Short Selling ETFs*

Frank Weikai Li[†] Qifei Zhu[‡]

Abstract

We provide novel evidence that arbitrageurs use exchange-traded funds (ETFs) as an avenue to circumvent short-sale constraints at the stock level. Using a large sample of U.S. equity ETF holdings, we document that shorting activity on ETFs rises with the difficulty of shorting underlying stocks. Stocks that are heavily shorted via their holding ETFs underperform those lightly shorted. The return predictability of ETF shorting is distinct from stock-level shorting measures, and is concentrated among stocks that face severe arbitrage constraints. These findings suggest that ETFs allow arbitrageurs to target overpriced stocks that are otherwise difficult to short.

JEL classification: G12, G14

Keywords: ETFs, Short Selling, Equity Lending, Limits to Arbitrage

*The paper is previously titled "Synethtic Shorting with ETFs". We are grateful to an anonymous referee and Jeffrey Pontiff (the editor) for insightful and constructive feedback. We also thank Yakov Amihud, Nicholas Barberis, Utpal Bhattacharya, Jesse Blocher, Ekkehart Boehmer, Lauren Cohen, Zhi Da, Brian Henderson (discussant), Harrison Hong, Zsuzsa Huszar, Bige Kahraman (discussant), Alexander Ljungqvist (discussant), Roger Loh, Abhiroop Mukherjee, Matthew Ringgenberg, Melissa Prado (discussant), Adam Reed, Matthew Richardson, Pedro Saffi, Rik Sen, Clemens Sialm, Wing Wah Tham, Sheridan Titman, Robert Whitelaw, K.C. John Wei, Kelsey Wei, Andrew (Di) Wu (discussant), Jialin Yu, Jianfeng Yu, Bohui Zhang, Chu Zhang, Xiaoyan Zhang and the seminar participants at NYU, UT Austin, UT Dallas, Penn State, University of Georgia, HKUST, PBCSF, CKGSB, SAIF, Singapore Management University, University of Hong Kong, National University of Singapore, Hong Kong PolyU, the WFA 2018. the FIRS 2017, the EFA 2017, the Fifth Annual USC Marshall Ph.D. Conference and the PhD forum of 29th Australasian Finance and Banking Conference for helpful comments. Ruxue Gong provides excellent research assistance. All errors are ours.

[†]Singapore Management University, Lee Kong Chian School of Business (Email: wkli@smu.edu.sg)

[‡]Nanyang Business School, Nanyang Technological University (Email: qifei.zhu@ntu.edu.sg)

The market for exchange-traded funds (ETFs) has been growing exponentially during the recent decade. Globally, assets under management accelerated from US$675 billion in 2008 to $5.8 trillion in 2020, and over 350 issuers offer more than 8,000 ETFs.[1] Around 10% of the market capitalization and 30% of the trading volume of securities traded on US stock exchanges are attributable to ETFs (Ben-David, Franzoni, and Moussawi (2017)).[2] As an investment vehicle, ETFs provide investors with a cost-efficient way to passively manage their assets (Madhavan (2014)). At the same time, academics and practitioners share their misgivings that the rise of ETFs increases underlying stock volatility, propagates shocks across their constituents, reduces informational efficiency, and may have contributed to the Flash Crash of May 2010.[3]

An important aspect of ETFs that has been largely under-studied is the associated short selling activities. Just as other exchange-traded securities, ETFs can be sold short. It is evident that the short selling activities of ETF products are highly active.[4] Figure 1 shows the aggregate level of ETF short interests compared to the size of the ETF market in our sample. The dollar value of ETF short selling exceeded $80 billion multiple times in our sample period, representing 10–40% of the corresponding ETFs' market capitalization. For comparison, short interests of US equity on average is about 4% of their market capitalization.

What drives ETF short selling activities? Firstly, some investors may use ETF shorting to hedge the risks borne by their long positions ("hedging"). Secondly, some short selling comes with the creation-redemption mechanism undertaken by ETFs' authorized participants, as documented by Evans, Moussawi, Pagano, and Sedunov (2018) ("operational"). Thirdly, some traders may want to bet against the future performance of the ETF or its underlying constituents ("directional betting"). For directional betting, one could bet against a whole

[1]https://www.blackrock.com/au/intermediaries/ishares/authorized-participants-and-market-makers.

[2]Usually, an ETF tracks a particular stock or bond index by physically holding all constituent securities (or a sample of them) as its underlying assets. A small fraction of ETFs track their indices using swaps or other derivatives. The latter group are out of the scope of our research in this paper.

[3]See for example, Ben-David, Franzoni, and Moussawi (2018), Da and Shive (2018), Israeli, Lee, and Sridharan (2017) and Madhavan (2012).

[4]During our sample period, 3.9% of equity ETFs have average short interest ratios above 20% and 10.64% have average short interest ratios above 10% of shares outstanding. For stocks, the corresponding figures are 0.92% and 5.46%, respectively.

market or against a subset of stocks using "synthetic shorting" with ETFs. In the latter case, traders who want to bet against a subset of stocks, but are unable or unwilling to do so directly, short the ETFs instead. Since short selling individual stocks are often constrained (Ljungqvist and Qian (2016)), shorting the ETF and hedging other stock constituents allows traders to obtain similar negative exposures to certain underlying stocks. In this paper, we term such combination trading strategy *synthetic shorting with ETFs* and we hypothesize that such a strategy is a significant driver of ETF short selling activities and has important implications for future performance of underlying securities.

Accounts by market participants corroborate this conjecture: For example, *MarketWatch* reports that "One hedge fund with which Weinhoffer is familiar was struggling to borrow a stock and instead shorted an ETF that contained the shares, ... The manager then took long positions in all the other stocks in the ETF."[5] *Barron's* recently runs a column piece titled exactly as "Synthetic Shorting with ETFs."[6] It points out that hedge funds frequently create synthetic shorting using ETFs, particularly when it is costly or outright impossible to borrow the target stock(s). In addition, some hedge funds prefer using ETFs to short underlying stocks so that their rivals cannot easily detect their trading strategies.

Our empirical analyses suggest that directional betting, especially synthetic shorting, is an important component of ETF shorting activities. First, we document that shorting activity on the ETF rises with the difficulty of shorting underlying constituents. Specifically, ETF short interest is high when the lendable supply of underlying constituents is low and the cost of shorting the underlying stocks is high. When we proxy the level of short-sale constraint of constituent stocks using idiosyncratic volatility (Pontiff (2006)) and Amihud (2002) illiquidity measure, we find that ETF short selling is more active when the underlying stocks are less liquid or more volatile.[7] These relationships are robust to controlling for ETF characteristics such as ETF size, turnover, and flow, and the intensity of operational shorting activities. It suggests that there is a (partial) substitution between stock-level shorting and ETF shorting, especially when an ETF's underlying stocks are difficult to short.

[5] "More equity hedge funds turn to shorting ETFs", *MarketWatch*, June 1, 2007

[6] "Synthetic Shorting with ETFs", Crystal Kim, *Barron's*, Feb 27, 2017

[7] In sharp constrast, short selling on stocks is more active when stock is more liquid and less volatile. See Table 2 of Hong, Li, Ni, Scheinkman, and Yan (2015) for evidence.

To the extent that short sellers use ETFs to express their bearish opinions on individual stocks, stocks that are heavily shorted via their holding ETFs should expect to earn negative returns in the future. Thus, in the second part of our empirical analyses, we examine the informativeness of ETF-level shorting for future returns. As a first check, we examine the predictive relation at the ETF level. We find that ETF short ratio negatively predict future ETF returns, suggesting that a non-trivial fraction of shorting activity in ETF is driven by directional betting. However, since ETF shorting can also be used for synthetic shorting on individual stocks, one should be able to more reliably glean information from ETF shorting activities by aggregating ETF short interests to the stock level.

One of the key innovations of our paper is that we construct a short interest ratio for each stock from the short interests of all ETFs holding that particular stock. This measure, which we call the *ETF-based short ratio*, reflects the collective shorting demand of that stock through short selling ETFs. The idea is that stocks in the intersection of several highly shorted ETFs are more likely to be the true targets of ETF short bets. We first corroborate our ETF-based short ratio by showing that stocks which are heavily shorted via ETFs have significantly lower hedge fund ownership, suggesting that the measure does reflect the bearish view of sophisticated investors. More importantly, we find that the ETF-based short ratio strongly forecasts stock returns, even after controlling for stock-level shorting measures. An equal-weighted, monthly rebalanced, long-short strategy that sells the decile of stocks that are most heavily shorted via their holding ETFs and buys the decile of stocks that are the most lightly shorted earns 59 basis points per month ($t = 3.08$) after adjusting for the Carhart (1997) four factors. A similar strategy with value weights earns a Carhart (1997) four-factor alpha of 41 bps per month ($t = 2.06$).[8] This is consistent with our hypothesis that ETFs are (at least partially) used by short sellers to bet against overvalued stocks that are otherwise difficult to short.

Brown, Davies, and Ringgenberg (2021) show that ETF flows signal non-fundamental demand shocks and negatively forecast future returns at both the ETF and stock level. To

[8]Adjusting for the recently proposed factor models–the Fama and French (2016) five-factor model, the Hou, Xue, and Zhang (2015) q-factor model, the Stambaugh and Yuan (2016) mispricing factor model or the Daniel, Hirshleifer, and Sun (2020) behavioral factors–does not affect the return spread of the long/short portfolio much.

the extent that ETF flows are correlated with its short interests, the return predictability we document in this paper might be driven by flow-induced price pressure. To address this concern, we first construct a measure of ETF-based flow for each stock by aggregating monthly ETF flows to stock level, and include it in Fama-MacBeth regressions. Consistent with Ben-David, Franzoni, and Moussawi (2018) and Brown, Davies, and Ringgenberg (2021), ETF flows prorated to stock level negatively predict next-month stock return. Importantly, we find the return predictability of ETF-based SR is robust to the inclusion of ETF-based flow. Second, we examine the persistence of the return predictability of ETF-based short ratio. We find ETF-based SR significantly predict future returns up to eight months, and its predictability is not reversed in the subsequent 12 months. The persistence in return predictability alleviates concerns that the short-run predictability is driven by flow-induced price pressure. Instead, it suggests that there are value-relevant information contained in ETF shorting that is slowly incorporated into stock prices. By contrast, ETF-based flow becomes insignificant when forecasting returns beyond the first month, suggesting that the sources of return predictability are different for these two measures.

If short sellers are more likely to switch to ETFs when their target stocks are difficult to short, the return predictability of the ETF-based short ratio should be more pronounced among stocks that have severe shorting constraints. To test this, we double-sort stocks first by proxies of short-sale constraints and then by their ETF-based short ratio. We find the return predictability of the ETF-based short ratio is indeed concentrated within the group of stocks that are subject to greater impediments to arbitrage. For example, the Carhart (1997) four-factor alpha of the long-short portfolio sorted on the ETF-based short ratio is a monthly 0.83% ($t = 2.57$) in the tercile with the lowest institutional ownership, while it is only -0.17% ($t = -1.19$) for stocks with high institutional ownership. The difference in monthly alphas between stocks with low and high institutional ownership is 1.0% ($t = 3.78$). In the same vein, the predictive power of the ETF-based short ratio is more pronounced for stocks that have lower lendable supply, higher borrowing cost, higher idiosyncratic volatility, and lower liquidity. This evidence supports our hypothesis that ETF short selling works as an alleviation mechanism for hard-to-short stocks.

In a Fama–MacBeth regression setting, we confirm that the ETF-based short ratio has incremental return predictability for future stock returns when we control for stock-level shorting variables. Importantly, the predictive power of the ETF-based short ratio is comparable in economic significance to well-known stock-level shorting activity measures such as stock-level short interest ratio or shorting costs. For example, a one-standard deviation increase in ETF-based short ratio is associated with a lower future monthly return of 16 basis points, while a one-standard deviation increase in stock short ratio decreases future return also by around 16 basis points. The return predictability of our ETF-based short ratio cannot be fully explained by betting on sector-level or factor-level trend, as it also predicts industry-adjusted and characteristics-adjusted excess stock return. In addition, ETF-based short ratio also negatively predicts firm-specific news, as measured by earnings announcements return and standardized unexpected earnings, consistent with the notion that ETF-based short ratio contains information about firm fundamentals. By contrast, we do not find any predictability of ETF-based flow for earnings news, which is consistent with our interpretation that ETF flows mainly capture investors' non-fundamental demand shock.

This paper contributes to several strands of the literature. First, it extends a large literature that examines the information content of short interests. Numerous studies document that a stock's short ratio is a strong contrarian predictor of future returns.[9] The common interpretation is that when investors have divergence of opinions and short-sale constraints are binding, the value of the stocks will only reflect the optimists' view, hence they are more likely to be overvalued (Miller (1977)).[10] Short interest is a proxy for the amount of negative information excluded from the market price (Figlewski (1981)). Different from the previous literature, our paper considers the information contained in ETF short interest for future stock returns. As pointed out by Chen, Hong, and Stein (2002), short interest is an insufficient proxy for the degree of overvaluation, as a stock with low short interest could

[9]See for example, Figlewski (1981), Dechow, Hutton, Meulbroek, and Sloan (2001), Desai, Ramesh, Thiagarajan, and Balachandran (2002), Asquith, Pathak, and Ritter (2005), Boehme, Danielsen, and Sorescu (2006), Cohen, Diether, and Malloy (2007), among others. There are also studies documenting a similar negative predictive relation at the aggregate market level (e.g., Rapach, Ringgenberg, and Zhou (2016) and Chen, Da, and Huang (2020)).

[10]A large literature explores the effects of heterogeneous beliefs on equilibrium asset prices when short-selling is constrained. e.g. Harrison and Kreps (1978), Scheinkman and Xiong (2003), Hong, Scheinkman, and Xiong (2006) and Duffie, Garleanu, and Pedersen (2002).

be extremely difficult to short. Consistent with this intuition, we document that the ETF-based short ratio contains incremental information for stock returns even after controlling for the stock's own short ratio, especially among those most constrained by the availability of lendable supply.

Our paper also contributes to the burgeoning literature that examines the impact of ETFs on financial markets and documents mixed evidence. Theoretically, Bhattacharya and O'Hara (2015) show that information feedback between ETFs and underlyings could cause propagation of shocks unrelated to fundamentals and market instability. Empirically, Ben-David, Franzoni, and Moussawi (2018) provide evidence that ETFs increase non-fundamental volatility of their underlying assets. Da and Shive (2018) document that higher ETF trading activity leads to excess return comovement among the constituent stocks.[11] On the bright side, several recent papers identify a positive effect of ETFs on the price efficiency of underlying securities. For example, Huang, O'Hara, and Zhong (2021) show that industry ETFs help improve firm-specific informational efficiency by facilitating industry risk hedging for informed investors. Glosten, Nallareddy, and Zou (2021) document that ETF activities increase systematic informational efficiency for stocks with weak information environments. Bhojraj, Mohanram, and Zhang (2020) find that sector ETFs are effective at transmitting industry information across firms. By highlighting that one benefit of ETFs is to facilitate short selling on overvalued underlying stocks, our paper contributes to the growing literature on the consequences of ETF investing on the stock market.

The rest of this paper is organized as follows. Section 1 details the institutional background of ETF shorting selling and compares it to stock-level short selling. Section 2 describes the various data we used in the analysis and presents summary statistics. Section 3 examines the cross-sectional determinants of ETF short interest and the return predictability of ETF short interest at the ETF level. In Section 4, we examine the informativeness of ETF short selling for stock returns. Section 5 concludes the paper.

[11]In addition, Israeli, Lee, and Sridharan (2017) find that an increase in ETF ownership is accompanied by a decline in pricing efficiency for the underlying component securities. Bhattacharya, Loos, Meyer, and Hackethal (2016) find that individuals investing in passive ETFs do not improve their portfolio performance, due to poor ETF timing and selection. Brown, Davies, and Ringgenberg (2021) show that non-fundamental demand shocks distort ETF prices and imposes a non-trivial costs on investors.

1 Institutional Background on ETF Short Selling

ETF short selling is prevalent despite its lack of attention from media and common investors. Leading practitioner books, such as Gastineau (2010), claim that "'[s]hort selling in the ETF marketplace is a large part of ETF trading volume, and ETF short positions are often so large relative to total ETF shares outstanding". Indeed, the most heavily shorted ETFs often have short ratios that are higher than 100%. In this section, we review some institutional background to ETF short selling practices.

1.1 The "Create-to-Lend" Mechanism

When a trader attempts to short sell ETF shares, there are two routes to take: She can ask her broker to borrow ETF shares directly from institutional investors or brokerage firms with lending programs. Alternatively, the broker can borrow or purchase[12] the underlying securities, turn them to an Authorized Participant (AP), then let the AP create new units of the ETF so that the broker can lend these shares to the short seller. This mechanism uses the creation–redemption feature of ETFs and is dubbed "create-to-lend". In creating new units, APs sometimes only need to deliver a representative sample of all stocks that the ETF holds. This mechanism potentially opens the door for easier access to shorting the ETF as opposed to shorting specific hard-to-borrow assets.

Some empirical evidence suggests that the create-to-lend mechanism is an important avenue for ETF short selling. For example, the total short interest of the S&P500 SPDR ETF is on average greater than its lendable supply, suggesting that some fraction of short selling is borrowed through creation (Karmaziene and Sokolovski (2019)). Since ETF Authorized Participants are often well-connected financial intermediaries, they are usually able to more easily locate underlying securities for borrowing. Asquith and Meulbroek (1995) and Danielsen and Sorescu (2001) cite several reasons why "ordinary" investors might face higher transaction costs in trying to establish short positions than brokers. The differential search costs in lending markets between prime brokers and traders is likely an important

[12]The broker would have to hedge her position by short selling the securities herself.

advantage for ETF short selling.[13]

The create-to-lend mechanism also makes ETF short selling difficult to be squeezed. In order to short squeeze an ETF, one must not only buy the shares of the ETF, but also deplete the lending supply of underlying stocks. Otherwise, short sellers could simply create additional ETF shares to answer the call. As a result, ETF short squeezes are "virtually unknown" (Gastineau (2010)).

1.2 Market Liquidity

ETF securities are usually more liquid than their underlying stocks. They are traded more frequently, have smaller bid-ask spread, and have shorter days-to-cover. Hong, Li, Ni, Scheinkman, and Yan (2015) argue that days-to-cover (DTC), defined as open short interest divided by average daily trading volume, is an important measure for the crowdedness of short sale trades. In a sense, it captures how fast arbitrageurs are able to exit their short positions when necessary. In our sample, the average days-to-cover for ETFs is about 2 days, which is significantly shorter than the average DTC for stocks. This means that short sellers would be able to cover their positions in reasonable speed and at reasonable costs should the market conditions turn against them. The short days-to-cover is an attractive feature for ETF short selling, especially when short trades are getting more crowded in the recent decade (Hanson and Sunderam (2014)).

1.3 Regulation

ETFs are also more lightly regulated in terms of short selling compared to stocks. Unlike stocks, ETFs have never been subject to the "uptick" rule. The uptick rule dictates that a short order must be placed above the last transaction price, or the "uptick". This rule has been shown to impede short selling activities (Alexander and Peterson (2008); Diether, Lee, and Werner (2009)). The fact that ETFs are not subject to the uptick rule allows traders to implement more flexible trading strategies using ETFs to form synthetic short positions

[13]Using the stock lending market data from Brazil, Chague, De-Losso, De Genaro, and Giovannetti (2017) find that well-connected borrowers pay lower lending fee, even for the same stock on the same day.

on the constituent stocks. During the 2008 Financial Crisis, the Securities and Exchange Commission (SEC) temporarily banned short sales in 797 financial stocks, but this ban list did not include any ETFs. Many market participants, as suggested by Karmaziene and Sokolovski (2019), circumvented the ban by short selling financial-sector ETFs instead.

1.4 Synthetic Shorting with ETFs

Given the advantages of ETF short selling discussed in the previous subsections, we argue that some short sellers would use ETFs to create "synthetic" short positions instead of directly shorting individual names. In order to do so, the trader would short the ETF that contains the target stocks, and enter long positions in all of or a sample of the ETF's underlying stocks that are not the targets. If the ETF is value weighted, a buy-and-hold synthetic shorting strategy would inversely track the performance of the target stocks.

In evaluating the synthetic shorting strategy against direct shorting, the short seller must trade off the benefits and costs of shorting via ETFs. The most direct cost is the lending fee for borrowing ETF shares. When we compare the average lending cost score ($DCBS$ in Markit dataset) for stocks and ETFs in our sample, the lending costs are roughly the same. This cost is partially offset by the management expense of the ETFs, which average about 44 basis points per annum. Another important source of costs are the transaction costs of entering into the long positions of the synthetic shorting strategy. Since an ETF typically has hundreds of underlying stocks, establishing long positions in all but a few stocks can incur a non-trivial amount of transaction costs for the trader.[14] On the other hand, if the short seller does not fully hedge, she bears the risk of price movement of the ETF itself. Should the ETF price unexpectedly appreciate, the short seller would suffer losses.

The static lending cost, however, may not be the only concern for short sellers. Engelberg, Reed, and Ringgenberg (2018) point out that an important consideration for short sellers is the risk of loan being called and, conditional on being called, whether the lending cost spikes at the inopportune moment. In the data, the distribution of stock-level lending cost score

[14] A recent paper by Frazzini, Israel, and Moskowitz (2012) estimate that the actual trading costs faced by real-world arbitrageurs are an order of magnitude smaller than previous studies suggest. The mean transaction costs are about 11 bp and 21 bp in large cap and small cap stocks, respectively.

is skewed to the right, both unconditionally and within-stock. In contrast, the distribution of ETF lending cost score is less skewed, suggesting less risks involving shorting ETFs. Moreover, since the stock lending market is highly fragmented and the stock-level lending fee data come from transactions where the short seller and the lender are successfully matched, it is possible that the observed lending cost is an under-estimation of the true lending cost. When a stock is highly difficult to short, supply is extremely limited and the intended short sellers fail to short the stock. In such cases, more liquid ETF securities can be the only alternative outlet for carrying out the short trade.

2 Data and Summary Statistics

2.1 Sample Construction

Our sample contains all U.S. domestic equity ETFs that physically replicate their indices.[15] To obtain a list of such ETFs, we start by intersecting all funds in the CRSP mutual fund database with ETF designation (etf_flag=F) with securities in CRSP monthly stock file with share code of 73. We then manually filter out non-domestic or non-equity ETFs by parsing the fund name.[16] To retrieve holdings of ETFs, we use Thomson Reuters Mutual Fund holdings database (S12) and supplement it by CRSP Mutual Fund holdings database.[17] We require an ETF to hold at least 20 common stocks. Our sample contains a total of 601 ETFs during the time span of 2002 to 2019.

Monthly short interest series of both ETFs and stocks comes from Compustat. Each month, U.S. exchanges report the level of short interest on the 15th of each month.[18] To form the short interest ratio (SR), we normalize short interest by total shares outstanding from CRSP. We obtain stock lending supply (lendable shares divided by total shares outstanding),

[15] Most ETFs in the U.S tend to physically replicate their underlying index. The Investment Act of 1940 requires ETFs to hold 80% of their assets in securities matching the fund's name.

[16] We search for terms in fund names such as "International", "World", "Ex-US", "Treasury", or "Municipal".

[17] Zhu (2020) documents that Thomson Reuters database fails to include a large fraction of newly-founded mutual funds and ETFs after 2008, while the data quality of CRSP Mutual Fund holdings database has been improving since 2007.

[18] After September 2007, short interest data are reported twice each month and we keep the last report of each month. Our results are not materially affected if we use mid-month report throughout our sample.

the stock lending utilization ratio (shares on loan divided by lendable shares), and stock lending fees from the Markit Securities Finance (formerly Data Explorer) database. Markit provides two variables that proxy for stock lending cost. The first variable, SAF, is the simple average fees of stock borrowing transactions from hedge funds in a given security, which is the difference between the risk-free rate and the rebate rate. SAF is only available for a stock to the extent that the stock is being shorted by a Markit client hedge fund. The second variable, $DCBS$ (Daily Cost of Borrowing Score), which covers all stocks, is a score from 1 to 10 created by Markit using their proprietary information. This score is intended to capture the cost of borrowing the stock: A score of 1 represents the cheapest to short and 10 represents the most difficult. The SAF variable is available after November 2006, while the $DCBS$ score is available after October 2003.[19]

We obtain monthly stock returns from the Center for Research in Security Prices (CRSP) and annual accounting data from Compustat. Our sample of stocks starts with all common stocks traded on NYSE, Amex, and NASDAQ. We adjust the stock returns by delisting. If a delisting return is missing and the delisting is performance-related, we set the delisting return to be -30% (Shumway (1997)). We remove stocks with month end price less than $3.

We use standard control variables in our empirical analysis. Stock market cap ($LnME$) is defined as the natural logarithm of a stock's market capitalization at the end of June in each year. Book-to-market ratio ($LnBM$) equals to the most recent fiscal year-end report of book value divided by the market capitalization at the end of calendar year t-1. Book value equals the value of common stockholders' equity, plus deferred taxes and investment tax credits, minus the book value of preferred stock. Momentum (MOM) is defined as the cumulative holding-period return from month t-12 to t-2, skipping the most recent month t-1. Short term reversal (REV) is the prior month's return. *Turnover* is the monthly trading volume over shares outstanding, averaged over past 12 months. Since the dealer nature of the NASDAQ market makes its turnover difficult to compare with the turnover observed on NYSE and AMEX, we follow Gao and Ritter (2010) by adjusting trading volume for NASDAQ stocks.[20] Institutional ownership (IO) is the sum of shares held by institutions

[19] See Saffi and Sigurdsson (2011) and Beneish, Lee, and Nichols (2015) for a detailed account of Markit equity lending database.

[20] Specifically, we divide NASDAQ volume by 2.0, 1.8, 1.6, and 1.0 for the periods before February 2001,

from 13F filings in each quarter divided by total shares outstanding. Idiosyncratic volatility ($IVOL$) is the standard deviation of the residuals from the regression of daily stock excess returns on Fama and French (1993) 3-factor returns within a month (Ang, Hodrick, Xing, and Zhang (2006)). Institutional ownership data is available from Thomson Reuters (formerly CDA/Spectrum) Institutional Holdings database (13F).

2.2 ETF Characteristics

ETFs are characterized by both ETF-level variables and the weighted-average characteristics of their underlying stocks. At the ETF level, our focus is the short ratio of the ETF, defined as open short interests divided by shares outstanding of the ETF. We are also interested in an ETF's market capitalization (CRSP Price * Total Shares Outstanding), turnover ratio (CRSP Volume/Total Shares Outstanding), past 12-month return, return volatility, expense ratio, and monthly flow. To calculate ETF flow, we obtain month-end ETF shares outstanding from Bloomberg.[21] The flow of ETF i in month t is calculated as:

$$Flow_{i,t} = \frac{SharesOutstanding_{i,t} - SharesOutstanding_{i,t-1}}{SharesOutstanding_{i,t-1}}. \quad (1)$$

As for the characteristics of underlying stocks, we aggregate the stock-level idiosyncratic volatility, market capitalization, book-to-market ratio, short interest ratio, lending supply, lending utilization, lending cost (DCBS score), and Amihud (2002) illiquidity measure by using a weighted average of these characteristics at the ETF level.

Panel A of Table 1 shows the summary statistics of ETF characteristics. Since we are interested in the short selling activities on ETFs, we further sort ETFs into quintile groups based on ETF short ratio and summarize the average characteristics for each group. The results are shown in Panel B of Table 1. One thing to note is that ETFs with the highest short ratio are significantly larger than ETFs with the lowest short ratio. This in part explains why the value-weighted short ratio of ETFs (about 15%) is much larger than the

between February 2001 and December 2001, between January 2002 and December 2003, and after January 2004, respectively.

[21] In a small number of cases where Bloomberg does not have shares outstanding information for a particular ETF-month, we supplement our data with the shares outstanding information from CRSP.

simple average ETF short ratio (about 4.8%).

2.3 ETF-Based Short Ratio for Stocks

The key innovation of our paper is to aggregate the information content in ETF short selling activities to the stock level and construct a variable that we call the "*ETF-based short ratio*" (or *ETF-based SR)*. Intuitively, if a stock is overvalued but difficult to short directly, traders can form "synthetic" short portfolios by combining short positions in ETFs and long positions in other ETF constituent stocks. Since we do not observe synthetic shorters' long positions, short interests on an ETF can be attributed to any subset of its underlying stocks. To more efficiently extract information from ETF short selling to underlying stocks, we examine the intersection of the short selling of all ETFs: if some stocks are held by several ETFs that all have high short ratios, it is more likely that these stocks are the targets of synthetic shorting strategies.

To this end, we first calculate the total value of short interests for each ETF-month, and we attribute this dollar value of short interests to its constituent stocks proportional to the value of stocks held by the ETF. Specifically, for Stock i during month t, the dollar value of short selling via ETF e equals

$$short\_value_{i,e,t} = short\_interest_{e,t} * P_{e,t} * \frac{shares\_held_{i,e,t-1} * P_{i,t}}{\sum_{j \in J_e} shares\_held_{j,e,t-1} * P_{j,t}} \tag{2}$$

where $P_{e,t}$ denotes the per share price of ETF e, $P_{i,t}$ denotes the per share price of Stock i, J_e denotes the set of stocks held by ETF e, and $shares\_held_{i,e,t-1}$ denotes the number of Stock i's shares held by ETF e at last quater end $t-1$.

Then, for Stock i, we aggregate the dollar value of short selling across all ETFs that hold Stock i during month t, and scale it by the market capitalization of Stock i:

$$\text{ETF-based } SR = \frac{\sum_{e \in E_i} short\_value_{i,e,t}}{shares\_outstanding_{i,t} * P_{i,t}} \tag{3}$$

where E_i is the set of ETFs that hold Stock i.

As a simple numerical example, suppose there are only three ETFs in the stock market.

ETF A has a short interest ratio of 5%, an NAV of \$100 million and holds 10% of its assets in stock x and 20% of its assets in stock y. ETF B has a short interest ratio of 10%, an NAV of \$200 million and holds 5% of its assets in stock x and 25% of its assets in stock z. ETF C has a short interest ratio of 20%, an NAV of \$ 50 million, and holds 15% of its assets in stock y and 5% of its assets in stock z. Further suppose that stock x has a market cap of \$50 million, stock y has a market cap of \$55 million, and stock z has a market cap of \$100 million. Then the ETF-based SR for stock x is $\frac{100*5\%*10\%+200*10\%*5\%+50*20\%*0\%}{50} = 3\%$. The ETF-based SR for stock y is $\frac{100*5\%*20\%+200*10\%*0\%+50*20\%*15\%}{55} = 5\%$. The ETF-based SR for stock z is $\frac{100*5\%*0\%+200*10\%*25\%+50*20\%*5\%}{100} = 5.5\%$.

While we do not have data on the long leg of the synthetic shorting strategy, we partially validate our ETF-based short ratio measure by examining its relation with hedge fund ownership. Hedge funds are among the most sophisticated investors in stock market and often conduct arbitrage trades. If the ETF-based short ratio captures well the actual short-selling intentions from synthetic shorting trades, one should find that hedge funds will collectively hold less of those stocks with high ETF-based short ratio. Table A.1 in the Internet Appendix shows that there is indeed a significantly negative cross-sectional relation between hedge fund ownership and ETF-based short ratio.[22] It indicates that stocks which are heavily shorted via ETFs are eschewed by the most sophisticated investors.

Another caveat about our ETF-based short ratio measure is that we use the ETF portfolio weight from the most recent quarter-end to approximate the exposure to stocks a trader would obtain when he shorts the ETF. This is not necessarily precise as an ETF's portfolio weight of stocks may change over the course of a quarter, especially if the ETF's creation/redemption baskets do not fully replicate the underlying holdings. However, as documented by Todorov (2021), equity ETFs tend to cover 85% to 95% of securities in their creation/redemption baskets, with the largest equity ETFs fully replicating.

[22]The classification of hedge funds from 13F filers follows the scheme of Maslennikov and Hund (2015). We thank Sergey Maslennikov for generously providing us with the hedge fund classification data.

2.4 Stock Characteristics

Table 2 presents the summary statistics of the stock characteristics. Panel A reports the time-series average of the cross-sectional means and standard deviations of the variables for the full sample. The average short interest ratio (SR) in our sample period is 4.73%. The median SR is around 2.98%. This means that while most stocks have low shorting activity, a small fraction of stocks are heavily shorted. Our key variable of interest, the ETF-based short ratio (ETF_sr), has a mean of 0.50% and median of 0.35% of a stock's market cap. The cross-sectional standard deviation of ETF_sr is 0.44%, indicating a large cross-sectional variation among stocks. The median annualized lending fee SAF is small: only 22.5 bps. However, the distribution of the lending fee is highly skewed to the right, with the mean annualized lending fee being 258 bps. This is consistent with the literature that although most stocks are easy to borrow, a small fraction of stocks with low lendable supply have high shorting costs (D'avolio (2002)). And these stocks are the most prone to overpricing induced by short-selling constraints. The average lendable supply is 17.28% of total shares outstanding, with a standard deviation of 9.73%. We also construct a ETF-based flow measure at the stock-level for the average flows received by holding ETFs:

$$\text{ETF-based flow}_{i,t} = \frac{\sum_{e \in E_i} (Flow_{e,t} * shares\_held_{i,e,t-1})}{SharesOutstanding_{i,t}}, \tag{4}$$

where $Flow_{e,t}$ is the flow of ETF e at month t, $shares\_held_{i,e,t-1}$ is the number of firm i's shares held by ETF e at end of previous month, and $SharesOutstanding_{i,t}$ is the number of shares outstanding for firm i at month t.

Table 2 Panel B reports the pairwise rank correlation among our variables where they overlap. As we can see, the rank correlation between a stock's ETF-based short ratio and its ETF-based flow is 0.19. This moderate positive correlation suggests that a stock's ETF-based short ratio is largely independent of its prorated ETF flows. The rank correlation between ETF-based short ratio and the stock's own short ratio is 0.49. Given the high correlation, in our later empirical analyses of the return predictability of ETF-based short ratio, we will control for existing shorting demand or cost measures at the stock level.

3 ETF-level Evidence for Synthetic Shorting

In this section, we examine the determinants of short interests at ETF level. If, as we hypothesize, ETF short selling provides an alternative mechanism for market participants to gain negative exposure to underlying stocks, we should expect ETF short interest to be associated with the tightness of the short-selling conditions of the underlying stocks. To this end, we use Fama-MacBeth regressions to show the relation between stock-level shorting constraints and the short interests of ETFs.

3.1 Determinants of ETF Short Selling

To understand how much ETF short selling represents traders' attempts to bet against underlying stocks, we run monthly Fama and MacBeth (1973) regressions of ETF short ratios on ETF characteristics and the characteristics of ETFs' underlying stocks. The goal is to control for other motives of ETF short selling, such as hedging or operational shorting (Evans, Moussawi, Pagano, and Sedunov (2018)), and to examine the relation between ETF shorting and the shorting supply/demand of underlying stocks.

To achieve this goal, we regress ETF short ratios, defined as the monthly short interests scaled by ETF shares outstanding, on a large set of explanatory variables: ETF characteristics include ETF turnover, the logarithm of ETF market capitalization (ME), ETFs' annual expense ratio, ETF monthly flows, ETFs' past 12-month return, ETF's monthly return volatility, and ETFs' lending fee scores. The characteristics of ETFs, which are relatively persistent, can be reasonably expected to capture the general hedging demand for shorting ETFs. Furthermore, Evans, Moussawi, Pagano, and Sedunov (2018) document that ETF authorized participants regularly engage in "operational shorting", in which APs absorb buy-sell order imbalance by holding the creation baskets and shorting the ETFs. To control for the operational shorting activities, we calculate the three-day buy-sell order imbalance for each ETF prior to the release of monthly short interests. We then take a maximum between zero and the order imbalance to proxy for the operational shorting as in Evans, Moussawi, Pagano, and Sedunov (2018). We expect order imbalances would positively correlate with ETF short ratios. Lastly, we control for the weighted average characteristics of ETFs' underlying stocks: the logarithm of stock market capitalization ($LnME$) and the book-to-market ratio (BM).[23]

Column (1) of Table 3 reports the effects of ETF and stock characteristics on the ETF short ratio. The turnover ratio of an ETF is a deciding factor of its short selling activity. To the extent that turnover ratio represents the trading liquidity of the ETF, this result is consistent with the intuition that high liquidity of ETFs attracts short sellers who are wary of short selling risks. The past returns of ETFs are negatively associated with ETF short ratios, indicating that investors tend to short sell underperforming ETFs. Consistent with operational shorting motives, month-end buy-sell order imbalance is positively associated with ETF short ratios. ETF monthly flow is negatively associated with short ratio, suggesting that short sellers on average avoid highly popular ETFs. In terms of the characteristics of underlying stocks, ETFs holding value stocks (high book-to-market ratios) have higher short ratios.

The baseline results in Column (1) suggest that several ETF-level characteristics can explain slightly less than 40 percent of the cross-sectional variations in ETF short ratios. Our next step is to examine whether the short-selling activities of underlying stocks are associated with short ratios of the corresponding ETFs. If some traders use ETFs to circumvent short-sale constraints of underlying stocks, then the shorting supply and demand of constituents should have explanatory power on ETF short ratios, above and beyond ETF characteristics.

Column (2) of Table 3 includes the weighted average stock idiosyncratic volatility and illiquidity measure (Amihud (2002)) as regressors. ETFs whose underlying stocks are more illiquid and volatile tend to have higher short ratios. To the extent that both illiquidity and idiosyncratic volatility are impediments to directly shorting underlying stocks, this is consistent with the notion of synthetic shorting with ETFs. In Column (3) of Table 3, we include the weighted-average short ratio of underlying stocks as a regressor. There is insignificant relation between the average stock short ratio and the ETF short ratio. This may be due to the fact short ratios are equilibrium quantities determined by both supply and demand, as pointed out by Cohen, Diether, and Malloy (2007). It is possible that the short interest of an ETF's underlying stocks is high because of ample lending supply. Hence

[23]We do not include stock past returns, since it is highly correlated with ETF past returns ($corr > .95$).

it is informative to separate out the supply of lendable shares and the demand for shorting underlying stocks.

In Column (4), we proxy for stock lending supply using lendable shares from Markit scaled by total shares outstanding. We find that lendable supply of underlying stocks is negatively correlated with the ETF-level short ratio. A one percentage point decrease in stock lending supply is associated with a 22.5 bps increase ($t = 3.64$) in ETF short ratio. In Column (5), we isolate the strength of shorting demand by using the stock lending utilization ratio. Our results suggest that demand for shorting underlying stocks is positively correlated with the ETF short ratio. A ten percentage point increase in the utilization ratio is associated with a 52.6 bp increase in the ETF short ratio ($t = 4.53$).

Finally, in Columns (6) and (7), we examine the relation between stock lending fee and ETF short ratios. The lending fee is proxied by the DCBS scores provided by Markit, ranging from 1 to 10. In Column (6), a one-notch increase in the DCBS score of underlying stocks is associated with a 1.20 percentage points increase in the ETF short ratio ($t = 5.79$). In Column (7), we calculate the average lending fee for the decile of most expensive-to-borrow stocks within an ETF's constituents. The coefficient on *top decile stock lending fee* is positive, indicating that the high lending fees to access a small subset of the underlying stocks within the ETF seem to be an important driver for ETF short-selling activities. The evidence is consistent with our hypothesis that ETF short selling is particularly attractive for synthetically shorting their most difficult-to-short constituents.

Overall, our findings on the determinants of ETF short selling indicate that ETF short selling is more active for ETFs whose underlying stocks are in high shorting demand and low lending supply. Under those circumstances, ETFs may become an alternative venue for traders to gain negative exposure to underlying stocks.

3.2 Does ETF Short-Selling Predict ETF Returns?

We provide further evidence for what motivate traders to short ETFs by examining the return predictability of ETF short ratios. The literature on stock short selling has shown that stock-level short interests negatively predict future stock returns (e.g., Figlewski

(1981); Dechow, Hutton, Meulbroek, and Sloan (2001); Asquith, Pathak, and Ritter (2005)). Moreover, Rapach, Ringgenberg, and Zhou (2016) document that short interests aggregated across stocks negatively forecast future market return. One would expect a similar inverse relationship between ETF short ratios and future ETF returns if most of ETF short selling is driven by directional betting. On the other hand, if most of ETF short selling is used for hedging or operational shorting, then one would expect ETF short ratios and future ETF returns to be uncorrelated.

To empirically investigate this question, we sort ETFs into quintile portfolios each month based on their short ratios and hold them over the next month. The ETF portfolio return is weighted either equally or by the market cap of the ETFs. A long-short portfolio is formed by taking a long position in the most lightly shorted ETFs (Quintile 1) and a short position in the most heavily shorted ETFs (Quintile 5). The return series runs from January 2002 to December 2019.

Table 4 reports the returns and alphas of five ETF portfolios and the long-short portfolio. In Panel A, returns are equal-weighted. The monthly return spread for the long-short portfolio that buys ETFs that are lightly shorted and sells ETFs that are heavily shorted is 5 basis points with a t-stat of 0.73. After adjusting for the Fama and French (1993) three factors and the Carhart (1997) four factors, the long-short abnormal return is about 16 to 17 bps per month with a t-stat exceeding 2.8. The significant profitability of this long-short strategy indicates that ETF-level short interests are informative for future ETF returns.

In Panel B of Table 4, returns are value-weighted, and the return predictability of ETF short ratio is similar to the equal-weighted strategy. The risk-adjusted returns of the long-short strategy is about 8 to 9 bps per month, and depending on the risk-adjustment, the t-statistics range from 1.84 to 2.07. This suggests that at least part of short positions on ETF is directional and contains value-relevant information for ETF return.[24]

Taken together, our empirical results echo the findings in the stock shorting literature: ETF short sellers seem to have superior information about future returns of the ETFs, or

[24] In untabulated tests, we double sort ETFs on short ratios and monthly flows, as recent studies document that ETF flows negatively predict ETF returns (Brown, Davies, and Ringgenberg (2021), Ben-David, Franzoni, and Moussawi (2018)). We find that the return predictability of ETF short ratios is independent of ETF flows.

at least some of their constituent stocks; and they are able to earn abnormal returns by short selling. This suggests that directional betting is the main component of ETF shorting activities. However, to further distinguish between short selling that bets against the whole sector/market and (synthetic) short selling that targets a subset of underlying stocks, one should examine ETFs jointly. In subsequent analyses, we aggregate ETF short interests to individual stock level and examine the return predictability of ETF shorting for stocks.

4 The Information Content of ETF Shorting for Individual Stocks

To the extent that arbitrageurs use ETFs to express their bearish opinions on individual stocks, stocks that are heavily shorted via their holding ETFs should expect to earn negative returns in the future. Moreover, if arbitrageurs are more likely to switch to ETFs when their target stocks become difficult to short, the return predictability of the ETF-based short ratio should be concentrated among stocks that have severe impediments to shorting. In this section, we test the return predictability of the ETF-based short ratio using both portfolio sorts and Fama-MacBeth regressions.

4.1 Portfolio Sorts

In this section, we show that stocks sorted on their ETF-based short ratio generate significant return spreads. We conduct the decile portfolio sorts as follows. At the end of each month, we sort stocks into deciles by their ETF-based short ratios. We then compute the average return of each decile portfolio over the next month, both equal-weighted and value-weighted. This gives us a time series of monthly returns for each decile. We use these time series to compute the average return of each decile over the period of January 2002 to December 2019. As we are most interested in the return spread between the two extreme portfolios, we also report the return to a long–short portfolio (i.e., a zero-investment portfolio that goes long the stocks in the lowest ETF-based short ratio decile and shorts the stocks in the highest decile). Table 5 reports the average excess return (and associated *t*-

statistics) of this long–short portfolio in the leftmost columns, with the Fama and French (1993) three-factor adjusted alphas in the middle, and the Carhart (1997) four-factor alphas in the rightmost column.

In Panel A of Table 5, the equal-weighted portfolio excess return decreases from 1.03% to 0.26% per month from the lowest decile to highest decile of ETF-based short ratio. The return spread for the long-short portfolio sorted on ETF-based short ratio is 0.77% per month, with a t-stat of 3.83. Adjusting for risk exposure to the Fama and French (1993) three factors reduces the long-short return spread to 0.58%, but the alpha is still highly significant with t-stat of more than 3. For Carhart (1997) four-factor adjusted alphas, the return spread is 0.59% per month, with a t-stat of 3.08. In Panel B, we see that the value-weighted results are weaker but are nonetheless statistically and economically significant across the board. For excess returns, the long-short portfolio generates a monthly 0.59% with a t-stat of 2.87. The figure decreases to 0.41% for the three- and four-factor alphas with a t-stat of around 2. So, regardless of the metric, stocks that are heavily shorted via ETFs underperform those lightly shorted. The economic magnitude is non-trivial given the fact that many well-documented asset pricing anomalies are no longer profitable in our sample period (Chordia, Subrahmanyam, and Tong (2014)).

In Table A.2 in the Internet Appendix, we report the factor loadings of the long and short leg, as well as the hedge portfolio, on the Carhart (1997) four factors. For both the equal-weighted and value-weighted portfolios, the hedge portfolio loads positively on the market factor, and load negatively on the value factor. The large positive loading on the market factor may explain why factor-adjusted alphas are smaller than the raw return spread.

Brown, Davies, and Ringgenberg (2021) show theoretically and empirically that ETF flows provide signals of non-fundamental demand shocks and negatively predict future returns for both ETFs and underlying stocks. To examine whether the return predictability of ETF-based short ratio is independent of ETF-based flow, we calculate residual ETF-based short ratio as the residual from monthly cross-sectional regression of ETF-based short ratio on ETF-based flow. We then sort stocks into decile portfolios based on the residual ETF-based short ratio and report the results in Table A.3 in the Internet Appendix. Results show that the residual ETF-based SR also generates significant return spreads. The equal-weighted long-short portfolio generates 0.51% ($t = 2.61$) four-factor alpha, while the value-weighted portfolio alpha is 0.47% with a t-stat of 2.03. This finding thus suggests that ETF short interests contain incremental return information beyond that contained in ETF flows.

4.2 Robustness of Portfolio Sorts

In Table A.4 in the Internet Appendix, we examine the robustness of our portfolio sorts. The first row shows the return spread when returns are weighted by past month gross return, as suggested by Asparouhova, Bessembinder, and Kalcheva (2013). The gross-return-weighted return spread is 0.63% ($t = 3.30$). The second row shows that our results are robust when we subtract the corresponding Fama-French 48 industry return from stock return. This suggests that there are stock-specific return information contained in ETF-based short ratio. In the third row, we augment the Carhart (1997) four-factors with the Pástor and Stambaugh (2003) liquidity factor, and the alphas remain significant. In Rows (4)-(7), we show the results survive when we use the Fama and French (2016) five-factor model, the Stambaugh and Yuan (2016) mispricing factor model, the Hou, Xue, and Zhang (2015) q-factor model, and the Daniel, Hirshleifer, and Sun (2020) short- and long-horizon behavioral factors to calculate alphas. Our results survive when we exclude stocks that have a price less than $5 or whose market cap is below the size threshold of NYSE bottom decile in the prior month end, as shown in rows (8) and (9), respectively. The tenth row shows that the long-short portfolio alphas are still highly significant if we skip a month between when we sort stocks and when we calculate strategy returns. The eleventh row reports the long-short alpha when we form decile portfolios based on the ETF-based short ratio after purging out the effect of stock's own short ratio.[25] In the last row of Table A.4, we report the portfolio alphas based on quintile sorts of ETF-based short ratio. Both the equal-weighted and value-weighted portfolios generate highly significant alphas. In sum, across almost all the specifications, stocks that are heavily shorted via ETFs significantly underperform those lightly shorted.

[25]Specifically, each month we run a cross-sectional regression of the ETF-based short ratio on stocks' own short ratio and take the regression residual as our sorting variable.

4.3 Two-Way Sorts on Limits to Arbitrage and ETF-Based Short Ratio

Having established the return predictability of the ETF-based short ratio through univariate portfolio sorts, we next examine whether the return predictability varies across stocks with differential degree of short-sale constraints. If, as we hypothesized, some arbitrageurs short ETFs to circumvent short-selling constraints at the stock level, the return predictability of ETF-based short ratio should be stronger among such difficult-to-short stocks.

To test this, we conduct sequential double sorts first on limits-to-arbitrage proxy and then on ETF-based short ratio. At the end of each month, all stocks are sorted into terciles based on a specific proxy for limits to arbitrage, and within each tercile, we further sort the stocks into quintiles based on their ETF-based short ratio. Returns are equally weighted within each portfolio. We use multiple measures of arbitrage frictions, including lendable supply, institutional ownership, lending fee, idiosyncratic volatility, and Amihud (2002) illiquidity. The first three measures more accurately capture the constraints in the equity lending market, while the latter two measures are proxies for more general arbitrage frictions.

Table 6 report the monthly Carhart (1997) four-factor alphas for equal-weighted portfolios. In Panel A, we use the lendable supply as a proxy for limits to arbitrage, which directly measures the tightness of the equity lending market. Consistent with our hypothesis, the return spread on the ETF-based short ratio is much higher among stocks with low lendable supply. Specifically, the four-factor alpha is 0.68% (t = 2.32) in the lowest lendable supply tercile. The figures are only 0.26% and -0.12% for the other two terciles, and are no longer significant. The difference in alphas between stocks with low and high lendable supply is 0.80% (t = 3.67). In Panel B, we show that the same pattern is observed when we use institutional ownership as a proxy for short-sale constraints (Nagel (2005)). Because institutional investors actively participate in stock lending programs, the fraction of shares owned by institutional investors is highly correlated with actual lending supply.[26] Consis-

[26]The average cross-sectional correlation between institutional ownership and lendable supply is 0.79 in our sample.

tent with our hypothesis, the four-factor alpha based on ETF_sr is 0.83% ($t = 2.57$) in the tercile with the lowest institutional ownership, while it decreases to -0.17% for stocks with high institutional ownership. The difference in alphas between stocks with low and high institutional ownership is 1.0% ($t = 3.78$).

In Panel C, we use the stock lending cost as a direct proxy for short-sale constraints (Jones and Lamont (2002); Drechsler and Drechsler (2014)). Following the literature, we sort stocks into two groups based on whether a stock's DCBS score is above or below 2. As we can see, the return predictability of ETF-based short ratio is significantly amplified among stocks with elevated borrowing cost. The monthly four-factor alpha is 1.38% ($t = 3.63$) among stocks with DCBS scores above 2, and only -0.09% among stocks with DCBS scores equal or below 2. In addition, the difference in long-short portfolio alpha across stocks with different level of lending costs is mainly driven by the short leg. Specifically, for stocks with high shorting costs, those most heavilty shorted through holding ETFs have monthly alpha of -178 bps. This further supports our hypothesis that ETFs are used by arbitrageurs to target overvalued stocks that are costly to short directly.

Pontiff (2006) argues that stocks with high idiosyncratic volatility are more costly to arbitrage. Duan, Hu, and McLean (2010) and Stambaugh, Yu, and Yuan (2015) provide empirical evidence supporting this argument. Panel D of Table 6 reports the double sorting results when we use idiosyncratic volatility as a proxy for limits to arbitrage. The monthly return spread is 0.83% ($t = 2.53$) for stocks in the highest tercile of idiosyncratic volatility, and is close to zero and become insignificant for stocks with lower idiosyncratic volatility. In Panel E, we use Amihud (2002) illiquidity as proxy for arbitrage friction and find similar results. In Table A.5 in the Internet Appendix, we conduct the same set of analyses for value-weighted portfolios and find very similar evidence that ETF-based short ratio more strongly predict returns of stocks with high short-sale constraints.

Overall, the stronger return predictability of ETF-based short ratio among stocks with lower lendable supply, higher lending cost, higher idiosyncratic volatility, and lower liquidity is consistent with our hypothesis that short sellers effectively use ETFs to create synthetic short positions on stocks that are costly to short directly.

4.4 Fama-MacBeth Regressions

We now test our main hypothesis using the Fama and MacBeth (1973) regression methodology. One advantage of this methodology is that it allows us to examine the predictive power of ETF-based short ratio while controlling for other known predictors of cross-sectional stock returns. This is important because, as shown in Table 2, ETF-based short ratio is correlated with some of these predictors. We conduct the Fama-MacBeth regressions in the usual way. Each month, starting from January 2002 and ending in December 2019, we run a cross-sectional regression of stock returns on lagged ETF-based short ratio and a set of control variables known to predict stock returns, including the natural logarithm of the market value of equity ($LnME$), the natural logarithm of the book-to-market ratio ($LnBM$), returns from the prior month (REV), returns from the prior 12-month period excluding month t-1 (MOM), institutional ownership (IO), and idiosyncratic volatility ($IVOL$).

Table 7 reports the time-series averages of the coefficients on the independent variables, and the t-statistics are Newey-West adjusted with four lags to control for heteroskedasticity and autocorrelation.[27] We only include ETF-based short ratio (ETF_sr) in Column (1) as a baseline and it attracts a negative coefficient of -0.345 ($t = -2.92$). This is consistent with our portfolio sorting results in which stocks that are heavily shorted via ETFs have lower future returns. In Column (2), we add the usual controls including firm size, book-to-market ratio, past 1-month return, and past 12-month returns. The coefficient of ETF_sr is -0.345 with a t-stat of -2.46. In Column (3), we further add institutional ownership, idiosyncratic volatility, and ETF-based flow in the regression. The coefficient of ETF_sr increases to -0.456 with a t-stat of -2.63. The economic magnitude is also reasonable. The difference of ETF-based short ratio between the lowest and highest decile portfolio is 1.15%, which implies a monthly return spread of 52 bps between the two extreme portfolios. The economic magnitude estimated from Fama-MacBeth regression is in line with our portfolio sorting results. For the control variables, the sign of most coefficients is consistent with the literature, except for momentum, which attracts an insignificant coefficient.[28] We find

[27]The results are robust if we use Newey-West Standard Errors with 12 lags or no lags.

[28]This is due to the 2009 momentum crash, see Daniel and Moskowitz (2016). The coefficient on Momentum becomes positive once we exclude year 2009 from our sample.

ETF-based flow also negatively predict next-month stock return, consistent with Ben-David, Franzoni, and Moussawi (2018) and Brown, Davies, and Ringgenberg (2021).

Our ETF-based short ratio is constructed as the dollar value of short interests on a stock via ETFs over the stock's market capitalization. As a robustness test, we construct an alternative version of ETF-based SR (ETF_sr2) by replacing the denominator in equation (3) with the dollar value held by all ETFs for a stock. Column (4) of Table 7 shows that this alternative ETF-based short ratio also negatively predict future returns with a t-stat of -2.12.

We conduct several additional tests under the Fama-MacBeth regression framework and report the results in Table A.6 in the Internet Appendix. First, one would expect the synthetic shorting strategy to be more effective when the stocks targeted by short sellers have larger portfolio weights within the shorted ETFs. For these stocks, taking the long positions in non-targeted stocks within the ETFs is less costly. To test this conjecture, we create a dummy variable $Highweight$ which equals to one for a stock belonging to the top quintile in terms of its average weight within the holding ETFs, where average weight is defined as the average weight a stock has across all of its holding ETFs. Column (1) reports the Fama-MacBeth regression of next-month returns on the interaction between ETF-based SR and the dummy $Highweight$ ($ETFsr\_Highweight$). Consistent with our hypothesis, the coefficient on this interaction term $ETFsr\_Highweight$ is significantly negative, while ETF_sr itself becomes insignificant. Secondly, we run Fama-MacBeth regression of future stock return on ETF-based short ratio for two subperiods: one from 2002 to 2010 and another from 2011 to 2019. Columns (2) and (3) of Table A.6 show that the return predictability of ETF-based is significant in both subperiods.

Our paper argues that ETF short interests contain incremental information about stock-level expected return beyond what contained in a stock's own short ratio, which drives the observed return predictability. A plausible alternative explanation is that the return predictability of ETF-based short ratio is due to short sellers using ETFs to bet on sector-level or factor-level trend. To distinguish between the two explanations, we run Fama-MacBeth regression of stocks' excess return on ETF-based SR, where excess return is computed after subtracting the corresponding industry or characteristic-matched portfolio return from individual stock returns. Industry return is calculated based on Fama-French 48 industry classification and characteristic-matched portfolio return is calculated following Daniel, Grinblatt, Titman, and Wermers (1997). Columns (4) and (5) of Table A.6 report the results. As we can see, the coefficients on ETF_sr are still significantly negative, with economic magnitude similar to our baseline result. To further address the concern that the industry or factor definition we use may not correspond to the true sectors or factors tracked by ETFs, in Column (6), we use the dependent variable as stock returns adjusted by the weighted average returns of all ETFs holding the stock. The coefficient on ETF_sr is still significantly negative, with a t-stat of -2.20. This result further supports our hypothesis that at least some investors short ETFs to trade on stock-specific information instead of sector-level or factor-level trends.

In Table A.7 in the Internet Appendix, we examine the persistence of the return predictability of ETF-based short ratio. The dependent variables in Columns (1)-(12) correspond to 1-month to 12-month ahead stock returns. The predictive power of ETF-based short ratio gradually decays from -0.456 to -0.051, and is significant for forecasting returns up to eight months into the future. Importantly, we do not observe any reversal in the return predictability of ETF_sr for subsequent 12 months. The persistence of the return predictability of ETF-based short ratio alleviates concerns that the short-run predictability is driven by price pressure. By contrast, the coefficient of ETF-based flow becomes insignificant when forecasting returns beyond the first month, suggesting that the source of return predictability are different for ETF_sr and ETF-based flow. While ETF short interests contain value-relevant information that is slowly incorporated into stock prices, flow-driven price pressure reverses quickly and hence does not predict long-horizon returns in our sample.

4.5 Controlling for Stock-Level Short-Selling Measures

A large literature on short selling documents that stock-level short interest is a strong negative predictor of future returns. Several recent papers find that in addition to shorting demand, lending supply and borrowing costs also negatively predict stock returns.[29] To test whether our ETF-based short ratio contains incremental predictive power, we control for various measures of stock-level shorting activities in Fama-MacBeth regressions. The result is reported in Table 8. In Column (1), we add the stock's own short interest ratio (SR) in the regression. Consistent with prior literature, SR is a strong contrarian predictor of future returns, with a coefficient of -0.03 and t-stat of -3.67. The coefficient on ETF-based short ratio (ETF_sr), however, survives with a coefficient of -0.37 ($t = -2.43$). This suggests that information extracted from ETF short interests is not fully absorbed by the stock's own short interests. The return predictability of ETF-based short ratio is economically comparable to that of the stocks' own short ratio. A one-standard deviation change in ETF-based short ratio translates into 16 bps of monthly stock return, while the figure is also 16 bps for stock's own short ratio. In Column (2), we add the stocks' lending cost measure ($DCBS$) in the Fama-MacBeth regression. Consistent with Engelberg, Evans, Leonard, Reed, and Ringgenberg (2020), $DCBS$ negatively and strongly predicts future returns with a t-stat of -7.63. More importantly, however, the coefficient on ETF_sr is not much affected and still highly significant with a t-stat of -3.52.

In Columns (3) and (4), we control for the stocks' utilization ratios and SIO, respectively. $Utilization$ is the ratio of shares borrowed to shares made available by Markit lenders. SIO is the short interest ratio scaled by institutional ownership. These two variables measure the tightness of the securities lending market by taking the intersection of shorting demand and supply. A stock that is highly shorted despite its low supply of lendable shares means the stock is more likely facing binding short-sale constraints. As we can see, the coefficients on the $Utilization$ and SIO are indeed significantly negative. However, ETF-based short ratio continues to significantly predict returns. In Column (5), we control for stocks' lendable supply ($Supply$) and the return predictability of ETF_sr still holds. In the last column, we show the return predictability of ETF_sr still holds when we control for both the stock's own short ratio and lending cost $DCBS$.

[29] See for example, Drechsler and Drechsler (2014) and Beneish, Lee, and Nichols (2015), among others.

4.6 Interaction Between ETF Short Selling and Stock Shorting Constraints

In this subsection, we test the cross-sectional prediction that ETF-based short ratio should have more pronounced return predictability among more difficult-to-short stocks. We do so by running Fama-MacBeth regressions of returns on ETF-based short ratio (ETF_sr) and its interaction with variables indicating tightening short-sale constraints. The results are reported in Table 9.

In Column (1), we create a dummy variable $LowIO$, which is equal to one when the stock is in the bottom quintile of institutional ownership in the cross-section. Our controls always include stock's own short ratio, so any return predictability associated with a stock's own short interests will be absorbed. Our variable of interest is $ETFsr\_lowIO$, the interaction between ETF_sr and the $LowIO$ dummy. As we can see, the coefficient on $ETFsr\_LowIO$ is -0.82 with a t-stat of -4.78. The coefficient on ETF_sr itself is negative and marginally significant, which suggests that the negative return predictability of the ETF-based short ratio is concentrated among stocks with greater short-sale constraints. This coefficient implies that the predictive power of ETF-based short ratio increases by more than two times for stocks in the bottom quintile of institutional ownership compared to stocks outside this group. In Column (2), we interact ETF_sr with a dummy $Lowsupply$, which indicates whether a stock is in the bottom quintile of lendable supply. The coefficient on this interaction term is negative with a t-stat of -1.85. In Column (3), we use a stock's utilization ratio to a proxy for the tightness of the equity lending market. We create a $Highutil$ dummy that equals to one when the stock is in the top quintile of utilization in the cross-section. The coefficient on the interaction between ETF_sr and $Highutil$ is again negative, although statistically insignificant. Column (4) reports the result when we use lending fee as proxy for shorting constraints. The variable $Highfee$ is a dummy equal to one when the stock has a $DCBS$ score greater than 2. The coefficient on this interaction term is -0.627 with a t-stat of -2.64. In Column (5), our proxy for frictions in the shorting market is stock turnover, as Hong, Li, Ni, Scheinkman, and Yan (2015) point out that short sellers are reluctant to take large short positions in low-turnover stocks. Consistent with this intuition, ETF-based short ratio exerts stronger return predictability among stocks with lower turnover. Our last proxy for short-sale constraints is whether the stock has any exchange-traded put option, as previous studies find that put options allow short sellers to express negative views through trading on the option market (Boehme, Danielsen, and Sorescu (2006); Danielsen and Sorescu (2001)). The variable $Noput$ is a dummy that equals one when the stock does not have any put option traded in a given month.[30] Supporting our hypothesis, Column (6) shows that the return predictability of ETF-based short ratio is significantly more pronounced for the subset of stocks without an exchange-traded put option.

4.7 ETF Short Selling and Earnings Surprises

The return predictability result suggests that ETF-level short interests reveal short sellers' bearish views on individual stocks, and their consensus opinions are correct on average. This raises the question of what kind of information short sellers possess when synthetically shorting stocks through ETFs. As prior studies (Dechow, Hutton, Meulbroek, and Sloan (2001)) find that short sellers (partially) derive their information advantage through analyzing firm fundamentals, we conjecture that their trades should predict future firm earnings news in the same direction as they predict future stock returns. In this section, we test this conjecture by using two proxies of firm earnings news.

Our first measure of earnings news is the cumulative abnormal returns $CAR(-1, +1)$ in a three-day window around quarterly earnings announcement date, which we extract from the Institutional Brokers' Estimate System (I/B/E/S) database. Our second proxy of earnings surprises is the standardized unexpected earnings (SUE), computed as change in split-adjusted quarterly earnings per share from its value four quarters ago divided by stock price one month prior to earnings announcements. We run Fama-MacBeth regression of earnings surprises on ETF-based short ratio and control for the usual stock characteristics that are observed prior to the earnings announcement month. We also control for lagged earnings surprises following the literature. Table 10 reports the results. In Columns (1) and (2), abnormal return is calculated as daily stock return minus return on the CRSP value-

[30]Stock-level option information are from Option Metrics database.

weighted index return. In Columns (3) and (4), abnormal return is calculated as daily stock return minus the return on the characteristics-matched portfolio following Daniel, Grinblatt, Titman, and Wermers (1997).

Across Columns (1) to (4), we find ETF-based short ratio negatively predict earnings announcement returns, even after controlling for stock's own short ratio. The economic magnitude of this predictability is also substantial. For example, Column (2) shows that the coefficient on ETF_sr is -0.214 ($t = -2.11$), suggesting that return spread between two extreme decile portfolios sorted on ETF_sr during the three-day earnings announcement window is 0.24%, compared to a monthly return spread of around 0.6% including all trading days. This implies that about 40% of the return predictability of ETF-based short ratio is concentrated on a three-day window around quarterly earnings announcement, which represents only 5% of all trading days. The fact that abnormal return is concentrated on earnings announcement days also makes our findings difficult to square with risk-based explanations (LaPorta, Lakonishok, Shleifer, and Vishny (1997); Engelberg, McLean, and Pontiff (2018)). In Columns (5) and (6) of Table 10, we find similar results when using SUE as dependent variable. By contrast, we find the coefficients of ETF-based flow are insignificant across all specifications. The finding is consistent with our interpretation that ETF flows mainly capture investors' non-fundamental demand shock, which should have little predictability for earnings news of the underlying stocks.

5 Conclusion

ETFs have become an important asset class in recent years. In this paper, we provide novel evidence that arbitrageurs use ETFs as an avenue to circumvent short-sale constraints at the stock level. ETFs are more liquid, and are not subject to the "uptick rule". In addition, new ETF shares could be created for the sole purpose of short selling. For these reasons, ETFs could be used by arbitrageurs to create synthetic short positions on stocks that are otherwise costly to short. Consistent with this hypothesis, we find that shorting activities on ETFs increase with the difficulty of shorting their underlying constituents.

We then construct a stock-level short ratio by aggregating the short interests of all ETFs holding a stock, which reflects the collective shorting demand on this stock through short selling ETFs. We find that this ETF-based short ratio strongly predicts future returns, even after controlling for stock-level shorting measures. Moreover, the return predictability of the ETF-based short ratio is concentrated among stocks that face the most severe shorting constraints. These findings corroborate our conjecture that ETF short selling is used by traders as a partial substitute of shorting underlying stocks.

References

Alexander, G. J., and M. A. Peterson, 2008, "The effect of price tests on trader behavior and market quality: An analysis of Reg SHO," *Journal of Financial Markets*, 11(1), 84–111.

Amihud, Y., 2002, "Illiquidity and stock returns: cross-section and time-series effects," *Journal of Financial Markets*, 5(1), 31–56.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, "The cross-section of volatility and expected returns," *The Journal of Finance*, 61(1), 259–299.

Asparouhova, E., H. Bessembinder, and I. Kalcheva, 2013, "Noisy prices and inference regarding returns," *The Journal of Finance*, 68(2), 665–714.

Asquith, P., and L. K. Meulbroek, 1995, *An empirical investigation of short interest*. Division of Research, Harvard Business School.

Asquith, P., P. A. Pathak, and J. R. Ritter, 2005, "Short interest, institutional ownership, and stock returns," *Journal of Financial Economics*, 78(2), 243–276.

Ben-David, I., F. Franzoni, and R. Moussawi, 2017, "Exchange-traded funds," *Annual Review of Financial Economics*, 9, 169–189.

——— , 2018, "Do ETFs increase volatility?," *The Journal of Finance*, 73(6), 2471–2535.

Beneish, M. D., C. M. Lee, and D. Nichols, 2015, "In short supply: Short-sellers and stock returns," *Journal of Accounting and Economics*, 60(2), 33–57.

Bhattacharya, A., and M. O'Hara, 2015, "Can ETFs Increase Market Fragility? Effect of Information Linkages in ETF Markets," *Effect Of Information Linkages in ETF Markets (October 10, 2015)*.

Bhattacharya, U., B. Loos, S. Meyer, and A. Hackethal, 2016, "Abusing ETFs," *Review of Finance*, 21(3), 1217–1250.

Bhojraj, S., P. Mohanram, and S. Zhang, 2020, "ETFs and information transfer across firms," *Journal of Accounting and Economics*, 70(2-3), 101336.

Boehme, R. D., B. R. Danielsen, and S. M. Sorescu, 2006, "Short-sale constraints, differences of opinion, and overvaluation," *Journal of Financial and Quantitative Analysis*, 41(02), 455–487.

Brown, D. C., S. W. Davies, and M. C. Ringgenberg, 2021, "ETF arbitrage, non-fundamental demand, and return predictability," *Review of Finance*, 25(4), 937–972.

Carhart, M. M., 1997, "On persistence in mutual fund performance," *The Journal of finance*, 52(1), 57–82.

Chague, F., R. De-Losso, A. De Genaro, and B. Giovannetti, 2017, "Well-connected short-sellers pay lower loan fees: A market-wide analysis," *Journal of Financial Economics*.

Chen, J., H. Hong, and J. C. Stein, 2002, "Breadth of ownership and stock returns," *Journal of financial Economics*, 66(2), 171–205.

Chen, Y., Z. Da, and D. Huang, 2020, "Short Selling Efficiency," in *Proceedings of Paris December 2020 Finance Meeting EUROFIDAI-ESSEC*.

Chordia, T., A. Subrahmanyam, and Q. Tong, 2014, "Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?," *Journal of Accounting and Economics*, 58(1), 41–58.

Cohen, L., K. B. Diether, and C. J. Malloy, 2007, "Supply and demand shifts in the shorting market," *The Journal of Finance*, 62(5), 2061–2096.

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

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997, "Measuring mutual fund performance with characteristic-based benchmarks," *The Journal of Finance*, 52(3), 1035–1058.

Daniel, K., D. Hirshleifer, and L. Sun, 2020, "Short-and long-horizon behavioral factors," *The review of financial studies*, 33(4), 1673–1736.

Daniel, K., and T. J. Moskowitz, 2016, "Momentum crashes," *Journal of Financial Economics*, 122(2), 221–247.

Danielsen, B. R., and S. M. Sorescu, 2001, "Why do option introductions depress stock prices? A study of diminishing short sale constraints," *Journal of Financial and Quantitative Analysis*, 36(04), 451–484.

D'avolio, G., 2002, "The market for borrowing stock," *Journal of Financial Economics*, 66(2), 271–306.

Dechow, P. M., A. P. Hutton, L. Meulbroek, and R. G. Sloan, 2001, "Short-sellers, fundamental analysis, and stock returns," *Journal of Financial Economics*, 61(1), 77–106.

Desai, H., K. Ramesh, S. R. Thiagarajan, and B. V. Balachandran, 2002, "An investigation of the informational role of short interest in the Nasdaq market," *The Journal of Finance*, 57(5), 2263–2287.

Diether, K. B., K.-H. Lee, and I. M. Werner, 2009, "It's SHO Time! Short-Sale Price Tests and Market Quality," *The Journal of Finance*, 64(1), 37–73.

Drechsler, I., and Q. F. Drechsler, 2014, "The shorting premium and asset pricing anomalies," working paper, National Bureau of Economic Research.

Duan, Y., G. Hu, and R. D. McLean, 2010, "Costly arbitrage and idiosyncratic risk: evidence from short sellers," *Journal of Financial Intermediation*, 19(4), 564–579.

Duffie, D., N. Garleanu, and L. H. Pedersen, 2002, "Securities lending, shorting, and pricing," *Journal of Financial Economics*, 66(2), 307–339.

Engelberg, J., R. D. McLean, and J. Pontiff, 2018, "Anomalies and news," *The Journal of Finance*, 73(5), 1971–2001.

Engelberg, J. E., R. B. Evans, G. Leonard, A. V. Reed, and M. C. Ringgenberg, 2020, "The Loan Fee Anomaly: A Short Seller's Best Ideas," *Available at SSRN 3707166*.

Engelberg, J. E., A. V. Reed, and M. C. Ringgenberg, 2018, "Short-selling risk," *The Journal of Finance*, 73(2), 755–786.

Evans, R. B., R. Moussawi, M. S. Pagano, and J. Sedunov, 2018, "ETF Short Interest and Failures-to-Deliver: Naked Short-Selling or Operational Shorting?," .

Fama, E. F., and K. R. French, 1993, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33(1), 3–56.

——— , 2016, "Dissecting anomalies with a five-factor model," *Review of Financial Studies*, 29(1), 69–103.

Fama, E. F., and J. D. MacBeth, 1973, "Risk, return, and equilibrium: Empirical tests," *The Journal of Political Economy*, pp. 607–636.

Figlewski, S., 1981, "The informational effects of restrictions on short sales: some empirical evidence," *Journal of Financial and Quantitative Analysis*, 16(04), 463–476.

Frazzini, A., R. Israel, and T. J. Moskowitz, 2012, "Trading costs of asset pricing anomalies," *Fama-Miller Working Paper*, pp. 14–05.

Gao, X., and J. R. Ritter, 2010, "The marketing of seasoned equity offerings," *Journal of Financial Economics*, 97(1), 33–52.

Gastineau, G. L., 2010, *The exchange-traded funds manual*, vol. 186. John Wiley & Sons.

Glosten, L., S. Nallareddy, and Y. Zou, 2021, "ETF activity and informational efficiency of underlying securities," *Management Science*, 67(1), 22–47.

Hanson, S. G., and A. Sunderam, 2014, "The growth and limits of arbitrage: Evidence from short interest," *Review of Financial Studies*, 27(4), 1238–1286.

Harrison, J. M., and D. M. Kreps, 1978, "Speculative investor behavior in a stock market with heterogeneous expectations," *The Quarterly Journal of Economics*, pp. 323–336.

Hong, H., W. Li, S. X. Ni, J. A. Scheinkman, and P. Yan, 2015, "Days to cover and stock returns," working paper, National Bureau of Economic Research.

Hong, H., J. Scheinkman, and W. Xiong, 2006, "Asset float and speculative bubbles," *The Journal of Finance*, 61(3), 1073–1117.

Hou, K., C. Xue, and L. Zhang, 2015, "Digesting Anomalies: An Investment Approach," *Review of Financial Studies*, 28(3), 650–705.

Huang, S., M. O'Hara, and Z. Zhong, 2021, "Innovation and informed trading: Evidence from industry ETFs," *The Review of Financial Studies*, 34(3), 1280–1316.

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

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

Karmaziene, E., and V. Sokolovski, 2019, "Short selling equity exchange traded funds and its effect on stock market liquidity," *Journal of Financial and Quantitative Analysis*, pp. 1–54.

LaPorta, R., J. Lakonishok, A. Shleifer, and R. Vishny, 1997, "Good News for Value Stocks:

Further Evidence on Market Efficiency," *Journal of Finance*, 52(2).

Ljungqvist, A., and W. Qian, 2016, "How constraining are limits to arbitrage?," *The Review of Financial Studies*, 29(8), 1975–2028.

Madhavan, A., 2012, "Exchange-traded funds, market structure, and the flash crash," *Financial Analysts Journal*, 68(4), 20–35.

——— , 2014, "Exchange-Traded Funds: An Overview of Institutions, Trading, and Impacts," *Annual Review of Financial Economics*, 6, 311–341.

Miller, E. M., 1977, "Risk, uncertainty, and divergence of opinion," *The Journal of Finance*, 32(4), 1151–1168.

Nagel, S., 2005, "Short sales, institutional investors and the cross-section of stock returns," *Journal of Financial Economics*, 78(2), 277–309.

Pástor, L., and R. F. Stambaugh, 2003, "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy*, 111(3), 642–685.

Pontiff, J., 2006, "Costly arbitrage and the myth of idiosyncratic risk," *Journal of Accounting and Economics*, 42(1), 35–52.

Rapach, D. E., M. C. Ringgenberg, and G. Zhou, 2016, "Short interest and aggregate stock returns," *Journal of Financial Economics*, 121(1), 46–65.

Saffi, P. A., and K. Sigurdsson, 2011, "Price Efficiency and Short Selling," *The Review of Financial Studies*, pp. 821–852.

Scheinkman, J. A., and W. Xiong, 2003, "Overconfidence and speculative bubbles," *Journal of Political Economy*, 111(6), 1183–1220.

Shumway, T., 1997, "The delisting bias in CRSP data," *The Journal of Finance*, 52(1), 327–340.

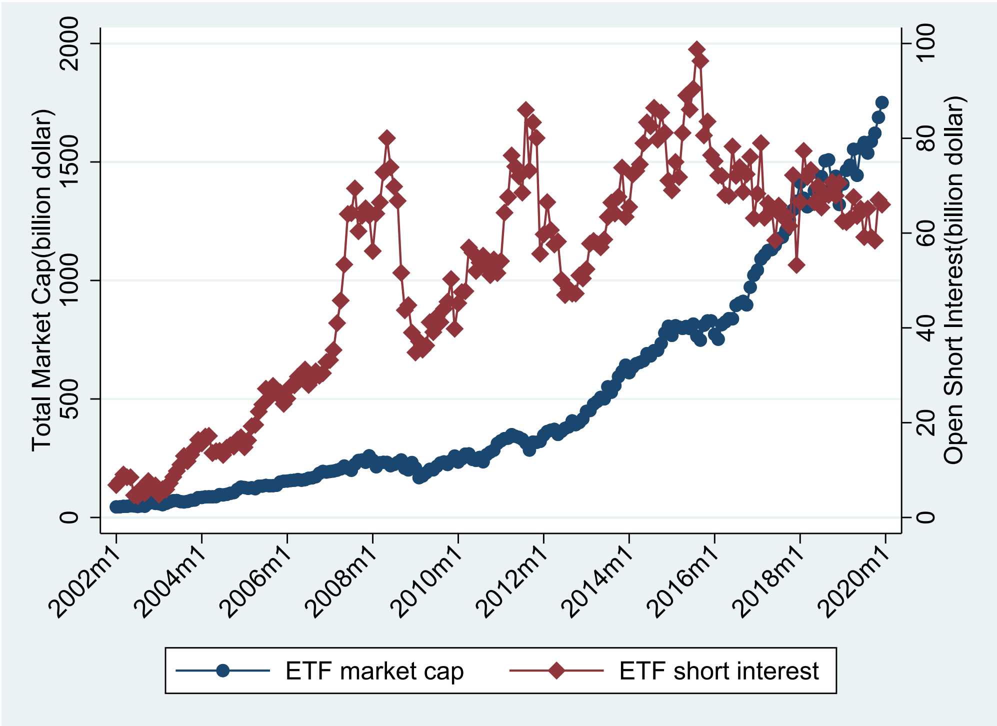
Stambaugh, R. F., J. Yu, and Y. Yuan, 2015, "Arbitrage asymmetry and the idiosyncratic volatility puzzle," *The Journal of Finance*.

Stambaugh, R. F., and Y. Yuan, 2016, "Mispricing factors," *The Review of Financial Studies*, 30(4), 1270–1315.

Todorov, K., 2021, "The anatomy of bond ETF arbitrage," *BIS Quarterly Review*.

Zhu, Q., 2020, “The missing new funds,” *Management Science*, 66(3), 1193–1204.

Figure 1: **ETF Total Market Capitalization and Total Short-Selling Value**



This figure shows the total market capitalization of all physically replicating domestic equity ETFs and the aggregate value of their mid-month open short interests. The total market capitalization is measured in billions of US dollars and is shown on the left axis. The value of open short interest is measured in billions of US dollars and is shown on the right axis.

Table 1: **Summary Statistics: ETF characteristics**

This table shows the characteristics of ETFs and the weighted-average characteristics of underlying stocks. Panel A reports the average, the median, the 25th percentile, the 75th percentile, and the standard deviation of characteristics in the pooled full sample. In Panel B, we sort ETFs into five portfolios for each month based on their short ratio. Characteristics are first averaged within each portfolio-month and then averaged across months.

	Panel A: Full Sample				
	Mean	Median	P25	P75	Std
Short Ratio(%)	4.79	0.93	0.29	2.99	10.74
Market Cap(mn)	1,078	127	26	770.31	3,800
Monthly Flow (%)	2.60	0	-0.15	3.69	13.51
12-month Return(%)	9.35	10.17	-0.21	19.96	20.75
Return Volatility(%)	5.30	4.67	3.17	6.57	2.93
Turnover Ratio(%)	2.17	0.76	0.43	1.50	5.51
Number of Stocks	419	219	98	479	538
Weighted-average characteristics of underlying stocks					
Market Cap(mn)	113,126	51,398	3,697	152,722	167,332
Book-to-Market	0.43	0.06	0.02	0.37	0.88
Idiosyncratic Volatility(%)	1.51	1.35	1.09	1.74	0.60
Short Ratio(%)	4.03	3.40	2.19	5.51	2.25
Lending Supply(%)	26.09	26.29	23.85	29.40	6.55
Lending Utilization(%)	12.62	10.56	6.22	16.53	8.85
Lending Cost	1.14	1.06	1.01	1.20	0.14
Amihud Illiquidity(%)	0.24	0.03	0.01	0.15	0.73
	Panel B: ETFs Sorted by Short Ratio				
	Low SR	2	3	4	High SR
Short Ratio(%)	0.17	0.82	1.30	3.68	19.34
Market Cap(mn)	690	1,340	1,165	1,905	2,488
12-month Return(%)	9.08	10.57	11.29	12.08	11.23
Return Volatility(%)	4.60	4.54	4.69	5.05	5.72
Turnover Ratio(%)	0.87	0.94	1.15	1.69	7.82
Number of Stocks	391	482	445	447	582
Weighted-average characteristics of underlying stocks					
Market Cap(mn)	107,034	114,066	109,016	102,601	94,546
Book-to-Market	0.38	0.34	0.35	0.40	0.59
Idiosyncratic Volatility(%)	1.41	1.40	1.42	1.47	1.57
Short Ratio(%)	3.75	3.69	3.84	4.00	4.41
Lending Supply(%)	23.73	23.66	23.81	23.97	23.66
Lending Utilization(%)	13.34	12.80	13.18	13.58	15.16
Lending Cost	1.04	1.04	1.13	1.08	1.24
Amihud Illiquidity(%)	0.22	0.20	0.19	0.21	0.36

Table 2: **Summary Statistics: Stock Characteristics**

This table presents the descriptive statistics of stock-level variables. Panel A reports the summary statistics for the full sample. Panel B reports the pairwise rank correlations among our variables where they overlap. We first calculate the summary statistics of each variable in each cross-section and then calculate the time-series average of cross-sectional statistics. The ETF-based short ratio (ETF_sr) is calculated as the dollar value of a stock shorted via its holding ETFs divided by the stock's market capitalization. ETF-based flow is monthly ETF flows aggregated to stock level. Short interest ratio (SR) is the number of shares shorted over total shares outstanding. LnME is the natural log of firm's market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of the book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative returns from month t-12 to t-2. Institutional ownership (IO) is the sum of shares held by institutions from 13F filings in each quarter divided by the total shares outstanding. Turnover is the monthly trading volume over shares outstanding, averaged over past 12 months. IVOL is the idiosyncratic volatility, calculated following Ang, Hodrick, Xing, and Zhang (2006). DCBS is a score from 1 to 10 created by Markit using their proprietary information and is intended to capture the cost of borrowing the stock. SAF is the annualized lending fee. Lendable shares (supply) is the shares held and made available to lend by Markit lenders divided by total shares outstanding. Utilization is the ratio of shares borrowed to shares made available by Markit lenders. The overall sample period is from January 2002 to December 2019.

Panel A: Summary Statistics

	Mean	Median	P25	P75	Std.
ETF_sr	0.50%	0.35%	0.05%	0.97%	0.44%
ETF-based flow	0.07%	0.05%	-0.02%	0.15%	0.11%
SR	4.73%	2.98%	1.16%	6.23%	5.36%
IO	61.38%	67.46%	38.33%	85.36%	29.55%
LnME	6.58	6.49	5.23	7.81	1.86
LnBM	-0.705	-0.613	-1.169	-0.168	0.809
MOM	0.140	0.077	-0.129	0.314	0.461
Turnover	17.55%	13.32%	6.90%	22.48%	16.27%
IVOL	0.022	0.018	0.012	0.028	0.015
DCBS	1.517	1.002	1.001	1.226	1.348
SAF	258.2	22.5	12.5	30.8	1053.2
Supply	17.28%	18.10%	9.05%	24.82%	9.73%
Utilization	25.45%	12.35%	4.29%	30.91%	37.16%

Table 2 Continued

Panel B: Pairwise Correlations

	ETF_sr	ETF-based flow	SR	IO	LnME	LnBM	MOM	Turnover	IVOL	DCBS	Ln(SAF)	Supply	Utilization
ETF_sr	1.00												
ETF-based flow	0.19	1.00											
SR	0.49	0.11	1.00										
IO	0.46	0.22	0.45	1.00									
LnME	0.18	0.22	0.26	0.55	1.00								
LnBM	-0.12	-0.04	-0.25	-0.15	-0.28	1.00							
MOM	0.04	0.03	-0.04	0.08	0.23	-0.03	1.00						
Turnover	0.29	0.14	0.67	0.54	0.43	-0.27	0.00	1.00					
IVOL	-0.11	-0.13	0.08	-0.23	-0.51	0.02	-0.17	0.09	1.00				
DCBS	-0.26	-0.14	0.02	-0.39	-0.39	-0.02	-0.15	-0.05	0.33	1.00			
Ln(SAF)	-0.05	-0.12	0.20	-0.21	-0.35	-0.02	-0.13	0.08	0.33	0.52	1.00		
Supply	0.56	0.25	0.43	0.76	0.58	-0.10	0.10	0.51	-0.30	-0.45	-0.31	1.00	
Utilization	0.25	-0.01	0.69	0.08	-0.01	-0.21	-0.12	0.41	0.21	0.31	0.36	-0.01	1.00

Table 3: **Cross-Sectional Determinants of ETF Short Ratio**

This table reports the results from the Fama and MacBeth (1973) regression of the monthly ETF short interest ratio on ETF characteristics and constituent stocks' characteristics. ETF Return and ETF Return Vol are the 12-month mean and volatility of the ETF monthly return. ETF Buy-Sell Imbalance is maximum of zero and cumulative three-day ETF intraday buy-sell imbalance at month end. ETF monthly flow is the monthly change in ETF's shares outstanding divided by lagged shares outstanding. Stock-level characteristics are weighted-averaged within an ETF's quarterly holdings. Stock IVol is stock idiosyncratic volatilities per Ang, Hodrick, Xing, and Zhang (2006). Stock short ratio is the mid-month open short interests divided by shares outstanding. Short Supply is the total lendable shares from Markit divided by shares outstanding. Lending fee is a score from 1 to 10 created by Markit to capture the difficulty of shorting a stock. Illiquidity is the Amihud (2002) illiquidity measure. The standard errors are adjusted using Newey-West method. ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ETF Turnover	0.801***	0.808***	0.803***	0.800***	0.807***	0.799***	0.801***
	(14.24)	(14.23)	(14.32)	(14.28)	(14.34)	(14.30)	(14.41)
ETF Ln(ME)	-0.000982***	-0.000924**	-0.000996***	-0.000868**	-0.000980***	-0.000896***	-0.000968***
	(-2.92)	(-2.51)	(-3.02)	(-2.53)	(-2.92)	(-2.67)	(-2.97)
ETF Return	-0.0242***	-0.0259***	-0.0218**	-0.0202**	-0.0244***	-0.0226**	-0.0243**
	(-2.74)	(-2.73)	(-2.38)	(-2.15)	(-2.62)	(-2.32)	(-2.50)
ETF Return Vol	0.120*	0.0409	-0.0439	-0.0233	-0.0120	0.0119	0.0205
	(1.88)	(0.54)	(-0.68)	(-0.35)	(-0.19)	(0.18)	(0.31)
ETF Lending Fee Score	-0.00488***	-0.00493***	-0.00494***	-0.00506***	-0.00508***	-0.00501***	-0.00498***
	(-5.92)	(-6.15)	(-5.95)	(-5.95)	(-6.00)	(-6.03)	(-6.17)
ETF Buy-Sell Imbalance	0.207***	0.195***	0.192***	0.194***	0.206***	0.189***	0.202***
	(3.02)	(2.97)	(2.76)	(2.83)	(3.04)	(2.78)	(2.96)
ETF Monthly Flow	-0.0204***	-0.0211***	-0.0229***	-0.0235***	-0.0221***	-0.0223***	-0.0210***
	(-3.00)	(-3.21)	(-3.30)	(-3.47)	(-3.25)	(-3.29)	(-3.16)
Stock Return IVol		1.382***	2.221***	2.149***	1.707***	1.912***	1.518***
		(4.67)	(7.66)	(7.58)	(6.77)	(6.73)	(5.49)
Stock Illiquidity		1.067***					
		(4.59)					
Stock Short Ratio			-0.00844				
			(-0.14)				
Stock Lending Supply				-0.225***			
				(-3.64)			
Stock Lending Utilization					0.0526***		
					(4.53)		
Stock Lending Fee Score						0.0120***	0.00766***
						(5.79)	(3.75)
Top Decile Stock Lending Fee							0.00347***
							(5.01)
Stock Ln(ME)	-0.000362	0.00110**	0.00190***	0.000642	0.00296***	0.00117***	0.00114**
	(-0.95)	(2.59)	(3.48)	(1.51)	(5.11)	(2.62)	(2.51)
Stock BM	0.00479***	0.000463	0.00327***	0.00232**	0.00238**	0.00159	0.00115
	(4.48)	(0.40)	(3.21)	(2.32)	(2.41)	(1.56)	(1.08)
Observations	22167	22167	22167	22167	22167	22167	22167
R^2	0.399	0.431	0.434	0.435	0.429	0.441	0.452

Table 4: **ETF Portfolio Returns Sorted on ETF Short Ratio**

This table reports the monthly average returns, Fama and French (1993) 3-factor alpha, and Fama and French (1993) and Carhart (1997) 4-factor alpha (in percentage) for each of the five quintile portfolios formed by ETF funds, as well as the long-short portfolio (Low-High). At the end of each month, all ETF funds are sorted into quintiles based on their mid-month reported short ratio, and a long-short portfolio is formed by buying the lowest quintile and shorting the highest quintile portfolio. Portfolio returns are computed over the next month. Panel A reports results for equally weighted portfolios, and Panel B shows results for value-weighted portfolios. The sample runs from January 2002 to December 2019.

Panel A: Equal-weighted Quintile Portfolio Returns and Alphas

	Mean	t-stat	FF(93)	t-stat	Carhart(97)	t-stat
Low	0.69	2.27	-0.02	-0.52	-0.03	-0.55
2	0.69	2.29	-0.03	-0.65	-0.03	-0.67
3	0.71	2.28	-0.03	-0.63	-0.04	-0.82
4	0.68	2.08	-0.10	-1.99	-0.11	-2.16
High	0.64	1.81	-0.19	-3.24	-0.20	-3.46
Low - High	0.05	0.73	0.16	2.81	0.17	2.98

Panel B: Value-weighted Quintile Portfolio Returns and Alphas

	Mean	t-stat	FF(93)	t-stat	Carhart(97)	t-stat
Low	0.74	2.54	0.03	0.70	0.04	0.80
2	0.63	2.17	-0.08	-2.29	-0.08	-2.36
3	0.72	2.48	0.01	0.19	-0.00	-0.03
4	0.71	2.31	-0.03	-0.45	-0.04	-0.64
High	0.69	2.29	-0.05	-2.57	-0.05	-2.86
Low - High	0.05	1.09	0.08	1.84	0.09	2.07

Table 5: **Stock Portfolio Returns Sorted on ETF-based Short Ratio**

This table reports the monthly average excess returns (Exret), Fama and French (1993) 3-factor alpha, and Fama and French (1993) and Carhart (1997) 4-factor alpha (in percentage) for each of the 10 decile portfolios, as well as the long-short portfolio (Low-High). At the end of each month, all stocks are sorted into deciles based on their ETF-based stock short ratio (ETF_sr) and a long-short portfolio is formed by buying the lowest decile and shorting the highest decile portfolio. Portfolio returns are computed over the next month. Panel A reports results for equally weighted portfolios and Panel B shows results for value-weighted portfolios. The sample runs from January 2002 to December 2019.

Panel A: Equal-weighted Decile Portfolio Returns and Alphas

	Exret	t-stat	FF(93)	t-stat	Carhart(97)	t-stat
Low	1.03	2.23	0.19	2.15	0.21	2.57
2	0.95	2.23	0.12	0.69	0.17	1.06
3	0.73	1.84	-0.06	-0.40	-0.03	-0.19
4	0.67	1.80	-0.12	-1.19	-0.09	-1.01
5	0.85	2.36	0.07	0.79	0.09	1.22
6	0.68	1.46	-0.18	-1.55	-0.15	-1.41
7	0.93	2.01	0.09	1.03	0.12	1.48
8	0.91	2.22	0.16	0.81	0.21	1.13
9	0.87	1.91	0.08	0.83	0.10	0.95
High	0.26	0.60	-0.39	-2.29	-0.38	-2.24
Low - High	**0.77**	3.83	**0.58**	3.01	**0.59**	3.08

Panel B: Value-weighted Decile Portfolio Returns and Alphas

	Exret	t-stat	FF(93)	t-stat	Carhart(97)	t-stat
Low	0.97	2.39	0.15	1.52	0.13	1.39
2	0.94	2.29	0.01	0.05	0.03	0.19
3	0.77	2.06	-0.08	-0.72	-0.08	-0.70
4	0.75	2.26	-0.05	-0.62	-0.04	-0.58
5	0.77	2.66	0.06	1.23	0.06	1.20
6	0.62	1.61	-0.15	-1.11	-0.17	-1.26
7	0.93	2.22	0.10	1.32	0.10	1.23
8	0.78	2.31	0.07	0.56	0.09	0.72
9	0.74	1.78	0.01	0.08	0.00	-0.01
High	0.38	0.93	-0.26	-1.40	-0.28	-1.53
Low - High	**0.59**	2.87	**0.41**	2.07	**0.41**	2.06

Table 6: **Two-way sorts on Short-Sale Constraints and ETF-based Short Ratio**

This table reports monthly Carhart (1997) 4-factor alphas (in percentages) sorted on proxies of short-sales constraints and ETF-based stock short ratios (ETF_sr). At the end of each month, all the stocks are sorted into terciles based on a proxy for short-sale constraints (except for lending fee), and within each tercile the stocks are further sorted into quintiles based on their ETF-based short ratios. For lending fee measure, we sort stocks into two groups based on whether a stock's DCBS score is above 2. Returns are equally weighted within each portfolio. We use lendable supply, institutional ownership, lending fee, idiosyncratic volatility, and Amihud (2002) illiquidity as proxies for short-sale constraints in Panels A, B, C, D, and E, respectively. In the bottom row of each panel, we report the difference of abnormal return based on ETF-based SR in high and low short-sale constrained groups. The overall sample runs from January 2002 to December 2019. The lending fee and lendable supply sample is from January 2004 to December 2019.

Panel A: Lendable Supply and ETF-based SR

	ETF-based Stock Short ratio					
Lendable Supply	Low	2	3	4	High	Low-High
Low	0.36	0.12	-0.09	-0.10	-0.32	0.68
	(1.73)	(0.59)	(-0.46)	(-0.68)	(-1.90)	(2.32)
2	0.21	0.04	0.17	0.12	-0.05	0.26
	(1.59)	(0.56)	(1.71)	(1.24)	(-0.37)	(1.09)
High	-0.03	0.10	-0.19	0.14	0.09	-0.12
	(-0.30)	(1.00)	(-1.75)	(1.70)	(0.71)	(-0.76)
Low Supply Sample -						0.80
High Supply Sample						(3.67)

Panel B: Inst.Ownership and ETF-based SR

	ETF-based Stock Short ratio					
Inst. Ownership	Low	2	3	4	High	Low-High
Low	0.29	0.17	-0.20	-0.11	-0.53	0.83
	(1.37)	(0.88)	(-1.09)	(-1.02)	(-2.90)	(2.57)
2	0.19	0.10	0.32	0.14	0.14	0.05
	(1.33)	(1.34)	(2.84)	(1.38)	(1.02)	(0.26)
High	-0.04	0.05	-0.17	0.07	0.14	-0.17
	(-0.31)	(0.45)	(-1.52)	(0.90)	(1.64)	(-1.19)
Low IO Sample -						1.00
High IO Sample						(3.78)

Panel C: Lending fee and ETF-based SR

	ETF-based Stock Short ratio					
Lending Fee	Low	2	3	4	5	Low-High
Low	0.04	0.11	0.36	0.20	0.13	-0.09
	(0.53)	(1.53)	(2.72)	(2.66)	(1.08)	(-0.68)
High	-0.38	-0.79	-0.32	-1.31	-1.76	1.38
	(-1.22)	(-2.84)	(-1.01)	(-5.61)	(-6.53)	(3.63)
High Fee Sample -						1.46
Low Fee Sample						(4.11)

Table 6 Continued

Panel D: Idiosyncratic Vol. and ETF-based SR

	ETF-based Stock Short ratio					
Idiosyncratic Vol	Low	2	3	4	High	Low-High
Low	0.36	0.21	0.16	0.18	0.31	0.04
	(4.36)	(3.77)	(2.88)	(2.27)	(2.37)	(0.25)
2	0.21	0.09	-0.14	0.22	0.02	0.19
	(1.51)	(0.78)	(-1.35)	(2.92)	(0.15)	(0.85)
High	0.12	-0.25	-0.22	-0.06	-0.71	0.83
	(0.44)	(-1.05)	(-1.22)	(-0.41)	(-3.74)	(2.53)
High IVOL Sample -						0.79
Low IVOL Sample						(3.00)

Panel E: Amihud Illiquidity and ETF-based SR

	ETF-based Stock Short ratio					
Illiquidity	Low	2	3	4	5	Low-High
Low	-0.01	0.04	0.11	0.09	-0.10	0.10
	(-0.07)	(0.60)	(1.76)	(1.69)	(-1.43)	(0.88)
2	0.00	-0.06	0.25	0.19	-0.26	0.26
	(-0.04)	(-0.54)	(3.09)	(2.10)	(-1.59)	(1.05)
High	0.38	0.22	-0.11	0.07	-0.26	0.64
	(1.63)	(0.96)	(-0.46)	(0.43)	(-1.39)	(1.94)
High Illiquidity Sample -						0.54
Low Illiquidity Sample						(1.78)

Table 7: **Fama-MacBeth Regression: Baseline**

This table reports the results from the Fama and MacBeth (1973) regression of monthly stock returns on ETF-based short ratio (ETF_sr). Size (LnME) is the natural log of a firm's market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of the book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative returns from month t-12 to t-2. The short term reversal measure (REV) is the lagged monthly return. Institutional ownership (IO) is the sum of shares held by institutions from 13F filings in each quarter divided by the total shares outstanding. IVOL is the idiosyncratic volatility, calculated following Ang, Hodrick, Xing, and Zhang (2006). ETF-based flow is calculated as ETF monthly flows aggregated to stock level. ETF_sr2 is the dollar value of short interests on the stock via ETFs divided by the total value of a stock held by all the ETFs. All t-statistics are Newey-West adjusted with four lags to control for heteroskedasticity and autocorrelation. ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively. The sample runs from January 2002 to December 2019.

	(1)	(2)	(3)	(4)
ETF_sr	-0.3452***	-0.3452**	-0.4556***	
	(-2.92)	(-2.46)	(-2.63)	
ETF_sr2				-0.0145**
				(-2.12)
LnME		-0.0005	-0.0009*	-0.0008*
		(-0.84)	(-1.80)	(-1.74)
LnBM		-0.0001	-0.0003	-0.0002
		(-0.12)	(-0.42)	(-0.39)
REV		-0.0116***	-0.0102***	-0.0103***
		(-3.47)	(-2.77)	(-3.34)
MOM		-0.0012	-0.0012	-0.0011
		(-0.38)	(-0.39)	(-0.39)
IO			0.0073**	0.0059***
			(2.35)	(2.60)
IVOL			-0.1005**	-0.0959***
			(-2.53)	(-3.11)
ETF-based flow			-1.1876***	-1.2657***
			(-2.61)	(-3.66)
Constant	0.0069	0.0073	0.0130**	0.0125**
	(1.52)	(1.01)	(2.06)	(2.08)
Ave.R-sq	0.010	0.037	0.044	0.044
N.of Obs.	557245	536070	533740	533740

Table 8: **Fama-MacBeth Regression: Controlling for Stock-level Shorting Variables**

This table reports the results from the Fama and MacBeth (1973) regression of monthly stock returns on ETF-based short ratios (ETF_sr) and controlling for stock-level shorting variables. Size (LnME) is the natural log of a firm's market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of the book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative returns from month t-12 to t-2. The short term reversal measure (REV) is the lagged monthly return. Institutional ownership (IO) is the sum of shares held by institutions from 13F filings in each quarter divided by the total shares outstanding. IVOL is the idiosyncratic volatility, calculated following Ang, Hodrick, Xing, and Zhang (2006). ETF-based flow is ETF monthly flows aggregated to stock level. Short interest ratio (SR) is the number of shares shorted over total shares outstanding. DCBS is a score from 1 to 10 created by Markit using their proprietary information and is intended to capture the cost of borrowing the stock. SIO is the short interest ratio (SR) divided by institutional ownership. Lendable shares (supply) is the shares held and made available to lend by Markit lenders divided by total shares outstanding. Utilization is the ratio of shares borrowed to shares made available by Markit lenders. All t-statistics are Newey-West adjusted with four lags to control for heteroskedasticity and autocorrelation. ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively. The sample runs from January 2002 to December 2019.

	(1)	(2)	(3)	(4)	(5)	(6)
ETF_sr	-0.3709**	-0.5877***	-0.3795**	-0.4131**	-0.3804**	-0.5551***
	(-2.43)	(-3.52)	(-2.10)	(-2.40)	(-2.15)	(-3.70)
LnME	-0.0010**	-0.0013***	-0.0010**	-0.0009*	-0.0009*	-0.0013***
	(-1.97)	(-2.62)	(-2.07)	(-1.76)	(-1.67)	(-2.91)
LnBM	-0.0005	-0.0011	-0.0009	-0.0005	-0.0003	-0.0012*
	(-0.70)	(-1.37)	(-1.23)	(-0.73)	(-0.45)	(-1.79)
REV	-0.0111***	-0.0135***	-0.0110***	-0.0108***	-0.0107***	-0.0143***
	(-3.00)	(-3.48)	(-3.01)	(-2.93)	(-2.91)	(-4.10)
MOM	-0.0013	-0.0018	-0.0017	-0.0012	-0.0017	-0.0020
	(-0.43)	(-0.57)	(-0.54)	(-0.39)	(-0.55)	(-0.72)
IO	0.0093***	0.0039	0.0070**	0.0065**	0.0082**	0.0056**
	(3.00)	(1.32)	(2.22)	(2.03)	(2.20)	(2.18)
IVOL	-0.0854**	-0.0474	-0.0381	-0.0857**	-0.1104***	-0.0483*
	(-2.24)	(-1.20)	(-0.89)	(-2.21)	(-2.76)	(-1.71)
ETF-based flow	-1.0429**	-0.9404*	-1.1052**	-1.1562**	-1.0449**	-0.7382*
	(-2.04)	(-1.85)	(-2.45)	(-2.53)	(-2.33)	(-1.66)
SR	-0.0300***					-0.0213***
	(-3.67)					(-2.89)
DCBS		-0.0031***				-0.0027***
		(-7.63)				(-4.95)
Utilization			-0.0133***			
			(-3.46)			
SIO				-0.0073***		
				(-3.89)		
Supply					0.0008	
					(0.09)	
Constant	0.0127**	0.0210***	0.0141**	0.0134**	0.0133**	0.0204***
	(2.01)	(3.47)	(2.32)	(2.13)	(2.07)	(3.46)
Ave.R-sq	0.046	0.048	0.049	0.046	0.045	0.050
N.of Obs.	533740	520538	514826	533738	528500	520538

Table 9: **Fama-MacBeth Regression: Interaction between ETF-based Short Ratio and Short-Sale Constraint Measures**

This table reports the results from the Fama and MacBeth (1973) regression of monthly stock returns on ETF-based stock short ratios (ETF_sr) and its interaction with several variables that indicate tightening short-sale constraints. LowIO (Lowsupply) is a dummy equal to one when the stock is in the bottom quintile of institutional ownership ratio (lendable supply) in the cross-section. Highutil is a dummy equal to one when the stock is in the top quintile of utilization in the cross-section. Highfee is a dummy equal to one when the DCBS score is larger than 2. Lowturn is a dummy equal to one when the stock is in the bottom quintile of past 12-month average turnover in the cross-section. Noput is a dummy equal to one when the stock has no exchange-traded put option in the month. Other control variables are the same as in the previous tables. All t-statistics are Newey-West adjusted with four lags to control for heteroskedasticity and autocorrelation. ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively. The sample runs from January 2002 to December 2019.

	(1)	(2)	(3)	(4)	(5)	(6)
ETF_sr	-0.3193*	-0.4182**	-0.4641***	-0.2778	-0.3416**	-0.2243
	(-1.72)	(-2.24)	(-2.91)	(-0.37)	(-2.02)	(-1.34)
ETFsr_lowIO	-0.8245***					
	(-4.78)					
LowIO	-0.0055***					
	(-2.61)					
ETFsr_lowsupply		-0.8481*				
		(-1.85)				
Lowsupply		-0.0020				
		(-1.14)				
ETFsr_highutil			-0.1525			
			(-0.75)			
Highutil			-0.0071**			
			(-2.36)			
ETFsr_highfee				-0.6273***		
				(-2.64)		
Highfee				-0.0193		
				(-1.64)		
ETFsr_lowturn					-0.3348*	
					(-1.75)	
Lowturn					0.0012	
					(0.47)	
ETFsr_Noput						-0.3683**
						(-2.13)
Noput						0.0018
						(0.45)
SR	-0.0259***	-0.0302***	-0.0284***	-0.0186**	-0.0302***	-0.0292***
	(-3.19)	(-3.74)	(-3.08)	(-2.31)	(-3.74)	(-2.97)
LnME	-0.0009*	-0.0012*	-0.0010*	-0.0032	-0.0010*	-0.0007
	(-1.66)	(-1.95)	(-1.83)	(-1.46)	(-1.76)	(-1.11)
LnBM	-0.0005	-0.0005	-0.0003	-0.0018	-0.0006	-0.0000
	(-0.69)	(-0.74)	(-0.47)	(-1.49)	(-0.90)	(-0.00)
REV	-0.0112***	-0.0116***	-0.0141***	-0.0129***	-0.0113***	-0.0111***
	(-2.87)	(-2.90)	(-3.59)	(-3.06)	(-3.61)	(-3.39)
MOM	-0.0019	-0.0020	-0.0019	-0.0040	-0.0011	-0.0015
	(-0.59)	(-0.61)	(-0.57)	(-0.97)	(-0.44)	(-0.56)
IO		0.0067**	0.0072*	-0.0090	0.0095***	0.0095***
		(2.21)	(1.94)	(-0.61)	(3.81)	(3.60)
IVOL	-0.1029**	-0.1123**	-0.0869**	-0.2343	-0.0792***	-0.0796**
	(-2.54)	(-2.49)	(-2.13)	(-1.48)	(-2.72)	(-2.45)
ETF-based flow	-0.9168**	-1.0174**	-0.7130	-1.7340***	-1.0030***	-0.9912***
	(-2.09)	(-2.39)	(-1.17)	(-2.73)	(-3.04)	(-2.95)
Constant	0.0191**	0.0171**	0.0164**	0.0515	0.0119*	0.0106*
	(2.51)	(2.16)	(2.18)	(1.42)	(1.88)	(1.66)
Ave.R-sq	0.047	0.049	0.049	0.049	0.050	0.050
N.of Obs.	533740	533740	533740	533740	533740	533740

Table 10: **Fama-MacBeth Regression: Predicting Earnings Surprises**

This table reports the results of the Fama and MacBeth (1973) regression of earnings surprises on ETF-based short ratio (ETF_sr). In Columns (1)-(4), earnings surprise is measured as three-day cumulative abnormal returns around quarterly earnings announcement. In Columns (1) and (2), abnormal return is calculated as daily stock return minus return on the CRSP value-weighted portfolio return. In Columns (3) and (4), abnormal return is calculated as daily stock return minus the return on the characteristics matched portfolio following Daniel, Grinblatt, Titman, and Wermers (1997). In Columns (5) and (6), earnings surprises is measured as standardized unexpected earnings (SUE), which is computed as change in split-adjusted quarterly earnings per share from its value four quarters ago divided by stock price one month before earnings announcements. Short interest ratio (SR) is the number of shares shorted over total shares outstanding. Lagged CAR is the three-day cumulative abnormal return around previous quarter's earnings announcement. Lagged SUE is the standardized unexpected earnings from previous quarter. Size (LnME) is the natural log of a firm's market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of the book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative returns from month t-12 to t-2. The short term reversal measure (REV) is the lagged monthly return. Institutional ownership (IO) is the sum of shares held by institutions from 13F filings in each quarter divided by the total shares outstanding. IVOL is the idiosyncratic volatility, calculated following Ang, Hodrick, Xing, and Zhang (2006). ETF-based flow is ETF monthly flows aggregated to stock level. All t-statistics are Newey-West adjusted with four lags to control for heteroskedasticity and autocorrelation. ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively. The sample runs from the 1st quarter of 2002 to 4th quarter of 2019.

	Dep.Var= Market-adjusted CAR(-1, +1)		Dep.Var= DGTW-adjusted CAR(-1, +1)		Dep.Var= SUE	
ETF_sr	-0.2185***	-0.2139**	-0.1527**	-0.1666*	-1.0215***	-0.7986**
	(-2.96)	(-2.11)	(-2.54)	(-1.74)	(-3.50)	(-2.64)
SR		-0.0577**		-0.0461**		-0.0171
		(-2.03)		(-2.10)		(-0.39)
Lagged CAR	-0.0551**	-0.0307***	-0.0480**	-0.0277***		
	(-2.02)	(-3.22)	(-2.15)	(-3.83)		
Lagged SUE					0.1519	0.0696
					(1.44)	(1.56)
LnME	0.0004	0.0006	0.0002	0.0005	-0.0001	0.0000
	(0.44)	(0.61)	(0.30)	(0.55)	(-0.04)	(0.01)
LnBM	0.0031	0.0032	0.0029	0.0032	-0.0089**	-0.0089**
	(1.40)	(1.32)	(1.59)	(1.45)	(-2.47)	(-2.47)
REV	-0.0100	0.0181***	-0.0027	0.0191***	0.0172	0.0037
	(-0.46)	(3.38)	(-0.16)	(3.48)	(0.64)	(0.10)
MOM	-0.0052	0.0065	-0.0035	0.0060	0.0183	0.0204*
	(-1.07)	(1.14)	(-1.11)	(1.04)	(1.41)	(1.75)
IO	0.0052*	0.0091***	0.0053**	0.0076***	-0.0183***	-0.0153***
	(1.71)	(6.40)	(2.45)	(4.74)	(-3.08)	(-3.21)
IVOL	-0.6922	-0.1813***	-0.5731	-0.1768***	0.0323	0.0458
	(-1.55)	(-6.38)	(-1.63)	(-6.11)	(0.16)	(0.23)
ETF-based flow	-0.6024	-0.5679	-0.5513	-0.4926	-0.2620	-0.2530
	(-0.83)	(-0.76)	(-0.87)	(-0.76)	(-0.43)	(-0.44)
Constant	0.0120**	0.0022	0.0103***	0.0028	0.0180	0.0146
	(2.37)	(0.53)	(3.00)	(0.85)	(1.34)	(1.25)
Ave.R-sq	0.071	0.074	0.069	0.073	0.166	0.168
N.of Obs.	142435	142435	141103	141103	139072	139072

Appendix

Table A.1: **ETF-based Short Ratio and Hedge Fund Ownership**

This table reports the results from the Fama-MacBeth (1973) regression of ETF-based Short Ratio on stock characteristics. The dependent variable is ETF-based short ratio, calculated as the dollar value of stock shorted via its holding ETFs divided by the stock's market capitalization. HFO is the quarterly hedge fund ownership, defined as the sum of shares held by all hedge funds reported at each quarter divided by the total number of shares outstanding. Size (LnME) is the natural log of firm's market capitalization at the end of the June of each year. Book-to-market (LnBM) is the natural log of the book-to-market ratio. The cases with negative book value are deleted. Momentum (MOM) is defined as the cumulative returns from month t-12 to t-2. IO is the institutional ownership ratio. IVOL is the idiosyncratic volatility following Ang, Hodrick, Xing, and Zhang (2006). Turnover is the monthly trading volume over shares outstanding, averaged over past 12 months. All t-statistics are Newey-West adjusted with four lags to control for heteroskedasticity and autocorrelation. ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively. The sample runs from the 1st quarter of 2002 to 4th quarter of 2019.

	(1) Dep.Var=ETF_sr
HFO	-0.0057***
	(-11.13)
LnME	-0.0010***
	(-14.12)
LnBM	-0.0005***
	(-11.07)
MOM	0.0005***
	(5.45)
IO	0.0094***
	(18.93)
IVOL	-0.0371***
	(-11.66)
Turnover	0.0010***
	(6.47)
Constant	0.0064***
	(14.55)
Ave.R-sq	0.266
N.of Obs.	190634

Table A.2: **Factor Loadings of Portfolios Sorted on ETF-based Short Ratio**

This table reports the factor loadings of the two extreme decile portfolios sorted on ETF-based short ratio. MKTRF, SMB, HML, and UMD stand for the market factor, size factor, value factor, and the momentum factor, respectively. Columns (1)-(3) report results for the equal-weighted portfolios, and Columns (4)-(6) report results for the value-weighted portfolios. The sample runs from January 2002 to December 2019.

	Equal-weighted Portfolio			Value-weighted Portfolio		
	(1) Low	(2) High	(3) Low-High	(4) Low	(5) High	(6) Low-High
Alpha	0.0021**	-0.0038**	0.0059***	0.0013	-0.0028	0.0041**
	(2.57)	(-2.24)	(3.08)	(1.39)	(-1.53)	(2.06)
Mktrf	1.0680***	0.8665***	0.2015***	1.0808***	0.8830***	0.1978***
	(48.07)	(18.51)	(3.83)	(43.36)	(17.40)	(3.64)
SMB	1.0524***	1.0182***	0.0342	0.8359***	0.8442***	-0.0083
	(27.58)	(12.66)	(0.38)	(19.53)	(9.69)	(-0.09)
HML	0.2129***	0.4288***	-0.2159***	0.1667***	0.4036***	-0.2368***
	(6.13)	(5.86)	(-2.62)	(4.28)	(5.09)	(-2.79)
UMD	-0.1083***	-0.0415	-0.0668	0.1010***	0.1068**	-0.0058
	(-5.28)	(-0.96)	(-1.37)	(4.39)	(2.28)	(-0.11)
Adj.R-sq	0.971	0.856	0.133	0.956	0.808	0.108
N.of Obs.	216	216	216	216	216	216

Table A.3: **Portfolio Returns Sorted on Residual ETF-based Short Ratio**

This table reports the monthly Carhart (1997) 4-factor alpha (in percentage) for each of the 10 decile portfolios, as well as the long-short portfolio (Low-High). At the end of each month, all stocks are sorted into deciles based on their residual ETF-based stock short ratio (residual ETF_sr) and a long-short portfolio is formed by buying the lowest decile and shorting the highest decile portfolio. Portfolio returns are computed over the next month. Residual ETF-based short ratio is the residual from cross-sectional regression of ETF-based short ratio on ETF-based flow. ETF-based flow is calculated as the monthly ETF flows aggregated to stock level. The sample runs from January 2002 to December 2019.

	EW	t-stat	VW	t-stat
Low	0.18	1.13	0.12	0.69
2	0.03	0.22	0.13	0.92
3	0.15	1.27	0.10	0.68
4	0.19	1.69	0.06	0.51
5	-0.03	-0.26	-0.05	-0.52
6	0.16	1.35	0.17	1.91
7	0.02	0.14	0.11	0.92
8	0.03	0.21	0.05	0.60
9	-0.26	-1.80	-0.10	-0.83
High	-0.33	-1.59	-0.35	-2.03
Low - High	**0.51**	2.61	**0.47**	2.03

Table A.4: **Robustness of Portfolio Sorts**

This table reports robustness tests for decile portfolio sorts based on ETF-based short ratio. In the first set of robustness tests, we report the Carhart (1997) 4-factor alpha of gross return-weighted portfolio returns in which the weights are 1 + the stock's lagged monthly return, following Asparouhova, Bessembinder, and Kalcheva (2013). The second set of robustness tests shows Carhart (1997) alpha using Fama-French 48 industry-adjusted return. The third row shows the alpha using Pástor and Stambaugh (2003) liquidity factor augmented with the Fama-French factors and the momentum factor. In the fourth set of tests, we report the alphas using the Fama and French (2016) Five Factor model. In the fifth, sixth and seventh set of tests, we report the alphas using the Stambaugh and Yuan (2016) Mispricing Factors model, the Hou, Xue, and Zhang (2015) Q-factor model and the Daniel, Hirshleifer, and Sun (2020) short- and long-horizon behavioral factors. In the eighth and ninth set of analyses, we exclude stocks with a price lower than $5 and stocks whose market cap is below the size treshold of NYSE bottom decile. In the tenth row, we skip a month between the moment at which ETF-based SR is constructed and the moment at which we start measuring returns. In the eleventh row, we first regress ETF-based SR on stock's own SR and form decile portfolios based on the residual ETF-based SR. In the last row, we conduct quintile portfolio sorts based on ETF-based SR and report the four-factor alpha. Column (1) reports the results for equal-weighted portfolio, and Column (2) reports for the value-weighted portfolio. The sample runs from January 2002 to December 2019.

	EW	VW
Gross return-weighed portfolio	0.633 (3.30)	N/A
FF48 Industry-adjusted	0.572 (2.03)	0.361 (1.72)
FF + Cahart + PS Factor	0.589 (3.14)	0.405 (2.10)
FF5 factor (Fama and French, 2015)	0.460 (2.34)	0.313 (1.85)
Mispricing factors (Stambaugh and Yuan, 2017)	0.584 (2.51)	0.409 (1.97)
Q-factor (Hou, Xue, and Zhang, 2015)	0.634 (3.07)	0.439 (2.08)
Behavioral Factors (Daniel, Hirshleifer, and Sun, 2020)	0.688 (3.46)	0.521 (2.54)
Exclude Price<=$5	0.371 (2.42)	0.311 (1.63)
Exclude Microcap Stocks	0.485 (2.54)	0.431 (1.88)
Skip a month	0.443 (2.58)	0.362 (1.96)
ETF-based SR after removing stock's own SR	0.594 (2.86)	0.411 (1.99)
Quintile Portfolio Sorts	0.431 (3.82)	0.337 (2.65)

Table A.5: **Two-way sorts on Short-Sale Constraints and ETF-based Short Ratio**

This table reports monthly Carhart (1997) 4-factor alphas (in percentages) sorted on proxies of short-sales constraints and ETF-based stock short ratios (ETF_sr). At the end of each month, all the stocks are sorted into terciles based on a proxy for short-sale constraints (except for lending fee), and within each tercile the stocks are further sorted into quintiles based on their ETF-based short ratios. For lending fee measure, we sort stocks into two groups based on whether a stock's DCBS score is above 2. Returns are value weighted within each portfolio. We use lendable supply, institutional ownership, lending fee, idiosyncratic volatility, and Amihud (2002) illiquidity as proxies for short-sale constraints in Panels A, B, C, D, and E, respectively. In the bottom row of each panel, we report the difference of abnormal return based on ETF-based SR in high and low short-sale constrained groups. The overall sample runs from January 2002 to December 2019. The lending fee and lendable supply sample is from January 2004 to December 2019.

Panel A: Lendable Supply and ETF-based SR

	ETF-based Stock Short ratio					
Lendable Supply	Low	2	3	4	High	Low-High
Low	0.15	-0.12	-0.39	0.10	-0.29	0.44
	(0.90)	(-0.48)	(-1.84)	(0.48)	(-1.77)	(1.99)
2	0.14	0.01	0.08	0.17	-0.20	0.34
	(1.10)	(0.21)	(0.64)	(1.50)	(-1.14)	(1.39)
High	-0.11	0.08	-0.26	0.13	0.04	-0.15
	(-0.95)	(0.84)	(-2.03)	(1.67)	(0.33)	(-0.93)
Low Supply Sample -						0.59
High Supply Sample						(2.76)

Panel B: Inst.Ownership and ETF-based SR

	ETF-based Stock Short ratio					
Inst. Ownership	Low	2	3	4	High	Low-High
Low	0.18	-0.05	-0.24	0.01	-0.41	0.59
	(0.89)	(-0.19)	(-1.28)	(0.07)	(-2.22)	(2.22)
2	-0.08	0.06	-0.06	0.06	-0.05	-0.02
	(-0.64)	(1.03)	(-0.40)	(0.65)	(-0.33)	(-0.11)
High	-0.05	-0.05	-0.14	0.05	0.10	-0.15
	(-0.46)	(-0.46)	(-0.97)	(0.74)	(1.13)	(-1.03)
Low IO Sample -						0.74
High IO Sample						(2.82)

Panel C: Lending fee and ETF-based SR

	ETF-based Stock Short ratio					
Lending Fee	Low	2	3	4	5	Low-High
Low	-0.10	0.01	0.07	0.13	0.00	-0.10
	(-1.42)	(0.24)	(0.72)	(1.66)	(0.00)	(-0.73)
High	-0.29	-1.08	-0.93	-0.87	-1.51	1.22
	(-0.65)	(-2.59)	(-2.18)	(-3.30)	(-4.62)	(2.50)
High Fee Sample -						1.33
Low Fee Sample						(2.50)

Table A.5 Continued

Panel D: Idiosyncratic Vol. and ETF-based SR

	ETF-based Stock Short ratio					
Idiosyncratic Vol	Low	2	3	4	High	Low-High
Low	-0.16	-0.02	-0.28	0.10	-0.08	-0.08
	(-0.95)	(-0.13)	(-1.57)	(1.14)	(-0.45)	(-0.31)
2	0.25	0.11	0.08	0.14	0.21	0.04
	(2.89)	(1.83)	(1.33)	(1.23)	(1.40)	(0.26)
High	0.05	-0.61	-0.12	-0.02	-0.56	0.61
	(0.15)	(-1.89)	(-0.59)	(-0.12)	(-2.76)	(1.78)
High IVOL Sample -						0.69
Low IVOL Sample						(1.72)

Panel E: Amihud Illiquidity and ETF-based SR

	ETF-based Stock Short ratio					
Illiquidity	Low	2	3	4	5	Low-High
Low	0.07	-0.07	0.02	0.07	-0.16	0.23
	(0.83)	(-0.92)	(0.32)	(1.30)	(-1.27)	(1.50)
2	0.05	-0.03	0.22	0.19	-0.17	0.22
	(0.44)	(-0.32)	(2.74)	(2.33)	(-0.99)	(0.95)
High	0.28	0.37	-0.08	0.07	-0.22	0.50
	(1.38)	(1.76)	(-0.34)	(0.46)	(-1.18)	(1.57)
High Illiquidity Sample -						0.27
Low Illiquidity Sample						(0.72)

Table A.6: **Fama-MacBeth Regression: Additional Analysis**

This table reports the results from the Fama and MacBeth (1973) regression of monthly stock returns on ETF-based stock short ratios (ETF_sr). In Column (1), we interact ETF-based SR with a highweight dummy. The highweight dummy equals to one for stocks in the top quintile of weight with the ETF, where weight is defined as the average weight a stock has across all its holding ETFs. In column (2) and (3), the sample period is from 2002 to 2010, and from 2011 to 2019, respectively. The dependent variable is the Fama-French 48 industry-adjusted and characteristic-adjusted return (Daniel, Grinblatt, Titman, and Wermers (1997)) in columns (4) and (5), respectively. In Column (6), the dependent variable is the stock-level monthly returns adjusted by the weighted average returns of ETFs holding the stock. The control variables are the same as in the previous tables. All t-statistics are Newey-West adjusted with four lags to control for heteroskedasticity and autocorrelation. ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively. The sample runs from January 2002 to December 2019.

	Interaction with stock weight in ETFs	Subperiod		Dep.Var = Industry-adjusted stock return	Dep.Var = DGTW-adjusted stock return	Dep.Var = ETF-adjusted stock return
		2002-2010	2011-2019			
	(1)	(2)	(3)	(4)	(5)	(6)
ETF_sr	-0.3710	-0.6384**	-0.3417**	-0.3060**	-0.4064***	-0.3904**
	(-0.91)	(-2.21)	(-2.34)	(-2.11)	(-2.63)	(-2.20)
ETFsr_Highweight	-0.3930**					
	(-2.01)					
Highweight	0.0012					
	(0.60)					
LnME	-0.0008*	-0.0009*	-0.0009	-0.0001	-0.0010***	-0.0007
	(-1.93)	(-1.73)	(-0.98)	(-0.25)	(-2.95)	(-1.60)
LnBM	-0.0004	-0.0003	-0.0003	0.0005	-0.0002	-0.0001
	(-0.59)	(-0.30)	(-0.29)	(0.62)	(-0.44)	(-0.14)
REV	-0.0107***	-0.0107**	-0.0096	-0.0091***	-0.0110***	-0.0099***
	(-2.76)	(-2.53)	(-1.43)	(-2.77)	(-3.09)	(-2.65)
MOM	-0.0020	0.0009	-0.0040	-0.0024	-0.0015	-0.0021
	(-0.61)	(0.38)	(-0.65)	(-0.89)	(-0.70)	(-0.66)
IO	0.0067**	0.0059***	0.0093	0.0061	0.0060**	0.0065*
	(2.08)	(4.63)	(1.35)	(1.55)	(2.06)	(1.94)
IVOL	-0.1104***	-0.0969*	-0.1053	-0.1134**	-0.1154***	-0.1455***
	(-2.71)	(-1.97)	(-1.54)	(-2.23)	(-3.01)	(-3.20)
ETF-based flow	-1.3201***	-0.8884*	-1.5892**	-0.9994***	-0.5165	-0.6434
	(-3.12)	(-1.81)	(-2.29)	(-3.05)	(-1.15)	(-1.46)
Constant	0.0128**	0.0166***	0.0083	-0.0019	0.0047	0.0031
	(2.11)	(2.80)	(0.72)	(-0.32)	(1.53)	(0.96)
Ave.R-sq	0.045	0.046	0.041	0.039	0.029	0.041
N.of Obs.	533740	284650	249090	492233	492233	492220

Table A.7: **Fama-MacBeth Regression: 1- to 12-months ahead return predictability**

This table reports the results from the Fama and MacBeth (1973) regression of monthly stock returns on ETF-based stock short ratios (ETF_sr). The dependent variables are monthly stock returns from 1 month to 12 months ahead. The control variables are the same as in the previous tables. All t-statistics are Newey-West adjusted with four lags to control for heteroskedasticity and autocorrelation. ***, **, and * represent significance levels of 1%, 5%, and 10%, respectively. The sample runs from January 2002 to December 2019.

	(1) 1m	(2) 2m	(3) 3m	(4) 4m	(5) 5m	(6) 6m	(7) 7m	(8) 8m	(9) 9m	(10) 10m	(11) 11m	(12) 12m
ETF_sr	-0.4556***	-0.4291***	-0.2938***	-0.3727***	-0.3300***	-0.2461***	-0.2551**	-0.2445**	-0.1761	-0.1834	-0.1138	-0.0514
	(-2.63)	(-2.57)	(-4.20)	(-4.13)	(-4.47)	(-2.96)	(-2.44)	(-2.04)	(-1.64)	(-1.52)	(-1.25)	(-0.49)
LnME	-0.0009*	-0.0009**	-0.0004	-0.0012	-0.0003	-0.0005	-0.0004	-0.0007*	0.0002	-0.0008*	-0.0000	-0.0005
	(-1.80)	(-2.16)	(-0.79)	(-1.59)	(-0.73)	(-1.17)	(-1.16)	(-1.90)	(0.19)	(-1.79)	(-0.03)	(-1.17)
LnBM	-0.0003	-0.0016	-0.0008	-0.0018	-0.0009	-0.0016*	-0.0013	-0.0022*	-0.0011	-0.0014	-0.0007	-0.0005
	(-0.42)	(-1.29)	(-1.12)	(-1.59)	(-1.11)	(-1.76)	(-1.51)	(-1.78)	(-1.35)	(-1.57)	(-0.69)	(-0.57)
REV	-0.0102***	0.0003	0.0082*	-0.0018	0.0087	-0.0027	-0.0013	-0.0031	-0.0044	-0.0063	0.0092	0.0016
	(-2.77)	(0.05)	(1.91)	(-0.38)	(1.53)	(-0.46)	(-0.29)	(-0.55)	(-0.77)	(-0.95)	(1.50)	(0.37)
MOM	-0.0012	-0.0030	0.0002	-0.0012	-0.0026	-0.0024	-0.0029	-0.0022	-0.0033	-0.0030	-0.0012	-0.0023
	(-0.39)	(-0.80)	(0.05)	(-0.33)	(-0.80)	(-0.88)	(-1.33)	(-1.13)	(-1.03)	(-1.03)	(-0.83)	(-1.08)
IO	0.0073**	0.0057	0.0079***	0.0049	0.0077***	0.0048	0.0069**	0.0003	0.0037	0.0036	0.0072**	0.0054*
	(2.35)	(1.56)	(2.77)	(1.34)	(2.82)	(1.47)	(2.60)	(0.04)	(0.87)	(0.80)	(2.30)	(1.70)
IVOL	-0.1005**	-0.1688**	-0.0205	-0.1181*	-0.0906*	-0.1507*	0.0414	-0.0737	0.0568	-0.0745	0.0509	-0.0335
	(-2.53)	(-2.60)	(-0.32)	(-1.86)	(-1.86)	(-1.74)	(0.58)	(-0.88)	(0.92)	(-0.92)	(0.89)	(-0.71)
ETF-based flow	-1.1876***	-0.0122	0.1929	0.0481	0.5705	0.1161	-0.2109	0.0097	1.3181	0.5086	-2.4026***	-0.7167
	(-2.61)	(-0.03)	(0.46)	(0.06)	(0.88)	(0.22)	(-0.35)	(0.02)	(1.57)	(0.56)	(-2.79)	(-1.10)
Constant	0.0130**	0.0155***	0.0038	0.0124	0.0048	0.0096	0.0069	0.0147**	0.0037	0.0131**	0.0038	0.0088*
	(2.06)	(2.70)	(0.56)	(1.49)	(0.86)	(1.58)	(1.21)	(2.03)	(0.49)	(2.08)	(0.59)	(1.81)
Ave.R-sq	0.044	0.045	0.044	0.042	0.040	0.040	0.034	0.037	0.035	0.035	0.034	0.038
N.of Obs.	533740	531143	528481	525758	523018	520291	517532	514761	511980	509222	506458	503713