational Shorts, % of All Affiliated ETF Market Cap, *excluding individual ETF Operational Shorts and Volume*				0.108***
				(6.176)
Market-Wide Operational Shorts, % of ETF Market Cap, *excluding Affiliated Lead Market Marker ETF Operational Shorts and Volume*				0.0372
				(1.171)
Maximum Rolling R-Squared with Available Futures Contracts at (t-1)	-0.00280***	-0.00298***	0.00601***	0.00501***
	(-4.056)	(-4.337)	(5.477)	(4.955)
Available Options Dummy at (t-1)	-0.000829***	-0.000868***	0.00214***	0.00218***
	(-3.397)	(-3.785)	(4.298)	(4.280)
Observations	2,307,010	2,307,615	2,307,010	2,307,615
R-squared	0.125	0.126	0.158	0.157
ETF Fixed Effects	Yes	Yes	Yes	Yes
Date Fixed Effects	No	No	No	No

Table 11 – The Effect of Market Makers' Leverage on FTDs and Operational Shorting: This table displays Ordinary Least Squares (OLS) regression results. The dependent variables are ETF-related *Failures-to-Deliver (FTD)* and *Operational Shorting*. Independent Variables include the inverse of *Affiliated Lead Market Maker Capital Constraints*; *Market-Wide Capital Adequacy*; the 15-day lagged *log(Market Cap)* for the ETF; 15-day lagged *Average Share Turnover*, normalized by shares outstanding; lagged *Maximum Rolling R-Squared with Available Futures Contracts*; and the lagged *Available Options Dummy*. A complete list of variable names, sources, and definitions is provided in Appendix B. The sample period is March, 22 2004 – December 31, 2016, and t-statistics based on standard errors clustered at the ETF and date level are in parentheses. All specifications include ETF fixed effects, and the standard errors are clustered at the stock and date levels. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Fail-to-Deliver Shares / Shares Outstanding, at day (t)			Operational Shorting / Shares Oustanding, at day (t)		
Affiliated Lead Market Maker Capital Constraints	9.30e-09***	1.00e-08***	1.00e-08***	2.28e-08***	2.44e-08***	2.44e-08***
	(4.578)	(4.182)	(4.182)	(4.208)	(3.471)	(3.471)
Market-Wide Capital Constraints		-3.08e-09	-3.08e-09		2.27e-09	2.27e-09
		(-1.162)	(-1.162)		(0.170)	(0.170)
log (Market Cap), at (t-15)	1.89e-10***	1.87e-10***	1.87e-10***	5.46e-10***	3.35e-10	3.35e-10
	(3.827)	(3.663)	(3.663)	(2.721)	(1.381)	(1.381)
Average Share Turnover, as % of Shares Outstanding, at	1.58e-09*	1.34e-09	1.34e-09	9.38e-10	4.09e-10	4.09e-10
(t-15)	(1.707)	(1.361)	(1.361)	(1.003)	(0.418)	(0.418)
Maximum Rolling R-Squared with Available Futures		1.72e-10	1.72e-10		3.68e-09***	3.68e-09***
Contracts at (t-1)		(0.542)	(0.542)		(2.852)	(2.852)
Available Options Dummy at (t-1)		1.07e-10	1.07e-10		1.75e-09**	1.75e-09**
		(1.013)	(1.013)		(2.216)	(2.216)
ETF Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Date Fixed Effects	No	No	No	No	No	No
Observations	1,042,546	973,516	973,516	1,042,546	973,516	973,516
R-squared	0.188	0.188	0.188	0.166	0.167	0.167

Figure 1 – ETF Short Interest: This figure displays graphically the aggregate ETF short interest in billions of dollars and as a percentage of the overall short interest in U.S. Equity Markets on a daily basis from January 2003 to December 2016.

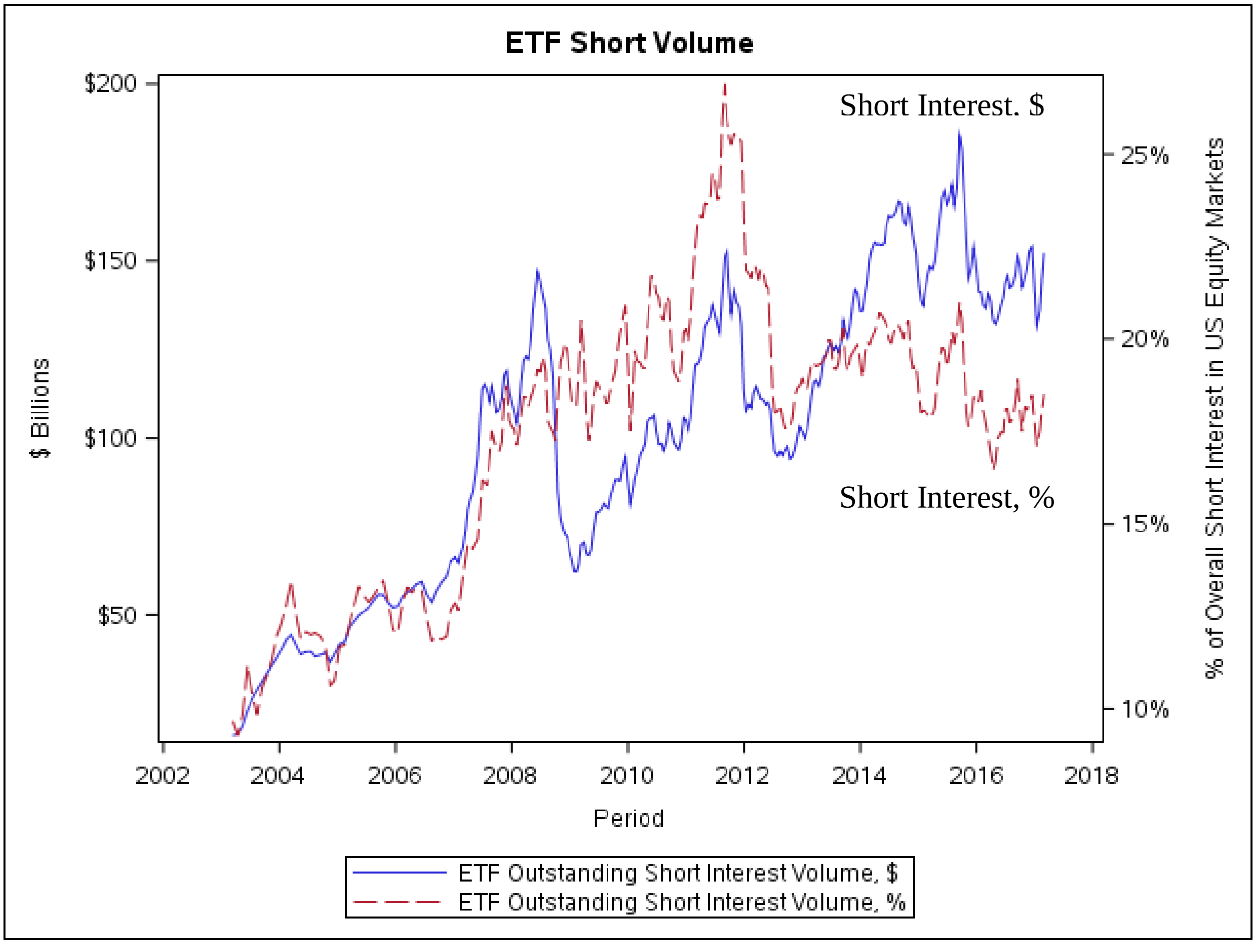


Figure 2 – Failure-to-Deliver (FTD) Activity of ETFs and Common Stocks: This figure displays graphically the average dollar volume of ETF and common stock FTDs on a daily basis from March, 22 2004 – December 31, 2016. We include only the rolling average daily FTD volume of stocks and ETFs in this graph, as they comprise the vast majority of total FTDs in the financial system.

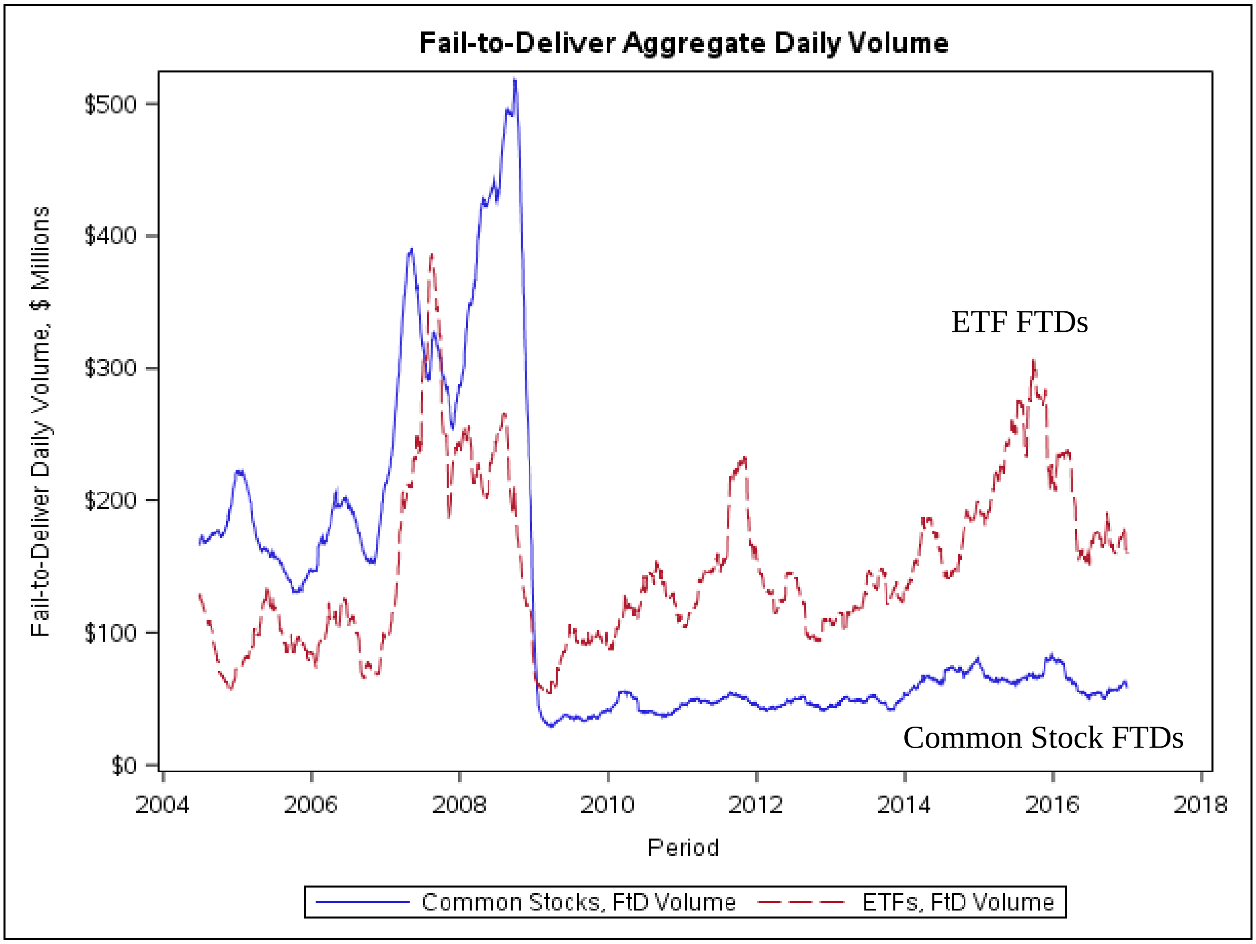


Figure 3 – ETF Settlement Failure Timeline: This figure displays the key events during a settlement failure for an ETF. Time *t* represents the time when an operational short is established. Dates *t+i*, where *i* is between 1 and 6, represent *i* days after the operational short position is established.

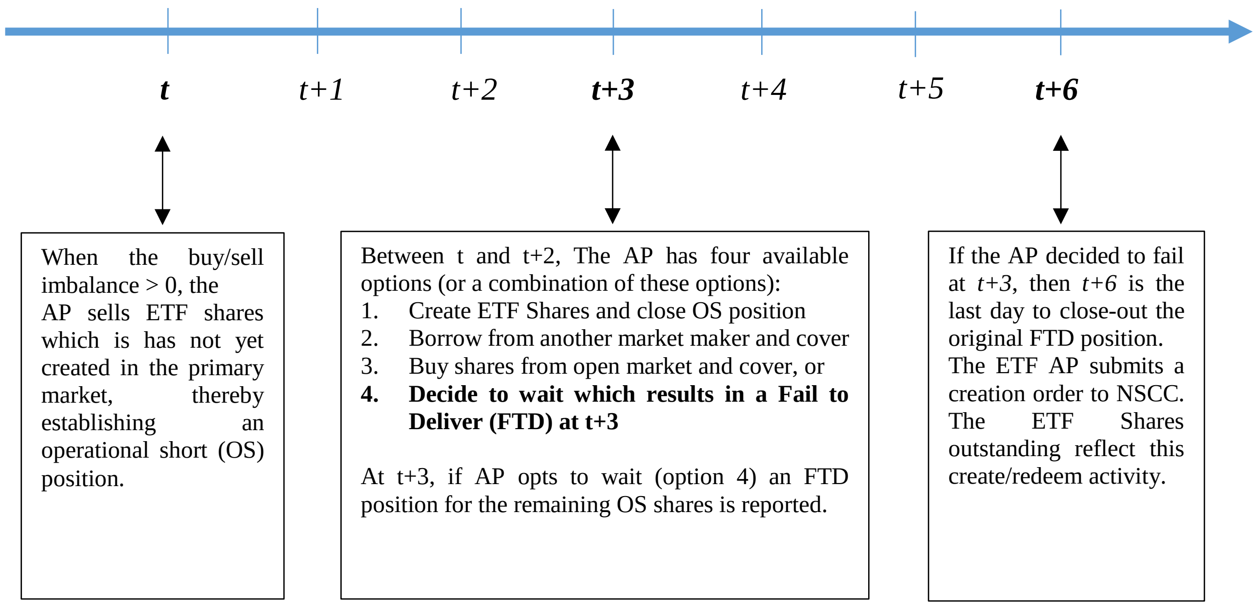


Figure 4 – Operational Shorting and Failure-to-Deliver (FTD) Activity of ETFs: This figure displays graphically the rolling-average daily dollar value of Operational Shorting activity and FTDs for ETFs from March, 22 2004 – December 31, 2016.

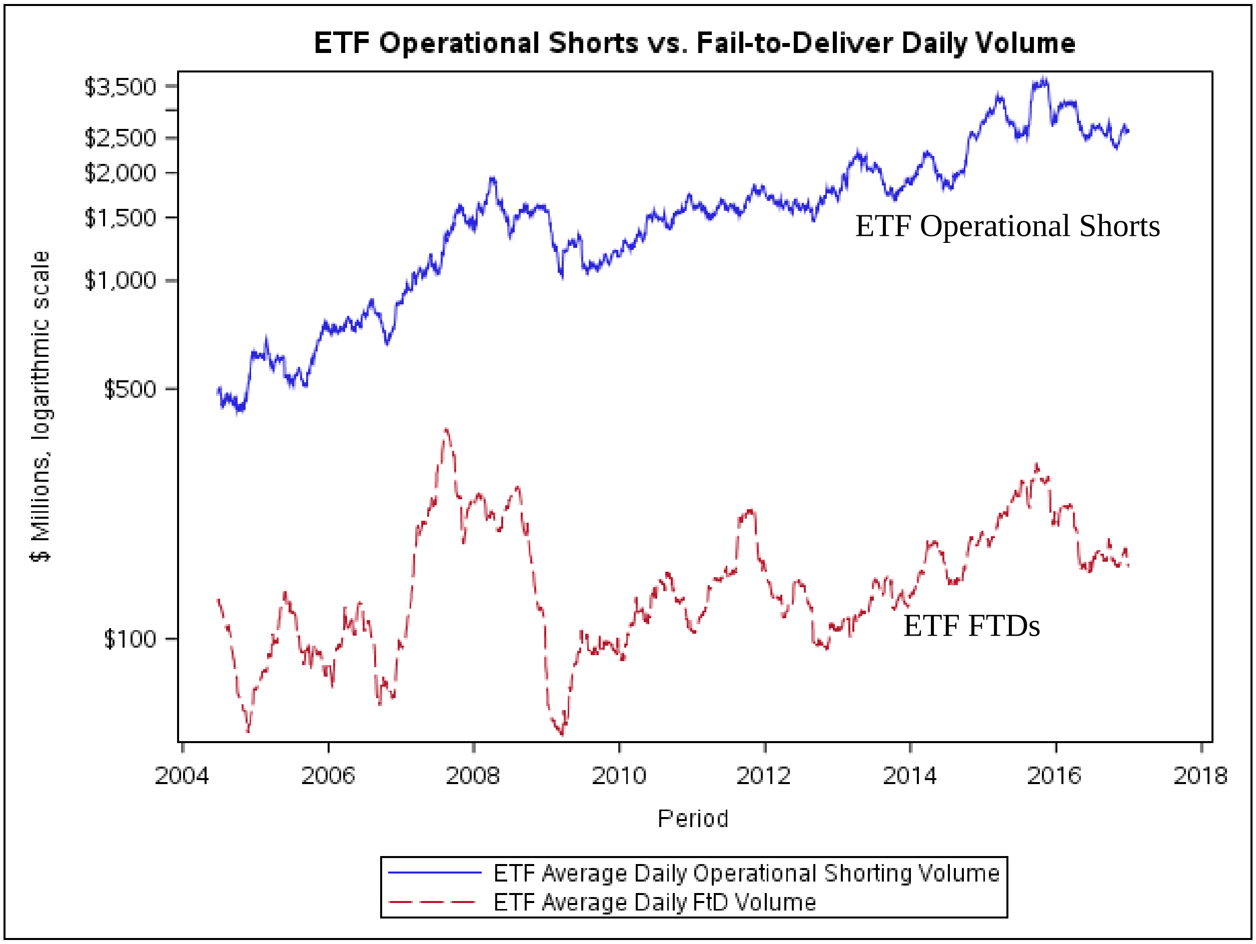
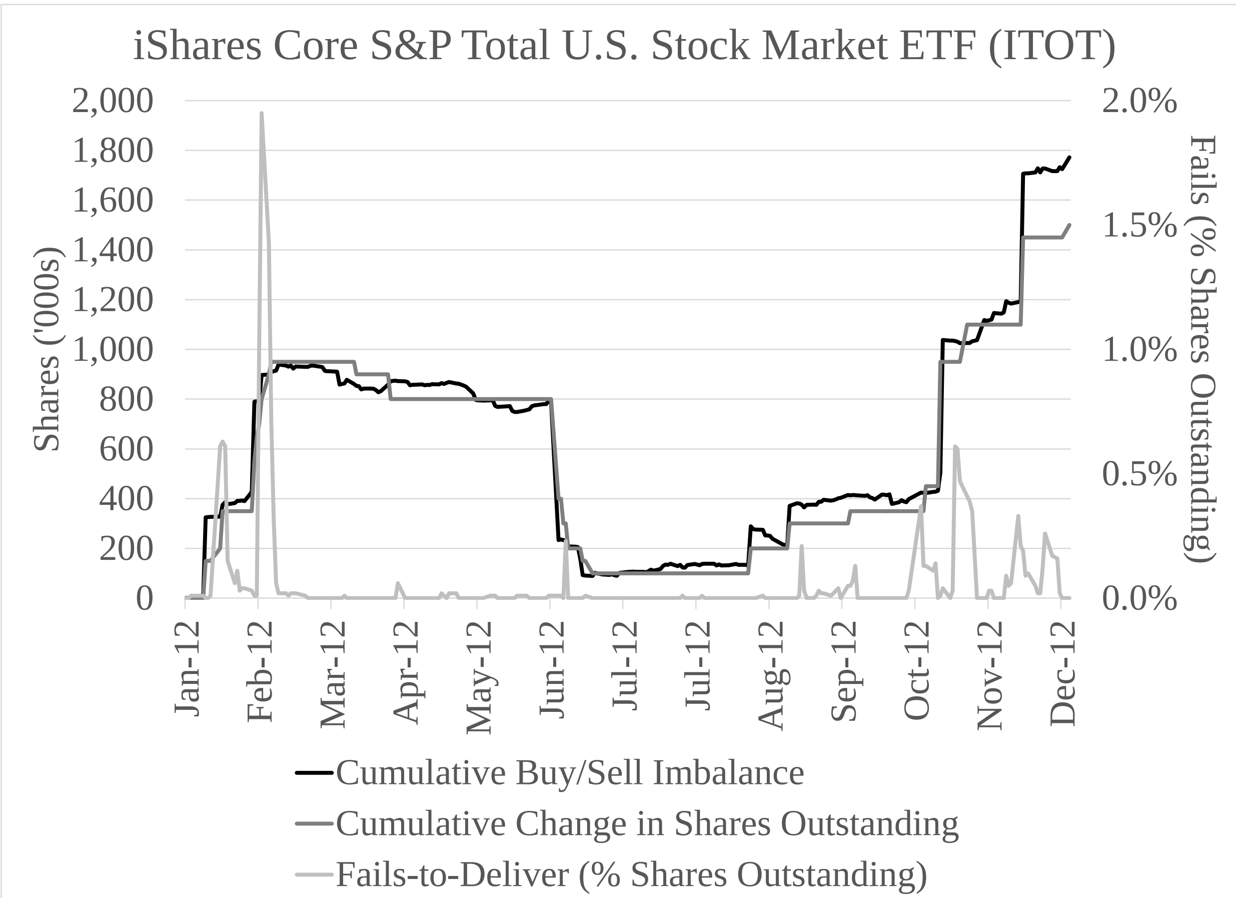


Figure 5 – An Example: ITOT – iShares Core S&P Total U.S. Stock Market ETF: This figure displays the cumulative buy-sell imbalance and the cumulative change in shares outstanding (in 1,000s of shares indexed by the left vertical axis) for the iShares Core S&P Total U.S. Stock Market ETF (ticker:ITOT) over the year 2012. Both the buy-sell imbalance and the change in shares outstanding values are set equal to 0 at the beginning of 2012 and are cumulative from that point forward. The figure also plots the ITOT failures-to-deliver as a percentage of total shares outstanding (in % indexed by the right vertical axis).



Appendix A: Numerical Example of the Value of Waiting to Deliver ETF Shares

To illustrate the incentive a risk-neutral AP might have to wait and deliver shares at a later date (e.g., at T+6 days) rather than immediately creating new ETF shares to cover a short position related to an arbitrage opportunity, we have developed the following numerical example. In this model, we formulate estimates of the profit potential for two alternative strategies to cover a hypothetical short position of 100 shares: 1) sell ETF shares at time t=0 at the current market price, P_0, and then immediately place a creation unit order for 100 shares with the ETF plan sponsor by purchasing the underlying securities in the ETF basket at the current Net Asset Value (NAV_0), or 2) sell ETF shares at time t=0 at the current market price, P_0, and then enter a long futures position on the underlying ETF at t=0 with a futures price of F_0 to hedge and "lock in" an arbitrage profit today between the ETF's current market price (P_0) and the futures price, F_0. However, in this second strategy, the AP will then *wait* until t=6 to place a creation unit order for, ideally, *less than* 100 shares (thus avoiding some of the costs associated with creating these new ETF shares).[35] We refer to the first strategy as the "Short and Create" method and the second strategy as the "Short and Hedge, then Create" approach.

In order to formalize the payoffs to these two strategies, we present the following formulas:

Short and Create's profit: $\pi' = \{(P_0 - NAV_0) - (f + \lambda)\}OIB_0$ (A1)

Short and Hedge, then Create's profit: $\pi = \{(P_0 - F_0) - (f + \lambda)(1 - \gamma) - c\}OIB_0$ (A2)

where,

f = the creation unit fee (expressed as a dollar amount per ETF share),

c = the cost to hedge in the futures market (expressed as a dollar amount per ETF share),

OIB_0 = the number of shares the AP initially shorts to offset the positive buy-sell order imbalance caused by other traders' excess demand for the ETF's shares at t=0, and

[35] In this set-up, we abstract away from fixed, minimum creation unit sizes and allow the AP to create ETF shares for whatever the exact amount of shares the AP has shorted. In addition, for simplicity, we assume that the explicit transaction cost for the AP to trade the ETF shares is zero (i.e., the AP does not incur any commission / brokerage costs to buy or sell the ETF).

λ = the "market impact" cost purchasing shares of the underlying basket of securities held by the ETF. This is also expressed as a dollar amount per ETF share and represents a linear cost for trading the underlying basket related to the AP's initial short position (OIB_0). One can view this as a cost paid to liquidity providers in the underlying securities to compensate them for their risk in trading with more informed traders, as in a Kyle (1985) model, or to cover inventory holding and order processing costs. For simplicity, we use a linear relation but a function that is convex in OIB_0 (e.g., a quadratic term) could also be used to increase the market impact costs for larger AP short positions. This alternative function would only favor waiting to deliver even further and thus we use the simpler, more conservative linear relation which allows the Short and Create strategy a better chance of out-performing the Short and Hedge, then Create strategy.

γ is the percentage of shares from the AP's short position that is expected to reverse over the 6-day waiting period. This "order reversal" parameter is a key determinant of the trade-off between the profit potentials for the two competing strategies. If $\gamma = 0$, then the AP will have to incur the market impact and creation costs on 100% of the short position and thus will cause the Short and Hedge, then Create strategy to be more costly than the Short and Create method. However, if $\gamma = 1.00$, then all of the order flow reverses over the 6-day period and the AP can simply purchase the ETF shares in the secondary market to cover the initial short position without having to incur the creation fee and market impact costs associated with creating some ETF shares by buying shares in the underlying basket of securities.

$F_0 = NAV_0 \cdot (1 + {}^{R}/_{365})^{(T)}$ is the futures price at t=0 which, for simplicity, is based solely on the ETF's NAV_0 and the daily risk-free rate (*R / 365*). This contract is assumed to expire exactly in T=6 days so that the futures price converges to the ETF's NAV at t=6 and the arbitrage opportunity disappears at that time as well (i.e., $F_6 = NAV_6 = P_6$ so that no arbitrage exists between the futures, NAV, and ETF prices).[36]

[36] These assumptions about convergence to the same price at t=6 are made to simplify the calculations but the main insights of the model would remain unchanged if we were to allow for some divergence in these prices at the time of delivery.

Since the AP is risk-neutral, the difference between the above two payoffs equals what we call the "Value of Waiting."

$$\pi - \pi' = \{(NAV_0 - F_0) + (f + \lambda)\gamma - c\} \cdot OIB_0 \;=\; (\{NAV_0 - F_0 - c\} \cdot OIB_0) + (f + \lambda) \cdot OIB_0 \cdot \gamma$$

(A3)

The second equality in the above equation re-arranges the variables so that one can see that the Value of Waiting is a linear function with the first term representing a constant $((\{NAV_0 - F_0 - c\} \cdot OIB_0))$. The first term can be viewed as a constant because all of these parameters are known at t=0. The second term includes a slope $((f + \lambda) \cdot OIB_0)$ and a single independent variable (γ). Similarly, the slope term is also known at t=0. Thus, the only unknown variable in the above model is the percentage of shares which will reverse over the course of the 6-day waiting period (γ). Although this percentage could be forecasted by the AP with varying degrees of accuracy, it is not known with certainty at t=0 because market conditions and investor actions can cause γ to fluctuate over the 6-day window.

Based on Equation (A3) presented above, we create a numerical example by assuming specific values for the model's parameters and then varying the level of γ between 0 and 1.00.[37] Figure A1 displays the trade-off between the two trading strategies and shows that the Short and Create strategy is more profitable whenever γ is below 0.169 (i.e., less than 16.9% of the order flow reverses). In contrast, the Short and Hedge, then Create strategy is more profitable above this break-even value of γ. Thus, when γ is greater than 0.169, the AP will have an incentive to use a long futures position to hedge the initial short position and then wait to create ETF shares for only the portion that does not reverse (i.e., for (1 - γ) of

---

[37] We assume the following values: P_0 = \$12.00 per share, NAV_0 = \$10.00 per share, F_0 = \$10.003 per share, λ = \$0.01 per share, *c* = \$0.0001 per share, *f* = \$.01 per share, *R* = .02 (i.e., 2% per year), and γ varies from 0.00 to 1.00. We also assume that the ETF's market price, NAV, and futures price all converge to \$11.00. For example, at γ = 0.40 (i.e., 40% of the order imbalance reverses), the Value of Waiting favors the Short and Hedge, then Create Strategy with a 6-day return of +0.23% in excess of the alternative Short and Create strategy. This gain is computed as a percentage of the Short and Create strategy's profit. On annualized basis, this represents a 15.20% return associated with waiting. As Figure A1 illustrates, the Value of Waiting varies greatly from -0.17% to +0.84% over the interval of γ = 0.00 to 1.00.

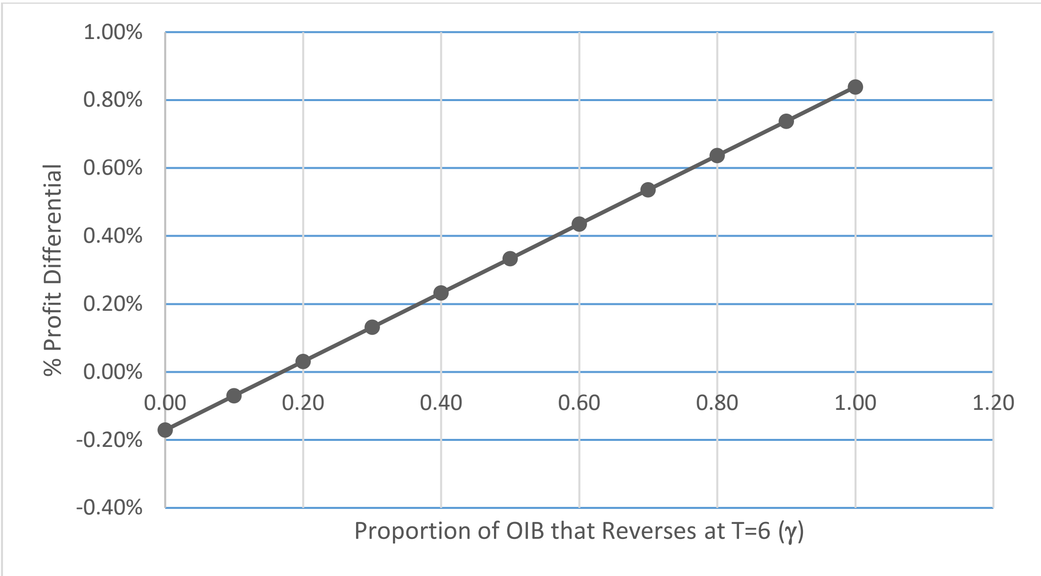
OIB_0). In effect, by waiting, the AP can avoid incurring the creation fee and market impact costs ($f + \lambda$) for that portion (γ) of the initial short position (OIB_0).

Figure A1 shows there is a clear trade-off between the two trading strategies and that the predictability of reversals in order imbalances can dictate which approach is most profitable for a specific ETF within a particular set of market conditions. Since we observe in our empirical results a large degree of operational shorting and FTD activity within ETF markets, one can surmise that the incentives to wait are more likely to outweigh the incentives to immediately create new shares to cover an AP's shorting activity. Thus, the "Value to Waiting" appears to be quite large for many APs in the U.S. ETF market. So, even though the numerical example presented here is fairly straightforward, it captures the main factors affecting the AP's decision-making process. Interestingly, our results are consistent with Nutz and Scheinkman's (2017) continuous-time model of trading among risk-neutral agents with heterogeneous beliefs when there are positive, convex costs of carrying a long position. In their model, the risky asset's supply and the associated carrying costs can interact to create situations where the "option to delay" (i.e., to wait and trade at a more favorable price in the future) affects the pricing of the asset.

One could also extend the above model in several ways. For example, although the trade-off outlined here is linear, the relationship could be nonlinear if AP's are assumed to be risk-averse and/or market impact costs are convex in the level of order imbalances. Also, another extension of the above model could incorporate order flow volatility as an alternative variable to describe the AP's uncertainty in terms of whether to choose to wait and deliver at T+6. For example, rather than use the order reversal parameter (γ), we could use the variance of order flow as another factor that affects the AP's choice between the two strategies discussed above. The above extensions are beyond the scope of the current analysis but, even if incorporated, the insights of basic model outlined here related to the trade-off between costs and benefits of the two strategies would remain intact.

Figure A1. The Value of Waiting

The chart below displays the trade-off between the payoffs to the Short and Create vs. the Short and Hedge, then Create trading strategies. The net payoff values are determined by Equation (A3) and the parameter assumptions described in Appendix A, as well as variations in the percentage of the initial order imbalance (OIB_0) that reverses over time (γ). Positive values indicate that there is an incentive for APs to wait and deliver ETF shares at the end of the 6-day trading window. Negative values represent levels of γ where the AP should not wait to deliver the shares and instead pursue the Short and Create strategy. The *% Profit Differential* is expressed as a percentage of the Short and Create strategy's profit level. The break-even point where the two strategies yield the same profit occurs when $\gamma = 0.169$ based on the model's parameter assumptions.



Appendix B: Definitions of Key Variables in the Analysis

This table presents definitions and sources for key variables used in our analysis.

Dependent Variables	Definition	Source
Short Interest/Shares Outstanding	The number of shares for the ETF sold short, divided by the total number of the ETF's shares outstanding.	Compustat; Bloomberg
Fail-to-Deliver/Shares Outstanding	The number of ETF shares not delivered on time, divided by the total number of the ETF's shares outstanding.	NSCC via the SEC: http://www.sec.gov/foia/docs/failsdata.htm
Net Create/Redeem Activity	The change in the ETF's shares outstanding from *t-1* to *t*.	Bloomberg
ETF Order Imbalance	The difference between buy- and sell-orders for the ETF.	NYSE TAQ database
Operational Shorting/Shares Outstanding	The buy/sell imbalance for trading the the ETF minus the change in share creation for the ETF, normalized by the ETF's shares outstanding.	NYSE TAQ database; Bloomberg
1-Month Forward Looking ETF Return	The percentage change in the price of the ETF from *t+1* to *t+22*.	CRSP
Mispricing Change	The difference between the ETF market price and NAV as a percentage of the ETF price.	Bloomberg
Absolute Mispricing Change	The absolute value of the difference between the ETF market price and NAV as a percentage of the ETF price.	Bloomberg
Average Intraday NBBO Spread of Underlying Stocks in ETF Basket	The average intraday national best bid and offer (NBBO) spread of stocks in the ETF basket, weighted by the size of the trade that immediately follows this NBBO quote	NYSE TAQ database
Average Intraday Second-by-Second Return Volatility of Underlying Stocks in ETF Basket	The intraday volatility of stocks in the ETF basket, calculated using second-by-second returns, computed from the last traded price recorded in each second	NYSE TAQ database
Financial Stress Index	An index compiled by the St. Louis Federal Reserve that combines 18 different indicators of financial system stress.	St. Louis Federal Reserve – FRED Database
Independent Variables		
log(Market Cap)	The natural logarithm of the ETF's market capitalization	Bloomberg
Share Turnover/Shares Outstanding	The volume of ETF shares traded each day, normalized by total ETF shares outstanding	Bloomberg and CRSP
Daily Cost of Borrow Score	The daily cost of borrowing based on a decile rank score of lending fee, where 100 equals the highest securities borrowing cost	Markit Securities Finance Database (formerly Data Xplorers)
Available Options Dummy	A proxy for the ability to use the ETF options markets to hedge a long or short exposure of an ETF	OptionMetrics
ln(Creation Unit Dollar Size)	The natural log of the dollar value of the size of the creation of a new ETF unit	ETF Global database

Creation Unit Fee per Share	The fee per share of creating a new ETF unit	ETF Global database
Maximum Rolling R-Squared with Available Futures Contracts	The roll assumption used in constructing the daily futures returns is the 'last-trading-day' or 'end-to-end roll' method	Quandl
Discount (in absolute value)	The absolute value of ETF mispricing, conditional on negative ETF mispricing	Bloomberg
Reversal Proxy	The past 22-day return of the ETF	CRSP
Momentum Proxy	The past 12-month return of the ETF with one month reversal	CRSP
Institutional Ownership	The total shares of the ETF owned by institutions, normalized by total shares outstanding	Thomson-Reuters 13F Database
Idiosyncratic Volatility	The standard deviation of the residuals from a 200-day rolling regression of excess returns on Fama-French 4 factor model	CRSP
Average ETF Ownership in Underlying Stocks in ETF Basket	The average ETF ownership of stocks, calculated over all stocks held by the ETF	Thomson-Reuters Mutual Fund Ownership database, Blomberg
Affiliated Lead Market Maker Capital Constraints	The ratio of net capital required to adjusted net capital for the ETF's affiliated lead market maker	CFTC's Futures Commission Merchants Financial Reports
Market-Wide Capital Constraints	The ratio of net capital required to adjusted net capital for all market participants, net of the ETF's affiliated lead market maker	CFTC's Futures Commission Merchants Financial Reports