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: November, 2018

Abstract

We identify an alternative source of ETF shorting related to the market maker liquidity provision and creation/redemption activities. This "operational shorting" arises due to a regulatory exemption, allowing ETF market makers to satisfy excess demand in secondary markets by selling ETF shares that have not yet been created. We find that operational shorting is associated with improved liquidity and greater price efficiency in the underlying securities held by an ETF, and with short-term return reversals consistent with liquidity supplying motives rather than informed trading. Delayed ETF creation to cover operational shorts results in failures to deliver and is found to be a valuable option in the presence of a liquidity mismatch between the ETF and the underlying securities. Operational shorting can lead, however, to increased counterparty risk and trading linkages between liquidity providers. We document a commonality in operational shorting across ETFs that share the same lead market maker and find that financial leverage can amplify this commonality.

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

JEL Codes: G1, G12, G14, G23

We are thankful for helpful comments and feedback from Aleksandar Andonov, Itzhak Ben-David, Jesse Blocher, Ethan Chiang, Francois Cocquemas, Shaun Davies, Darrell Duffie, Austin Gerig, Wes Gray, Jeff Harris, Pankaj Jain, Bryan Johanson, Qiang Kang, Dolly King, Stewart Mayhew, Tom McInish, Suchi Mishra, Jason Morrison, Adam Reed, Matt Ringgenberg, Vivek Sharma, James Simpson, and Jack Vogel, as well as participants at the Chicago Financial Institutions Conference, European Finance Association Annual Meeting, FINRA-Columbia Market Structure Conference, 5th Luxembourg Asset Management Summit, Financial Management Association Annual Meeting, Philly Five Research Symposium, Nasdaq – Villanova Synapse 2018 Conference, Southern Finance Association Annual Meeting, VSB Mid-Atlantic Research Conference in Finance, and seminar participants at the Federal Reserve Board, Federal Reserve Bank of Richmond (and its Charlotte and Baltimore branches), Florida International University, 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

U.S. exchange traded funds (ETFs) manage in excess of $3.5 trillion[1] and account for over 30% of the dollar volume of U.S. trading.[2] Central to the rise of ETFs is the arbitrage mechanism underlying liquidity provision in this market. Market makers and other institutions categorized as "Authorized Participants" (APs), can arbitrage the difference in price between the ETF itself and the basket of securities underlying that ETF.[3] In the process, APs create new shares of the ETF when there is excess demand for the ETF (i.e., the ETF price is greater than the net asset value of the basket of underlying securities, commonly referred to as the NAV). Conversely, these APs redeem shares when there is excess supply (i.e., selling pressures driving the ETF price to be less than the NAV). In these situations, we show that the classic definition of arbitrage as a set of nearly instantaneous transactions (e.g. buy the underlying basket of securities, swap for ETF shares, and sell ETF shares) does not necessarily apply. The AP's unique ability to sell ETF shares they do not have yet – to be 'operationally short' those shares – creates new, unexplored short-selling incentives for APs.

Our study relates to a growing academic literature that examines this ETF arbitrage mechanism and its economic implications.[4] While a key assumption in this literature is that arbitrage activities require the AP to trade in both the ETF and the underlying basket of securities, a unique exemption from SEC delivery requirements (Rule 204) allows APs/market makers to bypass trading in the underlying market altogether:

> "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

---

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

[2] Robin Wigglesworth, *Financial Times*, "ETFs are eating the US stock market", 1/24/17.

[3] 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. Many APs (but not all) are market makers and vice versa. The Reg SHO Rule 204 exemption and the trading dynamics described in our paper pertain to market makers and to APs who are market makers. Antoniewicz and Heinrichs (2015) reports that an ETF has, on average, around 5 APs that are active, and are registered market makers with obligations to provide continuous buy and sell quotes for ETF shares on secondary markets. We assume that an ETF market maker is also an authorized participant or has an agent with AP agreement with the ETF sponsor, and so we refer to such a market maker interchangeably in the paper as AP or MM.

[4] For example, see Ben-David, Franzoni and Moussawi (2018), Da and Shive (2018), Israeli, Lee, and Sridharan, (2017), Malamud (2015), Pan and Zeng (2016), Bhattarcharya and O'Hara (2016), and Glosten, Nallareddy, and Zou (2017).

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..." [5]

Under this market maker exemption, an AP can sell new ETF shares that are not yet created to satisfy a bullish order imbalance,[6] while opting to delay the physical share creation – that is, purchasing the basket of underlying securities and swapping that basket for the corresponding number of ETF shares – either until a later date or not at all, if the trade imbalance reverses (i.e. the excess demand for ETF shares shifts to excess supply). Put another way, the AP has the option to sell ETF shares and then fail to deliver them at settlement.[7]

We propose a simple and novel way to measure the use of this exemption, and show that our measure is consistent with the economics behind the proposed mechanism. When the AP sells ETF shares but postpones their creation and delivery, that delay is effectively a form of short-selling. The AP owes or is short the ETF shares until they ultimately deliver those shares to the investor who purchased them in the secondary market. We call this practice '*operational shorting*' and measure it by comparing two quantitites: a) the buy-sell trade imbalance (measured using signed intraday trade data) of an ETF to proxy for the purchase/sale of ETF shares by investors and b) changes in the daily shares outstanding of an ETF to proxy for the delayed, or non-contemporaneous, net share creation activity. If the buy-sell trade imbalance is

---

[5] 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.

[6] While we focus on the creation process, a similar option exists with respect to redemption: the AP could purchase shares of the ETF without ever redeeming them for the underlying. We focus on creation instead of redemption for three reasons. First, stockpiling shares of the ETF constitutes inventory risk for the AP, whereas the sale of ETF shares which have yet to be created represents a delivery or counterparty risk for the investor who purchased the shares. Second, since the sale ETF creation delivery option constitutes both a short-sale and if delivery takes place after T+3, a failure-to-deliver (FTD), we have a point of validation for our operational shorting measure. Third, given the dramatic increase in ETF assets over our sample period, the scenario of excess demand for the ETF is more prevalent than an excess supply scenario.

[7] In a letter to investors dated March 13 2019, a leading ETF market maker discussed our paper's findings acknowledging that "Market makers don't always immediately create new ETF shares in order to settle a short sale. Use of the Reg SHO exemption likely is responsible for high levels of short interest and high fail-to-deliver rates. Short selling by market makers using the Reg SHO exemption isn't "informed" trading, and it does help to prevent liquidity-demand in ETFs from creating volatility in the prices of the underlyings."

positive at a given point in time but there is no contemporaneous creation of the ETF shares, then the AP is operationally short those shares because they have yet to create and deliver them to investors. When we examine the determinants of operational shorting, we find that operational shorting in an ETF's shares is driven by: 1) a higher liquidity mismatch with the ETF's underlying basket of securities and 2) the presence of efficient hedges. These 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.

The use of this market maker exemption, or operational shorting, has four important implications which we explore in this paper. First, it creates a "disconnect" between trading in the ETF and trading in the underlying securities. On one hand, this disconnect could help mitigate the transmission of volatility in the ETF price to volatility in the underlying. On the other hand, it could impact the information transmission mechanism between the ETF and the underlying, thus altering price efficiency in these markets. We test these two competing hypotheses by examining the relation between ETF operational shorting activities, stock volatility, and the best bid and offer spreads on an intraday basis. Consistent with Ben-David, Franzoni, and Moussawi (2018) and a growing literature on ETFs,[8] 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*."[9] 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.

[8] 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.

[9] As we discuss in later sections, this evidence is consistent with prior literature (Fotak, Raman, and Yadav (2014), and Merrick, Naik, and Yadav (2005)) arguing that settlement failures can serve as an "important release valve" that removes any binding constraints on market participants' ability to supply liquidity and perform valuable arbitrage activities.

Second, we test to see if operational shorting increases ETF short interest and failures-to-deliver (FTDs).[10] Recent enforcement actions[11] and the increasing role played by ETFs in short-selling activity have been interpreted by academics, regulators, and practitioners as an indication of naked or abusive short-selling practices.[12] However, if the exercise of this market maker exemption is the source of the high short interest and FTD numbers, not only does it mitigate the concern about abusive short-selling, it may be an indication the exemption is being used as intended by regulators – to enhance liquidity. We find evidence that operational shorting is associated strongly with both short interest and FTDs. Figure 1 shows a high correlation between operational shorting and FTDs at an aggregate level. Repeating the analysis at the ETF level and controlling for 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-related imbalance that is greater than the number of shares created).

Third, if market making activities by the AP contribute strongly to the high observed short interest and FTDs, this has important implications for the information content of such short selling indicators. 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. Operational shorting, in

[10] The overwhelming percentage of FTDs across all securities that are associated with ETFs (78% in 2016) suggests potentially even greater economic implications for the ETF creation option.

[11] 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.

[12] 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.

contrast, is motivated by liquidity provision and does not necessarily have the same predictive power for future returns.[13] In analyzing the relation between operational shorting and future return, we find that operational shorting is negatively related to the following week's ETF return, but there is no statistically significant relation between operational shorting and the future return on the underlying securities. This result, combined with the strong positive relation between operational shorting and concurrent ETF returns, suggests that APs observe contemporaneous price pressures due to excess demand of ETF shares and anticipate the short-term ETF price reversal but their operational shorting is not informative about the value of the underlying securities. We find further evidence that operational shorting is liquidity driven by splitting the sample. The statistically significant negative relation 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. These results have important implications for the extant short-selling literature because they underscore 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.

Fourth, the common incentives of APs to exercise this exemption have implications for counterparty risk and potential financial contagion. If the use of this exemption is correlated within and across APs, it has the potential to increase the possibility of financial contagion through a commonality in

[13] Our paper is also not the first to examine the liquidity implications of short sales. Focusing on intraday equity short-selling, Comerton-Forde, Jones, and Putnins (2016) separate "liquidity-demanding" from "liquidity-supplying" short sales and find return and liquidity implications similar to our paper. Boehmer, Jones and Zhang (2008) examine the information content of shorts, finding cross-sectional differences in the future return implications of different sources of short-selling (e.g. institutional non-program shorts). This exemption to provide liquidity with regard to security creation, while novel, is not unprecedented in other fields. For example, in Edwards and Hanley (2010), the "Green Shoe" or overallotment option granted to underwriters following an IPO and the option granted to specialists on the NYSE to sell securities short.

the liquidity provision activities of an inter-connected network of ETF APs and market makers.[14] This mechanism can be important, because ETFs, as hybrid investment vehicles, form an essential nexus among several areas of the financial system.[15] To assess the potential for operational shorting to serve as a contagion mechanism, we examine the intra- and inter-firm correlation in operational shorting/FTD activity. We find that increases in operational shorting and FTDs for one ETF are related both to the operational shorting/FTDs of other ETFs that are traded by the same AP and operational shorting/FTDs for ETFs traded by other APs. We also examine the impact of leverage on this phenomenon and find that the financial leverage of an AP is positively related to operational shorting and FTDs. Overall, this presents a possible scenario for contagion. A spike in operational shorting/FTDs, coinciding with a drop in liquidity in the ETF by an AP in one ETF, could create a ripple effect to other ETFs within the same AP, and to other APs making markets in a common set of ETFs. This ripple effect, magnified by how close each AP is to their regulatory leverage constraint, could reveberate 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. Although we find operational shorting increases the liquidity of ETFs on average, these results suggest it does so at the cost of greater inter-connection within and between APs, an effect magnified by financial leverage.

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 main results for operational shorting. Section five presents the benefits and costs of operational shorting, and section six concludes.

[14] 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 participants' overall activity. Antoniewicz and Heinrichs (2015) report similar numbers of active APs and APs registered as market makers, suggestive that the lead market maker would be a viable proxy for an active AP.

[15] 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. 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.

2. ETF Market Making and Operational Shorting

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.[16] Similar to stocks and closed end funds, ETF shares trade on exchanges, and this 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 market makers, 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.[17] 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.[18] 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 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

---

[16] 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.

[17] 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. Antoniewicz and Heinrichs (2015) reports that an ETF has, on average, around 5 APs that are active, and are registered market makers with obligations to provide continuous buy and sell quotes for ETF shares on secondary markets. We assume that an ETF market maker is also an authorized participant or has an agent with AP agreement with the ETF sponsor, and so we refer to such a market maker interchangeably in the paper as AP or MM.

[18] 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.

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 securities. Note that this redemption process is generally 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 ETF Arbitrage and Operational Shorting

As discussed earlier, an important objective of APs in the primary ETF market is to arbitrage the difference between the ETF's market price and its NAV, or the price of the underlying securities that comprise the ETF basket. As demand for the ETF from investors in the secondary market grows, the ETF's market price should increase, potentially creating a 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.[19]

APs and ETF market makers are able to accommodate intraday demand using the flexibility of the multi-day settlement window. They can short sell ETF shares to meet excess buying demand in the intraday market and then hedge their exposure until they backfill supply at a later date with the creation/redemption

[19] As *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.

process and cover their short positions. Therefore, 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 call an *operationally short* position.

2.3 Incentives to Operationally Short

There are a number of operational reasons why an AP might want to delay creation/operationally short. 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 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 trading costs, the ETF's 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.

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 block for 50,000 shares or 2 blocks for 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 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 in the secondary market. 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 the trade-off between the costs and benefits of an AP covering a short position either immediately or waiting for up to 6 days is 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).

While the high level of ETF short interest and settlement failures, 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. 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 occurs.[20] This can happen due to operational shorting, as part of bona fide market making activity, as well as by directional shorting via 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.

While the literature on equity FTDs 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. (2018) demonstrate that ETFs may have consequences for the volatility of financial markets. Furthermore, in contrast to earlier findings, Stratmann and Wellborn (2012) find that ETF-related FTDs Granger-cause higher stock market volatility and lower future returns which can ultimately lead to increased market instability. Additional institutional detail and a more in-depth review of the FTD literature can be found in Sections 1 and 2 of Appendix C.

[20] While a shortened T+2 settlement cycle was implemented for most securities on September 5, 2017 (SEC's final rule that amended Exchange Act Rule 15c6-1 to shorten the settlement cycle to t+2: https://www.sec.gov/rules/final/2017/34-80295.pdf), T+3 was the settlement cycle during most of our sample period.

2.4 Measuring 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 initiated by informed traders, that can also result in FTDs. To understand the intuition behind the measure, 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, APs "acting as market makers or agents to market makers" might submit a create order on that date and have three trading days (until t+3) to deliver the basket of underlying to complete the creation.[21] 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 2 contains an illustrative example of how the cumulative buy-sell trade imbalance, the 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 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 – indexed on right side axis) occurs in ITOT shares. It would appear that APs and market makers are accommodating the

[21] 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."

demand, but the delay in creating them generates the FTDs observed. While Figure 2 focuses on a single example, sections 3 and 4 of Appendix C examine the daily dynamics of buy/sell imbalances, creation/redemption activity and FTDs and show the insights of Figure 2 apply more generally across our entire sample.

The operational shorting measure we propose 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 might result due to the lagged response of APs/market makers to the excess demand. The formula for our measure of operational shorting is:

$$\begin{aligned} &\mathit{Operational\ Shorting} \\ &= \frac{\max[0, \left(\text{Cumulative Buy/Sell Imbalance}(t-3, t-1) - \Delta\text{Shares Outstanding}(t-1, t)\right)]}{\text{Shares Outstanding}(t\text{-}3)} \end{aligned} \tag{1}$$

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).[22] 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 for bona fide market making by an AP. We take the maximum of the buy-sell trade imbalance and 0 to ensure our measure captures only buy imbalances. We then subtract the daily net create/redeem activity, which is computed as the change in ETF shares outstanding from t-1 to t, because it is at time-t when prior short sales are expected to be covered. We normalize the result by dividing by the number of ETF shares outstanding at the start of this rolling window. 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.

[22] NBBO, which stands for the national best bid and offer, is obtained from the NYSE TAQ Daily (Millisecond Feed) Database. More details on the detailed signing algorithm are in the next section and in Appendix D.

3. Data

Because ETFs sit at the intersection of many different markets, our empirical analysis requires data from a number of different sources. A complete listing of variables, definitions, and sources is provided in Appendix B. The FTD data[23] are from the SEC's website and are made available to the SEC by National Securities Clearing Corporation's (NSCC).[24] 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.[25] We collect these data from March 22, 2004, which is the beginning of the FTD dataset, through December 31, 2016.[26]

We supplement the SEC data with additional variables from other sources. We merge the data with 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 volatility*. We gather ETF characteristics from the 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. We collect the ETF holdings of underlying stocks from the Thomson-Reuters Mutual Fund Ownership and CRSP Mutual Fund Holdings Databases. Buy and sell trade volume information, the intraday National Best Bid and Ask (NBBO) spread, and the intraday return volatility are calculated from the NYSE Daily TAQ database (Millisecond Feed). Short interest data are

---

[23] The FTD data can be downloaded from the following SEC page: http://www.sec.gov/foia/docs/failsdata.htm.

[24] 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 information.

[25] 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 only the aggregated outstanding balance of fails at a given day.

[26] 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.

extracted from Compustat on a biweekly basis, and represent 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 the 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 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 (2018) and extract the ETF shares outstanding data from Bloomberg because they are not not reported accurately in CRSP and Compustat.[27]

In order to compute the daily buy-sell imbalances in ETF shares, we need to first sign the ETF trades as buys and sells. For this, we use the NYSE 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). For each trade, we compute the national best bid and offer (NBBO) quote at the beginning of each millisecond. Then, we compare the trade price for all trades occurring during a millisecond to the prevailing best bid and best offer at the beginning of this millisecond. 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 sells on the bid prices.[28] 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

[27] Bloomberg sources the ETF shares outstanding data directly from ETF sponsors, administrators, and custodians for most ETFs. Bloomberg provides the new shares outstanding information reflecting accepted create/redeem orders in the after-market hours on the transaction date. While Bloomberg generally 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.

[28] 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.

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. Thus, our modified method takes into consideration the idea that buys are more likely to be executed at the ask price, and sells at the bid price, and whenever an outside trade is observed during that millisecond, then the algorithm adjusts dynamically and 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 sells at 4:00 pm to construct our buy and sell volume for the day. Appendix D provides a detailed discussion of our novel methodology to classify trades and provides several empirical tests that evaluate and compare our method to Lee and Ready (1991) and Holden and Jacobsen (2014).

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 99^{th} 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 Operational Shorting as a Driver of ETF Short Interest and Failures-to-Deliver

While the literature has used short interest and FTDs as measures of informed, directional short-selling, operational shorting may be an important component of ETF short interest and FTDs but the motivation behind operational shorting has very different implications than directional shorting. In Table 2, we first test whether or not operational shorting is an important driver of overall ETF short-selling activity. To address this question 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.

To capture alternative motivations for short-selling, we include the lagged *Short Interest Ratio* and the *Daily Cost of Borrow Score*. 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. The *Daily Cost of Borrow Score*, measures the cost of borrowing the ETF shares based on a rank score of securities lending fees (1 to 10, where 10 corresponds to ETFs with the highest borrowing costs). Similar to including lagged short interest, the inclusion of the *Daily Cost of Borrow Score* captures both directional and operational motivations for short-selling. As it becomes more expensive to borrow and short ETF shares, directional short-sellers and liquidity-providing APs become more likely to fail-to-deliver. In this case, APs are also less likely to borrow shares to hedge their position. While including these two variables may bias our results towards the null hypothesis that operational short-selling does not play a role in short interest or FTDs, including them will also give a more accurate assessment of the role of operational shorting in contributing to overall ETF shorting activity.

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 (2012) related to the effects of ETF liquidity and options. We control for the ETF's liquidity by including its size (*log of Market Cap*) and trading volume (*Share Turnover*). We expect the ETF's asset size to be negatively related to FTDs because larger funds are typically more liquid and thus it is easier to locate shares. Having controlled for the size of a given ETF, 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*) that equals 1 if options are traded on the ETF. All regressions used in our analysis include ETF and date fixed effects, and standard errors are clustered by ETF and date.

Table 2 shows that the coefficients on all variables 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* is a statistically significant determinant of short interest and FTDs. In regression (6), comparing the coefficients on *Operational Shorting* and lagged *Short Interest*, which are both denominated by shares outstanding, we see that the coefficient on *Operational Shorting* is economically larger. 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 short-selling that is due to liquidity provision, as measured by our *Operational Shorting* variable.

4.2 Incentives behind AP's Operational Shorting Activity

We next examine in Table 3 the determinants of operational shorting activity to confirm the ETF market making incentives that we describe in Section 2. The variables we include 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*).[29] We include ETF and date fixed effects and we also cluster standard errors by ETF and date. The number of observations decreases across the different specifications due to data availability issues.[30]

---

[29] Another incentive variable of interest is the ETF management fee, which, according to anecdotal evidence collected from conversations with several ETF market makers, might represent an important incentive for operational shorting. ETF market makers can capture this management fee in their operational shorting position especially when using a close hedge on the underlying basket. This variable is however subsumed in the ETF fixed effects as most ETFs do not have meaningful variation in their expense ratios during the bulk of our sample period.

[30] The *Creation Unit* variables are only available for a subset of ETFs. Similarly, the *Maximum Rolling R-Squared*

In this analysis, we examine the hypothesis that operational shorting activity is driven by AP's incentives to capture arbitrage profits resulting from buying pressure on ETF shares and the subsequent mispricing of the ETF relative to the NAV or the fundamental value of the underlying basket. To capture these arbitrage opportunities, we include the ETF *Mispricing* variable. This measure captures times when the ETF market price is at a premium (positive *Mispricing*) and a discount (negative *Mispricing*) relative to the NAV. However, 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 (*Premium* and *Discount*) which are equal to the *absolute* value of the mispricing variable only when the mispricing is positive or 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* to have a positive relation to *Operational Shorting*.

Additionally, we expect that the availability of futures and options to hedge the underlying are important determinants of operational shorting. These hedges would shield operationally short APs 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. *Maximum Rolling R-squred with Available Futures Contracts* measures how well available futures contracts correlate with the returns of the ETF's stated benchmark index. For each date, we regress the previous 252 days of ETF NAV returns on the futures return from the S&P 500-mini, the S&P MidCap 400-mini, and the Russell 2000-mini contracts.[31] The maximum R^2 across these three regressions is the value assigned to *Maximum Futures R-*

can only be computed for those ETFs that 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 intraday average relative spread variable are available.

[31] We collect the futures data from the Quandl website, 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.

squared. If an AP or other investor wanted to hedge their exposure to an ETF, this variable captures the suitability of using futures on one of the three equity indexes as a reliable hedging vehicle.[32] Options listed on the ETF would also facilitate the hedging of ETF-specific risk. We expect a positive relation between these hedging-related variables and *Operational Shorting* because the presence of hedging instruments can allow an AP to provide more liquidity when they can use the futures and/or option markets to hedge their 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 Pan 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. APs should 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 and frictions: *Creation Unit Fee* (for a single creation unit transaction) and *ln(Creation Unit Dollar Size)*. As discuss in Section 2, we expect higher creation unit sizes and fees to encourage APs to engage in more operational shorting in order to minimize these costs.

Regression (2) in Table 3 includes four main independent variables: the ETF liquidity-related control variables as well as the two proxies for ETF-specific transaction costs or frictions. We find in

---

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)."

[32] 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 later, 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).

model (2) that 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 waiting 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, thus saving on the capital outlay required to accumulate the requisite shares in the underlying securities. However, in our full model, regression (7), we find that these institutional frictions are no longer significant, as other factors are stronger determinants of operational shorting activity.

Regression (3) adds proxies for the availability of 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. Additionally, arbitrage profits are an important driver of AP's market making and related operational shorting activity, as expected. 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). Models (4), (5), and (7) confirm our expectation that a 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 Model (7)). In addition, consistent with the notion that greater liquidity mismatches are positively related to operational shorting, models (5), (6) and (7) show that *Proxy for Liquidity Mismatch* is positively and significantly related to *Operational Shorting* (although the relation is statistically not as strong as the *ETF Mispricing* variable).

Taken together, the results in model (7) suggest that smaller, actively traded ETFs that have greater potential profits from capturing the ETF premium, with available hedging alternatives, and with larger liquidity mismatches have a greater propensity for 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 delay could alleviate demands for liquidity in the underlying securities market while also increasing FTDs of ETF shares. The results from Table 3 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. Our findings in Table 3 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.

4.3 ETF Operational Shorting and Liquidity Provision

Our previous results show that operational shorting is an important component of ETF short interest and FTDs and there is a well-established literature documenting the relationship between increased short-selling activity and future underperformance using short interest, FTDs, and a variety of other short-selling measures.[33] 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. At the same time, Comerton-Forde, Jones, and Putnins (2016) find evidence not only of liquidity demanding, informed short-selling, but also liquidity-supplying shorts. Even though the focus of Comerton-Forde, Jones, and Putnins (2016) is on individual equities, their evidence that liquidity-supplying shorts are strongly contrarian in nature and improve market quality, provides a useful comparison for 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 this caveat, examining whether or not operational shorting activity is consistent with ETF liquidity provision is an important verification that indeed these short-sales represent legitimate market making activities.

[33] 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 similar: constrained short-selling is associated with over-valuation.

In this section, we examine the motivation behind operational shorting by regressing ETF and underlying returns on our measure. Having sold ETF shares to satisfy excess demand and the corresponding mispricing, APs must then choose to create and deliver the ETF shares or to operationally short. If APs believe that the original mispricing was liquidity driven, then they are more likely to operationally short in anticipation of the expected reversal. Otherwise, APs would be more inclined to purchase the underlying, create the ETF shares and deliver them to the buyer. In thinking about the decision to operationally short, the AP would also likely take into account the liquidity of the underlying relative to the ETF. When the ETF is more liquid than the underlying (i.e., a high liquidity mismatch), it is more likely to attract liquidity investors and result in a reversal.

Using return data for both ETFs and the funds' underlying assets, we examine the impact of operational shorting on both contemporaneous and future returns. While our measure of operational shorting is anchored along the three-day settlement window, and is measured on a rolling basis, we run our return regression analysis over non-overlapping discrete weekly intervals, comparable to the timing of the operational shorting measure plus the additional settlement time afforded APs by their exemption. For each ETF, and on each day, we first compute the daily risk-adjusted excess returns using the Fama-French four factor model. We then regress the cumulative weekly excess return on a number of ETF characteristics.

The key independent variables in these regressions are *Create Orders – Weekly* and *Operational Shorting - Weekly*. *Create Orders - Weekly* is the change in ETF shares oustanding during the week if shares are created, and zero otherwise. *Operational Shorting - Weekly* is the excess demand (total positive buy-sell trade imbalance) of ETF shares during the week minus all create orders during this week, and zero otherwise.

Table 4 reports the results of this test. Panel A provides the descriptive statistics for all variables used in this weekly return analysis. *Total Return* and *Fama-French 4-Factor Excess Return* are reported in percentage terms and are winsorized, along with all other variables used in the regressions in Panel B. These two variables average 7.3 bps and -6.4 bps during the sample period, with a maximum return of 12.35% and 7.17% respectively. *Operational Shorting - Weekly* is comparable in magnitude to *Create Orders -*

Weekly (with averages of 1.49% and 1.59%, respectively, and standard deviations of 5.0% and 5.8%).

Panel B of Table 4 reports the regression results. At the top of each column, the table specifies if the dependent variable is calculated using *ETF* or *NAV* returns, if it is total return (*Ret*) or Fama-French 4-Factor Excess Return (*FF4* α), and whether it is measured over the concurrent week (*t*) or the following week (*t+1*). The regression specifications also differ by the sample used which is indicated in the last two rows of the table. Specifications (1) through (3) use the entire sample of ETFs and (4) through (6) use the sample of non-U.S. equity ETFs (e.g. fixed income and foreign equity ETFs). Specifications (7) through (10) use the domestic equity ETF subsample and in specifications (9) and (10), further split this sample into ETFs with high and low liquidity mismatches, where *High* indicates the ETF was more liquid than the underlying assets as measured by the difference in intraday spreads between the two.

In regression (1) of Panel B for Table 4, we find that ETF *Operational Shorting* is related to higher contemporaneous ETF returns consistent with the evidence from Table 3 of mispricing arbitrage profits as a primary motivation behind operational shorting. Using total returns as the measure of interest, specification (4) finds similar results for non-U.S. equity ETFs. In specifications (2), (5) and (7), we find that *Operational Shorting* is negatively related to future ETF returns for all types of ETFs, suggesting return reversals following operational shorting by APs. This finding documents the contrarian nature of operational shorting as APs respond to investors' buying demand. While this evidence is consistent with the liquidity provision motivation for operational shorting, it stands in sharp contrast to the insignificant evidence related to create orders. Economically, an average increase in operational shorting (around 1.49% of shares outstanding) is associated with a 1.98% contemporaneous change in Fama-French 4-Factor Excess Return during the same week. This corresponds to an increase that is equivalent to 1.04 times the standard deviation of excess returns, 18% of which reverses in the following week (or about a -0.35% total reversal in weekly returns, on average) for all ETFs. Reversals are 50% stronger for non-U.S. equity ETFs, and 14% stronger for U.S. equity ETFs with larger liquidity mismatches.

Interestingly, price pressure on the ETF shares does not translate into price pressure on the underlying basket of stocks, consistent with operational shorting mitigating the transmission of the ETF liqudity shocks to the underlying securities. In regressions (3), (6) and (8) we repeat the analysis using returns on the underlying securities (*NAV*) instead of ETF market prices to calculate our performance measures. Whether we use four factor alphas for the overall sample or the U.S. equity subsample (i.e., (3) and (8)) or total returns for the non-U.S. equity subsample (6), we find the same results: operational shorting activity in the previous week has no predictive power for returns on the underlying securities.

To better understand what drives this ETF return predictability, we turn to the subsample regression analysis in specificiations (9) and (10). These two specifications sub-divide the U.S. equity ETF sample into 'Low' and 'High' liquidity mismatches. As before, we measure the liquidity mismatch as the difference between the the average intraday spread of the ETF basket of securities and the ETF's intraday spread. In the 'Low' liquidity mismatch sample, where the ETF and underlying securities exhibit similar liquidity to their underlying basket, we find no predictive power of operational shorting for returns. In the 'High' liquidity mismatch subsample, however, where the ETF is more liquid than the underlying basket, we find, as expected, a statistically significant, negative relation between operational shorting and future returns.

Overall, this evidence of liquidity provision is compelling: while the operational shorting activity of APs has some predictive power for a return reversal in ETF shares, it has no predictive power for the return on the underlying basket of securities held by the ETF. In other words, APs 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-driven demand. We also find that this effect is concentrated in the subsample where the underlying assets are less liquid than the ETF, where liquidity trades in the ETF are more likely relative to the underlying. In unreported results, we also find that the predictive power of operational shorting on ETF returns does not persist for returns further in the future (e.g., two or more weeks later), further suggesting that the effect of operational shorting is temporary due to liquidity provision rather than "directional" short-selling based on longer-lasting "permanent" changes in the ETF's fundamental value.

We interpret our evidence of operational shorting and its related effects on future price reversals as an indication of a market maker's ability to separate uninformed, liquidity-motivated order flow from informed order flow-- the latter flows being more likely to lead to price discovery. Our evidence also suggests that APs act strategically depending on how they interpret the order flow, and make profits not only from the mispricing arbitrage but also from potential reversals following liquidity-driven price pressures. These results are consistent with Kyle's (1985) 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. If an AP observes informed buying activity, there is no value in waiting and instead the AP has an incentive to cover the short ETF position quickly by purchasing the underlying securities and issuing a create order. 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 by enhancing price discovery.

5. The Benefits and Costs to the Financial System of Operational Shorting

In the previous section we find that APs profit from arbitrage opportunities and by taking contrarian positions when there are price pressures due to excess liquidity demand. Our return results from the previous section also suggest that operational shorting may act as a buffer for the underlying basket of securities. In this section, we explore the implications of operational shorting for financial markets more generally by analyzing the impact of operational shorting on ETF mispricing and the liquidity and volatility of underlying securities. We also examine the potential for operational shorting activity to serve as a mechanism for financial contagion through an analysis of the correlation of operational shorting activity within and across APs, as well as the role of constrained leverage at the AP in operational shorting activity.

5.1 The Effects of Operational Shorting on ETF Mispricing

In Table 5, we examine the impact of *Operational Shorting* 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 (*Mispricing Change)* 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.

Across all eight specifications, we find that the level of *Operational Shorting* has a strong negative relation with ETF mispricing. Specifications (1) and (2) show that *Operational Shorting* is negatively related to the contemporaneous (signed) ETF mispricing variable, while models (3) and (4) show that the lagged *Operational Shorting* (at *t-1*) variable also leads to a reduction in ETF mispricing. This result is in line with our prior finding that operational shorting is incentivized by arbitrage opportunities in the form of ETF mispricing to reduce the size of these mispricing oportunities (consistent with recent papers on ETF arbitrage such as Brown, Davies, and Ringgenberg (2017)). Models (5) through (8) repeat these tests with the *Absolute Mispricing Change* as the depedent variable, 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.

5.2 The Effects of Operational Shorting on Underlying Securities

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 "operational shorting" on common stocks that are held by those ETFs. Fotak, Raman, and Yadav (2014), 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.[34] Accordingly, we follow Fotak et al. (2014), and use average spreads and

[34] 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

intraday volatility as measures of the liquidity and market quality of trading in individual securities in Table 6.[35] Our first measure, the *Average Intraday Spread* of the underlying stocks is computed in two steps: first, for each stock and on 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. The results using this measure are shown in Panel A. 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, aggregated at the ETF level in a similar manner to the spread measure. The results using this measure are shown in Panel B.

The basic structure of the empirical model is similar to the one used in Table 5 and includes contemporaneous and lagged values of *Operational Shorting*, which is our main variable 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 (2018) 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 any persistence in volatility and spread measures, as well as address potential reverse causality concerns.

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.

[35] In Table 6, our sample is restricted to U.S. equity-only ETFs because these are the only funds that we can reliably 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).

Panel A of Table 6 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 (up to three days) 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 lower liquidity and 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; Saglam, Tuzun, and Wermers (2019); Agarwal, Hanouna, Moussawi, and Stahel (2017)).[36]

In Panel B of Table 6, a similar set of regressions are run on the intraday volatility of the underlying stocks held by U.S. equity ETFs. We find that operational shorting is also negatively related to intraday return volatility after controlling for 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, 2018). We interpret our findings to suggest 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 spreads and the volatility of underlying stocks suggests that liquidity in the underlying stocks held by ETFs improves as operational

[36] 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. Agarwal, Hanouna, Moussawi, and Stahel (2017) find that ETF arbitrage mechanism exacerbates the co-movement in the liquidity of underlying stocks, and that the effect of ETF ownership on liquidity commonality is independent from that of the ownership by index mutual funds, active mutual funds, and other institutional investors.

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)). 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 could 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.

Panel C of Table 6 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 a variance ratio of 1.0 if the trading in these securities is 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 6 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 underlying securities held by ETFs improves as markets engage in higher levels of operational shorting activity. Consistent with the results in Table 4 related to when APs actively create new ETF units, it appears that ETF market makers could be observing the order flow and using these patterns to identify when ETF prices are likely to be subject to transient liquidity shocks. In turn, the market makers could be taking action in both the ETF and the underlying securities that absorbs these liquidity shocks from being transmitted to underlying basket prices, thus reducing the noise in stock prices and making the market for these securities more informationally efficient.

Overall, the evidence in Tables 5 and 6 suggests that operational shorting is used by market makers to profit from contrarian positions against liquidity traders and, as a consequence, such liquidity provision dampens the potentially adverse effects of ETFs on the volatility and liquidity of underlying stocks in their baskets, as well as enhances these stocks' price efficiency. Operational shorting thus acts 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. Such results represent an additional incentive for APs to delay covering these short positions, especially in ETFs with less liquid underlying baskets.

5.3 Operational Shorting and Financial Contagion

While operational shorting may enhance ETF liquidity and underlying price efficiency, 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."[37]

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

5.3.1 Potential Spillover Effects of FTDs and Operational Shorting Activity

Given that FTDs represent a delay in payment from one party to another, increases in the overall volume of FTDs might impair some market participants' ability to meet their other obligations in a timely way, leading to greater counterparty risk. In the case of operational shorting, when an AP delays share creation in one ETF due to a lack of liquidity, for example, the AP may also delay share creation in other ETFs for which they make markets. Similarly, operational shorting by one AP may may be correlated with operational shorting by other APs that are trading in the same or similar ETFs. In such a scenario, widespread delays in ETF share creation or, equivalently, increases in operational shorting, could spread across the market of APs and ETFs. Counterparty risk arises from such commonality despite the DTCC's role in settlement and clearing due to the fact that some inter-broker transactions are exempt from DTCC settlement.[38] Therefore, we expect counterparty risk to be more pronounced across market makers with higher operational shorting activity relative to the reported FTDs.

---

[37] 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.

[38] See SEC Rule 19b-4 regarding certain trades not subject to be reported to the DTCC between unaffiliated brokers because "they necessarily take place at the end of the day, after the common client has reviewed its end of day positions and has instructed the clearing brokers as to which positions it will move for custody purposes". See the proposal http://www.dtcc.com/~/media/Files/Downloads/legal/rule-filings/2015/nscc/SR-NSCC-2015-004.pdf and approval order in the Federal Register: Release No. 34-76462; File No. SR-NSCC-2015-004: http://www.dtcc.com/~/media/Files/Downloads/legal/rule-filings/2015/nscc/SR-NSCC-2015-004-Approval-Order-34-76462.pdf. Unfortunately, DTCC does not provide publicly the data pertaining to the fraction of daily volume that is not subject to DTCC settlement, which is likely to be higher in ETFs due to their higher non-ATS OTC volume (i.e., trades that are neither executed on exchanges nor in dark pools).

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, excluding the FTDs or operational shorts in the particular 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.[39]

In Table 7, 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 (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

[39] 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.

because these FTDs have been shown in our analysis to be more highly correlated with liquidity provision activities (which are non-directional in nature).

Regressions (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.

5.3.2 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.[40] We measure a broker's leverage by examining the firm's capital constraints. This is done by computing the following ratio:

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

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

This ratio is closely monitored by the CFTC because it measures a market maker's capital (the denominator) relative to the CFTC's net capital requirements (the numerator). 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 *Capital Constraints* as an independent variable in the regression to estimate the effect of financial leverage on FTDs and operational shorting.

Table 8 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, provides evidence of one potential channel 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.

6. Conclusion

This study is the first comprehensive analysis of how the liquidity provision activities of ETF authorized participants can have important effects on short-selling, failures to deliver, ETF and underlying security returns and market quality, as well as counterparty risk and potential market-wide contagion. We first propose, test, and find evidence that ETF short interest and FTDs are driven by "operational shorting," a manifestation of market making efforts when APs/market makers are faced with excess buying demand from investors.

We also document that operational shorting is a contrarian form of liquidity provision, as it coincides with high price pressure and predicts strong ETF return reversals in the following week, especially for ETFs that have greater liquidity mismatches. Operational shorting has no predictive power for returns on the underlying securities, however, inconsistent with an informed, directional motive for short-selling. Consistent with this liquidity provision explanation, we also find that operational shorting activities result in lower mispricing between the ETF and its basket's underlying securities, and act as a buffer that reduces the adverse effects of ETF liquidity shocks.

Our evidence, while emphasizing the unique create/redeem mechansim of ETFs, is consistent with prior literature on liquidity-supplying shorts by market makers and their impact on market efficiency and liquidity. 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 and liquidity provision. 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 operational shorting can enhance ETF liquidity provision, it may generate a negative externality in terms of higher levels of FTDs that spillover from one ETF to another within the same AP, and from one AP to another, because these firms make markets for ETFs with similar underlying securities. We find that increased levels of FTDs and operational shorting in the lead market maker's other ETFs can have a contagion-like effect. APs exhibit both within-firm and across-firm commonality in exercising this option to delay creation and that APs facing greater financial leverage constraints are more likely to exercise the option. 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 increase risk within the overall market.

Given that the above results have concentrated 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, rather than excess buying pressure. 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, since hedges are likely to be more expensive in that case, especially at times when heavy outflows from ETFs coincide with stressful market conditions. Thus, additional research into this asymmetry between and operational buying versus operational shorting is warranted.

References

Agarwal, V., P. Hanouna, R. Moussawi, and C. Stahel, 2017, Do ETFs Increase the Commonality in Liquidity of Underlying Stocks? Working paper.

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, 2018, Do ETFs Increase Volatility? *Journal of Finance* 73 (6), forthcoming.

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.

Comerton-Forde, C., Jones, C.M., Putnins, T.J., 2016, Shorting at close range: A tale of two types, *Journal of Financial Economics* 121, 546-568.

Da, Z., and S. Shive, 2018, Exchange Traded Funds and Asset Return Correlations, *European Financial Management* 24 (1), 136-168.

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.

Edwards, A.K., and K. Weiss Hanley, 2010, Short selling in initial public offerings, *Journal of Financial Economics* 98, 21-39.

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.

Glosten, L. R., S. Nallareddy, and Y. Zou, 2017, ETF Activity and Informational Efficiency of Underlying Securities. Working paper, Columbia Business School.

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

Holden, C., and S. Jacobsen, 2014, Liquidity Measurement Problems in Fast, Competitive Markets: Expensive and Cheap Solutions, *Journal of Finance* 69 (4), 1747 – 1785.

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.

Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46 (2), 733-746.

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.

Murayev, D., and J. Picard, 2016, Does Trade Clustering Reduce Trading Costs? Evidence from Periodicity in Algorithmic Trading, 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.

Saglam, M., T. Tuzun, and R. Wermers, 2019, Do ETFs increase liquidity? Working paper, University of Cincinnati.

Stratmann, T., and J. Welborn, 2012, Exchange-Traded Funds, Fails-to-Deliver, and Market Volatility, Working paper, George Mason University.

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

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 U.S. 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. We provide a complete list of variable names, sources, and definitions in Appendix B.

Sample	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 – The Determimants 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. We provide a complete list of variable names, sources, and definitions 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 3 – 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*. We provide a complete list of variable names, sources, and definitions 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 4 – Operational Shorting and Contemporaneous/Future Returns: The dependent variable in these regressions are 1-week contemporaneous (t) or forward-looking (t+1) total returns (Ret) or Fama-French 4-factor risk-adjusted alphas (FF4 α) measured in percentage terms. This measure is based on the ETF or NAV price as indicated in the header. Independent variables are measured at week *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 where market capitalization is measured in millions of dollars, the ETF's *Average Share Turnover %* (as a percentage of ETF shares outstanding), and the ETF's *Amihud Illiquidity*. We provide a complete list of variable names, sources, and definitions in Appendix B. Specifications 1-3 include all ETFs, 4 non-U.S.-equity ETFs, 5 through 7 U.S.-equity ETFs. Specifications 6 and 7 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, with lower intraday spread than the average intraday spread for basket stocks. The sample period is March, 22 2004 – December 31, 2016, all variables are winsorized, and t-statistics 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: Descriptive Statistics of the weekly ETF return sample

Variable	N	Mean	Std	Min	Median	Max
Total Return (%)	572,629	0.073325	3.225214	-13.010820	0.123325	12.354650
Fama-French 4-Factor Excess Return (%)	559,555	-0.063528	1.906160	-7.881115	-0.000072	7.176298
NAV Return (%)	572,629	0.031369	3.176574	-12.792320	0.080483	12.252790
Fama-French 4-Factor NAV Excess Return (%)	559,555	-0.093955	1.882294	-7.938895	-0.017297	6.959463
Operational Shorting - Weekly, scaled by Shares Outstanding	572,629	0.014866	0.049982	0.000000	0.000000	0.442983
Create Orders - Weekly, scaled by Shares Outstanding	572,629	0.015939	0.058057	0.000000	0.000000	0.500000
log (Market Cap)	572,626	-2.368839	2.230165	-10.290330	-2.415137	5.418783
Average Share Turnover	571,583	0.038239	0.105723	0.000000	0.009128	3.216983
Amihud Illiquidity Measure	563,492	0.164901	0.499989	0.000001	0.009850	13.367430

Panel B: Weekly return results for All ETFs, U.S. Equity ETFs, and all other non U.S. Equity ETFs (foreign equity and fixed income ETFs)

	Weekly Return									
	ETF	ETF	NAV	ETF	ETF	NAV	ETF	NAV	ETF	ETF
	FF4 α (t)	FF4 α (t+1)	FF4 α (t+1)	Ret (t)	Ret (t+1)	Ret (t+1)	FF4 α (t+1)	FF4 α (t+1)	FF4 α (t+1)	FF4 α (t+1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Operational Shorting - Weekly (t)	1.331***	-0.232***	-0.027	1.396***	-0.324***	-0.152	-0.250***	-0.073	-0.031	-0.265***
	(11.70)	(-3.03)	(-0.35)	(6.40)	(-3.06)	(-1.49)	(-2.75)	(-0.81)	(-0.13)	(-2.78)
Create Orders - Weekly (t)	-0.076	-0.140*	0.004	0.0033***	0.0028**	0.007***	-0.091	-0.085	-0.154	-0.028
	(-0.77)	(-1.87)	(0.06)	(3.45)	(2.05)	(7.57)	(-1.08)	(-1.12)	(-0.79)	(-0.29)
log (Market Cap), at (t-1)	-0.022**	-0.039***	-0.033***	-0.041***	-0.061***	-0.057***	-0.032***	-0.032***	-0.045***	-0.044***
	(-2.57)	(-4.08)	(-4.02)	(-3.21)	(-4.77)	(-4.66)	(-3.78)	(-4.06)	(-3.82)	(-3.72)
Average Share Turnover (t-1)	-0.097	-0.188**	-0.018	-0.089	-0.111	-0.042	0.194	0.126	0.090	0.058
	(-1.05)	(-2.00)	(-0.19)	(-0.45)	(-0.59)	(-0.22)	(1.02)	(0.67)	(0.39)	(0.28)
Amihud Illiquidity Measure (t-1)	0.024	0.040	0.013	0.085**	0.0471	0.0315	-0.001	0.004	-0.043	0.025
	(0.99)	(1.49)	(0.87)	(2.46)	(1.43)	(1.07)	(-0.05)	(0.32)	(-0.95)	(0.78)
Observations	551,252	550,664	550,664	256,612	255,592	255,592	222,161	222,161	60,958	158,914
R-squared	0.077	0.082	0.086	0.426	0.427	0.414	0.073	0.047	0.085	0.089
ETF & Date Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ETF & Date Clustering	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
ETF Sample	All	All	All		Non US-Equity		US-Equity	US-Equity	US-Equity	US-Equity
Liquidity Mismatch (ETF vs Underlying)									Low	High

Table 5 – 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*. We provide a complete list of variable names, sources, and definitions 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 6 – 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. We provide a complete list of variable names, sources, and definitions 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: The Effect of Operational Shorting on the 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 6 – Effects of ETF Operational Shorting on the Liquidity of the Underlying Securities: *(continued)*

Panel B: The Effect of Operational Shorting on the 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 6 – Effects of ETF Operational Shorting on the Liquidity of the Underlying Securities: *(continued)*

Panel C: The Effect of Operational Shorting on the 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 7 – 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. We provide a complete list of variable names, sources, and definitions 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 8 – 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*. We provide a complete list of variable names, sources, and definitions 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	1.58e-09*	1.34e-09	1.34e-09	9.38e-10	4.09e-10	4.09e-10
Outstanding, at (t-15)	(1.707)	(1.361)	(1.361)	(1.003)	(0.418)	(0.418)
Maximum Rolling R-Squared with Available		1.72e-10	1.72e-10		3.68e-09***	3.68e-09***
Futures 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 – 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.

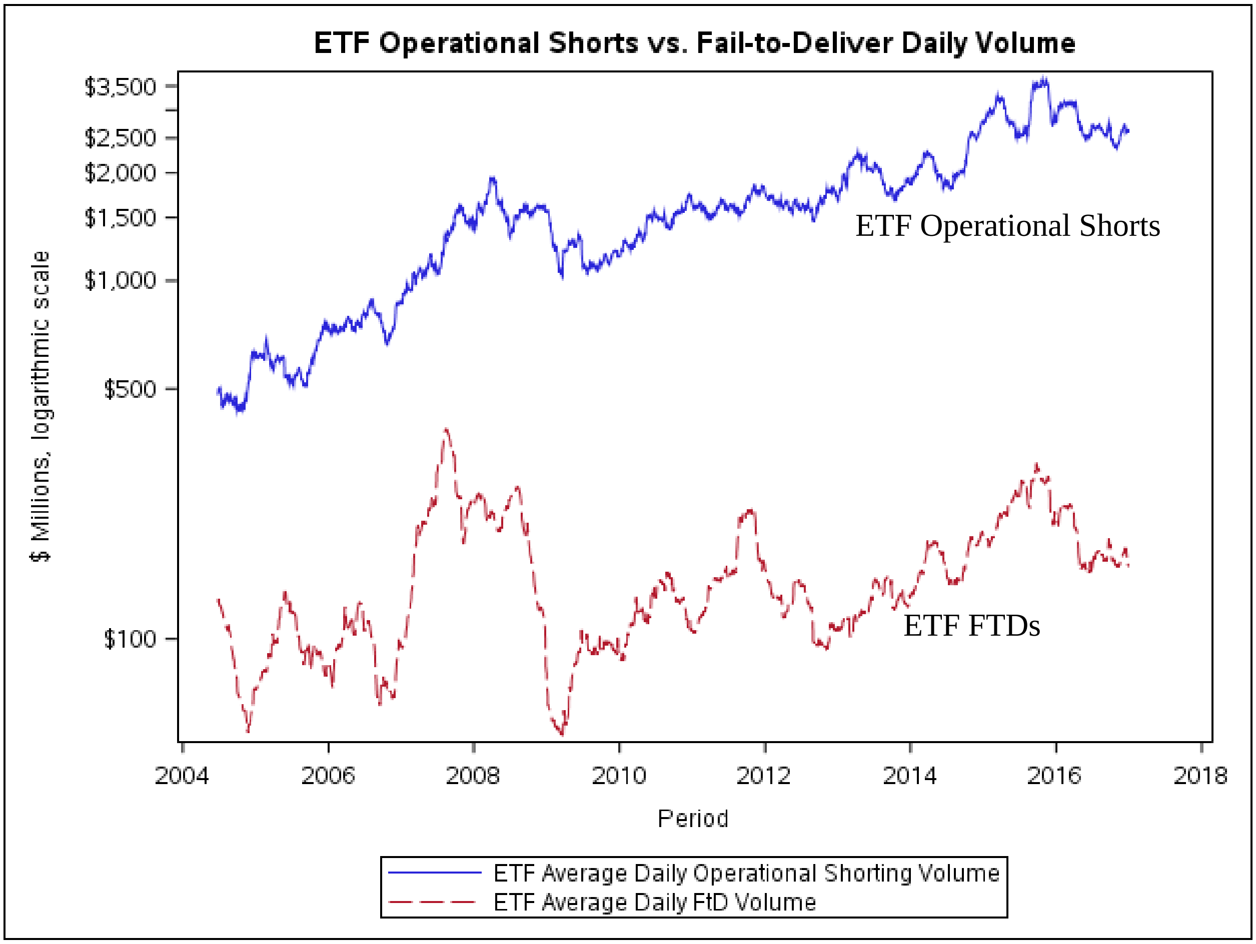
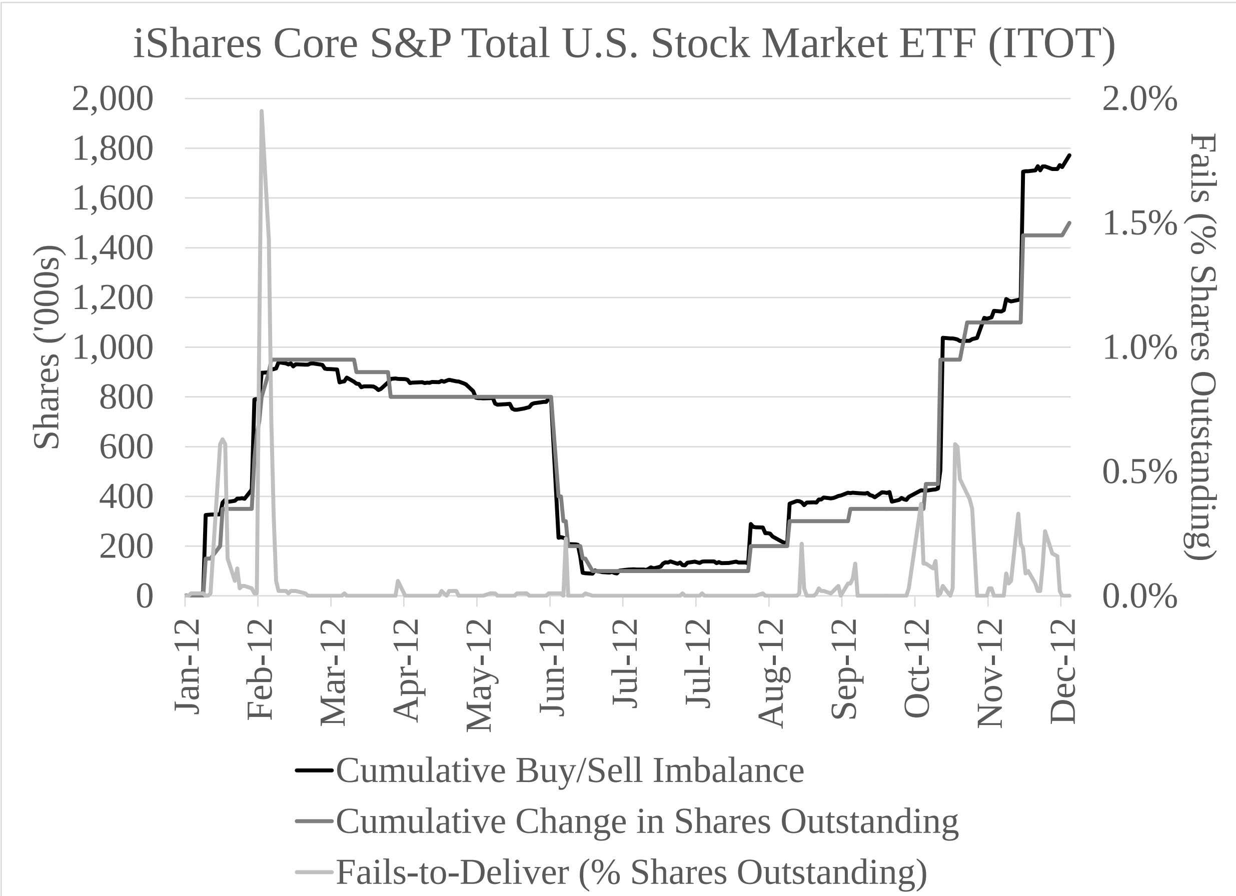


Figure 2 – An Example: ITOT – iShares Core S&P Total U.S. Stock Market ETF: This figure displays the cumulative buy-sell imbalance and the cumulative change in shares outstanding (in 1,000s of shares indexed by the left vertical axis) for the iShares Core S&P Total U.S. Stock Market ETF (ticker:ITOT) 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).[41] 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

λ = the "market impact" cost purchasing shares of the underlying basket of securities held by the ETF.

[41] 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).

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).[42]

Since the AP is risk-neutral, the difference between the above two payoffs equals what we call the

[42] 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.

"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.[43] 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 OIB_0). In effect, by waiting, the AP can avoid incurring the creation fee and market impact costs ($f + \lambda$)

[43] 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.

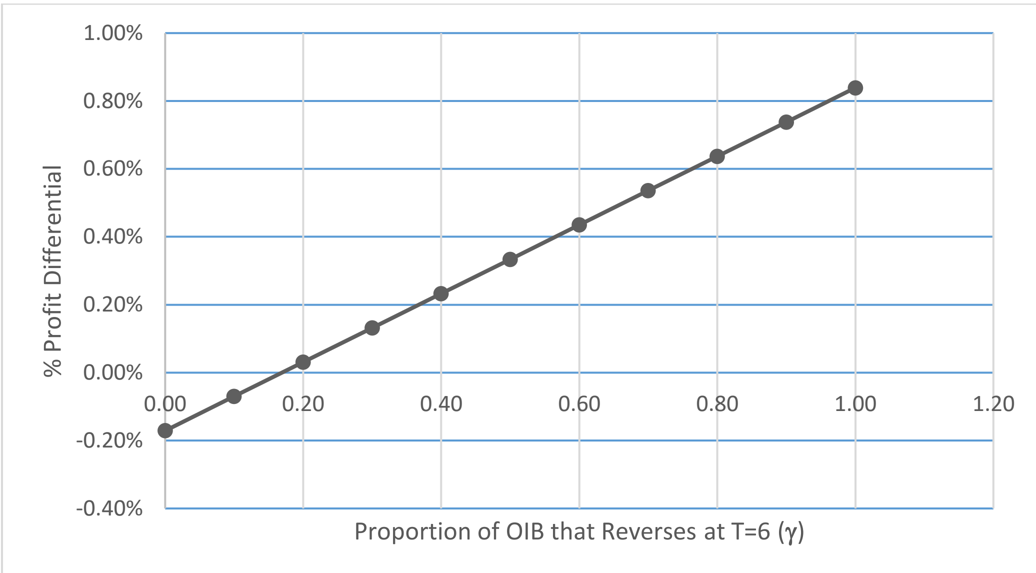
for that portion (γ) of the initial short position (OIB_0).

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 A.1 – 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, Thomson-Reuters Mutual Fund Ownership database, CRSP Mutual Fund 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, Thomson-Reuters Mutual Fund Ownership database, CRSP Mutual Fund Database
Weekly Fama and French 4-Factor Excess Return	Sum of the daily excess return over the week, computed using the lagged betas from a Fama and French four factor model over the previous 200 days.	CRSP
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, CRSP Mutual Fund Database, Blomberg
Liquidity Mismatch	The difference between the average intraday spread of the stocks in the ETF basket, and the ETF's intraday spread. Intraday spread are computed using spreads at the national best bid and offer (NBBO) and averaged over the entire day using the size of the trade that immediately follows this NBBO quote as weights. The average basket spread weighted by the ETF holding in each stock in the basket.	TAQ, Thomson-Reuters Mutual Fund Ownership database, CRSP Mutual Fund 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

Appendix C: Additional Failure-to-Deliver Background and Summary Statistics

1. Background on 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.[44] Figure C.1 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. Failure-to-deliver can occur with any type of security, and Table C.1 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 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

[44] 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.

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.[45] 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. FTD Summary Statistics

Table C.1 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 C.2 provides a graphical representation of the FTD volume for these two security types. The total dollar volume of FTDs increased until it reached 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 C.1 shows the average dollar value of ETF-related FTDs now represents 78.5% of all FTDs (up from 29.5% in 2008).

3. The Persistence of ETF Net Creation Activity and Trade Imbalances

[45] "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).

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:

$$Net\ Creation\ Activity_t\ \ or\ ETF\ Order\ Imbalance_t = \alpha_0 + \ \alpha_1 Controls + \sum_{n=0}^{8} \beta_n ETF\ Order\ Imbalance_{t-n} + \sum_{n=0}^{8} \gamma_n Net\ Creation\ Activity_{t-n} + \ \epsilon_t \qquad \text{(C1)}$$

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 C.2. 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*.[46] *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

[46] 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.

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 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 C.2 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 C.2 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. 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 C.3 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 C.3, 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 C.3 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 in Table C.2. 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 C.3 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.

Table C.1 – 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

Panel 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

Panel 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 C.2 – 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	0.00530***	0.00525***	0.00422***	0.00350***	0.00512***	0.00343***
Outstanding, at (t-15)	(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 C.3 – 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	0.300***	0.00262***	0.292***	0.00369***	0.295***	0.00323***
Outstanding, at (t-15)	(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

Figure C.1 – 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.

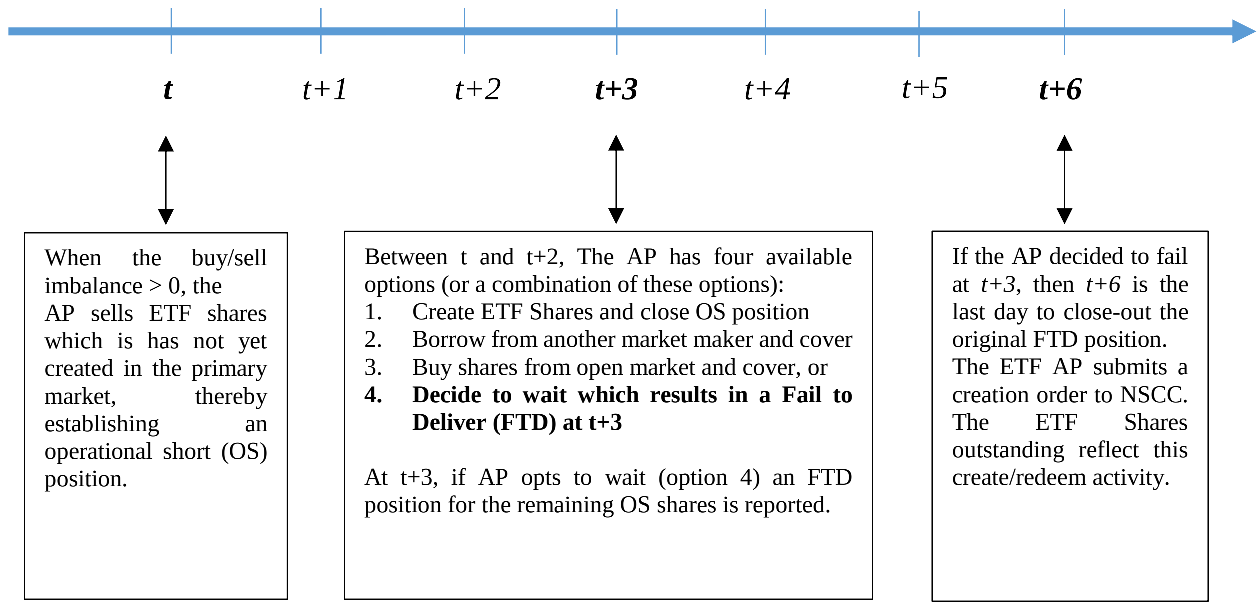
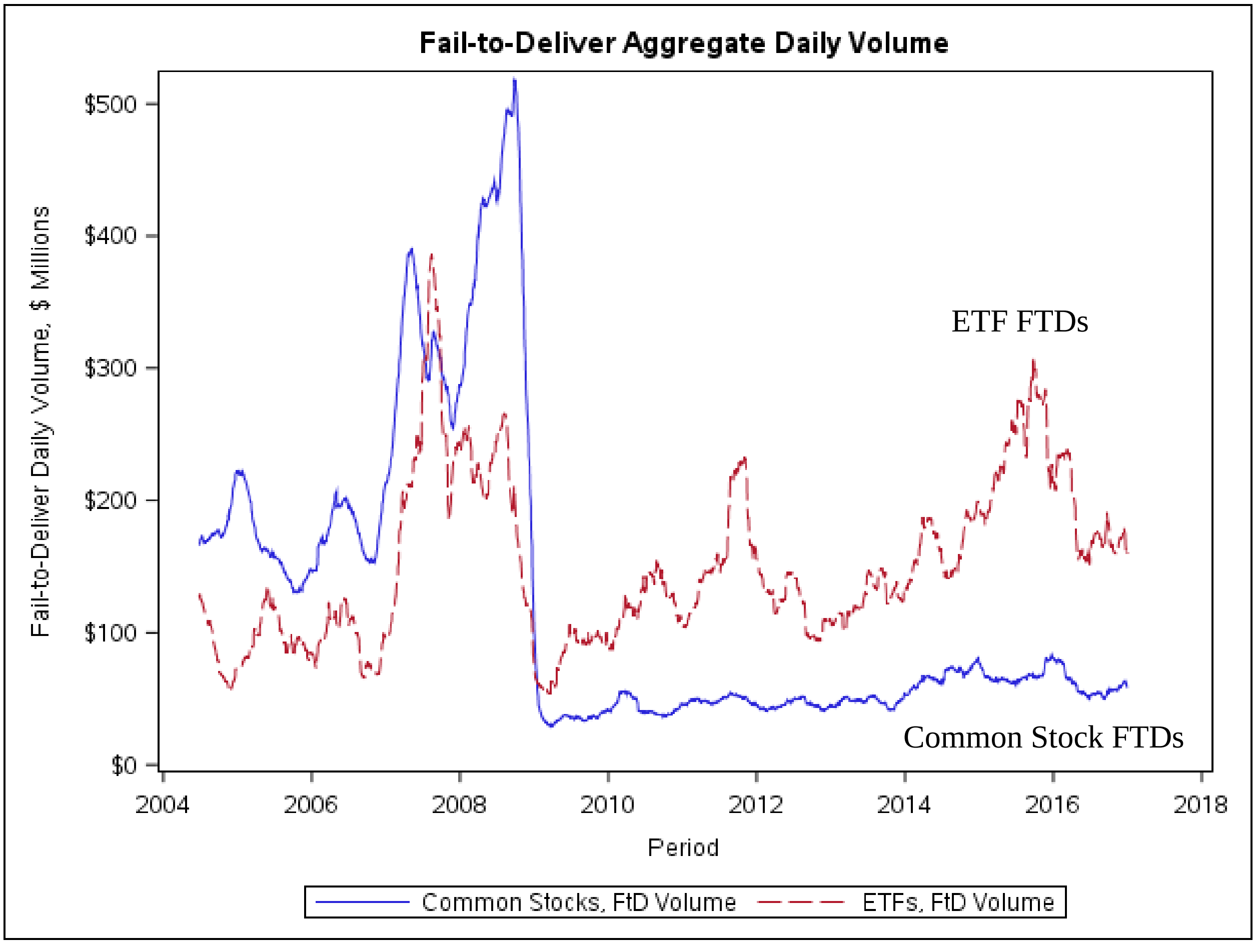


Figure C.2 – Failure-to-Deliver (FTD) Activity of ETFs and Common Stocks: This figure displays graphically the average dollar volume of ETF and common stock FTDs on a daily basis from March, 22 2004 – December 31, 2016. We include only the rolling average daily FTD volume of stocks and ETFs in this graph,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



Appendix D: Trade Signing Algorithm; Discussion and Evaluation Results

We propose an improved methodology to compute the daily buy and sell volumes in our paper. Lee and Ready (1991) is among the first papers to advocate using first a quote test and then a tick test to sign a trade. This method classifies a trade as a buy if the trade price is above the national best bid and offer (NBBO) midpoint, and as a sell if the executed price is below the prevailing NBBO quote midpoint. If the trade price is equal to the quote midpoint, the method instead uses the tick test. More recently, Ellis, Michaely, and O'Hara (2000) criticize using the quote midpoint as reference as they document significant clustering of buys on the offer price and sales on the bid prices. They conclude that the quote test is less accurate when the trades are not executed at the ask or the bid. This finding is especially important in recent periods in which we observe fast trading and a clustering of multiple trades within the most granular timestamp of the NYSE TAQ. This phenomenon makes it more difficult to accurately match the prevailing NBBO quote to each trade. Recently, researchers have matched trades with the prevailing NBBO at the beginning of each second when using the TAQ Monthly feed (with second timestamp granularity) or at the beginning of each millisecond when using the TAQ Daily feed (with millisecond timestamp granularity). In a more recent paper, Holden and Jacobsen (2014) introduce a new method -- the Interpolated Time -- in which they order multiple trades that are clustered within the smallest time interval (a second) and match them to the ordered quotes in the NBBO data in the same second, in order to make an educated guess about the quote that is likely to precede each of the trades clustered with a one second interval. However, this interpolated time method does not take into consideration the fact that clustered trades, as well as quotes, are not uniformly distributed within a second (or a millisecond when available), due to the nature of high frequency trading (see, e.g., Murayev and Picard (2016)).

We suggest that Ellis, Michaely and O'Hara's (2000) argument is especially valid when the Lee and Ready algorithm fails to take into consideration trades executed outside the prevailing quote. In Table D.1, we report the fraction of daily trades for all stocks that are clustered within a second across all trading days in 2014. Our results show that more than 75% of all trades (or share volume) for all stocks belong to multiple trades within a second. This proportion increases to over 85% when using stock market capitalizations as weights, as larger and more liquid stocks are more likely to trade more frequently within a second interval. When looking at the fraction of trades within a second that are executed at prices outside the matched NBBO quote at the beginning of the second, our results suggest that over 13% of stock share volume, and over 15% of ETF share volume have prices outside the matched NBBO. These values suggest that those NBBOs are likely stale and should not be used to in a quote test to sign the trades after the first reported outside during this second (these proportions are closer to 20% for ETFs when using market cap as weights). Importantly, this outside trade frequency represents the lower bound of the proportion of incorrectly classified trades, especially in instances when the NBBO spread is large enough that trades matched to stale NBBOs are not flagged as outside trades. We repeat this analysis using millisecond timestamps, which are available in the TAQ Daily feed for trades and quotes, and find substantial clustering even within a millisecond, and a significant number of outside trades when matching each trade during a millisecond to the prevailing NBBO quote at the beginning of the millisecond.

We thus propose a modified classification algorithm that combines the insights of Lee and Ready (1991) and Ellis, Michaely, and O'Hara (2000). We first construct the NBBO quote at the beginning of each millisecond using the NBBO classifiers in the quote file and the NBBO addendum data which is provided in the NYSE TAQ daily (millisecond) database.[47] We then match the NBBO quote at the beginning of each millisecond to all trades in the millisecond. 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. Therefore, we compare the trade price for all trades occurring during a millisecond to the matched NBBO if the price of the trades has not crossed the prevailing quote. Once an executed trade price crosses the prevailing NBBO within a millisecond, we stop using the quote test. Instead, and for the remaining trades during this millisecond, we rely on the tick test as it is likely that the

[47] The TAQ daily feed provides millisecond timestamp until July 24, 2015, microseconds timestamp from July 27, 2015, and nanoseconds timestamp (for Nasdaq trades (UTP)) starting in October 24, 2016. We use the most granular time stamp when available.

quote test is not accurate due to a stale NBBO, 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 method takes into consideration that buys (sells) are more likely to be executed at the ask (bid), and whenever an outside trade is observed during that millisecond, then the algorithm adjusts dynamically and 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 sells at 4:00 pm to construct our buy and sell volume for the day.

To evaluate the effectiveness of our proposed method, we directly compare it to Lee and Ready (1991) and Holden and Jacobsen (2014). We first use the WRDS Intraday Indicators dataset to extract the buy and sell volume using the Lee and Ready (1991) and Holden and Jacobsen (2014) methods for all stocks in 2014, which is the last year of the data.[48] We then construct our buy and sell volumes following the methodology described above and using the TAQ Monthly feed (second timestamp) where we match each trade to the prevailing NBBO quote at the start of the second. We compute the trade imbalance ratios for each of the three methods as the difference of the buy and sell volumes divided by the total volume. To proxy for the true trade imbalance, we construct the buy and sell volume following the Lee and Ready (1991) algorithm but using the TAQ Daily feed which provides the millisecond timestamp for all trades. We expect the millisecond matching of quotes to yield more accurate buy and sell volume classifications than all the remaining classifications that used the second-level timestamp to classify the same trades. We then compute the Pearson correlations between all these trade imbalance measures and for all stocks in all trading days during 2014, and present the results in Table D.2. We find that our method, despite using quotes at the beginning of each second, has the highest correlation with the trade imbalance constructed using the millisecond feed. Despite using the same Lee and Ready (1991) method with the second-level timestamps, our modified method has twice as large a correlation with the Lee and Ready method that uses the millisecond timestamp. We conclude that our modified algorithm -- that dynamically switches from the

[48] WRDS Intraday Indicators Dataset (IID) is constructed using the TAQ Monthly feed (with second timestamp). IID uses original codes provided by Holden and Jacobsen (2014) to construct the interpolated time buy and sell volumes. We are thankful to Jun Wu, Research Director at WRDS, for helping us run this test of various classification methods.

quote test to the tick test when needed -- is superior to the static Lee and Ready (1991) method and to the computationally intensive Holden and Jacobsen (2014) method. Therefore, we believe that using our modified algorithm to classify trades in the TAQ Daily feed with millisecond timestamps provides a more accurate buy and sell volume classification. The ITOT example included in Figure 2 is another testament that our trade classification algorithm, while far from perfect, can properly classify trades and provide meaningful trade imbalance patterns, even for highly liquid ETFs that are traded at higher frequencies.[49]

[49] Spearman correlations yield similar results. When using market capitalization of the securities as weights, the correlation of our method (# 4) becomes 75% versus 11% for Lee and Ready (1991) (# 2) and 24% for Holden and Jacobsen (2014) method (# 3).

Table D.1 – Trade Clustering and Outside Trades over Second vs. Millisecond Time Intervals: This table displays the statistics of all trades for all common stocks and ETFs using the NYSE TAQ Millisecond database for all trading days in 2014. For each security, we compute the fraction of daily trades that are clustered within one second and one millisecond time intervals, and then compute market cap-weighted and equal-weighted averages across all stocks and ETFs. Then, using the National Best Bid and Offer (NBBO) indicator in the quotes file and the supplementary NBBO datasets, we compute the NBBO at the start of each time interval, and the percentage of trades that are executed at a price outside the bid and ask of the prevailing NBBO quote.

Panel A: Fraction of Trades Clustered within One Second vs. One Millisecond – NYSE TAQ Database

	% of Trades Clustered within a second vs. millisecond in NYSE TAQ Database			
	% of multiple trades within a second		% of multiple trades within a millisecond	
# of Trades	EW	VW	EW	VW
Common Stocks	75.33%	87.57%	42.68%	54.90%
ETFs	56.28%	75.45%	29.11%	46.75%
Share Volume	EW	VW	EW	VW
Common Stocks	76.30%	87.61%	38.77%	50.00%
ETFs	60.44%	78.90%	30.84%	47.89%

Panel B: Fraction of Trades with Price Executed Outside the Prevailing NBBO at the Start of Each Second vs. Each Millisecond – NYSE TAQ Database

	% of Trades with Price Outside the Prevailing NBBO at the start of the second vs. millisecond in NYSE TAQ Database			
	% of trades outside the NBBO at the beg of the second		% of trades outside the NBBO at the beg of the millisecond	
# of Trades	EW	VW	EW	VW
Common Stocks	10.50%	13.86%	3.67%	4.16%
ETFs	10.52%	11.71%	1.73%	2.72%
Share Volume	EW	VW	EW	VW
Common Stocks	13.05%	16.05%	4.53%	6.08%
ETFs	15.28%	20.37%	3.73%	8.27%

Table D.2 – Correlation of Trade Classification Algorithms: This table displays the Pearson correlations of our trade classification algorithm with the Lee and Ready (1991) and Holden and Jacobsen (2014) interpolated quote classification methods. To run a horse race between the three methods, especially in instances with clustered trades within certain time intervals, we first compute buys and sells according to each method using the quote at the beginning of each second (rows (2) - (4)), and then compare the outcome of each method with a proxy for true buy and sell classification that uses the Lee and Ready (1991) method at the millisecond level – a more granular time interval frequency. Within this frequency, we expect the Lee and Ready (1991) method to yield a more accurate matching of trades with their prevailing NBBO quotes. We then construct the trade imbalance from each method as the difference between buy volume and sell volume divided by total volume. The table below presents the correlation between the proxy for true trade imbalance (row (1)) and the three methods: Lee and Ready (1991) using the quote test for the entire second (row (2)), Holden and Jacobsen (2014) employing interpolated time quote matching with trades within each second (row (3)) and our method, that dynamically switches from the quote test to the tick test on the first occurrence of an outside trade for all remaining trades within a second (row (4)).

	(1)	(2)	(3)	(4)
(1) Lee and Ready (1991) using Daily Millisecond TAQ feed with millisecond-level NBBOs	1			
(2) Lee and Ready using monthly feed with second- level NBBOs	42.9%	1		
(3) Interpolated Holden and Jacobsen (2014) using monthly feed with second-level NBBOs	59.5%	51.8%	1	
(4) Our measure with second-level NBBOs using modified Lee & Ready / Ellis, Michaely, & O'Hara approach	86.1%	57.7%	46.6%	1