Operational Shorting and ETF Liquidity Provision

Richard B. Evans
University of Virginia

Rabih Moussawi
Michael S. Pagano
John Sedunov
Villanova University

June 2022

Abstract

Due to a regulatory exemption, ETF market makers can satisfy excess demand in secondary markets by selling ETF shares that have not yet been created. While this ability to "operationally short" is not unique to ETFs, it plays a more prominent role in ETF liquidity provision, and results in elevated ETF failures-to-deliver. We propose a novel measure for "operational shorting" and show it is associated with improved liquidity and greater price efficiency in the ETF underlying securities. Higher retail trading activity and short-term return reversals are also consistent with liquidity-supplying motives rather than informed trading. Consequently, delayed ETF creation to cover operational shorts is found to be a valuable option in the presence of retail trading and liquidity mismatches between the ETF and its underlying securities.

Keywords: Exchange-Traded Funds, Authorized Participants, Short-Selling, Failure to Deliver, Liquidity, Retail Trading, Market Making

JEL Codes: G1, G12, G14, G23

This paper was originally titled "ETF Short Interest and Failures-to-Deliver: Naked Short-Selling or Operational Shorting?" 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, Francesco Franzoni, Jeff Harris, Pankaj Jain, Bryan Johanson, Qiang Kang, Dolly King, Stewart Mayhew, Tom McInish, Suchi Mishra, Jason Morrison, Steve Oh, 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 Banks of Richmond, Charlotte and Baltimore, 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

With over $7.2 trillion in assets[1] and accounting for 37% of U.S. dollar-trading volume,[2] exchange traded funds (ETFs) have played an increasingly important role in U.S. markets. Because the creation/redemption aspect of ETF liquidity provision directly connects trading in ETFs to trading in their underlying securities, recent papers have examined the relationship between the two, including Da and Shive (2018), Hamm (2014), Agarwal, Hanouna, Moussawi, and Stahel (2017), and Israeli, Lee, and Sridharan (2017). However, this literature has yet to examine one of the most important mechanisms related to the dynamics of providing liquidity to an ETF and its underlying securities: what we refer to as "operational shorting."

While creation and redemption, the process of exchanging blocks of ETF shares for baskets of the underlying securities and vice-versa, is a focal point of the existing ETF literature, most ETF liquidity provision does not translate into this specific arbitrage transaction. For example, the Investment Company Institute estimates that $1.00 of total ETF trading translates to only $0.15 of creation/redemption activity with the rest corresponding to secondary market activity.[3] Market makers (MMs) and authorized participants (APs) accommodate the remaining $0.85 of ETF liquidity provision without resorting to creation/redemption activity via a number of different ways, including operational shorting.[4] Building on Comerton-Forde, Jones, and Putnins (2016), we define operational shorting as the liquidity-supplying short sale of ETF shares by APs/MMs in the course of liquidity provision. Due to a unique exemption from SEC delivery requirements (Rule 204), market makers can sell ETF shares to satisfy a bullish order imbalance but postpone their creation and delivery. 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. APs/market makers can legally delay delivery of those shares for

[1] Investment Company Institute (ICI) Factbook, 2022, https://www.icifactbook.org/pdf/2022_factbook.pdf.
[2] April Joyner, *Reuters*, "Investors eye cracks in $4.4 trillion U.S. ETF market as sell-off rages", 3/24/20.
[3] See Figure 4.5 in page 76 of the 2022 ICI Factbook.
[4] An authorized participant (AP) is typically a market maker (MM) 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. The ICI Factbook 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 an AP agreement with the ETF sponsor, and so we refer to such a market maker interchangeably in the paper as AP or MM.

three additional trading days beyond standard clearing times, thus generating a failure-to-deliver (FTD).[5] Put another way, the AP has the option to sell short ETF shares and then fail to deliver them at settlement.[6]

We examine the importance of operational shorting to liquidity provision and the pricing of ETFs. We first propose a measure of operational shorting. To provide validation for our proposed measure, we test both the determinants of the measure and its relationship with ETF short interest and FTDs. We then explore how controlling for operational shorting affects previously documented results regarding ETF pricing and liquidity as well as the relationship between the ETF and its underlying securities.

Our proposed measure of operational shorting captures discrepancies between primary market activity, the creation/redemption of ETF shares by APs, and secondary market trading activity in ETF shares. Specifically, the measure compares two quantities: 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 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.[7]

To provide validation of our novel operational shorting measure, we first examine its relationship with various ETF short interest and volume measures. We find evidence that operational shorting is the most important determinant, statistically and economically, of the bi-weekly ETF short interest and daily short

[5] The SEC delivery requirements (Rule 204) provide an exemption for APs and 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 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..." 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.

[6] 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."

[7] Using Securities-on-Loan data from the Markit Securities Finance database as a proxy for directional shorting, we infer snapshots of operational shorting whenever short interest biweekly data are available and find that operational shorting represents approximately two-thirds of the overall short interest activity during our sample period.

volume levels. Even after controlling for directional shorting activity, we find that our measure exhibits the strongest statistical relationship of all included explanatory variables. 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). These results suggest that operational shorting is a primary driver of daily ETF shorting activity above and beyond directional shorting.

While the strong relationship between our proposed measure and short interest controlling for directional shorting provides important validation of our measure, a more compelling test is to examine the relationship between our measure and failures-to-deliver (FTDs).[8] Because the SEC has effectively prohibited FTDs for directional shorting but allowed it for operational shorting (i.e., SEC Rule 204), we would expect a significant relationship between the two if our measure captures operational shorting activity. Figure 1 shows a high correlation between our measure of operational shorting and FTDs at an aggregate level. We confirm this statistically and economically significant relationship at the ETF level and controlling for other potential determinants. Similar to the short interest and short volume results, our measure operational shorting exhibits the strongest statistical relationship with ETF FTDs of all determinants.

As a second validation test and to better understand the underlying economics, we also examine the determinants of operational shorting. The market maker's exemption is provided for the express purpose of liquidity provision so if operational shorting is being used as regulators intended, then its use by these market participants should be consistent with liquidity provision and result in improved liquidity. When we examine the determinants of operational shorting, we find that it is driven by: 1) a higher liquidity mismatch with the ETF's underlying basket of securities and 2) the presence of efficient hedges. Further, we follow Boehmer, Jones, Zhang, and Zhang (2021) to construct a proxy for daily retail ETF volume, which is more likely to be uninformed and liquidity driven. Indeed, we find that APs/MMs are operationally short in the presence of increased retail trading and when facing greater liquidity mismatches between the ETF and its underlying portfolio. In addition to further validating our measure of operational shorting, these results provide important insight into the rationale behind why APs wait and delay the (costly) assembly of the basket and creation of

[8] See Appendix A for more background and literature related to FTDs in U.S. financial markets.

new ETF shares until a future date, consistent with liquidity provision.

With ample evidence confirming our measure of operational shorting, we revisit the three primary issues examined in the prior ETF literature: 1) ETF trading/arbitrage and returns (e.g., Brown, Davies, and Ringgenberg (2021)); 2) the mispricing/tracking error associated with ETF investments (e.g., Bae and Kim (2020)); and 3) the impact of ETFs on the volatility and liquidity of the underlying securities (e.g., Ben-David, Franzoni, and Moussawi (2018)). In all three settings, we highlight the important role that operational shorting plays in these economic questions that has not been addressed by the prior literature.

In first revisiting the relationship between ETF trading activity and returns, we build on the prior result of Brown, Davies, and Ringgenberg (2021) that ETF arbitrage activity (i.e., creation and redemption) has predictive power for future returns. They interpret this predictability as AP creations/redemptions accommodating non-fundamental demand, thereby moving prices from their fundamental values, which upon reversion creates the observed return predictability. Recognizing that only a small fraction of secondary ETF trades result in primary market creation/redemption activity, and that APs can use operational shorting to similarly respond to non-fundamental demand, we examine the return predictability of both creations and operational shorting. Additionally, we separately analyze return predictability based on ETF market price and NAV (underlying security) returns. Consistent with Brown, Davies and Ringgenberg (2021), we find a negative relationship between ETF creation orders and future returns. Above and beyond the predictability associated with this arbitrage activity, however, we find that operational shorting negatively predicts future ETF returns and the effect is both statistically and economically more significant than the arbitrage activity itself.

Most interestingly, our results suggest that operational shorting has no predictive power for ETF NAV/underlying returns. Combined with the strong positive relation between operational shorting, retail order flow, and concurrent ETF returns, our return results suggest that APs observe contemporaneous price pressures due to excess demand for ETF shares and anticipate a short-term ETF price reversal (i.e., predicts ETF price returns), but their operational shorting activity is not informative about the value of the underlying securities. This stands in stark contrast to a long literature on the predictive power of individual equity short selling on returns due to the informed nature of short sellers.[9] This lack of NAV/underlying return predictability is also consistent with operational shorting that is positively related to uninformed retail trades. At over 20% of daily trading volume, these relatively uninformed retail trades can ultimately lower adverse selection costs and help enhance liquidity provision by ETF APs/market makers.

We find further evidence that the return predictability of operational shorting is driven by uninformed liquidity provision by conditioning on the liquidity characteristics of the underlying securities. The statistically significant negative relation between operational shorting and future 1-day 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 emphasize the need to separate the different motivations for ETF short selling: directional/information vs. operational/liquidity provision. This bifurcation has important implications for the extant short-selling literature. 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. In sum, our results suggest that any examination of ETF short selling or arbitrage activity needs to separately account for operational shorting. Not only is it a significant predictor in the ETF market price return analyses, it is the most statistically significant predictor in 6 of the 7 specifications.

Second, we revisit the determinants of ETF mispricing/tracking error once again accounting for operational shorting. Because the liquidity provision mechanism underlying ETFs is so dramatically different from index funds, mispricing of ETFs (i.e., market price) relative to their underlying securities (i.e., NAV) and the standard deviation of that mispricing over time (i.e., tracking error) have been an important focus of the

[9] Our paper is not the first to examine the liquidity implications of short sales. Focusing on intraday equity short-selling rather than ETFs, 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, Edwards and Hanley (2010) identify a similar regulatory provision in the "Green Shoe" or overallotment option granted to underwriters following an IPO and the associated exemption granted to designated market makers on the NYSE to sell securities short.

early academic literature.[10] More recently, Bae and Kim (2020) examined the relationship between ETF liquidity, underlying liquidity and tracking error, finding that illiquid ETFs and ETFs with illiquid underlying securities have larger tracking errors. In examining the relationship between changes to mispricing, its absolute value as a proxy for tracking error, and operational shorting, we uncover an important insight that stands in contrast to this result. Other than lagged mispricing, operational shorting is the strongest statistical predictor of decreased ETF mispricing across all included variables in our analysis. Combined with the ETF return reversal results, this decrease in ETF mispricing represents a return to the APs for their liquidity-supplying operational shorting. This insight is especially important because in analyzing the determinants, we find that APs are more likely to operationally short when there is a larger liquidity mismatch (i.e., the ETF is more liquid than the underlying). Together, these results suggest that operational shorting is an important and potentially confounding determinant of mispricing arbitrage and should be accounted for.

Third, by now taking into account operational shorting, we revisit perhaps the most examined aspect of the ETF literature:[11] the impact of ETF trading on the liquidity of the underlying securities. We measure the liquidity of the underlying securities using intraday best bid and offer spreads, second-by-second return volatility, and intraday variance ratios. Using all three measures, we find evidence consistent with Ben-David, Franzoni, and Moussawi (2018) and a growing literature that ETF ownership is positively associated with higher intraday spreads and greater volatility of the ETF's underlying basket of securities. However, we also show that operational shorting is negatively related to the basket's intraday spreads and volatility, thus acting as a "*release valve*."[12] Our results suggest that APs provide liquidity to the ETF market without (or before) entering the market for the underlying stocks to accommodate a sudden surge in buying demand through

[10] For example, Elton, Gruber, Comer, and Li (2002), Blume and Edelen (2004), Gastineau (2002, 2004), and Agapova (2011).

[11] For example, Sullivan and Xiong (2012), Da and Shive (2018), Hamm (2014), Bhattacharya and O'Hara (2016), Agarwal, Hanouna, Moussawi, and Stahel (2017), Dannhauser (2017), Israeli, Lee, and Sridharan (2017), Ben-David, Franzoni, and Moussawi (2018), Box, Davis, Evans, and Lynch (2021) and Chinco and Fos (2021). See Ben-David, Franzoni, and Moussawi (2017) for a survey of ETF literature.

[12] 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. For example, Pagano, Sedunov, and Velthuis (2021) find that Robinhood traders' contrarian, "buy-the-dip" strategies can help improve market quality during normal market conditions, but this effect was reversed during the stressful conditions of the Covid-19 pandemic in Spring 2020.

increases in operational shorting. We additionally find that operational shorting is associated with variance ratios closer to 1, thus suggesting higher price efficiency and lower intraday return reversals. Therefore, our evidence suggests that, consistent with liquidity provision, operational shorting acts 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.

While the primary contributions of the paper are the introduction and validation of an operational shorting measure as well as the evidence that it is an essential variable in analyzing key ETF research questions, the paper has one additional contribution: insight into what is driving high levels of ETF FTDs and short interest. In contrast to Jain and Jain's (2015) findings for FTDs in U.S. common stocks, we find that FTDs in U.S. ETFs are growing and, as a percentage of market capitalization, ETF-related FTDs are disproportionately larger than those in equities markets and now represent over 80% of FTDs across all securities in U.S. financial markets. Further, ETFs exhibit higher levels of short interest on average when compared to other publicly traded equities.[13] This dramatic increase in ETF FTDs, and short interest more generally, has attracted the attention of market participants and regulators who are concerned that this represents an increase in "naked" short selling. In particular, the concern is that this naked short selling is also "directional" in nature (i.e. speculative bets on future declines in ETF prices). In 2015, the SEC solicited comments from market participants on "topics related to the listing and trading of exchange-traded products" and several of the comment letters raised similar concerns about the short-selling of ETFs.[14] Recent enforcement actions by FINRA and Nasdaq related to "naked" ETF short positions underscore the possibility of improper ETF short-selling.[15] We provide compelling evidence that the high levels of ETF FTDs and short interest do not represent

[13] Figure 2 shows that as ETFs have grown, so has the short-selling activity in ETFs. As of June 2020, the aggregate dollar value of ETF short interest was upwards of $183 billion, accounting for 19.5% of the overall dollar value of short interest in U.S. equity markets, while constituting just under 11% of the total U.S. equity market capitalization.

[14] e.g., See this quote from an SEC comment letter (File number S7-11-15) in 2015 "Short selling is extreme in many ETFs. The lending markets are not being properly utilized to accommodate the selling, causing systemic risk from undisclosed leverage in the financial system (more shares sold than exist) for the benefit of very few while creating risks for all stakeholders, including taxpayers." https://www.sec.gov/comments/s7-11-15/s71115-19.pdf

[15] In March of 2016, FINRA and Nasdaq fined Wedbush Securities, an ETF authorized participant, 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 achieve short exposure would be to redeem or sell shares they did not own ("naked" redemption/short-selling), and subsequently fail-to-deliver (FTD) those

abusive short-selling, but rather represent the exercise of the option afforded APs to operationally short ETFs for liquidity provision.

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 the validation tests for our operational shorting measure. Section five then revisits the three primary issues examined in the prior ETF literature (return predictability, mispricing, and liquidity) but accounting for operational shorting, and section six concludes.

2. Institutional Details: 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. ETFs can also be traded intraday and swapped for their underlying securities via "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 help foster 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 provide 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. For U.S. equity ETFs, such transactions are typically in kind, and a creation

---

shares to Wedbush. As Thomas Gira, the FINRA Head of Market Regulation, 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."

[16] Moussawi, Shen, and Velthuis (2021) document that ETFs are tax efficient as they take advantage of Section 852(b)(6) exemption of the US Internal Revenue Code of 1986 which exempts the distribution of capital gains when the appreciated stocks are handed "in-kind" to redeeming investors. In effect, ETF investors can defer the ETF portfolio's short- and long-term realized capital gains until they sell their ETF shares.

basket of securities is exchanged for a creation unit of ETF shares after the trading day has ended. 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 APs who are market makers can earn commissions and fees from customer orders as well as potential profits from ETF-common stock arbitrage due to the price discrepancy between the ETF share (market price) and the underlying basket (net asset value or NAV). Through these incentives, APs help keep the ETF prices in the secondary market aligned with their intrinsic values. ETF shares are created (in effect, increasing the supply of ETF shares outstanding) when the AP delivers a specific basket of underlying securities to the ETF sponsor (investment manager) in exchange for a block of ETF shares equivalent to one or more ETF creation units. Note that this creation process is generally the result of buying pressures on ETF shares in the open market that can cause a premium in ETF prices relative to NAV, which creates an arbitrage opportunity for APs to buy the cheaper constituent basket and create ETF shares with the fund that are worth more in such instances.[17]

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. However, the arbitrage process and the share creation do not occur at the same time. As demand for the ETF from investors in the secondary market grows, the ETF's market price should increase, potentially creating a more attractive intraday arbitrage between the market price and NAV. The AP provides liquidity by selling the more expensive ETF shares while assembling the basket of underlying at the NAV during the day and swaps it for ETF shares at the end of the day. Therefore, these two different legs of the trade (i.e., selling ETF shares and buying the underlying basket/creating the ETF shares) are not

[17] The redemption process (in effect, taken out of circulation and thus lowering the supply of shares outstanding) 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.

instantaneous.[18]

APs and ETF market makers can 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 the supply later with the creation/redemption 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 *operational short* position.

2.3 Incentives to Operationally Short

There are several 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. 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 motivations become even more compelling if an inexpensive and liquid hedge is available through the futures or options markets. In Appendix B, we provide a cursory example as well as a numerical model that illustrate the incentives of a risk-neutral AP 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. The model suggests that the option to delay has the potential to generate large, predictable profits for APs while also

[18] 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.

enhancing the ETF's liquidity (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 for 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 is recorded even if it is for bona fide market making purposes.[19] This can happen due to operational shorting, as part of classic liquidity provision activity, while directional shorting via naked short selling is technically prohibited by SEC Rule 204. Despite this regulatory record-keeping process for FTDs, APs and other market makers can still take up to T+6 days to deliver their shares without failing to fulfill their responsibilities.

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, Franzoni, and Moussawi (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 A.

2.4 Measuring Operational Shorting

We propose a simple measure to estimate operational shorting as 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

[19] 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.

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.[20] 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 3 contains an illustrative example of how the cumulative buy-sell trade imbalance, the change in shares outstanding, and fails-to-deliver might relate, which further motivates our measure. The figure shows these cumulative quantities for the iShares Core S&P Total U.S. Stock Market ETF (ticker: ITOT) over the entire year of 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 the excess demand, plausibly 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. APs and market makers appear to be accommodating the demand for ITOT shares, but the delay in creating them generates the FTDs observed. While Figure 3 focuses on a single example, sections 3 and 4 of Appendix A examine the daily dynamics of buy/sell trade imbalances, creation/redemption activity, and FTDs and show the insights of Figure 3 apply more generally across our entire sample.

The operational shorting measure we propose compares the cumulative *buy-sell trade 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. Note that our measure could also capture some trading that does not directly involve APs.[21] Nevertheless, we consider our

[20] 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."

[21] 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.

measure to be a reasonable proxy for the liquidity provision by APs/market makers because these firms are permitted by the SEC to provide liquidity by selling securities short (Comerton-Forde, Jones, and Putnins, 2016) and delay the delivery of these securities for a few more days. To the extent that these non-AP-related trades make our measure a noisier proxy, then this phenomenon should work against us finding any statistically significant effects. However, as we show later, our measure's impact on short interest, FTDs, returns, spreads, and short-term volatility are all consistent with APs/market makers engaging in operational shorting activity and not directional shorting behavior. The formula for our measure of operational shorting is therefore:

$$Operational\ Shorting = \frac{\max[0, \left(Buy/Sell\ Imbalance(t-3, t-1) - \Delta Shares\ Outstanding(t-1, t)\right)]}{Shares\ Outstanding(t\ \text{-}\ 4)}$$

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 over the past 3 days are expected to be covered to avoid a potential FTD. We normalize the result by dividing by the number of ETF shares outstanding at the start of this rolling window (end of (t-3) or beginning of (t-4)). 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.

In Appendix C, we also present robustness results using an alternative construction of operational

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.

[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 E. We exclude closing and opening auction volume when computing our buy-sell imbalance measure.

shorting by formulating the dependent variable on non-overlapping, discrete *weekly* intervals. Following similar logic as the daily operational shorting variable, we construct weekly operational shorting for each week separately as the cumulative ETF buy-sell demand for all days during the week in excess of the total actual ETF share creation over the week. This avoids any potential confounding issues related to using an operational shorting variable constructed over 3 days relative to 1 day for the independent variables. Like the daily operational shorting data, we only focus on excess ETF demand that is not met with share creation, and negative values are set to zero. Appendix C reports the results for the determinants of operational shorting as well as the return implications using the weekly sample, and the results are very similar and, in some cases, stronger than the results based on the daily sample.

3. Data

Because ETFs sit at the intersection of many different markets, our empirical analysis requires data from numerous sources. We provide a complete listing of variables, definitions, and sources in Appendix D. We identify ETFs using the CRSP Mutual Fund Database. 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.[23] We collect the ETF holdings of underlying stocks from the Thomson-Reuters Mutual Fund Ownership and CRSP Mutual Fund Holdings databases. Stock price, return, and volume data come from CRSP and are used to calculate variables such as *market capitalization, share turnover, illiquidity, risk-adjusted excess returns,* and *idiosyncratic volatility*.

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. 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) as described in Appendix E. 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

[23] While our sample includes all ETFs (including fixed income and leveraged ETFs), we frequently limit our sample to only equity ETFs, and find similar results throughout the paper.

outstanding data from Bloomberg because they are not reported accurately in CRSP and Compustat.[24]

We also use the TAQ database to identify daily retail volume, following Boehmer, Jones, Zhang, and Zhang (2021), as the sum of all off-exchange trades (Exchange code 'D' in TAQ, referring to the FINRA Trade Reporting Facility (TRF) feed) between 9:30am and 4:00 pm that received price improvement over the prevailing NBBO. Boehmer, Jones, Zhang, and Zhang (2021) find that retail orders are not typically executed directly on registered exchanges, but instead are re-routed to wholesalers or executed via internalization where orders are filled from the broker's own inventory. According to Boehmer, Jones, Zhang, and Zhang (2021), both wholesaler-rerouted trades and internalized trades are associated with small price improvements over the posted NBBO, mostly by only a fraction of a cent, to induce the re-routing of these retail orders, especially for internalizers who need to demonstrate best execution of their retail client orders, as part of SEC Rule 606.

Short interest data are 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. Since the short interest measure is biweekly, and to better associate our daily operational shorting variable with daily shorting activities by market makers, we use the daily short volume data, provided from individual exchanges, to capture the ratio of daily volume flagged with the short sale identifier. To do that, we combine and aggregate the daily short volume data for each ETF from the following daily short volume feeds: NYSE, ARCA, NASDAQ, BATS, FINRA's TRF, and ORF. The amount of daily short volume for each ETF is scaled by the total daily share volume for that ETF. We supplement these short interest and short volume data with daily information on securities lending demand and supply, utilization, lending fees, and Daily Cost of Borrow Score using the Markit Securities Finance database (formerly Data Xplorers).

The FTD data[25] are from the SEC's website and are made available to the SEC by National Securities Clearing Corporation's (NSCC).[26] The FTD database contains CUSIP numbers, issuer names, prices, and the

---

[24] 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 the same day the create/redeem orders are submitted and accepted, it might take several days for other data vendors and exchanges to reflect this information.

[25] The FTD data can be downloaded from: http://www.sec.gov/foia/docs/failsdata.htm.

[26] The National Securities Clearing Corporation (NSCC) is regulated by the SEC and is a subsidiary of the Depository

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 fails-to-deliver position was initiated.[27] We collect these data from March 22, 2004, which is the beginning of the FTD dataset, through December 31, 2016.[28]

Table 1 presents summary statistics for the key variables in our analysis. These data are computed daily 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 15.47%, and the 99th percentile of its distribution is equal to 85.69%. This variability is accompanied with a high daily ETF short volume activity, which averages 45.6% of daily ETF trading volume and may be a product of the operational shorting mechanism that we described above. Our proxy for retail volume is on average 0.74% of shares outstanding, while daily ETF volume during market hours represents on average 3.62% of shares outstanding. While total ETF volume and retail volume show increasing trends over the sample period, these numbers indicate that retail volume represents on average 20.44% of the total ETF volume during market hours. Moreover, we find that 0.61% of the average ETF's shares are considered failures (FTDs) at any given time. Lastly, the average value of our operational shorting measure is 1.16%, with a standard deviation of 3.59%.

4. Validating the Operational Shorting Measure

After explaining the methodology behind our proposed measure of operational shorting in section 2, we next validate our operational shorting measure in subsection 4.1 by examining its relationship with short

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.

[27] 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.

[28] 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.

interest, short volume, and FTDs. In subsection 4.2, we present the underlying economics of operational shorting and its relation to ETF arbitrage and hedging activity.

4.1 Operational Shorting as a Driver of ETF Short Interest and Failures-to-Deliver

While pervious 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, as noted earlier, the motivation behind operational shorting has very different implications than directional shorting. The primary reason for a failure-to-deliver to occur is the liquidity provision. Similarly, the literature has shown that short volume is strongly related to intraday liquidity provision by market makers (e.g., Comerton-Forde, Jones, and Putnins (2016)). In Table 2, we validate our proposed measure of operational shorting by first examining whether it is an important driver of overall ETF short-selling activity. To address this question, in Panel A, we examine the determinants of ETF biweekly short interest (specifications 1 to 3), daily FTDs (specifications 4 to 6), both scaled by shares outstanding, as well as the daily ETF short volume ratio (specifications 7 to 9), and we include our measure of operational shorting on the right-hand side. We then validate our operational shorting measure by comparing it to an indirect measure of operational shorting. Specifically, in Panel B, we take a measure of total ETF shorting activity, *Short Interest*, and subtract off a measure of directional (or non-liquidity provision) short-selling activity, using Markit's *Total Demand Quantity* variable, which is the number of shares on loan by Markit borrowers.[29] Therefore, the difference between short interest and the short-selling loan inventory represents an alternative proxy for operational shorting.

We include the same set of control variables in both panels. To identify 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

[29] *Total Demand Quantity* includes shares borrowed by Markit borrowers from non-Markit lenders, as well as shares loaned by Market lenders to non-Market borrowers, whereas *Beneficial Owner on Loan Quantity* includes only the shares on loan by Markit borrowers. We find a similar result using *Beneficial Owner on Loan Quantity* instead of *Total Demand Quantity*.

interest coincide with operational shorting activity. Including this variable 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. 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 the highest borrowing costs), and primarily captures directional short-selling 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 with easier to locate shares. Having controlled for the size of a given ETF, we expect ETF share volume to be positively related to FTDs because greater trading intensity increases the likelihood that some shares might not be delivered in a timely fashion. As a proxy for hedging alternatives, we 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.

Panel A of Table 2 shows that the coefficients on all variables have the expected sign and are statistically significant at the 1% level. The number of observations in the first three regressions is lower and reflects the biweekly nature of the short interest information. Regressions (2), (3), (5), (6), (8) and (9) confirm that greater trading volume, short-selling activity, and securities borrowing costs are related to increased ETF FTDs, ETF short interest, and daily ETF short volume levels. Of particular interest is the positive and statistically significant coefficient on *Operational Shorting* in regressions (3), (6), and (9). *Operational Shorting* is a statistically significant determinant of short interest, daily short volume, and FTDs. In regressions (6) and (9), comparing the coefficients on daily *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 volume, short interest, and FTDs all 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. In addition to providing validation for our proposed measure of operational shorting, this finding underscores the need to decompose the effects of short selling that might not necessarily be directional or informational in nature, especially ETF short selling that is due to inherent liquidity provision, as measured by our *Operational Shorting* variable.

Columns (1) – (3) of Table 2, Panel B replicate columns (7) – (9) of Panel A, while columns (4) – (6) introduce *Total Demand Quantity*, which is normalized by *Shares Outstanding.* Columns (7) – (9) construct our proxy for the level of operational shorting, by subtracting the *Total Demand Quantity* from regressions (4) – (6) from the *Short Interest* variable in regressions (1) – (3). We find that the signs and significance of the coefficients on our control variables are generally in line with the results shown in Panel A. Importantly, we find that our *Operational Shorting* measure remains positively and significantly related to the proxy for Operational Shorting. Taken together, the results in Panels A and B provide important validation for our measure of *Operational Shorting*, showing that it is a primary predictor of ETF FTD level, short volume, and the difference between short interest and short-selling loan inventory, an alternative proxy for operational shorting.

4.2 Incentives behind AP's Operational Shorting Activity

As additional validation of our proposed measure, we examine the determinants of operational shorting activity to test whether the ETF market making incentives described in section 2 are supported by the data. The results, shown in Table 3, includes ETF and date fixed effects in addition to ETF liquidity measures (*log(Market Cap)* and *Average Share Turnover*) as essential controls. Panel A contains the primary results and Panel B reports additional robustness checks using various ETF subsamples.

The first hypothesis examined in this table is whether operational shorting activity is driven by AP's arbitrage incentives. The measure of arbitrage opportunities, denoted as the ETF *Mispricing* variable, is defined as the percent difference between ETF price and NAV and measured at (t-4) before the rolling excess demand window for operational shorting, which starts at (t-3). This is an estimate of the potential arbitrage profits available to the AP and we see in specifications (2) through (8) of Panel A that there is a strong, positive relationship between operational shorting and mispricing. This finding is consistent with APs motivation to capture arbitrage profits.

The mispricing result suggests that APs are likely to operationally short to satisfy excess demand for the ETF. Such incentives are likely to be stronger when price pressures associated with this excess demand are attributed to uninformed trading. For this reason, we also control for *Retail Trading* as a percentage of shares outstanding averaged over the same 3-day period when operational shorting excess demand is observed (i.e., between (t-3) and (t-1)). In specification (5) of Panel A in Table 3, we see that retail trading volume is one of the most statistically significant determinants of operational shorting.

Additionally, we include proxies for the existence of hedging alternatives as we expect that the availability of futures and options to hedge the underlying are important determinants of operational shorting. These hedges would shield market makers and APs from unanticipated market swings, especially in the presence of informed trading that might be driving the excess demand in ETF shares. Market makers are also more inclined to hedge and delay creation to minimize transaction costs when the ETF underlying stocks are less liquid. In this case, hedging can preserve the option to wait until the market makers have a better gauge of the permanent component of the ETF order flow before committing to basket creation.

Our model includes multiple proxies for the availability of efficient hedges. *Maximum Rolling R-squared with Available Futures Contracts* measures how well 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.[30] The maximum R^2 across these three regressions is the value assigned to *Maximum Futures R-squared*. If an AP 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.[31] Options listed on the ETF would also facilitate the

[30] 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. 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)."

[31] As the example in Appendix B 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

hedging of ETF-specific risk. We expect a positive relation between these hedging-related variables and *Operational Shorting* because the presence of hedging instruments allows 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). The results in specifications (2) to (6) which include ETF and date fixed effects, as well as specifications (7) and (8) which include style (investment objective) and date fixed effects, illustrate the importance of within- and across-ETF variations in these hedge-related proxies to help explain the levels of ETF operational shorting.

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 additionally include a proxy for *Volume Mismatch,* defined as the log of the ratio between an ETF's 30-day average dollar volume and the average 30-day dollar volume for the fund's underlying securities. Both measures are computed only for U.S. Equity ETFs with available holdings data. 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, and when the underlying securities are relatively illiquid and do not trade as frequently as the ETF. In this case, it would be a valuable option for APs to delay creation before committing to gathering the less-liquid underlying basket of securities and incurring related transaction costs, perhaps waiting for excess demand to subside and ETF order flows to reverse in subsequent days. Regressions (4) and (6) of Panel A confirm our expectation that greater liquidity mismatches are more conducive and positively related to operational shorting.[32] Regressions (7) and (8) use style and date fixed effects to allow for cross-ETF variation, and

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).

[32] We run robustness checks with operational shorting constructed using different timing for the created shares, specifically using cumulative changes in shares outstanding between (t-2) and (t), and another robustness check with cumulative changes between (t-4) and (t) in order to match the trade imbalance period. The results are similar and stronger in certain cases. However, this approach leads to counting the same create order multiple times.

confirm all prior results that arbitrage, liquidity mismatch, and hedge availability are all important determinants of operational shorting activities.

Our model also includes controls for ETF-specific transaction costs and operational frictions: *Creation Unit Fee* (for a single creation unit transaction), and *ln(Creation Unit Dollar Size)*.[33] As discussed in Section 2, we expect larger ETF portfolios with higher creation unit sizes and fees to encourage APs to engage in more operational shorting in order to minimize these costs.[34] In Panel A of Table 3, the coefficients on *ln(Creation Unit Dollar Size)* and *Creation Unit Fee* in regression (3) 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. However, in our full model, regression (8), we find that these institutional frictions are less significant than in regression (3), as other factors are stronger determinants of operational shorting activity.

Panel B of Table 3 provides robustness checks using different ETF subsamples. We additionally control for average daily *ETF Volume* during market hours (i.e., between 9:30am and 4:00 pm) over the same period that average *Retail Volume* and operational shorting are observed. Retail trading volume as percent of shares outstanding remains strongly significant, along with the arbitrage opportunities variable (*Mispricing*) and hedging proxies. The results suggest that operational shorting is an important form of liquidity provision for different types of ETFs, and it is used by market makers to supply liquidity especially in the presence of less-liquid underlying securities and when the ETF order flow is associated with price pressures from uninformed retail trading.

Taken together, the results in Table 3 suggest that there is a greater propensity for operational shorting activity with smaller actively traded ETFs that have greater potential arbitrage opportunities, especially in the

[33] The historical *Creation Unit* and *Creation Fee* variables are only available for a subset of ETFs as provided by ETF Global data. Similarly, the *Maximum Rolling R-Squared* 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. For this reason, we provide the regression results with and without those variables.

[34] 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 closely aligned 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.

presence of retail order flow, readily available hedging alternatives, and larger liquidity mismatches. Since APs' arbitrage activities play an important role in the proper functioning of ETF markets, it is important to examine how operational shorting influences future ETF and NAV returns, and consequently future ETF mispricing. 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 B, 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.[35] More importantly, the consistency in the expected and observed economics provides important validation for our measure of operational shorting.

5. Operational Shorting and ETF/Underlying Returns, Mispricing and Liquidity

Given the evidence in section 4 validating our proposed measure of operational shorting, we now use that measure to revisit the three primary empirical questions examined in the prior ETF literature: the relationship between ETF trading/arbitrage and future returns in section 5.1; the mispricing/tracking error associated with ETF investments in section 5.2; and the impact of ETFs on the liquidity of the underlying securities in section 5.3. In all three settings, we highlight the important role that operational shorting plays in these economic questions, that has not been addressed by the prior literature.

5.1 ETF Operational Shorting and Future Returns

There is a well-established literature documenting the relationship between increased short-selling activity in individual stocks and future return underperformance using short interest, FTDs, and a variety of other short-selling measures.[36] One common interpretation of this strong predictive relation is that short sellers

---

[35] As noted earlier, we report robustness checks for Table 3 using an alternative version of operational shorting that uses non-overlapping periods measured over discrete weekly intervals. The results corroborate our evidence presented using the daily sample and are reported in Table C.1 of Appendix C, which contains Panel A and Panel B results that mirror both panels from Table 3.

[36] 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.

are informed, but constraints and borrowing costs prevent them from fully incorporating their information in market prices. Also, for ETFs specifically, Brown, Davies, and Ringgenberg (2021) find that ETF arbitrage activity (i.e., creation and redemption) has predictive power for future returns. They interpret this predictability result as evidence of a two-stage process. First, AP creations/redemptions accommodate non-fundamental demand, thereby moving prices of the ETF's underlying securities from their fundamental values. Second, over time the prices of the underlying securities correct, resulting in the observed return predictability.

As described earlier, the incentives associated with operational shorting, and its resulting return prediction differ from both of these prior results. First, in contrast to evidence of informed short selling for individual equities, operational shorting activity should facilitate the liquidity provision. The exception to Rule 204 for market makers is granted only when the operational short position is "attributable to bona fide market making activities." For individual equities, Comerton-Forde, Jones, and Putnins (2016) separate liquidity-demanding, informed short-selling from liquidity-supplying shorts. Their evidence that liquidity-supplying shorts are strongly contrarian in nature and improve market quality in contrast to informed shorts, provides a useful comparison for ETFs.

Second, Brown, Davies and Ringgenberg (2021) argue that their return prediction result is driven by the transmission of a non-fundamental demand shock to the underlying through arbitrage activity, which subsequently reverses. An alternative approach to satisfying this non-fundamental shock is operational shorting. Having sold ETF shares to satisfy excess demand and the corresponding mispricing, APs can create and deliver the ETF shares (i.e., the arbitrage event focused on by Brown, Davies and Ringgenberg (2021)) or they can 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 consider 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 along with their liquidity shocks and subsequent reversals. While this would also appear as return predictability, the non-fundamental demand shock would be transmitted to the ETF market price and not the prices of the underlying securities (i.e. the NAV).

To test for this unique role of operational shorting using return data for both ETF market prices and the NAV, we examine the impact of operational shorting on both contemporaneous and future returns. For equity ETFs and on each day, we also compute the daily risk-adjusted excess returns using the Fama-French four factor model, based on the ETF return and the return on the basket's NAV (which excludes the ETF mispricing component) as follows: first, we compute the four factor betas ($\beta_{mkt,\,t-1}$, $\beta_{smb,\,t-1}$, $\beta_{hml,\,t-1}$, and $\beta_{umd,\,t-1}$) of the ETF using the daily ETF returns (or NAV returns) and risk factors over a 200-day rolling window between (t-200) and (t-1). Then, using these betas and the risk factors values in day (t), we compute the daily ETF risk-adjusted excess return (or daily NAV excess return) as the difference between the daily return and benchmark return as described in the formula below:

$$Excess\ ret_t = ret_t - \left(r_{f,t} + \beta_{mkt,\,t-1}\left(r_{mkt,\,t} - r_{f,t}\right) + \beta_{smb\ t-1} SMB_t + \beta_{hml\ t-1} HML_t\ + \beta_{umd\ t-1} UMD_t\right)$$

We then regress the return and daily excess returns on several ETF characteristics. The key independent variables in these regressions are our *Operational Shorting* variable and the '*Create' Orders* measure. We use daily *Operational Shorting* as defined earlier, and we include *Create Orders* as simply the change in ETF shares outstanding between (t-1) and (t) if shares are created, and zero otherwise.

Table 4 reports the results of this test. Panel A provides the descriptive statistics for all variables used in this daily return analysis. *Total Return* and *Fama-French Four-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 0.3 bps and -1.8 bps during the sample period, with a maximum return of 11.85% and 5.35%, respectively. Panel B of Table 4 reports the regression results. At the top of each column, the table specifies whether the dependent variable is calculated using *ETF* or *NAV* returns, if it is the total return (*Ret*) or the Fama-French Four-Factor Excess Return (*FF4* α), and whether it is measured over the concurrent day (*t*) or the following day (*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, specifications (4) - (8) use the sample of U.S. equity ETFs, and the last two specifications (9) and (10) are based on non-U.S.-equity ETFs (e.g., fixed income and foreign equity ETFs). In specifications (7) and (8), we further split the domestic equity ETF subsample into ETFs with high vs. low liquidity mismatches, where *High* indicates the ETF was more liquid than the underlying securities, as measured by the difference in intraday spreads between the two. The number of observations reflect the different samples, and we only keep equity ETFs with available portfolio holdings data in order to compute the average underlying basket spreads used in the liquidity mismatch variable.

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 that price pressures leading to mispricing present important arbitrage opportunities that are a primary motivation behind operational shorting. Using daily excess returns as the measure of interest, specification (4) finds similar results for U.S. equity ETFs. In specifications (2), (5), and (9), we find that *Operational Shorting* is negatively related to future, next-day ETF returns for all types of ETFs, suggesting return reversals in the ETF market price following operational shorting by APs. This finding documents the contrarian nature of operational shorting as APs respond to investors' buying demand.[37] While this evidence is consistent with the liquidity provision motivation for operational shorting, it stands in sharp contrast to the insignificant and much weaker reversal evidence related to *Create Orders.* In other words, APs are less likely to submit create orders when they believe in future reversals in returns and order flows. Reversals also appear to be much stronger for U.S. equity ETFs with larger liquidity mismatches, as shown in specification (8).

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 non-fundamental demand shocks to the underlying securities. In regressions (3), (6) and (10), we repeat the analysis using returns on the underlying securities (*NAV*) instead of ETF market prices to calculate our performance measures. Whether we use NAV total returns for the overall sample or the NAV four factor excess returns for the U.S. equity subsample (i.e., specifications (3) and (6)) or total returns for the non-U.S.-equity subsample (i.e. specification (10)), we find the same results: operational shorting activity in the previous day has no predictive

[37] As described in more detail within Appendix B, APs typically hedge their short positions during the same day using derivatives or a large, liquid ETF such as the S&P 500 index SPDR ETF. This is done to establish a market-neutral position that protects the AP against the observed pattern of short-term gains and reversals in an ETF's market price.

power for the returns on the underlying securities.

To better understand what drives this ETF return predictability, we turn to the subsample regression analysis in specifications (7) and (8). 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 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, we find much weaker return predictive power of operational shorting. In the 'High' liquidity mismatch subsample, however, where the ETF is relatively much more liquid than the underlying basket, we find a statistically and economically significant negative relation between operational shorting and future returns that is more than twice as strong as the one observed for the ETFs with a 'Low' liquidity mismatch. Because the underlying securities are less liquid, it takes more time to assemble the basket and price discovery is likely to occur first in the more liquid ETF shares.

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, and where liquidity-driven trades in the ETF are more likely to occur relative to the underlying basket of securities.[38]

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 and liquidity-motivated order flow. Our evidence

[38] While our measure of operational shorting uses a three-day settlement window on a rolling basis, we report robustness checks for Table 4 using the weekly ETF sample based on the non-overlapping operational shorting variable that is measured over discrete weekly intervals. The weekly operational shorting measure is comparable to the timing of the operational shorting measure plus the additional settlement time afforded APs by their exemption. Like the approach described earlier in Table 3, the weekly return results are reported in Table C.2 of Appendix C and displays results in Panel A and Panel B that mirror both panels in Table 4. As shown in Table C.2, the weekly data provide even stronger evidence on the price pressures and return reversals around operational shorting activities. 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.

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.[39] 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.[40] 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 acting 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.

5.2 The Effects of Operational Shorting on ETF Mispricing

The unique liquidity provision mechanism for ETFs has added important functionality relative to existing investment products such as index funds, including intraday trading, short selling, and greater tax efficiency, to name a few. However, that same liquidity mechanism has raised greater concerns regarding the mispricing of ETFs relative to their underlying securities and the standard deviation of that mispricing (or tracking error), a primary focus of the early academic literature. As a more recent example of that literature, Bae and Kim (2020) show that illiquid ETFs and ETFs with illiquid underlying securities have larger tracking errors. At the same time, our prior results show that APs are more likely to operationally short when mispricing is higher, suggesting a possible mitigating effect.

In this section, we revisit the determinants of mispricing and tracking error, accounting for operational shorting. The results of this analysis are presented in Table 5. To gauge the effect of operational shorting on

[39] Our evidence is consistent with Ben-David et al. (2022) who showed that ETFs sponsors tend to increase the supply of new ETFs by introducing specialized ETFs at the time when the underlying stocks are overvalued due to retail trading and heightened investors sentiment.

[40] As noted earlier in the Introduction, our results are also consistent with the increasingly popular practice of market makers paying brokerage companies for retail ETF order flow (payment for order flow or PFOF) to help differentiate between informed and uninformed traders. https://www.wsj.com/articles/the-secretive-firm-set-to-expand-in-retail-options-two-sigma-securities-1494446194, and https://www.wsj.com/articles/why-free-trading-on-robinhood-isnt-really-free-1541772001. The four most popular retail brokerage companies - TD Ameritrade, Robinhood, E*Trade, and Charles Schwab - made more than $1.1 billion in the first six months of 2020 from selling retail order flow: https://www.wsj.com/articles/confetti-free-stocks-does-robinhoods-design-make-trading-too-easy-11597915801.

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 use *Absolute Mispricing Change,* which, similar to tracking error, allows us to test whether operational shorting activities are aimed to arbitrage away and shrink any mispricing opportunities. 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 the presence of arbitrage opportunities. Market makers are motivated by the profit incentives through price reversals and the mispricing arbitrage which represents the compensation for this liquidity provision via operational shorting and helps to quickly reduce the size of mispricing opportunities. Models (5) through (8) repeat these tests with the *Absolute Mispricing Change* as the dependent variable and confirm the significant negative relation between operational shorting and tracking error, suggesting that operational shorting is acting as an essential liquidity provision that serves to reduce tracking errors, consistent with the evidence in Bae and Kim (2020).

5.3 The Effects of Operational Shorting on Underlying Securities

One of the most examined aspects of ETF trading is its impact on the liquidity of the underlying securities. A growing academic literature provides evidence that ETF ownership is positively associated with higher intraday spreads and volatility of the ETF's underlying basket of securities. Ben-David, Franzoni, and Moussawi (2018), for example, find that liquidity shocks propagate through the arbitrage channel. However, our earlier evidence suggests that operational shorting is used in lieu of creation in order to accommodate liquidity, perhaps mitigating the impact of such liquidity shocks. Additionally, there is an earlier literature that suggests individual equity FTDs (i.e., Fotak et al. (2014), and Merrick, Naik, and Yadav (2005)) can serve as an "important release valve".[41]

[41] This notion of a "release valve" is also supported in terms of short-selling activity's impact on loosening institutional

Table 6 presents the results of our analysis. We follow Fotak et al. (2014) and use average spreads and intraday volatility as measures of the liquidity and market quality of trading in individual securities.[42] 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 every 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 price of last trade of 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 like 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 our analysis of liquidity and volatility effects using these underlying basket liquidity measures at the ETF level, controlling for ETF liquidity measures (e.g. ETF Share Turnover, *log*(ETF Market Cap), as well as the contemporaneous and lagged ETF levels of the dependent variable). Additionally, we control for the "liquidity level" effect of ETF ownership that is documented in earlier literature. For example, Ben-David, Franzoni, and Moussawi (2018) report causal evidence that links ETF ownership with increased volatility of underlying securities. To control for this effect, we construct a measure of average ownership by ETFs in the underlying basket of stocks using all the stocks in the ETF portfolio 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 one- to three-day lags of these ETF liquidity measures to control for any persistence in volatility and spread measures, as well as address potential

constraints and sharpening price discovery. For example, Chu, Hirshleifer, and Ma (2016) show that the introduction of Regulation SHO (which reduced short selling constraints) has led to a reduction in returns to asset pricing anomalies. The authors suggest that this increase in short selling ability has made arbitrage of asset pricing anomalies easier and thus has decreased the returns to these strategies. In effect, like FTDs, Regulation SHO acted as another form of release valve which can lead to increased market efficiency.

[42] 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).

reverse causality concerns.

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 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 previous literature, we also 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 Hamm (2014); Glosten, Nallareddy, and Zou (2021); Israeli, Lee, and Sridharan (2017); Dannhauser (2017); Saglam, Tuzun, and Wermers (2019); and Agarwal, Hanouna, Moussawi, and Stahel (2017)).[43]

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, and report that operational shorting is also negatively related to intraday return volatility. 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, Franzoni, and Moussawi, 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, 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 of the underlying stocks.

The negative relation between operational shorting and intraday spreads and the volatility of underlying stocks suggests that greater liquidity in the underlying stocks coincides with higher levels of

[43] 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.

operational shorting. Thus, an AP's operational shorting activity can be associated with 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. 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 and widening spreads. If, on the other hand, the AP does not engage in operational shorting, then it will transmit those liquidity shocks directly to the underlying securities, as we have described earlier. 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 profit from delaying 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 (2021), Xu, Yin, and Zhao (2019), 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 liquidity to the market in response to this informed order flow and acting as a buffer through operational shorting in response to liquidity-driven 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 three 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. Like O'Hara and Ye (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, either due to price pressures and reversals (negative autocorrelations), or due to positive feedback effects and autocorrelations. 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 intraday variance ratios, calculated on a daily basis using returns over all 15-second intervals during the day in order to measure the transitory component of stock prices between "intraday indicative value" dissemination intervals (typically 15 seconds).[44] Additional control variables are 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 from a random walk pattern. The results shown in this panel confirm that price efficiency in the underlying securities held by ETFs improves as APs and market makers engage in higher levels of operational shorting activity which absorbs ETF price pressures without transmitting these shocks to the underlying basket. Consistent with the results in Table 4 related to when APs actively create new ETF units, it appears that ETF market makers could be taking action in both the ETF and the underlying securities. This activity by APs can help absorb liquidity shocks and reduce the mean-reverting component and noise in stock prices in order to make 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 uninformed liquidity traders and, therefore, 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. This is consistent with the evidence in Comerton-Forde, Jones, and Putnins (2016), and with the notion that operational shorting is a potentially beneficial by-product of the liquidity provision from these market makers and APs.

6. Conclusion

While much of the academic literature on ETFs has focused on creation and redemption, only a small fraction of secondary market trading translates into this specific arbitrage trade. Instead, market makers and

[44] Prior to 2020 and according to the SEC, "Exchange Rules 14.11(c) and 14.11(i) relate to the listing and trading of Index Fund Shares and Managed Fund Shares on the Exchange. Among a number of other requirements, numerous sub-paragraphs of each of these rules require that an intraday estimate of the value of a share of each series (the "Intraday Indicative Value" or "IIV") of Index Fund Shares and Managed Fund Shares be disseminated and updated at least every 15 seconds." https://www.sec.gov/rules/sro/cboebzx/2020/34-88259.pdf.

authorized participants accommodate ETF liquidity provision in a number of different ways, including operational shorting. While market makers in many securities can operationally short-sell securities they do not own in order to provide liquidity, this particular SEC exemption is of particular importance for making markets in ETFs. Failures-to-deliver, an important indicator of operational shorting activity for ETFs, are higher for ETFs than all other security types (i.e., common stocks, OTC stocks, corporate bonds, ADRs, structured products, trusts and other securities), constituting approximately 80% of all fails.

We examine the importance of operational shorting to the liquidity provision and pricing of ETFs. We first propose a measure of operational shorting that compares the buy-sell trade imbalance to changes in the daily shares outstanding of an ETF. A positive buy-sell trade imbalance that is not accompanied by ETF share creation suggests the AP has yet to create those shares and deliver them to investors, our definition of operational shorting. Then, we validate our proposed measure by examining both its relationship to measures of short activity and its determinants. Finally, we re-visit three important previously documented results from the ETF literature (i.e., forecasting returns, ETF mispricing, and the impact of ETFs on their underlying securities) and find that our measure of operational shorting plays an important and previously unrecognized role. Overall, our results suggest that operational shorting plays an essential role in ETF liquidity provision and that in assessing ETF short-selling activity, distinguishing between directional and operational shorting activity is important in understanding the underlying economics.

While our focus in this paper has been on the short side of the liquidity provision in ETFs, 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 setting, 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 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. Therefore, additional research into this asymmetry between operational buying and operational shorting is warranted.

References

Agapova, A., 2011, Conventional mutual index funds versus exchange–traded funds, *Journal of Financial Markets* 14, 323–343.

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.

Bae, K., and D. Kim, 2020, Liquidity risk and exchange-traded fund returns, variances, and tracking errors, *Journal of Financial Economics* 138, 222-253.

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

Ben-David, I., F. Franzoni, B. Kim, and R. Moussawi, 2022, Competition for Attention in the ETF Space, *Review of Financial Studies*, forthcoming.

Ben-David, I., F. Franzoni, and R. Moussawi, 2017, Exchange Traded Funds (ETFs), *The Annual Review of Financial Economics*, 9 (6), 169-189.

Ben-David, I., F. Franzoni, and R. Moussawi, 2018, Do ETFs Increase Volatility? *The Journal of Finance* 73 (6), 2471-2535.

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

Blume, M., and R. Edelen, 2004, S&P 500 Indexers, Tracking Error, and Liquidity: a Complex Answer to Profiting, *Journal of Portfolio Management* 30, 37-46.

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.

Boehmer, E., C. Jones, X. Zhang, and X. Zhang, 2021, Tracking Retail Investor Activity, *The Journal of Finance* 76, 2249-2305.

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

Box, T., R. Davis, R. Evans, and A. Lynch, 2021, Intraday arbitrage between ETFs and their underlying portfolios, *Journal of Financial Economics*, 141 (3), 1078-1095.

Brown, D., S. Davies, and M. Ringgenberg, 2021, ETF Arbitrage and Return Predictability, *Review of Finance*, 25 (4), 937-972.

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, 2021, The Sound of Many Funds Rebalancing, *The Review of Asset Pricing Studies*, 11 (3), 502-551.

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

Comerton-Forde, C., C. Jones, and T. Putnins, 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, *The 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., R. Michaely, and M. O'Hara, 2001, The accuracy of trade classification rules: Evidence from Nasdaq, *Journal of Financial and Quantitative Analysis* 35, 529-551.

Elton, E., M. Gruber, G. Comer, and K. Li, 2002, Spiders: Where are the Bugs? *Journal of Business* 75, 453-473.

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

Evans, R., O. Karakaş, R. Moussawi, and M. Young, 2022, Phantom of the Opera: ETF Shorting and Shareholder Voting, *Working paper*.

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.

Gastineau, G., 2002, Equity Index Funds have Lost Their Way, *Journal of Portfolio Management* 28, 55-64.

Gastineau, G., 2004, The Benchmark Index ETF Performance Problem. A Simple Solution, *Journal of Portfolio Management* 30, 96-103.

Glosten, L. R., S. Nallareddy, and Y. Zou, 2021, ETF Activity and Informational Efficiency of Underlying Securities. *Management Science* 67, 22-47.

Hamm, S., 2014, The Effect of ETFs on Stock Liquidity, *Working paper*.

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

Huang, S., M. O'Hara, and Z. Zhong, 2021, Innovation and Informed Trading: Evidence from Industry ETFs, *Review of Financial Studies* 34 (3), 1280-1316.

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.

Jones, C., and O. Lamont, 2002, Short-sale constraints and stock returns, *Journal of Financial Economics* 66 (2-3), 207-239.

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

Lee, C., and M. Ready, 1991, Inferring trade direction from intraday data, *The 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.

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

Merrick, J., N. Naik, and P. Yadav, 2005, Strategic trading behavior and price distortion in a manipulated market: anatomy of a squeeze, *Journal of Financial Economics* 77 (1), 171-218.

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

Moussawi, R., K. Shen, and R. Velthuis, 2021, ETF Heartbeat Trades, Tax Efficiencies, and Clienteles: The Role of Taxes in the Flow Migration from Active Mutual Funds to ETFs, *Working paper*.

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

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

Ofek E., and M. Richardson, 2003, DotCom mania: The rise and fall of Internet stock prices, *The Journal of Finance* 58, 1113-1138.

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

Pagano, M.S., J. Sedunov, and R. Velthuis, 2021, How did Retail Investors Respond to the COVID-19 Pandemic? The Effect of Robinhood Brokerage Customers on Market Quality, *Finance Research Letters* 43.

Pan, K., and Y. Zeng, 2016, ETF arbitrage under liquidity mismatch, *Working paper*.

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*.

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

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.

Xu, L., X. Yin, and J. Zhao, 2019, The sidedness and informativeness of ETF trading and the market efficiency of their underlying indexes, *Pacific-Basin Finance Journal* 58, 101217.

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

Table 1 – Summary Statistics: This table presents daily 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 variable. We provide a complete list of variable names, sources, and definitions in Appendix D.

	Variable	Obs	Mean	Std.Dev.	p1	p25	p50	p75	p99
Entire ETF Sample	Fail-to-Deliver Shares, as % of Shares Outstanding	3,007,239	0.61%	3.98%	0.00%	0.00%	0.00%	0.11%	11.45%
	Operational Shorting, as % of Shares Outstanding	2,601,960	1.16%	3.59%	0.00%	0.00%	0.04%	0.74%	21.77%
	Net Create/Redeem Activity: log (1 + % change in Shares Outstanding)	3,006,045	0.14%	4.39%	-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%
	Short Volume as % of Daily Volume	2,098,960	45.64%	27.12%	0.00%	24.50%	45.43%	65.46%	100.00%
	Market Capitalization, $ million	3,007,054	$1,092.00	$5,394.00	$1.38	$16.81	$86.20	$428	$18,523
	Daily Share Turnover, % of Shares Outstanding, 60-day average	2,950,760	4.19%	10.17%	0.10%	0.62%	1.17%	2.80%	55.49%
	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	0.03%	0.57%	-2.33%	-0.12%	0.02%	0.18%	2.12%
	Maximum Rolling R-Squared with Available Futures Contracts	2,673,729	53.34%	29.15%	0.13%	30.47%	59.27%	77.39%	96.43%
	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	70,758	43,401	25,000	50,000	50,000	100,000	250,000
	Creation Unit Fee	931,999	$1,631	$3,030	$100	$500	$500	$1,400	$15,000
	Retail Volume, as % of Shares Outstanding	2,586,474	0.74%	1.84%	0.00%	0.10%	0.21%	0.49%	13.32%
	Daily Volume (during market hours), as % of Shares Oustanding	2,959,147	3.62%	8.83%	0.00%	0.29%	0.74%	2.28%	57.27%
	Bid-Ask Spread, at Close	2,956,434	0.36%	0.91%	0.01%	0.07%	0.15%	0.34%	3.54%
	Intraday NBBO Bid-Ask Spread, Trade Size Weighted	2,772,053	0.46%	6.03%	0.01%	0.06%	0.14%	0.29%	2.47%
	Intraday Volatility, using second-by-second intraday returns	2,703,755	0.03%	2.67%	0.00%	0.00%	0.01%	0.01%	0.05%
	Daily Cost of Borrow Score	1,768,565	3.20	1.52	1.00	2.00	3.00	4.00	7.00
	Indicative Fee	1,588,220	4.44%	3.86%	0.38%	1.75%	3.50%	6.00%	18.00%
	Short Interest Ratio	2,960,079	5.25%	15.47%	0.00%	0.28%	0.91%	3.29%	85.69%
US Equity ETF Sample	Intraday NBBO Bid-Ask Spread, Trade Size Weighted	1,072,682	0.46%	7.14%	0.01%	0.05%	0.11%	0.22%	1.93%
	Intraday Volatility, using second-by-second intraday returns	1,051,449	0.04%	3.75%	0.00%	0.00%	0.01%	0.01%	0.04%
	Daily Cost of Borrow Score	683,106	2.91	1.43	1.00	2.00	3.00	4.00	7.00
	Indicative Fee	611,415	3.74%	3.44%	0.38%	1.38%	3.00%	5.00%	15.00%
	Short Interest Ratio	1,107,382	6.02%	17.53%	0.00%	0.27%	0.86%	3.10%	105.30%
	Average Intraday NBBO Bid-Ask Spread for Underlying Basket Stocks	680,981	0.25%	2.10%	0.03%	0.05%	0.07%	0.11%	0.98%
	Average Intraday Volatility of Underlying Basket Stocks	680,971	0.02%	0.29%	0.01%	0.01%	0.02%	0.02%	0.06%
	Underlying Basket Stocks, Average Daily Cost of Borrow Score	639,591	1.10	0.20	1.00	1.01	1.03	1.12	1.98
	Underlying Basket Stocks, Average Indicative Fee	680,992	0.54%	0.40%	0.29%	0.39%	0.43%	0.54%	2.43%
	Underlying Basket Stocks, Average Short Interest Ratio	680,992	4.21%	2.29%	1.25%	2.43%	3.66%	5.48%	11.58%

Table 2 – The Determinants of ETF Short Interest and Failures-to-Deliver: This table displays regression results using ETF daily observations. The dependent variables in Panel A are *ETF Short Interest, ETF Fail-to-Deliver Shares*, and *ETF Daily Short Volume Ratio.* The first two dependent variables are normalized by total ETF *Shares Outstanding*, and the last dependent variable represents the ETF daily short volume as a fraction of daily ETF volume. The dependent variables in Panel B are *ETF Short Interest, ETF Total Demand Quantity,* and *ETF Short Interest net of Total Demand Quantity*. All three of these dependent variables are normalized by total ETF *Shares Outstanding*. *Short Interest* is available biweekly while *Fail-to-Deliver* and *Short Volume* are available on a daily frequency. The main independent variable is *Operational Shorting*, which is defined everyday as the excess demand (ETF buys – ETF sells) for ETF shares that exceeds ETF share creation activity, scaled by average ETF shares outstanding. Other ETF controls include the lagged *Short Interest Ratio* and *Daily Share Turnover,* as well as the *Daily Cost of Borrow Score* and an *Available Options Dummy* which is equal to 1 if we can find options with positive open interest in CBOE for the ETF. We provide a complete list of variable names, sources, and definitions in Appendix D. 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: ETF Short-Selling Activity

	Short Interest / Shares Outstanding (t)			Fail-to-Deliver Shares / Shares Outstanding (t)			Short Volume / Total Volume (t)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log (Market Cap) at (t-1)	-0.00203***	-0.00183***	-0.000878**	-0.00352***	-0.00326***	-0.00258***	-0.0103***	-0.00855***	-0.00328*
	(-5.708)	(-4.766)	(-2.213)	(-13.78)	(-8.799)	(-6.923)	(-6.668)	(-5.069)	(-1.958)
Share Turnover, as % of Shares Outstanding	0.0394***	0.0310***	0.0284***	0.0721***	0.0755***	0.0737***	0.0256*	0.0466***	0.0365***
at (t-1)	(5.427)	(4.577)	(4.228)	(7.650)	(6.381)	(6.185)	(1.786)	(3.386)	(2.632)
Short Interest Ratio, as % of Shares	0.697***	0.767***	0.767***	0.0469***	0.0332***	0.0322***	0.137***	0.103***	0.0946***
Outstanding at (t-1)	(37.34)	(43.58)	(43.52)	(8.947)	(6.807)	(6.655)	(11.61)	(9.573)	(8.813)
Daily Cost of Borrow Score at (t-1)		0.000558***	0.000536***		0.000421***	0.000408***		0.0104***	0.0101***
		(2.803)	(2.683)		(3.038)	(2.933)		(12.92)	(12.61)
Available Options Dummy at (t-1)		0.00258***	0.00230***		-0.00282***	-0.00300***		-0.0291***	-0.0309***
		(3.244)	(2.873)		(-4.645)	(-4.891)		(-7.398)	(-7.929)
Operational Shorting, as % of Shares			0.105***			0.0753***			0.544***
Outstanding at (t)			(7.493)			(9.613)			(15.72)
Constant	0.0212***	0.0185***	0.0127***	0.0165***	0.0171***	0.0129***	0.498***	0.489***	0.457***
	(10.41)	(7.388)	(4.929)	(14.36)	(9.171)	(6.893)	(66.26)	(54.53)	(51.47)
Observations	260,352	163,454	163,454	2,925,879	1,755,400	1,755,400	2,066,445	1,418,118	1,418,118
R-squared	0.787	0.848	0.849	0.100	0.125	0.129	0.118	0.136	0.141

Panel B: Total, Directional and Liquidity-Providing Short Sales

	Short Interest / Shares Outstanding (t)			Total Demand Quantity / Shares Outstanding (t)			(Short Interest - Total Demand Quantity) / Shares Outstanding (t)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log (Market Cap) at (t-1)	-0.00203***	-0.00183***	-0.000878**	-0.00072*	-0.00117***	-0.00108**	-0.000639	-0.000445	0.000328
	(-5.70)	(-4.76)	(-2.21)	(-1.93)	(-2.68)	(-2.46)	(-1.48)	(-0.90)	(0.66)
Share Turnover, as % of Shares	0.0394***	0.0310***	0.0284***	0.01427**	0.01521**	0.01499**	0.0149**	0.0148**	0.0130*
Outstanding at (t-1)	(5.43)	(4.58)	(4.23)	(2.47)	(2.50)	(2.47)	(2.11)	(2.02)	(1.77)
Short Interest Ratio, as % of Shares	0.697***	0.767***	0.767***	0.16326***	0.16653***	0.16647***	0.626***	0.629***	0.628***
Outstanding at (t-1)	(37.34)	(43.58)	(43.52)	(13.22)	(13.31)	(13.31)	(30.64)	(30.41)	(30.37)
Daily Cost of Borrow Score at (t-1)		0.000558***	0.000536***		0.00033	0.00033		0.000297	0.000286
		(2.79)	(2.68)		(1.35)	(1.35)		(1.24)	(1.20)
Available Options Dummy at (t-1)		0.00258***	0.00230***		0.00325***	0.00322***		-0.00125	-0.00148
		(3.24)	(2.87)		(3.15)	(3.12)		(-1.28)	(-1.52)
Operational Shorting, as % of Shares			0.105***			0.00897			0.0759***
Outstanding at (t-1)			(7.49)			(1.58)			(5.58)
Observations	260,352	163,454	163,454	141,250	123,114	123,114	141,250	123,114	123,114
R-squared	0.787	0.848	0.849	0.612	0.615	0.615	0.807	0.813	0.813

Table 3 – The Determinants of Operational Shorting: This table displays regression results using ETF daily observations where the dependent variable in both panels is the *Operational Shorting* variable, which is defined everyday as the excess demand (ETF buys – ETF sells) for ETF shares that exceeds ETF share creation activity, scaled by average ETF shares outstanding. In Panel A, the main independent variables are: the existence of a futures hedge – *Maximum Rolling R-Squared with Available Futures Contracts* and *Available Options Dummy*, the existence of retail trading activity – *Average 3-Day [(t-3), (t-2) and (t-1)] Retail Volume as % of shares outstanding*, the existence of arbitrage opportunities – *Mispricing* defined as the percent difference between ETF price and NAV, and the existence of liquidity or volume mismatches between the ETF and underlying securities. The *Proxy for Liquidity Mismatch* is defined as the difference between intraday spreads of the ETF's underlying securities and the ETF. The *Proxy for Volume Mismatch* is defined as the log of the ratio between ETF 30-day average dollar volume and the average 30-day dollar volume for the underlying securities. We control for operational determinants such as historical *Creation Unit Dollar Size*, and *Creation Unit Fee (per share)*. All independent variables are lagged and *Mispricing* is measured at (t-4) before the rolling excess demand window for Operational Shorting which starts at (t-3). All variables constructed using ETF portfolios are limited to U.S. Equity ETFs. We include other controls such as the lagged *log(Market Cap)*; *60-day Average Share Turnover*. We provide a complete list of variable names, sources, and definitions in Appendix D. 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. The first specifications include ETF and date fixed effects, and the last specifications include Style and date fixed effects. Panel B provides robustness checks using various subsamples of ETFs, where we additionally control for the *average daily volume between (t-3) and (t-1), as % of shares outstanding*. Standard errors are clustered at the ETF and date levels. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Determinants of Operational Shorting using Daily Sample

	Operational Shorting, as % of Shares Oustanding, at day (t)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log (Market Cap), at (t-1)	-0.0077***	-0.0082***	-0.0086***	-0.0097***	-0.0093***	-0.0102***	-0.0048***	-0.0091***
	(-14.43)	(-14.11)	(-10.94)	(-7.40)	(-14.84)	(-6.00)	(-9.12)	(-5.36)
Average Share Turnover, as % of Shares Outstanding, at (t-1)	0.0318***	0.0320***	0.0249***	0.0359***	-0.0058*	0.0131	-0.0024	0.0038
	(11.27)	(11.42)	(6.55)	(3.75)	(-1.93)	(0.99)	(-0.58)	(0.37)
Maximum Rolling R-Squared with Available Futures Contracts at (t-1)		0.0110***	0.0094***	0.0159***	0.0137***	0.0122***	0.0116***	0.0323***
		(6.84)	(4.70)	(4.55)	(7.24)	(3.53)	(4.15)	(3.55)
Available Options Dummy at (t-1)		0.0026***	0.0025***	0.0029***	0.0025***	0.0028**	0.0092***	0.0134***
		(3.91)	(3.00)	(2.84)	(3.63)	(2.22)	(4.97)	(4.36)
Mispricing at (t-4): % difference between ETF price and NAV		0.1293***	0.0956***	0.1423***	0.1057***	0.1554***	0.1576***	0.2597***
		(11.51)	(6.68)	(6.01)	(8.00)	(3.79)	(10.48)	(4.13)
Creation Unit Dollar Size, log, at (t-1)			0.0060***			0.0012		0.0067***
			(7.61)			(0.74)		(4.35)
Creation Unit Fee, per share, at (t-1)			0.0492***			0.0081		-0.0565
			(3.19)			(0.38)		(-1.59)
Liquidity Mismatch, at (t-1): Average Intraday Basket Spread - Intraday ETF Spread				0.1556***		0.1666		0.6047***
				(2.78)		(1.57)		(2.90)
Volume Mismatch, at (t-1): Log(30-Day ETF $ Volume / Implied 30-Day Basket $ Volume)				0.0097***		0.0086***		0.0138***
				(4.69)		(3.80)		(4.46)
3-Day Average Retail Volume, as % of Shares Outstsanding, at (t-1)					0.4442***	0.4111***	0.3361***	0.3074***
					(26.07)	(8.30)	(17.57)	(5.30)
Constant	0.0450***	0.0405***	-0.0450***	0.0481***	0.0460***	0.0370	0.0224***	-0.0703***
	(18.46)	(19.40)	(-4.31)	(8.47)	(18.47)	(1.54)	(17.18)	(-3.57)
Fixed Effects	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date	Style, Date	Style, Date
Observations	2,633,065	2,624,690	1,745,676	638,422	2,111,301	431,083	2,072,913	428,406
R-squared	0.161	0.163	0.198	0.226	0.220	0.294	0.095	0.157

Panel B: Robustness using Various ETF Subsamples

	Operational Shorting, as % of Shares Oustanding, at day (t)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log (Market Cap), at (t-1)	-0.0093***	-0.0092***	-0.0057***	-0.0103***	-0.0087***	-0.0037***	-0.0092***
	(-14.84)	(-14.65)	(-10.59)	(-11.04)	(-7.28)	(-7.31)	(-6.93)
Average Share Turnover, as % of Shares	-0.0058*	-0.0388***	-0.0232***	-0.0586***	-0.0355***	-0.0500***	-0.0324***
Outstanding, at (t-1)	(-1.93)	(-12.82)	(-7.46)	(-10.42)	(-5.59)	(-4.66)	(-5.10)
Maximum Rolling R-Squared with	0.0137***	0.0142***	0.0047***	0.0145***	0.0126***	0.0040*	0.0109***
Available Futures Contracts at (t-1)	(7.24)	(7.48)	(3.07)	(5.26)	(3.60)	(1.97)	(3.08)
Available Options Dummy at (t-1)	0.0025***	0.0024***	0.0005	0.0053***	0.0024**	-0.0002	0.0028**
	(3.63)	(3.41)	(1.23)	(4.54)	(2.46)	(-0.43)	(2.34)
Mispricing at (t-4): % difference between	0.1057***	0.0967***	0.0656***	0.0886***	0.1690***	0.1498***	0.1560***
ETF price and NAV	(8.00)	(7.43)	(5.66)	(4.92)	(8.65)	(6.68)	(6.89)
3-Day Average Retail Volume, as % of	0.4442***	0.2148***	0.1356***	0.1861***	0.3007***	0.3067***	0.2826***
Shares Outstsanding, at (t-1)	(26.07)	(12.50)	(6.44)	(8.20)	(9.79)	(7.67)	(9.20)
3-Day Average Daily Volume, as % of		0.1123***	0.0689***	0.1366***	0.1274***	0.1481***	0.1290***
Shares Outstsanding, at (t-1)		(18.52)	(11.61)	(11.24)	(10.38)	(5.07)	(12.03)
Constant	0.0460***	0.0444***	0.0378***	0.0530***	0.0410***	0.0217***	0.0440***
	(18.47)	(17.73)	(12.94)	(13.28)	(9.31)	(9.52)	(8.79)
Fixed Effects	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date
Observations	2,111,301	2,111,301	1,212,372	963,475	849,761	264,006	576,987
R-squared	0.220	0.232	0.138	0.279	0.226	0.141	0.247
Robustness - ETF Sample	Baseline Model	Control for Daily Volume between (t-3) and (t-1)	Above Median Size	Non-Equity ETFs	US-Equity ETFs	Equity ETFs + Low Liquidity Mismatch	Equity ETFs + High Liquidity Mismatch

Table 4 – Operational Shorting and Contemporaneous/Future Returns: The dependent variables in these regressions are the daily contemporaneous (t) or forward-looking (t+1) total returns (Ret) or Fama-French 4-factor risk-adjusted excess returns (FF4 α) measured in percentage terms (x100). This measure is based on the ETF or NAV price as indicated in the header. Independent variables are measured at day *t* and include the *Operational Shorting*, *Create Orders* as a percentage of ETF shares outstanding, *log(1+Market Cap)* of the ETF where market capitalization is measured in millions of dollars, the ETF *Average Share Turnover %* (as a percentage of ETF shares outstanding), and the ETF *Amihud Illiquidity*. We provide a complete list of variable names, sources, and definitions in Appendix D. In Panel B, specifications 1-3 include all ETFs, 4 through 8 are for U.S.-equity ETFs, and 9-10 are for non-U.S.-equity ETFs which include foreign equity and fixed income ETFs. Specifications 7 and 8 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, and 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 Daily ETF Return Sample

Variable	N	Mean	Std	Min	Median	Max
Total Return (%)	2,598,445	0.003	1.774	-12.939	0.038	11.845
Fama-French 4-Factor Excess Return (%)	2,598,445	-0.018	0.938	-5.769	-0.001	5.350
NAV Return (%)	2,598,445	-0.005	1.760	-13.602	0.022	11.894
Fama-French 4-Factor NAV Excess Return (%)	2,598,445	-0.024	1.039	-8.898	-0.001	7.776
Operational Shorting, scaled by Shares Outstanding	2,598,445	0.011	0.034	0.000	0.000	0.315
Create Orders, scaled by Shares Outstanding	2,598,445	0.003	0.012	0.000	0.000	0.099
log (1+Market Cap)	2,596,416	4.817	2.063	0.016	4.719	12.297
Daily Share Turnover, 60-day average	2,598,445	0.041	0.118	0.000	0.008	0.548
Amihud Illiquidity Measure, 60-day average	2,598,445	0.107	0.358	0.000	0.005	2.594

Panel B: Daily Return Results for All ETFs, U.S. Equity ETFs, and All Other Non-U.S.-Equity ETFs

	Daily Return									
	ETF Ret (t)	ETF Ret (t+1)	NAV Ret (t+1)	ETF FF4 α (t)	ETF FF4 α (t+1)	NAV FF4 α (t+1)	ETF FF4 α (t+1)	ETF FF4 α (t+1)	ETF Ret (t+1)	NAV Ret (t+1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Operational Shorting (t-1,t-3)	1.47206***	-0.51660***	-0.12175	0.44473***	-0.30335***	-0.01428	-0.18569***	-0.42241***	-0.38283***	0.09809
	(7.56)	(-4.77)	(-1.14)	(5.57)	(-4.87)	(-0.30)	(-2.72)	(-4.89)	(-5.39)	(1.54)
Create Orders (t)	-0.08002	-0.23684***	-0.10848	0.05043	-0.09521**	0.00157	-0.09782	-0.10103*	-0.15877***	0.02334
	(-0.96)	(-3.14)	(-1.38)	(1.20)	(-2.19)	(0.04)	(-1.63)	(-1.77)	(-2.80)	(0.42)
log (1+ Market Cap (t-1))	-0.00250	-0.01683***	-0.01334***	-0.00196	-0.00745***	-0.00585***	-0.00704***	-0.00891***	-0.01651***	-0.01096***
	(-0.55)	(-3.77)	(-2.81)	(-1.02)	(-3.94)	(-3.01)	(-2.95)	(-3.80)	(-5.22)	(-4.04)
Average Share Turnover (t-1)	-0.18419**	-0.14301*	-0.16163**	0.00548	0.03133	0.00693	0.09750**	-0.04004	-0.03428	-0.02821
	(-2.43)	(-1.91)	(-2.25)	(0.11)	(0.64)	(0.15)	(2.34)	(-0.60)	(-1.06)	(-0.90)
Amihud Illiquidity Measure (t-1)	0.01888***	0.01581**	0.01097*	0.00078	0.00251	0.01855***	-0.00940	0.00421	0.01793***	0.01274***
	(2.60)	(2.25)	(1.75)	(0.09)	(0.32)	(3.97)	(-0.40)	(0.65)	(3.35)	(2.73)
Observations	2,520,915	2,519,097	2,519,097	1,019,422	1,018,772	1,018,772	346,277	660,447	1,149,617	1,149,617
R-squared	0.244	0.245	0.209	0.064	0.063	0.034	0.060	0.078	0.366	0.304
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		US-Equity		US-Equity	US-Equity	Non US-Equity	
Liquidity Mismatch (ETF vs Underlying)							Low	High		

Table 5 – ETF Mispricing and Arbitrage Activity: This table displays regression results where the dependent variables are *Mispricing Change* and *Absolute Mispricing Change*, displayed on a daily basis in percentage terms. Independent variables include lagged dependent variables, contemporaneous and lagged *Operational Shorting;* lagged *log(Market Cap)*; 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 D. 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 (t), x100				Absolute Mispricing Change (t), x100			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Operational Shorting, as % of Shares	-0.31529***	-0.30753***			-0.16961***	-0.11762***		
Outstanding at (t)	(-14.57)	(-14.33)			(-6.78)	(-6.63)		
Operational Shorting, as % of Shares			-0.03312**	-0.18175***			-0.15005***	-0.09320***
Outstanding at (t-1)			(-2.41)	(-10.95)			(-6.11)	(-5.32)
log (Market Cap), at (t-1)	-0.00232*	-0.00259***	-0.00075	-0.00197**	-0.02438***	-0.01644***	-0.02426***	-0.01628***
	(-1.95)	(-2.62)	(-0.77)	(-2.28)	(-9.76)	(-9.66)	(-9.70)	(-9.56)
Average Share Turnover, as % of Shares	0.00653	0.00694	-0.00093	0.00261	-0.01738	-0.01158	-0.01786	-0.01223
Outstanding, at (t-1)	(0.92)	(1.00)	(-0.13)	(0.41)	(-0.85)	(-0.81)	(-0.87)	(-0.86)
Maximum Rolling R-Squared with		0.00630	0.00399	0.00691	-0.14986***	-0.09686***	-0.15001***	-0.09706***
Available Futures Contracts at (t-1)		(0.29)	(0.18)	(0.37)	(-7.07)	(-6.40)	(-7.08)	(-6.41)
Available Options Dummy at (t-1)		0.00055	0.00010	0.00060	-0.00940*	-0.00639*	-0.00943*	-0.00643*
		(0.26)	(0.05)	(0.32)	(-1.76)	(-1.75)	(-1.77)	(-1.76)
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 daily average of the *Intraday NBBO Spread* of underlying stocks held by U.S. equity-only ETFs. The dependent variable in Panel B is the daily average of the *Intraday Return Volatility* of underlying stocks. In Panel C, the dependent variable is *Intraday Variance Ratios* of underlying stocks. The key independent variable is *Operational Shorting*, measured at (t) or at (t-1). Additional independent variables include the lagged *Average ETF Ownership* in underlying stocks; *log(Market Cap)*; *Average Share Turnover* normalized by shares outstanding; lagged dependent variables as well as the contemporaneous and lagged ETF levels of the dependent variable (e.g. *Intraday NBBO Spread* of the ETF; *Intraday Return Volatility* of the ETF; and *Intraday Variance Ratio* of the ETF). We provide a complete list of variable names, sources, and definitions in Appendix D. 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), x100						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average ETF Ownership in Underlying Stocks in ETF Basket (t-1)	0.21621***	0.06654***			0.21603***	0.06649***	0.06653***
	(3.14)	(3.03)			(3.14)	(3.03)	(3.03)
Operational Shorting, as % of Shares Outstanding at (t-1)			-0.02583**	-0.00799*	-0.02345**	-0.00733**	
			(-2.11)	(-1.89)	(-2.40)	(-2.03)	
Operational Shorting, as % of Shares Outstanding at (t)							-0.00977***
							(-2.67)
log (ETF Market Cap), at (t-1)	0.49317**	0.23958***	0.46383**	0.22945**	0.49054**	0.23877***	0.23848***
	(2.41)	(2.60)	(2.24)	(2.49)	(2.40)	(2.60)	(2.59)
Average Share Turnover, as % of Shares Outstanding, at (t-1)	0.00417**	0.00089	0.00388**	0.00079	0.00414**	0.00089	0.00088
	(2.20)	(1.08)	(2.02)	(0.95)	(2.19)	(1.07)	(1.06)
Lagged Dependent Variable	No	Yes	No	Yes	No	Yes	Yes
Controls for Intraday ETF NBBO Spread	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)
Observations	837,347	837,333	837,906	837,333	837,347	837,333	837,333
R-squared	0.755	0.869	0.752	0.869	0.755	0.869	0.869

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), x100						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Average ETF Ownership in Underlying Stocks in ETF Basket (t-1)	0.01823***	0.00665***			0.01822***	0.00665***	0.00665***
	(2.79)	(2.78)			(2.78)	(2.77)	(2.78)
Operational Shorting, as % of Shares Outstanding at (t-1)			-0.00275**	-0.00104**	-0.00260***	-0.00100***	
			(-2.56)	(-2.51)	(-2.71)	(-2.73)	
Operational Shorting, as % of Shares Outstanding at (t)							-0.00097***
							(-2.71)
log (ETF Market Cap), at (t-1)	11.40274***	6.75895***	11.37129***	6.70402***	11.42790***	6.76923***	6.76956***
	(12.14)	(12.21)	(11.98)	(12.03)	(12.15)	(12.21)	(12.21)
Average Share Turnover, as % of Shares Outstanding, at (t-1)	0.06845***	-0.00042	0.06820***	-0.00116	0.06872***	-0.00031	-0.00026
	(10.17)	(-0.15)	(9.99)	(-0.39)	(10.18)	(-0.11)	(-0.09)
Lagged Dependent Variable	No	Yes	No	Yes	No	Yes	Yes
Controls for Intraday ETF Volatility	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)
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

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)
Average ETF Ownership in Underlying Stocks in ETF Basket (t-1)	0.06615**	0.03782**	0.06597**	0.03772**	0.06574**	0.03770**	0.03777**
	(2.09)	(2.08)	(2.08)	(2.07)	(2.07)	(2.06)	(2.07)
Operational Shorting, as % of Shares Outstanding at (t-1)			-0.02433***	-0.01333**	-0.02614***	-0.01461***	
			(-2.68)	(-2.43)	(-2.88)	(-2.67)	
Operational Shorting, as % of Shares Outstanding at (t)							-0.01509***
							(-2.61)
log (ETF Market Cap), at (t-1)	-0.00027	-0.00015	-0.00034	-0.00019	-0.00038	-0.00021	-0.00021
	(-0.56)	(-0.54)	(-0.69)	(-0.66)	(-0.78)	(-0.76)	(-0.76)
Average Share Turnover, as % of Shares Outstanding, at (t-1)	-0.00944	-0.00532	-0.00914	-0.00516	-0.00951	-0.00537	-0.00536
	(-1.41)	(-1.37)	(-1.37)	(-1.33)	(-1.42)	(-1.38)	(-1.38)
Lagged Dependent Variable	No	Yes	No	Yes	No	Yes	Yes
Controls for Intraday ETF Variance Ratio	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-3)	(t) - (t-5)	(t) - (t-5)	(t) - (t-5)
Observations	720,069	720,056	720,069	720,056	713,693	713,680	713,680
R-squared	0.799	0.836	0.799	0.836	0.799	0.836	0.836

Figure 1 – Operational Shorting and Failure-to-Deliver (FTD) Activity of ETFs: This figure displays the rolling-average daily dollar value of Operational Shorting activity and FTDs for ETFs from March 22, 2004 – December 31, 2016.

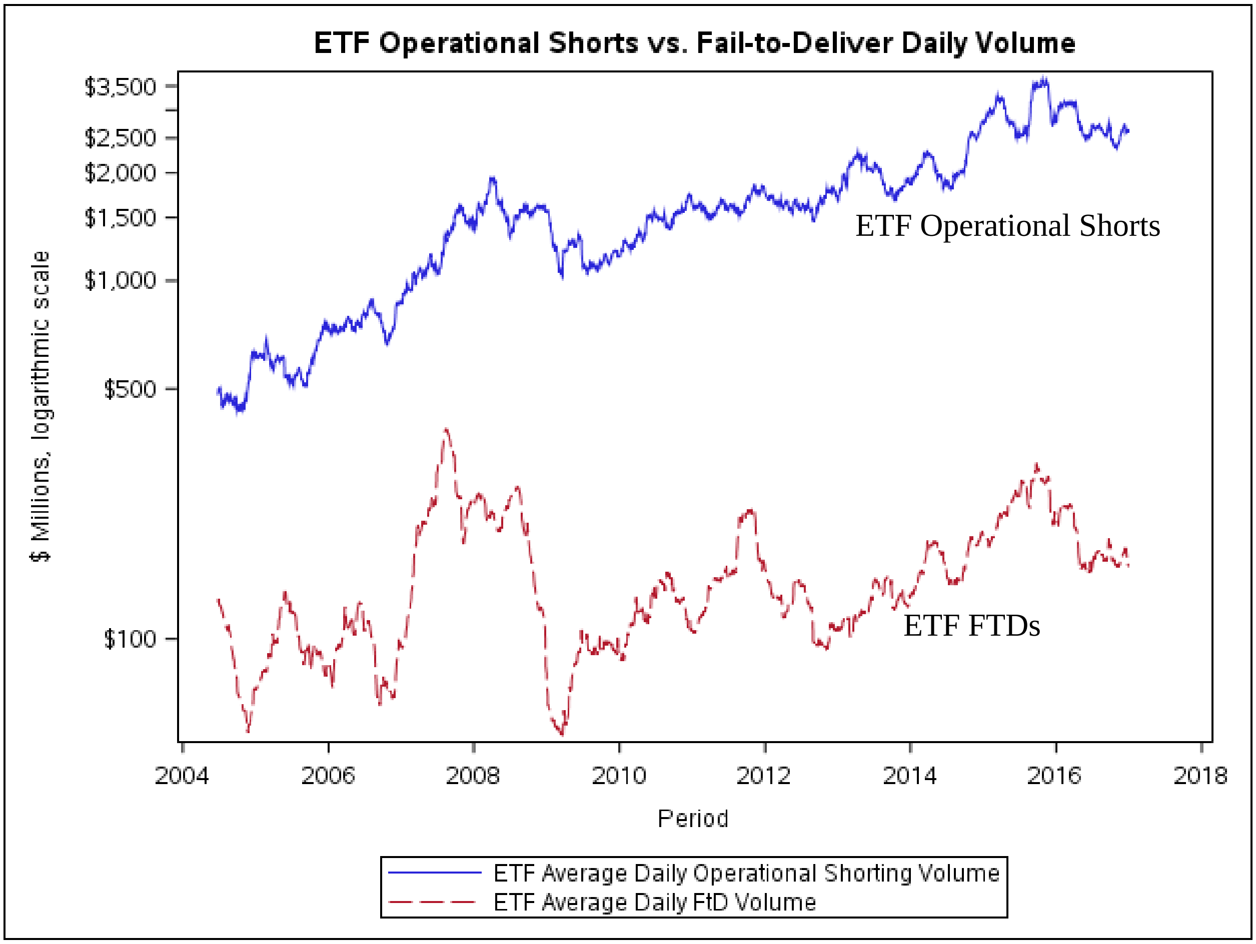


Figure 2 –Aggregate Short Interest of ETFs and Common Stocks: This figure displays the biweekly dollar value of short interest in ETFs and common stocks, as well as the fraction of ETF shorting in the period between January 2004 and June 2020.

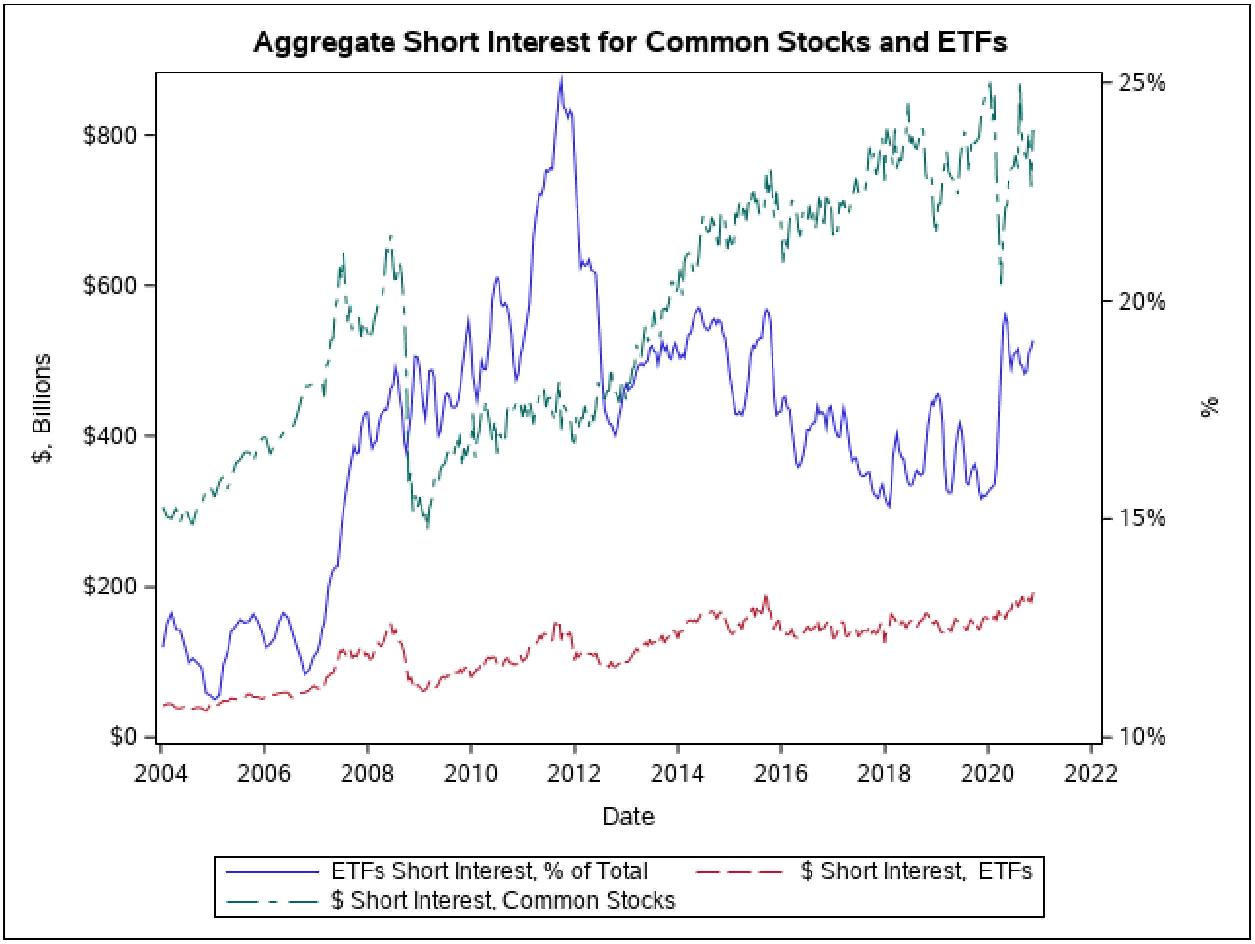
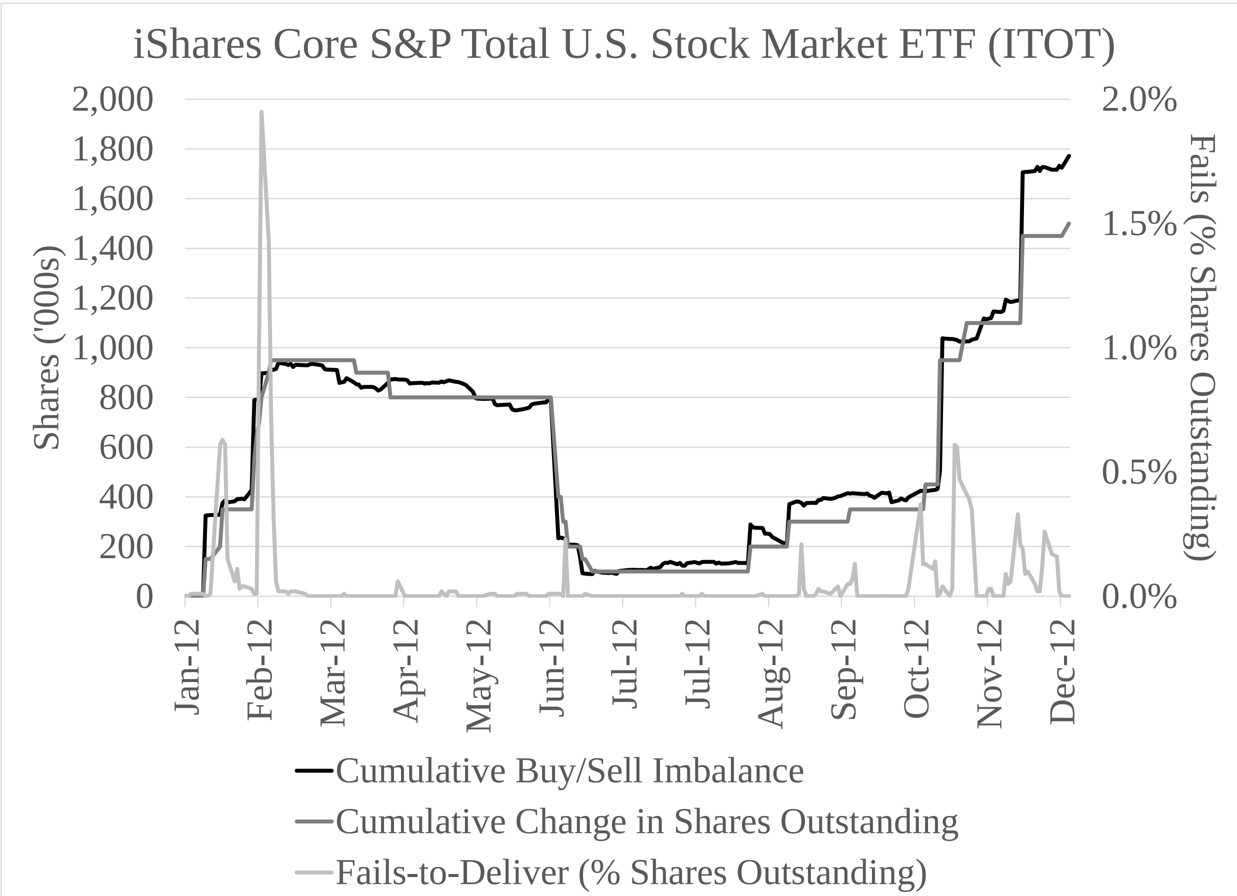


Figure 3 – An Example: ITOT – iShares Core S&P Total U.S. Stock Market ETF: This figure displays the cumulative buy-sell imbalance and the cumulative change in shares outstanding (in 1,000s of shares indexed by the left vertical axis) for the iShares Core S&P Total U.S. Stock Market ETF (ticker: ITOT) over the full year of 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).



Internet Appendix

Operational Shorting and ETF Liquidity Provision

<u>Appendix A:</u> Additional Discussion of Failure-to-Deliver 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.[45] Figure A.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 A.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

[45] 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.

valuation. They show that stocks with high levels of failures are more likely to be overvalued but this apparent trading opportunity is difficult to arbitrage due to the high costs of short selling in these relatively illiquid securities. Thus, less-liquid stocks can remain over-valued even in the presence of high levels of FTDs. In contrast, Jain and Jain (2015) report not only a significant decline in the level of equity FTDs but also a weakening in the relation between short selling activity and FTDs after the implementation of SEC Rules 203 and 204 in 2008-2009.

Additionally, Boni (2006) shows that FTDs were pervasive and persistent in U.S. equities during three settlement dates: September 2003, November 2003, and January 2004. This finding is consistent with market makers' incentive to "strategically fail" when borrowing costs are high. Boni's result suggests that one market participant's FTDs can spill over to other parts of the market and cause increased stress on the broader market. Using detailed data from a large options market maker, Evans et al. (2009) finds similar strategic failure behavior in U.S. equity options markets during 1998-1999. The authors observe that the use of FTDs is due to the relatively low cost of failing. They compute an FTD's cost as "the cost of a zero-rebate equity loan plus the expected incidence of buy-in costs" and find that it amounts to only 0.1 basis points in their sample.[46] 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 A.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 A.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

[46] "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).

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 A.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

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 B, 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:

$$Net\ Creation\ Activity_t \ \ or\ ETF\ Order\ Imbalance_t = \alpha_0 + \ \alpha_1 Controls + \sum_{n=0}^{8} \beta_n ETF\ Order\ Imbalance_{t-n} + \sum_{n=0}^{8} \gamma_n Net\ Creation\ Activity_{t-n} + \ \epsilon_t \qquad \text{(A1)}$$

Equation (A1) 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 (A1) are contained in Table A.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.[47] *Net Create/Redeem Activity* is constructed on a daily basis as the percentage change in the overall ETF shares outstanding. Since this variable is a percentage change that, like a stock's return, is bounded below at -100%, we construct our flows variable, *Net Create/Redeem Activity*, as the *log*(1 + % change in shares outstanding). This variable is likely to be more symmetrical for AP creation as well as redemption activities. After controlling for the two ETF liquidity variables, regressions (1) – (3) show that net creation activity is highly persistent with all of the net creation and order imbalance variables yielding positive and significant parameters at the 1% level. Thus, the prior sequence of net creation activity and order 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 A.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

[47] 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.

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 A.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 A.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 A.3, Panels A and B, we can see that the lagged value of *ETF Order Imbalance* at t-3 has the largest and most significant 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 surprising 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 A.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 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 A.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 A.3 shows that the time period between t-3 and t-1 is the most important in terms of economic and statistical significance. Consequently, we have formulated our definition of *Operational Shorting* 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 A.1 – Failures-to-Deliver (FTDs) Summary Statistics: This table presents summary statistics 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 A.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 D. 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 A.3 – The Effects of Net Creation Activity and Order Imbalances on Failures-to-Deliver and Short Interest: The dependent variable in Panel A is *Failure-to-Deliver*, and the dependent variable in Panel B is *Short Interest* (both measures are scaled by Total ETF Shares outstanding). Short interest data are biweekly and we carry forward short interest to run the regressions on daily order imbalances and create/redeem activity. 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 D. 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 A.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.

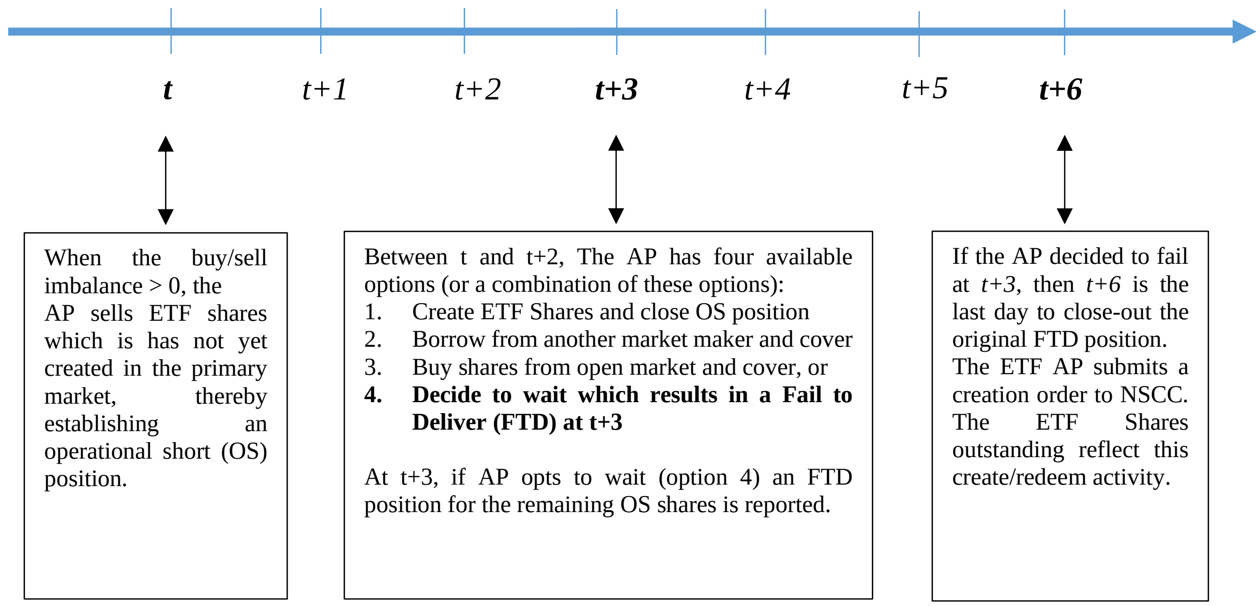
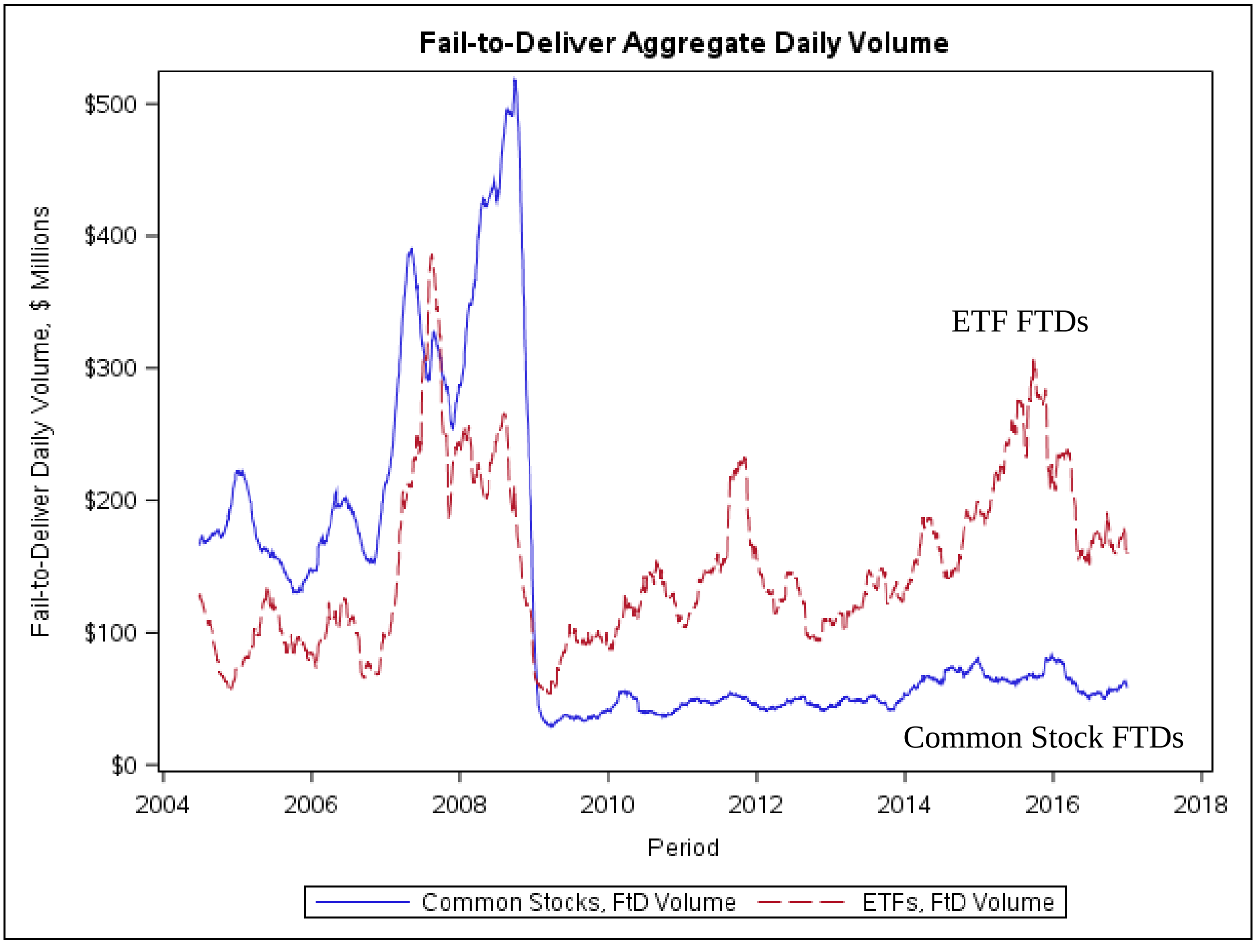


Figure A.2 – Failure-to-Deliver (FTD) Activity of ETFs and Common Stocks: This figure displays the average dollar volume of ETF and common stock FTDs on a daily basis from March 22, 2004 – December 31, 2016. We include only the rolling average daily FTD volume of stocks and ETFs in this graph, as they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



<u>Appendix B:</u> Numerical Example of the Value of Waiting to Deliver ETF Shares

As an example of the APs incentives for operational shorting, 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. In this example, however, 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, we provide a more explicit example that illustrates 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 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 place a creation unit order for execution of 100 shares with the ETF plan sponsor at the end of the trading day 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).[48] 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: $\pi' = \{(P_0 - NAV_0) - (f + \lambda)\}OIB_0$ *(B1)*

Short and Hedge, then Create's profit: $\pi = \{(P_0 - F_0) - (f + \lambda)(1 - \gamma) - c\}OIB_0$ (B2)

where,

f = the creation unit fee (for simplicity, it is expressed here as a dollar amount per ETF share but could be adapted to represent a fixed dollar amount),

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. 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 outperforming 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 γ = 0, then the AP will have to incur the market impact and

[48] 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).

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 γ = 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).[49]

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

$$\pi - \pi' = \{(NAV_0 - F_0) + (f + \lambda)\gamma - c\} \cdot OIB_0 = (\{NAV_0 - F_0 - c\} \cdot OIB_0) + (f + \lambda) \cdot OIB_0 \cdot \gamma \qquad \text{(B3)}$$

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 (B3) presented above, we create a numerical example by assuming specific

[49] 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.

values for the model's parameters and then varying the level of γ between 0 and 1.00.[50] Figure B.1 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 ($f + \lambda$) for that portion (γ) of the initial short position (OIB_0).

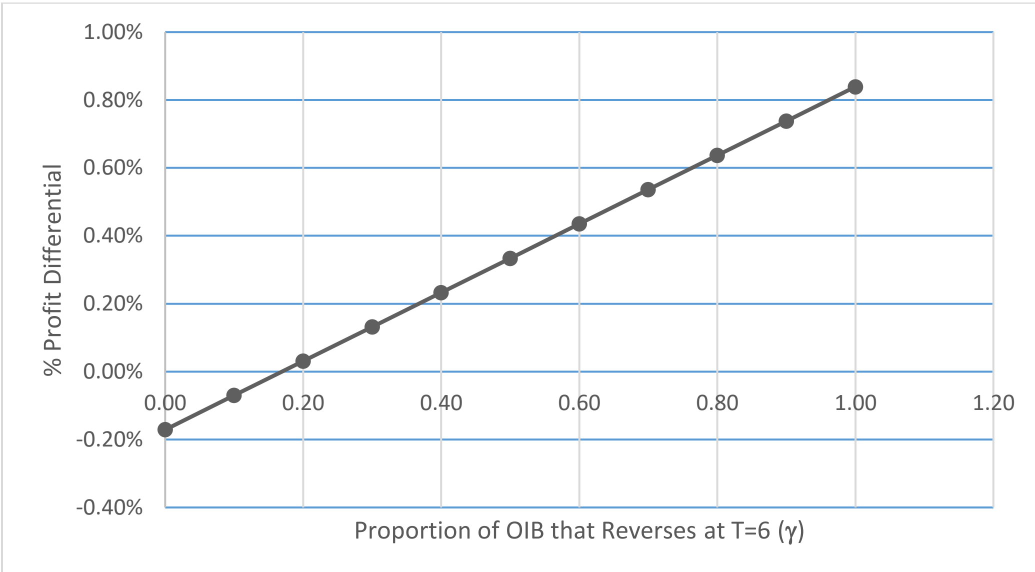
Figure B.1 shows there is a clear trade-off between the two trading strategies and that the predictability of reversals in order imbalances can dictate which approach is most profitable for a specific ETF within a particular set of market conditions. Since we observe in our empirical results a large 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 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.

[50] 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 B.1 illustrates, the Value of Waiting varies greatly from -0.17% to +0.84% over the interval of γ = 0.00 to 1.00.

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 APs 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 B.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 B, 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.



<u>Appendix C:</u> Robustness Checks using Weekly Operational Shorting Measure

While our main measure of operational shorting is anchored along the three-day settlement window, and is computed on a rolling basis, we present robustness results here using an alternative construction of operational shorting over non-overlapping, discrete weekly intervals. Following similar logic as the daily operational shorting, weekly operational shorting is constructed as the cumulative ETF buy-sell demand for all days during the week (ie. the sum of order imbalance from Monday to Friday in week (t)) in excess of the total actual ETF share creation over the week (change in shares outstanding from Friday in week (t-1) to Friday in week(t)) scaled by ETF shares outstanding during the prior week (i.e. on Friday in week (t-1)). Create Orders - Weekly is the change in ETF shares outstanding 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 C.1 presents the robustness checks for Table 3 which explores the determinants of operational shorting activity to confirm the ETF market making incentives but uses the weekly ETF sample with the weekly operational shorting measured over non-overlapping, discrete weekly intervals. Table C.1 includes Panel A and Panel B results mirroring both panels in Table 3. We use the drivers of operational shorting as well as the same controls as in Table 3, including proxies for mispricing, retail volume, the existence of futures or options hedge, liquidity and volume mismatches, and other operational determinants such as historical Creation Unit Dollar Size, and the Creation Unit Fee (per share). The results are very consistent and corroborate our evidence presented using the daily sample in Table 3.

Table C.2 represents an additional robustness test for Table 4 by providing the weekly return results along with a weekly operational shorting measure. For each ETF, we first compute the weekly risk-adjusted excess returns using the Fama-French four factor model. We then regress the cumulative weekly excess return on several 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 outstanding 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.

Panel A of Table C.2 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 7.5 bps and -6.2 bps during the sample period, with a maximum return of 9.85% and 5.88% respectively. Operational Shorting - Weekly is comparable in magnitude to Create Orders - Weekly (with averages of 1.4% and 1.5%, respectively, and standard deviations of 4.3% and 4.9%).

Panel B of Table C.2 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 the total return (*Ret*) or the 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. 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), we 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 C.2, we find that ETF *Operational Shorting* is related to higher contemporaneous ETF returns which is consistent with the evidence from Table C.1 and Table 3 related to arbitrage profits as a primary motivation for 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 next day 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 1.86% contemporaneous change in Fama-French 4-Factor Excess Return during the same week. This corresponds to an increase that is equivalent to 1.014 times the standard deviation of excess returns, 18.4% of which reverses in the following week (or about a -0.324% total reversal in weekly returns, on average) for all ETFs. Reversals are 33.7% stronger for non-U.S. equity ETFs, and 18.8% 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 liquidity 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. Regardless of 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 prior week has no predictive power for returns on the underlying securities, as measured by the NAV.

To better understand what drives short-term ETF returns, we turn to the subsample regression analysis in specifications (9) and (10). These two specifications sub-divide the U.S. equity ETF sample into 'Low' and 'High' liquidity mismatches. 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, our main results from Tables 3 and 4 are robust to an alternative definition of operational shorting which accounts for non-overlapping time intervals.

Table C.1 – Determinants of Operational Shorting; Robustness Tests using Weekly Sample: This table displays regression results using ETF weekly observations where the dependent variable in both panels is the *Operational Shorting*, constructed as the cumulative excess demand (ETF buys – ETF sells) for ETF shares that exceeds the ETF share creation activity during the week, scaled by lagged ETF shares outstanding. In Panel A, the main independent variables are: the existence of a futures hedge – *Maximum Rolling R-Squared with Available Futures Contracts* and *Available Options Dummy*, the existence of retail trading activity – *Average Daily Retail Volume during the week as % of shares outstanding* , the existence of arbitrage opportunities – *Mispricing* defined as the percent difference between ETF price and NAV, and the existence of a liquidity or volume mismatches between the ETF and underlying securities. The *Proxy for Liquidity Mismatch* is defined as the difference between intraday spreads of the ETF's underlying securities and the ETF. The *Proxy for Volume Mismatch* is defined as the log of the ratio between ETF 30-day average dollar volume and the average 30-day dollar volume for the underlying securities. We control for operational determinants such as historical *Creation Unit Dollar Size*, and *Creation Unit Fee (per share)*. All independent variables are lagged and *Mispricing* is constructed as the average daily mispricing during the week. All variables constructed using ETF portfolios are limited to U.S. Equity ETFs. We include other controls such as the lagged *log(Market Cap)*; *60-day Average Share Turnover*. We provide a complete list of variable names, sources, and definitions in Appendix D. 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. The first specifications include ETF and date fixed effects, and the last specifications include Style and date fixed effects. Panel B provides robustness checks using various subsamples of ETFs, where we additionally control for the *average daily volume during the week as % of shares outstanding*. Standard errors are clustered at the ETF and date levels. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A – Determinants of Operational Shorting using Weekly Sample

	Operational Shorting, as % of Shares Oustanding, at week (t)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log (Market Cap), at (t-1)	-0.0083***	-0.0087***	-0.0092***	-0.0113***	-0.0092***	-0.0131***	-0.0091***	-0.0122***
	(-11.80)	(-12.04)	(-9.71)	(-5.53)	(-12.64)	(-5.61)	(-8.69)	(-5.52)
Average Share Turnover, as % of Shares	0.1662***	0.1647***	0.1418***	0.1043***	0.1169***	0.0771***	0.1116***	0.0403
Outstanding, at (t-15)	(16.74)	(16.87)	(12.70)	(4.56)	(12.36)	(2.83)	(10.89)	(1.60)
Maximum Rolling R-Squared with Available		0.0252***	0.0213***	0.0340***	0.0300***	0.0295***	0.0304***	0.0700***
Futures Contracts at (t-1)		(9.50)	(6.42)	(4.07)	(10.62)	(3.83)	(8.67)	(3.92)
Available Options Dummy at (t-1)		0.0033	0.0057	-0.0007	0.0018	0.0019	0.0014	-0.0061
		(1.36)	(1.47)	(-0.22)	(0.67)	(0.41)	(0.90)	(-1.29)
Average Daily Mispricing (t): % difference		0.4409***	0.3545***	0.5449***	0.4158***	0.8254***	0.4091***	1.3845***
between ETF price and NAV		(10.84)	(6.11)	(4.57)	(9.92)	(4.29)	(8.82)	(4.02)
Creation Unit Dollar Size, log, at (t-1)			0.0077***			0.0028		0.0130***
			(5.87)			(1.08)		(4.26)
Creation Unit Fee, per share, at (t-1)			0.0357			-0.0318		-0.1342**
			(1.41)			(-0.52)		(-1.98)
Liquidity Mismatch, at (t-1): Average Intraday				0.6365***		0.5497**		3.6598***
Basket Spread - Intraday ETF Spread				(3.02)		(2.29)		(3.99)
Volume Mismatch, at (t-1): Log(30-Day ETF $				0.0192**		0.0198*		0.0490***
Volume / Implied 30-Day Basket $ Volume)				(2.33)		(1.85)		(3.88)
Average Daily Retail Volume, as % of Shares					0.7564***	0.7812***	0.7956***	0.2386
Outstsanding, at (t-1)					(13.81)	(6.88)	(12.94)	(1.34)
Constant	0.0480***	0.0329***	-0.0225**	0.0510***	0.0332***	0.0422**	0.0327***	-0.0663***
	(14.93)	(10.08)	(-2.41)	(5.65)	(9.52)	(2.25)	(10.01)	(-3.26)
Fixed Effects	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date	Style, Date	Style, Date
Observations	541,316	541,316	364,010	131,419	510,206	93,337	501,156	93,108
R-squared	0.189	0.191	0.228	0.257	0.203	0.317	0.208	0.158

Panel B – Robustness using Various ETF Subsamples

	Operational Shorting, as % of Shares Oustanding, at week (t)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log (Market Cap), at (t-1)	-0.0092***	-0.0090***	-0.0027***	-0.0097***	-0.0088***	-0.0027***	-0.0086***
	(-12.64)	(-12.13)	(-3.78)	(-8.85)	(-5.98)	(-5.16)	(-6.12)
Average Share Turnover, as % of Shares Outstanding, at (t-1)	0.1169***	0.0142	0.0311***	-0.0477**	0.0481*	0.0441*	0.0597***
	(12.36)	(1.23)	(2.93)	(-2.11)	(1.80)	(1.67)	(2.73)
Maximum Rolling R-Squared with Available Futures Contracts at (t-1)	0.0300***	0.0290***	0.0048*	0.0297***	0.0291***	0.0026	0.0259***
	(10.62)	(10.17)	(1.81)	(8.16)	(4.61)	(0.64)	(4.66)
Available Options Dummy at (t-1)	0.0018	0.0023	0.0009	0.0038	0.0004	-0.0011	0.0012
	(0.67)	(0.84)	(0.76)	(0.74)	(0.25)	(-1.13)	(0.55)
Average Daily Mispricing (t): % difference between ETF price and NAV	0.4158***	0.3976***	0.1081**	0.3225***	0.6280***	0.5435***	0.6199***
	(9.92)	(9.56)	(2.31)	(6.40)	(8.02)	(4.71)	(7.37)
Average Daily Retail Volume, as % of Shares Outstsanding, at (t)	0.7564***	0.2017***	0.1511*	0.0303	0.5822***	0.4529***	0.5161***
	(13.81)	(3.23)	(1.96)	(0.40)	(5.09)	(2.78)	(4.70)
Average Daily Volume, as % of Shares Outstsanding, at (t)		0.2771***	0.0751***	0.4128***	0.2055***	0.1911***	0.2143***
		(14.71)	(3.63)	(13.18)	(6.31)	(3.54)	(6.31)
Constant	0.0332***	0.0304***	0.0204***	0.0374***	0.0293***	0.0182***	0.0276***
	(9.52)	(8.58)	(4.28)	(6.32)	(5.94)	(6.05)	(5.34)
Fixed Effects	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date	ETF, Date
Observations	510,206	510,206	246,825	237,063	201,181	55,413	143,826
R-squared	0.203	0.209	0.105	0.253	0.205	0.100	0.235
Robustness - ETF Sample	Baseline Model	Control for Daily Volume Average during Week	Above Median Size	Non-Equity ETFs	US-Equity ETFs	Equity ETFs + Low Liquidity Mismatch	Equity ETFs + High Liquidity Mismatch

Table C.2 – 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 (x100). 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 during 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(1+Market Cap)* of the ETF where market capitalization is measured in millions of dollars, the ETF *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 D. In Panel B, specifications 1-3 include all ETFs, 4-6 non-U.S.-equity ETFs, 7 through 10 U.S.-equity ETFs. Specifications 9 and 10 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, and 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	0.075	3.091	-10.268	0.123	9.854
Fama-French 4-Factor Excess Return (%)	559,555	-0.062	1.834	-6.445	0.000	5.878
NAV Return (%)	572,629	0.032	3.044	-10.199	0.080	9.691
Fama-French 4-Factor NAV Excess Return (%)	559,555	-0.093	1.814	-6.610	-0.017	5.735
Operational Shorting - Weekly, scaled by Shares Outstanding	572,629	0.014	0.043	0.000	0.000	0.305
Create Orders - Weekly, scaled by Shares Outstanding	572,629	0.015	0.049	0.000	0.000	0.333
log (1+Market Cap)	572,626	4.592	2.135	0.881	4.504	9.782
Daily Share Turnover, 60-day average	571,583	0.035	0.079	0.000	0.009	0.508
Amihud Illiquidity Measure, 60-day average	563,492	0.152	0.386	0.000	0.010	2.379

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)	1.758***	-0.324***	0.037	2.655***	-0.593***	-0.221	-0.321***	-0.055	-0.044	-0.331***
	(13.88)	(-3.36)	(0.37)	(9.73)	(-3.49)	(-1.33)	(-2.80)	(-0.48)	(-0.15)	(-2.63)
Create Orders - Weekly (t)	-0.027	-0.187**	0.056	0.056	-0.387***	-0.051	-0.117	-0.084	-0.064	-0.110
	(-0.23)	(-2.08)	(0.57)	(0.26)	(-2.90)	(-0.42)	(-1.10)	(-0.86)	(-0.25)	(-0.95)
log (1 + Market Cap (t-1))	-0.019**	-0.035***	-0.033***	-0.030**	-0.068***	-0.059***	-0.031***	-0.031***	-0.043***	-0.033***
	(-2.23)	(-3.77)	(-3.51)	(-2.34)	(-4.89)	(-4.51)	(-3.22)	(-3.40)	(-3.67)	(-2.82)
Average Share Turnover (t-1)	-0.136	-0.075	-0.018	-0.351*	-0.148	-0.124	0.291	0.219	0.303	0.249
	(-1.10)	(-0.56)	(-0.12)	(-1.80)	(-0.70)	(-0.60)	(1.53)	(1.09)	(1.00)	(1.23)
Amihud Illiquidity Measure (t-1)	0.039	0.033	0.037**	0.087**	0.056	0.044	0.029	0.031	-0.022	0.040
	(1.62)	(1.21)	(2.19)	(2.58)	(1.62)	(1.43)	(0.78)	(1.54)	(-0.20)	(1.31)
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	0.077	0.088	0.417	0.427	0.414	0.079	0.049	0.089	0.086
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

<u>Appendix D:</u> 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
Total Demand Quantity	The quantity of ETF shares on loan by Markit borrowers.	Markit Securities Finance Database (formerly Data Xplorers)
Daily Short Volume Ratio	The number of daily short volume for each ETF scaled by the total share volume for that ETF, aggregated across various short volume reporting exchanges	NYSE, ARCA, NASDAQ, BATS, FINRA's TRF and ORF feeds
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 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. Averaged across all stocks in the ETF basket using holding size as weight.	TAQ, 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. Averaged across all stocks in the ETF basket using holding size as weight.	TAQ, Thomson-Reuters Mutual Fund Ownership database, CRSP Mutual Fund Database
Average Intraday Variance Ratio of Underlying Stocks	The intraday variance ratio of stocks in the ETF basket calculated as the absolute value of: the ratio of 15-second return variance divided by 3*(5-second return variance), minus 1. Averaged across all stocks in the ETF basket using holding size as weight.	TAQ, Thomson-Reuters Mutual Fund Ownership database, CRSP Mutual Fund Database

Daily Fama and French 4-Factor Excess Return	Daily excess return computed using the lagged betas from a Fama and French four factor model over the previous 200 days.	CRSP, Fama and French (1993) factors
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, Fama and French (1993) factors
Fail-to-Deliver Shares / Adjusted Net Capital	The number of ETF shares not delivered on time divided by the adjusted net capital.	SEC; CFTC's Futures Commission Merchants Financial Reports
Operational Shorting / Adjusted Net Capital	The buy/sell imbalance for trading the ETF minus the change in share creation for the ETF, divided by the adjusted net capital.	NYSE TAQ Database; CFTC's Futures Commission Merchants Financial Reports
Independent Variables		
Average Share Turnover/Shares Outstanding	The volume of ETF shares traded each day, normalized by total ETF shares outstanding, 60-day average	Bloomberg and CRSP
Amihud Illiquidity Measure	The absolute return divided by average daily volume in $ millions, 60-day average	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
Volume Mismatch	The log of the 30-Day ETF dollar volume divided by weighted average 30-Day dollar volume for all holdings in the ETF portfolio using holding dollar size as weight	CRSP, Thomson-Reuters Mutual Fund Ownership database, CRSP Mutual Fund Database
3-Day Average Retail Volume, as % of Shares Outstanding	3-day average of Retail volume, observed during market hours, and scaled by average shares outstanding over the prior month. Following Boehmer, Jones, Zhang, and Zhang (2021), we identify Retail volume as the sum of all off-exchange trades (Exchange code D in TAQ, referring to the FINRA TRF feed) that received price improvement over the prevailing NBBO.	TAQ, CRSP, Bloomberg
3-Day Average Daily Volume, as % of Shares Outstanding	3-day average of total daily volume observed during market hours, and scaled by average shares outstanding over the prior month.	CRSP, Bloomberg
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
Affiliated Lead Market Maker Fail-to-Deliver	The number of ETF shares not delivered on time as a percentage of all affiliated ETF market cap	SEC
Market-Wide Fail-to-Deliver	The number of ETF shares not delivered on time for all market participants as a percentage of all ETF market cap, net of the Affiliated Lead Market Maker	SEC
Affiliated Lead Market Maker Operational Shorts	The buy/sell imbalance for trading the Affiliated ETFs minus the change in share creation for the Affiliated ETFs, as a percentage of all affiliated ETF market cap.	NYSE TAQ Database
Market-Wide Operational Shorts	The buy/sell imbalance for trading the Affiliated ETFs minus the change in share creation for the Affiliated ETFs, as a percentage of ETF market cap, net of the affiliated Lead Market Maker	NYSE TAQ Database

<u>Appendix E:</u> Trade Signing Algorithm; Discussion and Evaluation Results

In this Appendix, we provide a detailed discussion of our 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). To compute the daily buy-sell imbalances in ETF shares, we need to first sign the ETF trades as buys and sells. 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 revised methodology that acknowledges the clustering of buys on the offer prices and sells on the bid prices.[51] Once an executed trade price crosses the prevailing NBBO within a millisecond, we stop using the quote test for the rest of the millisecond. Instead, and for the rest of the trades during this millisecond, we rely on the tick test, as it is likely that the quote test is not accurate, especially when there is intense high frequency algorithmic trading that is faster than the refresh rate of the quotes within a millisecond period. 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.

[51] 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.

1. Methodology Details

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)).

2. Horse Race and Empirical Results

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 E.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.[52] 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.

[52] 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.

Instead, and for the remaining trades during this millisecond, we rely on the tick test as it is likely that the 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.[53] 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 E.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

[53] 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.

the millisecond timestamp. We conclude that our modified algorithm -- that dynamically switches from the 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 3 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.[54]

[54] 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 E.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 E.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