The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXIII, NO. 6 • DECEMBER 2018

Do ETFs Increase Volatility?

ITZHAK BEN-DAVID, FRANCESCO FRANZONI, and RABIH MOUSSAWI*

ABSTRACT

Due to their low trading costs, exchange-traded funds (ETFs) are a potential catalyst for short-horizon liquidity traders. The liquidity shocks can propagate to the underlying securities through the arbitrage channel, and ETFs may increase the nonfundamental volatility of the securities in their baskets. We exploit exogenous changes in index membership and find that stocks with higher ETF ownership display significantly higher volatility. ETF ownership increases the negative autocorrelation in stock prices. The increase in volatility appears to introduce undiversifiable risk in prices because stocks with high ETF ownership earn a significant risk premium of up to 56 basis points monthly.

EXCHANGE-TRADED FUNDS (ETFS) ARE increasingly popular in financial markets. Introduced in the early 1990s, today this asset class boasts $2.5 trillion in assets under management (AUM) in the United States ($3.5 trillion globally), accounting for about 35% of the volume in U.S. equity markets. Increased access to liquidity and diversification is undoubtedly the greatest benefit for investors. Unlike traditional index funds, ETFs provide intraday liquidity, and thus they can satisfy high-frequency demand for trading. Moreover, investment

*Itzhak Ben-David is with the Fisher College of Business, The Ohio State University, and NBER. Francesco Franzoni is with Università della Svizzera italiana (USI) Lugano and the Swiss Finance Institute. Rabih Moussawi is with the Villanova School of Business, Villanova University, and Wharton Research Data Services (WRDS) at the Wharton School, the University of Pennsylvania. We are especially grateful to Andrew Ellul, Marco Di Maggio, Robin Greenwood (AFA discussant), and Martin Oehmke (NBER discussant). We thank Pierre Collin-Dufresne; Chris Downing; Vincent Fardeau; Thierry Foucault; Rik Frehen; Denys Glushkov; Jungsuk Han; Johan Hombert; Augustin Landier; David Mann; Rodolfo Martell; Massimo Massa; Albert Menkveld; Robert Nestor; Marco Pagano; Ludovic Phalippou; Anton Tonev; Tugkan Tuzun; Dimitri Vayanos; Scott Williamson; Hongjun Yan; and participants at seminars and conferences at the NBER Summer Institute (Asset Pricing), Toulouse School of Economics, Insead, HEC Paris, the Cambridge Judge Business School, Villanova University, USI Lugano, the 4th Paris Hedge Funds Conference, the 5th Paul Woolley Conference (London School of Economics), the 8th Csef-IGIER Symposium (Capri), the 5th Erasmus Liquidity Conference (Rotterdam), the 1st Luxembourg Asset Pricing Summit, the Center for Financial Policy Conference at the University of Maryland, Jacobs Levy's Quantitative Financial Research Conference at the Wharton School, the Geneva Conference on Liquidity and Arbitrage, the 20th Annual Conference of the Multinational Finance Society, the 7th Rothschild Caesarea Conference, the Swedish House of Finance, the FIRS conference (Toronto), and SAC Capital Advisors for helpful comments and suggestions. Ben-David acknowledges support from the Neil Klatskin Chair in Finance and Real Estate and from the Dice Center at the Fisher College of Business. Franzoni acknowledges support from the Swiss Finance Institute. An earlier version of this paper circulated under the title "ETFs, Arbitrage, and Shock Propagation." The authors do not have any material disclosure to make.

DOI: 10.1111/jofi.12727

strategies from which retail investors are typically precluded, such as using leverage and selling short, have become widely available due to ETFs.

One may wonder, however, whether the ease of trade that makes ETFs so popular has unintended consequences for the securities in the ETFs' baskets. The liquidity of ETFs likely attracts high-frequency demand. This demand can affect the prices of the underlying securities because ETFs and their baskets are tied by arbitrage.[1] The ETF-underlying securities may therefore be exposed to a new layer of demand shocks, which can make the prices of these securities more volatile. In this paper, we explore this conjecture and its implications for asset pricing.

To test this hypothesis, one must first understand the difference between the conjectured mechanism and the effect that other institutional investors can have on asset prices (e.g., Coval and Stafford (2007), Lou (2012), Vayanos and Woolley (2013)). Similar to the effect of mutual fund or hedge fund flows on stock prices, the demand for ETF shares exerts pressure on the prices of the underlying securities. What makes ETFs distinct is that they allow investors to access the market continuously and at a low trading cost. ETFs may therefore potentially attract more high-frequency demand than other institutional portfolios, including traditional index funds. Moreover, although active managers have discretion in the extent to which they track their benchmark, which gives them control over the price impact of their trades, arbitrage trading between ETFs and the underlying securities is mechanical and therefore may lead to a bigger price impact. Indeed, ETFs track specified indices, which means that ETF arbitrageurs have almost no discretion in terms of the timing and composition of their trades. As a result, although the channel of shock propagation that we conjecture is not qualitatively different from the price impact that other institutional investors could trigger as a response to flows, the frequency with which ETFs operate and the passive nature of their strategies make ETF arbitrage a distinct mechanism that warrants separate attention.

The natural alternative hypothesis to our conjecture is that ETFs provide a buffer for demand shocks that would otherwise hit the underlying securities' prices. Grossman (1989) makes this point about the introduction of futures, which in this respect are analogous to ETFs. He argues that the existence of futures entails additional market-making power to absorb the impact of liquidity shocks, which reduces volatility in the spot market (see also Danthine (1978), Turnovsky (1983)). To disentangle the two hypotheses, we need to determine the direction of the link between the presence of ETFs and stock-level volatility.

We start our empirical analysis by providing suggestive evidence that ETFs attract short-term investors. ETFs are significantly more liquid, on average, than the basket of underlying securities in terms of bid-ask spread, price

[1] If the ETF price deviates from the net asset value (NAV) of the portfolio holdings because of a demand shock, arbitrageurs trade the underlying securities in the same direction as the initial shock to the ETF price. As a result, the underlying securities can inherit the shocks that occur in the ETF market.

impact, and turnover. Consistent with theories positing that short-horizon clienteles self-select into assets with lower trading costs (Amihud and Mendelson (1986)), using trade-level data we find that institutions holding ETFs have a significantly shorter horizon than those holding the underlying securities.

We next conduct our main test of the effect of ETF ownership, the fraction of a stock's capitalization that is held by ETFs, on stock-level volatility. Using ordinary least squares (OLS) regressions, we find that this relation is positive and significant. In particular, a normal shock to ETF ownership shifts the volatility of the median stock in the S&P 500 to a place that is between the 55th and 64th percentiles.

Although the OLS regressions control for observable stock characteristics and include stock fixed effects, ETF ownership may be endogenous. To address this concern, we rely on the natural experiment provided by the annual reconstitution of the Russell indexes (Appel, Gormley, and Keim (2015), Chang, Hong, and Liskovich (2015)). Exploiting the mechanical rule that allocates stocks between the Russell 1000 and 2000 indexes, we use the index-switching event as an instrument for ETF ownership. This test confirms that the effect of ETF ownership on volatility is positive and statistically significant. Specifically, the shift in ETF ownership that follows the index reconstitution pushes the median stock's volatility up to the 65th percentile of the distribution. The instrumental variable (IV) estimates somewhat exceed their OLS counterparts, which suggests a negative omitted variable bias in the latter. However, the local nature of the experiment that we use for identification does not allow us to generalize these magnitudes to the full sample. We corroborate the validity of the exclusion restriction by showing that switching indexes has a stronger effect in months when ETF ownership is greater, and for stocks with a higher ratio of ETF ownership to mutual fund ownership (index or active) and hedge fund ownership.

Next, we take on the task of separating the hypothesis that the observed increase in volatility reflects the improvement in price discovery brought about by ETFs from the hypothesis that the increase in volatility arises from nonfundamental demand for liquidity, which ETFs attract. In the former case, the increase in volatility would be a result of stock prices responding more promptly to fundamental news, as in Andrei and Hasler (2015). In the latter case, the increase in volatility would amount to noise.

To separate these alternatives, we rely on the premise that liquidity shocks subsequently revert, whereas fundamental information leads to permanent price changes. First, we estimate predictive regressions of stock returns as a function of ETF flows at the stock-day level. We find that most of the contemporaneous stock-price effect of ETF flows reverts over the next 40 days, in line with the view that the demand shocks in the ETF market translate into nonfundamental price changes for the underlying securities. Second, using variance ratios, we find that ETF ownership leads stock prices to deviate from a random walk at the daily frequency and that the autocorrelation of returns becomes more negative. These findings support the conjecture that ETFs add noise to stock prices.

Next, we explore the channel linking ETF ownership to stock volatility. First, using trade-level data from Ancerno in OLS and IV regressions, we show that ETF ownership of stocks attracts traders with higher turnover. We interpret this evidence as consistent with the argument that ETFs attract a new clientele of investors, whose demand is passed on to underlying securities through the arbitrage channel. Second, we provide indirect evidence on the role of ETF arbitrage in increasing stock volatility. In particular, the effect of ETFs on volatility is weaker for stocks with higher limits of arbitrage (i.e., bid-ask spread and short-selling fees) and is stronger during times of more intense arbitrageur trading activity.

Finally, we study the implications of ETF ownership for expected returns. If the volatility that ETFs impound in stock prices is partly nondiversifiable, for example, because most listed stocks are included in ETF baskets, then investors may require compensation for bearing this risk. Consistent with this conjecture, we find that a long-short portfolio of the top minus the bottom quintile of stocks by ETF ownership earns a return premium of up to 56 basis points (bps) per month. This alpha survives in models that include up to seven risk factors and is confirmed in Fama and McBeth (1973) regressions that control for the price impact of ETF flows. Ultimately, the evidence suggests that ETFs are a new source of systematic risk.

Our study relates to several strands of the literature. A growing body of evidence shows that institutions have a role in nonfundamental demand shocks being impounded into asset prices because of flows from their investors (Coval and Stafford (2007), Greenwood and Thesmar (2011), Lou (2012), Basak and Pavlova (2013, 2016), Vayanos and Woolley (2013), Anton and Polk (2014), Hombert and Thesmar (2014), Lou and Polk (2014)). We highlight a previously unexplored channel: arbitrage activity between ETFs and the underlying baskets.[2]

Recent empirical literature also studies the effect of ETFs on asset prices. Da and Shive (2015) find that ETF ownership is associated with higher comovement of the underlying securities. This idea is consistent with our results: ETFs impound the same shocks into all of the stocks in their basket and therefore make them comove. Agarwal et al. (2017) document that ETF ownership also exacerbates the comovement in the liquidity of underlying stocks due to the underlying arbitrage mechanism. Israeli, Lee, and Sridharan (2015) argue that increased ETF ownership can lead to higher trading costs and lower benefits from information acquisition for the basket securities, a combination that results in less informative stock prices. Evans et al. (2017) show that ETF ownership increases intraday bid-ask spreads of the underlying securities, especially when the ETF authorized participants (APs) are more active in the ETF share creation and redemption process. Leippold, Su, and Ziegler

[2] Our paper indirectly relates to the rich literature on the effect of indexing (Shleifer (1986), Barberis, Shleifer, and Wurgler (2005), Greenwood (2005), Wurgler (2011), Chang, Hong, and Liskovich (2015)). The trigger for the effect we measure is trading in ETFs, as opposed to index reconstitution. Index membership matters only in defining the stocks that ETFs affect.

(2015) build a model in which the impact of ETFs on return correlations exceeds the effect of futures, and they find supportive empirical evidence in time series tests. Building on these studies, our paper provides causal evidence of the shock transmission channel and explores the implications for expected returns. Combining theory and empirics, Chinco and Fos (2016) describe the rebalancing cascades that follow from shocks to a nonfundamental stock price and test this theory in the context of ETFs. Bhattacharya and O'Hara (2017) show theoretically that ETFs can increase price fragility due to information linkages. Finally, Malamud (2015) develops a model in which ETFs can affect volatility through the liquidity shock transmission channel as well as through the time-varying risk premiums that investors require as compensation for taking on exposure to these shocks. Both of these claims are consistent with our empirical evidence.[3]

Our study relates to the long-running debate on the effect of derivatives on the quality of the underlying securities' prices. For example, Stein (1987) argues that imperfectly informed speculators in futures markets can destabilize spot prices. Grossman (1989) finds support for the alternative view that futures improve price efficiency.[4] We contribute to this debate by providing systematic evidence on ETFs, an asset class similar to futures but potentially more attractive to liquidity traders due to the lack of margin requirements and the absence of rollover risk. In December 2014, the AUM in ETFs tracking the S&P 500 surpassed the open interest in futures on the same index, suggesting that investors use ETFs as a way to achieve exposure to the stock market (Amery, 2015).

The study proceeds as follows. Section I provides institutional details and describes the data. Section II develops our hypotheses. Section III documents the effects of ETF ownership on stock volatility. Section IV explores the channel through which ETF ownership impacts stock volatility. Section V investigates the asset pricing implications of the impact of ETF ownership on stock volatility. Finally, Section VI concludes.

I. Institutional Details and Data Description

A. Institutional Details

ETFs are investment companies whose objective is to replicate the performance of an index, similarly to index mutual funds. Unlike index funds, however, ETFs are listed on an exchange and trade throughout the day. ETFs

[3] Ben-David, Franzoni, and Moussawi (2017) provide a detailed treatment of the growing literature on ETFs.

[4] Earlier studies that examine the impact of derivatives on volatility focus on futures. Bessembinder and Seguin (1992) find that high trading volume in the futures market is associated with lower equity volatility. They also find that unexpected futures trading volume is positively correlated with equity volatility. Chang, Cheng, and Pinegar (1999) document that the introduction of futures trading increased the volatility of stocks in the Nikkei index. Roll, Schwartz, and Subrahmanyam (2007) find evidence of Granger causality between prices in the futures and equity markets.

were first introduced in 1990 in Canada, and entered the United States in January 1993 with the issuance of Standard & Poor's Depository Receipts (SPDR), which track the S&P 500 (ticker: SPY). SPY is currently the largest ETF in the world, with about $224 billion in assets. As of the end of 2016, the number of ETFs and other exchange-traded products had exploded to more than 1,969 in the United States (6,625 globally), with these funds and products spanning various asset classes.

To illustrate the growing importance of ETFs in the ownership of common stocks, we present descriptive statistics for the S&P 500 and Russell 3000 universes in Table I. For S&P 500 stocks, the average fraction of a stock's capitalization held by ETFs has risen almost 50-fold, from 0.14% in 2000 to 7.05% in 2015. The statistics for Russell 3000 stocks paint a similar picture. The growth of ETFs appears even more dramatic when compared to that of index funds, which are the closest substitute, and active funds. Index fund average stock ownership, although higher in the early part of our sample, was remarkably stable over the sample period. At the end of the sample period, however, ETF ownership in the average stock was almost twice that of index funds. In terms of AUM in S&P 500 stocks, index funds grew by about 73%, which is substantial but orders of magnitude smaller than the growth in ETFs' AUM. Active funds, although still significantly larger than passive vehicles throughout the sample period, grew significantly less than ETFs.

ETFs are similar to futures in that they track an index. However, unlike futures, ETFs do not involve a rollover of the expiring contract, which can erode performance for investors with horizons beyond the short maturity of a futures contract. According to BlackRock (2014), the annualized rollover cost of a futures position in large-cap stocks (S&P 500, Euro Stoxx 50, FTSE 100) ranges from 0.9% to 1.4%. The total expense ratio for an ETF on the same indexes can be as low as 0.05% (e.g., the Vanguard S&P 500 ETF). Hence, ETFs provide a more cost-efficient way to track an index, especially for investors with uncertain trading horizons.

Turning next to closed-end funds, ETFs are traded in the secondary market by retail and institutional investors, similar to closed-end funds. However, unlike closed-end funds, new ETF shares can be created and redeemed. Because supply and demand in the secondary market determines the price of ETF shares, the ETF price can diverge from the value of the underlying securities (the NAV).[5] Some institutional investors—"authorized participants" (APs), that is, dealers that have signed an agreement with the ETF provider—can trade bundles of ETF shares (called "creation units," typically 50,000 shares) with the ETF sponsor. An AP can create new ETF shares by transferring to the ETF sponsor the securities underlying the ETF. These transactions constitute the primary market for ETFs. Similarly, the AP can redeem ETF shares and

[5] Unlike premiums and discounts in closed-end funds (e.g., Lee, Shleifer, and Thaler (1991), Pontiff (1996)), price divergence between the ETF and the NAV can be more easily arbitraged away due to the possibility of continuously creating and redeeming ETF shares. As a result, ETF premiums/discounts are orders of magnitude smaller than for closed-end funds.

Table I

ETF Ownership Statistics

The table presents descriptive statistics for stock ownership by ETFs, index mutual funds, and active mutual funds for the S&P 500 and Russell 3000 universes. For each year, across months and stocks, we average the number of institutions and the percentage of each stock owned by an institution type. We also report cumulative assets owned by these institutions in the stocks in our sample, as reported in Thomson-Reuters Mutual Fund Ownership data.

S&P 500									
	ETFs			Index Funds			Active Funds		
Year	Number of Funds	Percentage of Market Cap	Market Cap (in Dollars)	Number of Funds	Percentage of Market Cap	Market Cap (in Dollars)	Number of Funds	Percentage of Market Cap	Market Cap (in Dollars)
2000	4.5	0.14%	$16.6	111.6	2.61%	$318.6	240.5	15.22%	$1,872.1
2001	12.6	0.18%	$18.0	126.7	2.65%	$269.7	277.8	15.65%	$1,616.9
2002	14.2	0.46%	$39.6	137.6	2.76%	$239.9	288.4	16.07%	$1,416.2
2003	15.1	0.69%	$62.4	141.7	2.92%	$262.9	286.0	16.16%	$1,471.9
2004	18.4	0.86%	$90.6	142.4	3.23%	$341.2	284.3	16.83%	$1,804.4
2005	20.2	1.16%	$129.5	139.3	3.44%	$383.8	287.6	17.55%	$1,980.4
2006	21.5	1.29%	$157.2	130.4	3.64%	$444.4	281.2	18.10%	$2,231.2
2007	38.5	2.50%	$336.2	128.6	3.04%	$407.8	307.5	19.66%	$2,658.5
2008	54.2	4.18%	$430.6	131.1	2.28%	$237.5	345.2	21.18%	$2,214.3
2009	56.0	4.66%	$412.1	125.8	2.51%	$223.0	357.2	21.56%	$1,913.9
2010	54.7	4.77%	$506.6	113.7	2.60%	$275.9	337.2	21.44%	$2,281.3
2011	54.9	5.21%	$604.6	108.6	2.61%	$303.2	326.6	21.72%	$2,527.8
2012	51.5	5.58%	$717.6	99.9	2.73%	$351.2	299.8	20.90%	$2,695.8
2013	53.4	6.30%	$958.0	97.0	2.87%	$437.2	284.5	21.41%	$3,265.6
2014	47.7	6.47%	$1,140.7	91.7	2.90%	$512.0	260.1	20.96%	$3,697.0
2015	48.3	7.05%	$1,214.2	92.9	3.20%	$550.9	257.3	21.70%	$3,741.8

(Continued)

Table I—*Continued*

	Russell 3000								
	ETFs			Index Funds			Active Funds		
Year	Number of Funds	Percentage of Market Cap	Market Cap (in Dollars)	Number of Funds	Percentage of Market Cap	Market Cap (in Dollars)	Number of Funds	Percentage of Market Cap	Market Cap (in Dollars)
2000	2.2	0.12%	$17.5	40.4	2.33%	$343.1	80.3	16.44%	$2,456.3
2001	7.5	0.19%	$23.3	46.6	2.49%	$298.7	92.7	16.59%	$2,006.3
2002	8.1	0.48%	$48.9	51.5	2.65%	$276.5	97.4	17.00%	$1,777.8
2003	8.7	0.73%	$80.6	56.3	2.83%	$310.4	98.4	17.11%	$1,882.0
2004	10.6	0.93%	$121.2	57.2	3.15%	$413.4	102.6	17.95%	$2,361.5
2005	11.5	1.24%	$173.2	56.2	3.38%	$473.3	105.5	18.74%	$2,637.3
2006	12.5	1.39%	$210.8	55.4	3.66%	$555.7	105.1	19.14%	$2,932.8
2007	22.4	2.65%	$445.4	55.3	3.09%	$518.2	124.8	21.02%	$3,563.3
2008	31.0	4.34%	$555.8	57.9	2.26%	$291.7	144.0	22.54%	$2,941.7
2009	31.2	4.95%	$534.2	58.5	2.51%	$271.3	147.9	23.07%	$2,498.0
2010	30.1	5.14%	$677.5	52.2	2.59%	$339.5	139.8	23.07%	$3,050.8
2011	30.0	5.56%	$803.4	51.4	2.61%	$375.5	134.2	23.28%	$3,382.7
2012	28.2	5.94%	$942.1	49.1	2.71%	$427.0	122.9	22.40%	$3,572.6
2013	29.7	6.70%	$1,267.7	47.9	2.86%	$537.2	119.2	22.93%	$4,370.6
2014	27.6	6.83%	$1,502.4	45.4	2.89%	$632.2	109.8	22.21%	$4,917.8
2015	28.3	7.48%	$1,611.3	44.6	3.14%	$673.6	109.4	22.92%	$4,974.1

receive the underlying securities in exchange. For some funds, ETF shares can be created and redeemed in cash.[6,7]

If an ETF trades at a premium relative to the NAV, APs have an incentive to buy the underlying securities, submit them to the ETF sponsor, and ask for newly created ETF shares in exchange. The AP then sells the new supply of ETF shares on the secondary market. This process puts downward pressure on the ETF price and potentially leads to an increase in the NAV, reducing the premium. In the case of a discount, APs buy ETF units in the market and redeem them for the basket of underlying securities from the ETF sponsor. The APs can then sell the securities in the market. This process generates positive price pressure on the ETF and possibly negative pressure on the NAV, which reduces the discount.

ETF arbitrage also takes place continuously throughout the day and is carried out by hedge funds and high-frequency traders. These investors do not need to engage in primary market trades. Instead, they can turn to the secondary market, where they can buy the inexpensive asset and short sell the more expensive one between the ETF and the basket of underlying securities. They then hold the positions until prices converge, at which point they close the positions to realize a profit. Of course, the uncertainty involved in these profits does not qualify these trades as an arbitrage in a strict sense. ETF sponsors facilitate arbitrage activity by disseminating NAV values at a 15-second frequency throughout the trading day. They do so because the smooth functioning of arbitrage is what brings about the low tracking error of these instruments. Because of the low trading costs and availability of information, arbitraging ETFs against the NAV has become a popular trading strategy in recent years. According to some industry participants, statistical arbitrage accounts for 50% of the volume in the SPY, the most traded security in the United States, with $24 billion in average daily volume over the last three months of 2015. Over our sample period, secondary market volume in ETFs accounted for about 93% of total trading volume, which includes primary market creation and redemption activity.

The above institutional details, with some modifications, also apply to synthetic ETFs, which replicate the performance of the underlying index using total return swaps and other derivatives, and for which creation and redemption occur in cash. Secondary market arbitrage still involves transactions in the underlying securities. Similarly, the arbitrage process is an inherent characteristic of all types of ETFs. Hence, one should expect the effects we describe in this paper to hold for all types of underlying assets, irrespective of the replication mechanism.

[6] Creation and redemption in cash is especially common with ETFs on foreign or illiquid assets, for example, fixed income ETFs.

[7] Petajisto (2013) describes the fixed creation/redemption costs as ranging in absolute terms from $500 to $3,000 per creation/redemption transaction, irrespective of the number of units involved.

B. The Data

We use data from the Center for Research in Security Prices (CRSP), Compustat, Bloomberg, and OptionMetrics to identify ETFs traded on the major U.S. exchanges and to extract returns and prices. We first draw information from CRSP for all securities that have a historical share code of 73, which exclusively defines ETFs. We then screen all U.S.-traded securities in the Compustat XpressFeed and OptionMetrics data, identifying ETFs using the security type variables, and merge this sample with the CRSP ETF sample.

In our analysis, we focus on ETFs that are listed on U.S. exchanges and whose baskets contain U.S. stocks. Further, we restrict our sample to "plain vanilla" products that engage in physical replication, that is, those that hold the securities of the basket they aim to track. We omit from our sample leveraged and inverse-leveraged ETFs that use derivatives to deliver the performance of the index. These assets represent about 2% of ETFs according to BlackRock (2014). We also omit active ETFs that are below 1% of the AUM in the sector.[8]

We use the Thomson-Reuters Mutual Fund Ownership database as our source of ETF holdings data for our sample. Our final data set contains 454 distinct equity ETFs, which provide holdings information for 93% of all domestic equity ETFs in the United States between January 2000 and December 2015, with around 92% coverage of all U.S. equity ETFs in December 2015 and a maximum coverage of 97.1% of all domestic equity ETFs during our sample period. In the Internet Appendix,[9] we detail how we address the issue of unreliable Thomson holdings data starting in June 2015. We use total shares outstanding at day-end to compute the daily market capitalization of each ETF and to measure net share creations/redemptions (i.e., flows) for each ETF daily. Bloomberg is our primary source for shares outstanding, as it is more accurate. We complement Bloomberg with Compustat and OptionMetrics in the event of data gaps. As a dependent variable in our main tests, we compute daily stock volatility at the monthly frequency as the standard deviation of daily returns over a month. For some of our tests, we compute intraday volatility at a daily frequency using second-by-second data from the Trade and Quote (TAQ) database. We compute stock-level order imbalance at the daily frequency. We use the daily TAQ millisecond feed to classify each trade between 2007 and 2015, and the monthly TAQ feed at second-by-second granularity to classify trades between 2000 and 2006. Following Holden and Jacobsen (2014), we modify the Lee and Ready (1991) algorithm, according to the logic in Ellis, Michaely, and O'Hara (2000), to classify trades as buys or sells. (More details on our classification are provided in the Internet Appendix.)

We extract stock lending fees from the Markit Securities Finance (formerly Data Explorers) database. We use the average lending fee over the

[8] We restrict our sample to the following Lipper Objective Codes for broad-based U.S. equity: CA, EI, G, GI, MC, MR, SG, and SP. We also include sector funds that invest in U.S. companies with codes BM, CG, CS, FS, H, ID, NR, RE, TK, TL, S, and UT.

[9] The Internet Appendix is available in the online version of this article on the *Journal of Finance* website.

prior seven days. For some of our tests, we use Ancerno institutional trade-level data from firm Abel Noser.[10] Ancerno provides information on all of the trades of a portfolio during the period in which the manager reports to the database.

Table II reports summary statistics for the variables that we use in the analysis. Panel A presents summary statistics for the monthly stock-level sample used in our main regressions while Panel B reports the correlations for the same variables. Panel C presents summary statistics for the variables that we use in the return regressions at the daily frequency. Panel D presents statistics for the variables derived from the trade-level Ancerno data set. We describe these variables in more detail below and provide definitions in Table IA.I in the Internet Appendix.

II. Hypothesis Development and Evidence on Trading Horizon

A. Hypothesis Development

The main testable hypothesis of the paper is that ETFs are a catalyst for liquidity trading and that the ensuing price pressure propagates to the underlying securities via arbitrage. According to this hypothesis, stocks with higher ETF ownership should display higher volatility, all else being equal. We label this conjecture the *liquidity trading hypothesis*.

We provide a simple illustration of this channel. Imagine a situation in which the ETF price and the NAV of its portfolio are aligned at the level of the fundamental value, as in Panel A of Figure 1. Then a liquidity shock, that is, one that is unrelated to information about future cash flows, hits the ETF market. Arbitrageurs absorb the liquidity demand by shorting the ETF. Because they are risk-averse, as in Grossman and Miller (1988), arbitrageurs require compensation for the (negative) inventory in the ETF they are taking on. Hence, the ETF price has to rise (Figure 1, Panel B). To hedge their short ETF position, arbitrageurs take a long position in the securities in the ETF basket. To compensate market makers, the prices of the basket securities have to rise, as in Panel C of Figure 1. Eventually, when other sources of liquidity materialize, prices revert to fundamentals (Figure 1, Panel D). This prediction also emerges in Malamud's (2015) dynamic model of the ETF market.

For example, the trades of hedge funds that carry out high-frequency ETF arbitrage conform to the mechanism described in Figure 1. In addition, hedge funds can impound mispricing indirectly through their use of ETFs in statistical arbitrage. Suppose hedge funds short sell an overpriced stock and hedge the industry risk by going long in the corresponding sector ETF. This trade puts upward pressure on the ETF price (as in Figure 1, Panel B). Then ETF arbitrageurs transfer the price pressure to the securities in the ETF basket (as

[10] Previous studies, such as Puckett and Yan (2011) and Anand et al. (2012, 2013), argue convincingly for the representativeness of the institutions in Ancerno for the broader universe of institutional investors.

Table II

Summary Statistics

The table presents summary statistics for the variables used in the study. Panel A reports summary statistics for the stock-month sample. Panel B reports correlations for the same sample, and Panel C reports summary statistics for the stock-day sample. Panel D reports summary statistics for the Ancerno transaction-level data. The samples cover the period January 2000 to December 2015.

Panel A: Monthly Sample						
	S&P 500					
	N	Mean	*SD*	Min	Median	Max
Daily stock volatility (%)	81,243	2.010	1.400	0.581	1.600	11.000
Skewness	81,243	0.109	2.490	−8.900	0.114	8.900
Log(abs(VR 5 days/(5*1 day)))	81,243	−1.810	0.895	−3.850	−1.600	−0.582
ETF ownership	81,243	0.026	0.019	0.000	0.023	0.112
Index fund ownership	81,243	0.062	0.020	0.004	0.061	0.120
Active fund ownership	81,243	0.176	0.063	0.008	0.173	0.354
Hedge fund ownership	81,243	0.034	0.045	0.000	0.019	0.354
log(Mktcap ($m))	81,243	9.360	1.060	5.230	9.300	11.400
1/Price	81,243	0.039	0.040	0.004	0.029	0.556
Amihud ratio	81,243	0.000	0.001	0.000	0.000	0.036
Bid-ask spread (%)	81,243	0.207	0.462	0.000	0.048	3.700
Book-to-market	81,243	0.524	0.385	0.038	0.420	2.360
Past 12-month returns	81,243	0.113	0.366	−0.823	0.100	2.180
Gross profitability	81,243	0.301	0.224	−0.399	0.262	1.120

(Continued)

Table II—*Continued*

Panel A: Monthly Sample						
Russell 3000						
	N	Mean	*SD*	Min	Median	Max
Daily stock volatility (%)	418,748	2.610	1.760	0.581	2.120	11.000
Skewness	418,745	0.123	2.490	−8.900	0.130	8.900
Log(abs(VR 5 days/(5*1 day)))	418,748	−1.740	0.902	−3.850	−1.520	−0.582
ETF ownership	418,748	0.028	0.023	0.000	0.023	0.112
Index fund ownership	418,748	0.047	0.025	0.002	0.045	0.120
Active fund ownership	418,748	0.151	0.083	0.002	0.149	0.354
Hedge fund ownership	418,748	0.058	0.070	0.000	0.031	0.354
log(Mktcap ($m))	418,748	7.120	1.540	4.250	6.900	11.400
1/Price	418,748	0.073	0.083	0.004	0.047	0.556
Amihud ratio	418,748	0.015	0.037	0.000	0.002	0.278
Bid-ask spread (%)	418,748	0.293	0.527	0.000	0.113	3.700
Book-to-market	418,748	0.588	0.414	0.038	0.498	2.360
Past 12-month returns	418,748	0.133	0.485	−0.823	0.089	2.180
Gross profitability	418,748	0.293	0.258	−0.399	0.262	1.120

(*Continued*)

Table II—*Continued*

Panel B: Correlations

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Daily stock volatility (%)	(1)	1.000												
Skewness	(2)	−0.004	1.000											
Log(abs(VR 5 days/(5*1 day)))	(3)	0.027	0.000	1.000										
ETF ownership (%)	(4)	−0.074	−0.021	−0.022	1.000									
Index fund ownership (%)	(5)	−0.092	−0.007	−0.060	0.355	1.000								
Active fund ownership (%)	(6)	−0.102	−0.004	−0.097	0.160	0.428	1.000							
Hedge fund ownership (%)	(7)	0.071	0.008	−0.025	0.055	−0.005	−0.052	1.000						
log(Mktcap ($m))	(8)	−0.343	−0.001	−0.083	0.041	0.325	0.382	−0.147	1.000					
1/Price	(9)	0.402	0.001	0.040	−0.074	−0.169	−0.287	0.094	−0.459	1.000				
Amihud ratio	(10)	0.312	−0.003	0.104	−0.174	−0.262	−0.377	−0.035	−0.473	0.439	1.000			
Bid-ask spread (%)	(11)	0.273	0.000	0.058	−0.270	−0.212	−0.222	−0.084	−0.265	0.281	0.428	1.000		
Book-to-market	(12)	0.104	0.001	0.036	0.088	0.089	−0.164	−0.010	−0.227	0.234	0.190	0.132	1.000	
Past 12-month returns	(13)	−0.191	−0.007	−0.031	−0.009	−0.015	0.024	0.028	0.134	−0.240	−0.187	−0.147	−0.264	1.000
Gross profitability	(14)	−0.029	0.001	−0.030	−0.015	0.025	0.201	0.010	0.033	−0.061	−0.082	−0.054	−0.297	0.083

(*Continued*)

Table II—*Continued*

Panel C: Variables Used in Return Regressions (Daily Frequency)						
	S&P 500					
	N	Mean	*SD*	Min	Median	Max
Net flows ($ million)	1,501,051	0.250	40.700	−12.400	0.027	13.200
abs(Net flows) ($ million)	1,501,051	2.050	40.700	0.000	0.609	18.300
Net flows/Daily volume	1,501,056	0.002	0.027	−0.081	0.001	0.088
abs(Net flows)/Daily volume	1,501,056	0.015	0.022	0.000	0.007	0.111
NBBO size ($)	1,437,745	517	1813	54	295	3433
DGTW Ret[t] (%)	1,501,056	0.015	1.700	−4.820	−0.013	5.150
DGTW Ret[t, $t+4$] (%)	1,501,056	0.070	3.660	−10.000	−0.004	10.800
DGTW Ret[t, $t+9$] (%)	1,501,056	0.127	5.070	−13.500	0.018	14.700
DGTW Ret[t, $t+19$] (%)	1,501,056	0.209	7.090	−18.400	0.061	20.200
DGTW Ret[t, $t+39$] (%)	1,501,056	0.352	10.000	−25.400	0.136	28.500
abs(Net flows)$\times 10^6$/Mkt cap	1,542,880	0.012	0.016	0.000	0.006	0.078
sum(abs(Flows))$\times 10^6$/Mkt cap	1,542,880	0.018	0.020	0.000	0.012	0.096
Intraday price range (%)	1,542,880	0.028	0.021	0.006	0.022	0.111
Intraday volatility (%)	1,542,880	0.020	0.018	0.005	0.014	0.100
abs(Mispricing) (bps)	1,686,822	0.174	0.147	0.013	0.129	0.859
Net(mispricing) (bps)	1,660,396	−0.004	0.125	−0.499	0.006	0.369
Share lending fee (%)	1,029,618	0.213	1.110	0.000	0.099	72.400

(Continued)

Table II—*Continued*

Panel C: Variables Used in Return Regressions (Daily Frequency)						
Russell 3000						
	N	Mean	*SD*	Min	Median	Max
Net flows ($ million)	7,812,071	0.081	23.300	−4.600	0.000	5.130
abs(Net flows) ($ million)	7,812,071	0.617	23.300	0.000	0.091	7.770
Net flows/Daily volume	7,812,107	0.002	0.051	−0.166	0.000	0.176
abs(Net flows)/Daily volume	7,812,107	0.029	0.042	0.000	0.011	0.197
NBBO size ($)	7,733,982	226	1,057	16	100	1,886
DGTW Ret[t] (%)	7,812,107	0.007	2.070	−5.750	−0.038	6.360
DGTW Ret[t, $t+4$] (%)	7,812,107	0.021	4.410	−11.800	−0.095	13.000
DGTW Ret[t, $t+9$] (%)	7,812,107	0.026	6.090	−15.800	−0.138	17.400
DGTW Ret[t, $t+19$] (%)	7,812,107	0.021	8.540	−21.600	−0.206	23.800
DGTW Ret[t, $t+39$] (%)	7,812,107	0.007	12.100	−29.900	−0.359	33.900
abs(Net flows)$\times 10^6$/Mkt cap	8,546,816	0.017	0.022	0.000	0.008	0.094
sum(abs(Flows))$\times 10^6$/Mkt cap	8,546,816	0.022	0.026	0.000	0.013	0.108
Intraday price range (%)	8,546,297	0.036	0.028	0.006	0.028	0.143
Intraday volatility (%)	8,546,816	0.024	0.019	0.005	0.018	0.102
abs(Mispricing) (bps)	8,572,503	0.211	0.169	0.013	0.162	0.859
Net(mispricing) (bps)	8,454,831	−0.015	0.153	−0.499	−0.001	0.369
Share lending fee (%)	4,554,404	0.444	2.240	0.000	0.132	132.000

Panel D: Ancerno Data						
S&P 500						
	N	Mean	*SD*	Min	Median	Max
Price impact (bps)	28,291,346	3.757	98.373	−463.063	1.242	489.987
Turnover (%)	28,291,346	0.446	1.689	0.000	0.031	24.758
Russell 3000						
Price impact (bps)	54,250,928	5.355	114.126	−463.063	1.685	489.987
Turnover (%)	54,250,928	0.939	2.676	0.000	0.077	24.758

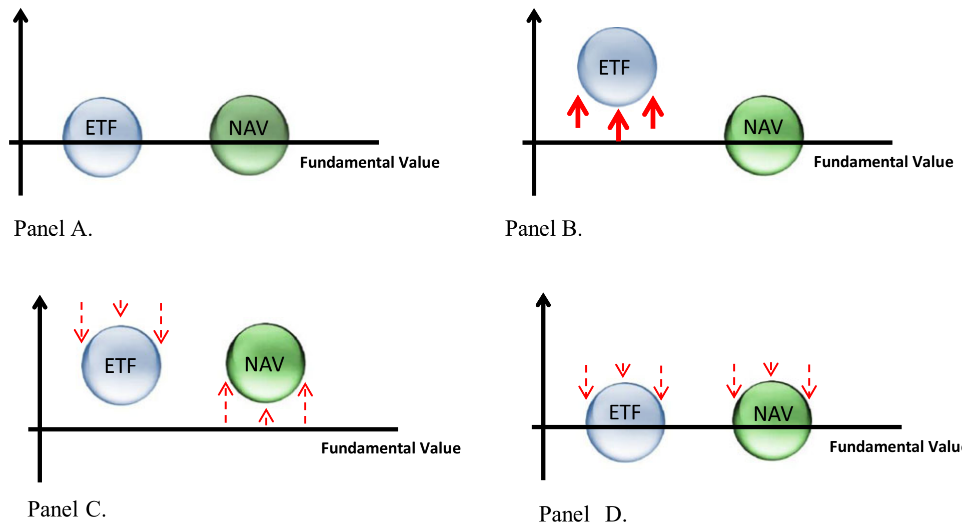
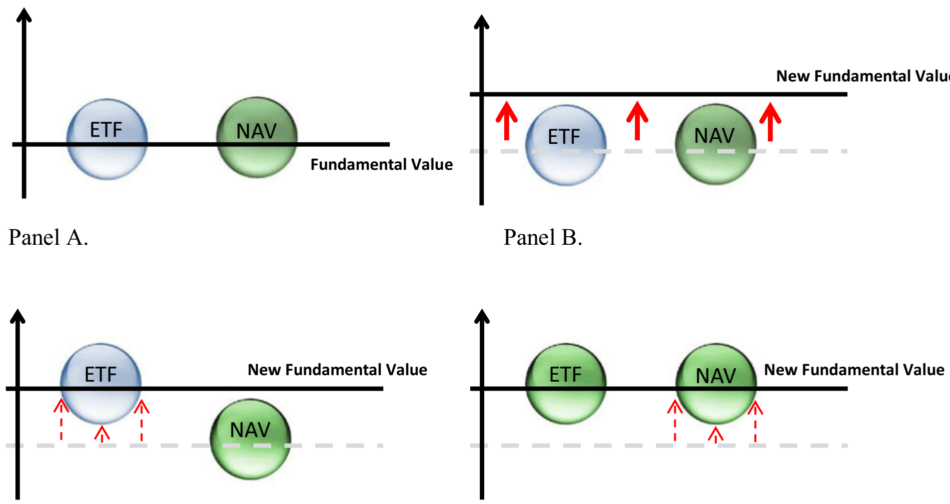


Figure 1. Illustration of the propagation of liquidity shocks via arbitrage. (Panel A) Initial equilibrium. (Panel B) Liquidity shock to ETF. (Panel C) Initial outcome of arbitrage: the shock is propagated to the NAV, and the ETF price starts reverting to the fundamental value. (Panel D) Equilibrium reestablished: after some time, both the ETF price and the NAV revert to the fundamental value. (Color figure can be viewed at wileyonlinelibrary.com)

in Figure 1, Panel C). Thus, ETFs can propagate mispricing to the underlying securities not only because they are traded directly by uninformed investors, but also because they participate in long-short strategies that involve other mispriced securities (similar to the propagation mechanism in Hong, Kubik, and Fishman (2012)).

The liquidity trading hypothesis predicts an increase in the demand shocks hitting the underlying securities because of increased ETF ownership. However, if the same liquidity traders merely shifted from trading a given stock to trading the ETFs holding that security, the volatility in the stock's price would not increase. To observe an increase in volatility, ETFs need to attract a new layer of liquidity demand. This scenario could arise if ETFs allow a new breed of short-horizon investors to express their liquidity demand, with more ease than using stocks. To explore this conjecture, in the first part of our analysis we compare the trading horizons of stock and ETF investors (Section II.B). Moreover, in Section IV.A we investigate the effect of ETF ownership on the investor churn ratio at the stock level.

The most natural alternative hypothesis to the liquidity trading hypothesis is that ETFs decrease volatility. Grossman (1989) makes such a claim about futures, which we extend by analogy to ETFs. The argument is that the introduction of a correlated asset class, such as ETFs, adds a new layer of market-making power that acts as a liquidity buffer. Investors who would have



Panel A. Panel B. Panel C. Panel D.

Figure 2. Illustration of the propagation of a fundamental shock with price discovery occurring in the ETF market. (Panel A) Initial equilibrium. (Panel B) Shock to fundamental value. (Panel C) Price discovery takes place in the ETF market. The ETF price moves to the new fundamental value. (Panel D) After a delay, the NAV catches up with the new fundamental. (Color figure can be viewed at wileyonlinelibrary.com)

originally satisfied their liquidity demand using the ETF-underlying stocks can now find satisfaction in the ETF market. Consequently, the liquidity shocks hitting the underlying securities are dampened and volatility decreases. We label this conjecture the *liquidity buffer hypothesis*. To separate these two hypotheses, we examine the causal effect of ETF ownership of stocks on stock-level volatility. Section III contains this analysis.

The liquidity trading hypothesis also differs from the view that ETFs are vehicles for improved price discovery (e.g., Glosten, Nallareddy, and Zou (2016), Madhavan and Sobczyk (2016)). Under this hypothesis, the starting point is a fundamental shock, that is, a permanent change in the value of the underlying securities. If ETFs facilitate price discovery, then ETF prices adjust immediately to the new information, while the prices of the underlying securities remain temporarily fixed ("stale pricing"). We illustrate this situation in Figure 2. A shock to the fundamental value of the ETF components (Figure 2, Panel B) perturbs the initial equilibrium (Figure 2, Panel A). The ETF price moves first because of price discovery (Figure 2, Panel C), and the prices of the underlying securities move with a delay because of stale pricing (Figure 2, Panel D). In this case, the presence of ETFs allows prices to react more quickly to fundamentals. As a result, a positive link could exist between ETF ownership and volatility, but the increased volatility would result from the faster impounding of fundamental information into prices. We label this conjecture the *price discovery hypothesis*.[11]

To disentangle the liquidity trading hypothesis from the price discovery hypothesis, we need to test whether ETFs are associated with increased mean reversion in prices, which follows from the propagation of liquidity shocks (as per Figure 1). Analogously, one can measure the effect of ETF ownership on price efficiency. Finally, to corroborate the liquidity trading hypothesis, we need to show that nonfundamentally motivated ETF trading has an effect on stock volatility. Sections IV and V address these issues.

B. ETFs versus Stocks: Liquidity, Investor Types, and Trading Horizon

The liquidity trading hypothesis posits that ETFs increase the volatility of the securities in their baskets because of the propagation of liquidity shocks. This conjecture requires the condition that, in a world without ETFs (i.e., the counterfactual world), less liquidity trading in the underlying securities would occur. Although we directly test this conjecture in Section IV.A, here we provide suggestive evidence in its support by contrasting the liquidity of ETFs with that of their portfolio constituents and comparing investor turnover in the two asset classes.

Motivated by the theory in Subrahmanyam (1991), one can argue that the bid-ask spreads on ETFs are low on average because of less severe information asymmetry than in stocks. Indeed, investors with stock-level private information are more likely to trade individual securities, and thus market makers impose higher bid-ask spreads to overcome adverse selection. By contrast, investors who place uninformed directional bets or trades for hedging purposes are more likely to trade entire baskets, such as ETFs. As a result, ETF spreads are less likely to contain an adverse-selection premium.

Table III, Panel A, presents systematic evidence on the difference in liquidity between ETFs and the underlying portfolios along three dimensions: the percentage of the bid-ask spread, the Amihud (2002) measure of price impact, and daily turnover. For all of the ETFs in our sample, we compute the average of each liquidity measure across all the stocks in the basket in a given quarter. Then, to replicate the strategy of an investor who allocates funds to all ETFs according to their market capitalization, we take the value-weighted mean of these measures across all ETFs in a given quarter. The table reports the time series average of these means in the 64 quarters of the sample, along with the results of tests for the statistical significance of their difference. Along all three dimensions, the average ETF is significantly more liquid than its basket

[11] Theoretical underpinning for this hypothesis can be found, for example, in Amihud and Mendelson (1987), who provide a simple model in which the volatility of trading prices is positively related to the speed with which prices adjust to fundamentals. In addition, Andrei and Hasler (2015) prove theoretically and empirically that investor attention increases the sensitivity of prices to fundamentals and therefore volatility. If ETF arbitrage makes prices adjust more promptly to fundamentals, or if stocks in ETFs are exposed to higher investor attention, the fundamental volatility of the underlying securities can increase.

Table III

ETFs versus Stocks: Liquidity, Institutional Ownership, Churn Ratio

The table reports statistics for ETF- and stock-level liquidity, investor turnover, and ownership. The sample includes S&P 500 stocks and ETFs. Panel A reports the security-level liquidity measures (bid-ask spread, the Amihud (2002) ratio, and daily turnover). For all ETFs in our sample, we compute the average measure of liquidity across the stocks in the basket in a given quarter. We then value-weight the ETF- and basket-level measures across all ETFs at the quarter level using ETF market capitalization. Panel B presents information about institutional ownership averaged across all 64 quarters, in the first quarter of the sample, and in the last quarter of the sample. Ownership averages are presented for the ETFs in our sample and all of the stocks in the CRSP data set. Churn ratio is computed as the sum of quarterly absolute changes in dollar holdings over average AUM. Panel C uses trade-level data from Ancerno, aggregated at the manager-stock-quarter level to present regressions of the churn ratio and adjusted churn ratio on an ETF indicator. Summary statistics for Panel C are presented in Table IA.III in the Internet Appendix. t-statistics for the test of the null hypothesis that the difference is equal to zero are presented in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample covers the period Q1:2000 to Q4:2015.

Panel A: Liquidity and Investors' Churn Ratio Measures

Liquidity Measures	Quarters	ETFs	Stocks	Difference	t-Statistics
Bid-ask spread	64	0.002	0.004	−0.002	(−3.611)
Amihud ratio	64	0.002	0.007	−0.004	(−7.408)
Daily turnover	64	0.085	0.010	0.075	−13.682

Panel B: Types of Institutional Ownership

	Common Stocks				ETFs			
	Ownership				Ownership			
Type of Institution	Total	Q1:2000	Q4:2015	Churn Ratio	Average	Q1:2000	Q4:2015	Churn Ratio
Bank and trust	1.7%	1.5%	1.9%	17.9%	3.3%	2.5%	4.8%	31.2%
Insurance company	0.8%	0.8%	0.9%	11.3%	0.5%	1.1%	0.4%	37.6%
Investment advisor/investment company	38.9%	32.6%	43.0%	18.5%	16.1%	14.2%	25.2%	45.1%
Investment advisor/hedge fund	14.7%	11.0%	15.6%	26.7%	8.7%	13.7%	6.3%	48.5%
Hedge fund	2.8%	0.9%	4.2%	74.8%	1.9%	1.0%	0.7%	100.8%
Pension fund	3.7%	2.7%	3.4%	13.3%	1.4%	1.6%	1.1%	58.5%
Research firm	2.0%	0.8%	2.9%	39.7%	17.6%	7.4%	17.6%	51.0%
Other institutions	0.7%	0.4%	0.7%	23.7%	1.3%	0.8%	0.1%	62.7%
All institutions	65.4%	50.9%	72.4%	22.6%	50.7%	42.2%	56.3%	47.8%

Panel C: Churn Ratio Regressions

Dependent Variable:	Churn Ratio			Adjusted Churn Ratio		
	(1)	(2)	(3)	(4)	(5)	(6)
Security is ETF	0.194***	0.197***	0.146***	0.571***	0.570***	0.372***
	(13.764)	(14.584)	(11.060)	(10.258)	(10.142)	(7.246)
Date fixed effects	No	Yes	Yes	No	Yes	Yes
Manager fixed effects	No	No	Yes	No	No	Yes
Observations	32,371	32,371	32,329	32,371	32,371	32,329
Adjusted-R^2	0.154	0.168	0.358	0.098	0.101	0.396

stocks. In particular, the bid-ask spread is lower by about 20 bps, price impact as measured by the Amihud (2002) ratio is also significantly lower for ETFs, and ETF turnover is about 7.5% higher.

A corollary of the conjecture that ETFs are more liquid than the underlying baskets is that ETF investors should display higher turnover. This prediction stems, for example, from Amihud and Mendelson's (1986) clientele effect, whereby short-horizon investors choose to trade in more liquid securities. The evidence in Table III, Panel B, which compares the churn ratio of different institutional classes in common stocks and ETFs, supports this conjecture.[12] The churn ratio comes from Cella, Ellul, and Giannetti (2013), who compute the institutional investor-level churn ratio as the sum of quarterly absolute changes in dollar holdings over average AUM. For all investor classes, the churn ratio is significantly larger for ETFs than for stocks.

The panel also reports shares held by each group as a fraction of total shares outstanding. Two observations are worth nothing. First, throughout the sample period, the institutional ownership of ETFs is far smaller (at 50.7% on average) than the institutional ownership of stocks (at 65.4% on average). Based on Stambaugh's (2014) argument that uninformed traders are mostly present among retail investors, this evidence suggests a higher density of liquidity traders in the ETF investor base. Second, research firms, which include broker-dealers, have greater ETF ownership (17.6%) than stock ownership (2%). This class of owners, along with hedge funds, corresponds to ETF arbitrageurs and market makers (including APs).

The churn ratio statistics in Table III, Panel B, may underestimate the extent of investor turnover because the quarterly frequency at which they are computed conceals intraquarter trading. To overcome this limitation, we rely on the Ancerno data set, which contains institutional trade-level data. We compute two measures of the manager-level churn ratio, noting that Ancerno does not provide information on AUM. For each manager- and security-quarter, our first measure compares the total dollar value of buy and sell trades with the absolute value of the difference between buy and sell trades. Hence, for manager k, security i, and quarter q, the churn ratio is

$$Churn\ Ratio_{k,i,q} = 1 - \frac{\left|\$Buy\ Trades_{k,i,q} - \$\ Sell\ Trades_{k,i,q}\right|}{\$Buy\ Trades_{k,i,q} + \$\ Sell\ Trades_{k,i,q}}. \quad (1)$$

This variable ranges between zero and one: zero if the trades in the quarter are in only one direction (either buy or sell) and one if buying activity equals selling activity in that security over the quarter. We average this variable across all of the securities traded by a manager in a given quarter, for ETFs and stocks, to obtain the manager-level churn ratio.

Notice that this definition of the churn ratio yields the same number when the manager sells and buys the security only once during the quarter as when

[12] Institutional category definitions from Thomson O.P. (Ownership and Profiles) are presented in Table IA.II in the Internet Appendix.

the manager flips the security several times. To take into account the number of times the trade in the security changes from buy to sell or from sell to buy during the quarter, we modify this variable as follows:

$$Adjusted\ Churn\ Ratio_{k,i,q} = Churn\ Ratio_{k,i,q} * \frac{N_{k,i,q} + 1}{2}, \tag{2}$$

where $N_{k,i,q}$ is the number of times during the quarter that the direction of a trade differs from that of the preceding trade.[13] Table IA.III in the Internet Appendix provides summary statistics for the two measures of the churn ratio, splitting the manager-quarter-level sample between stocks and ETFs. Across subsamples, we note that both measures of the churn ratio give a significantly higher score for ETFs. The raw churn ratio is 0.105 for stocks and 0.299 for ETFs. The adjusted churn ratio is 0.137 for stocks and 0.450 for ETFs. In Table III, Panel C, we regress the two churn ratio measures on a dummy for whether the security is an ETF. Standard errors are clustered at the quarter and manager levels. Across all specifications, including different types of fixed effects, the difference in churn ratio between ETFs and stocks is statistically significant.

Overall, the empirical evidence in this section suggests that institutions trade ETFs at a substantially higher frequency than stocks. This result provides initial support for the main assumption behind the liquidity trading hypothesis, namely, that ETFs are a catalyst for short-horizon investors. In Section IV.A, we test whether this clientele of high-turnover investors is passed on to the underlying securities.

III. The Effect of ETF Ownership on Volatility

A. ETF Ownership and Volatility: OLS Regressions

To separate the liquidity trading from the liquidity buffer hypothesis, we test whether ETF ownership leads to higher volatility of the underlying securities. ETF ownership of stock i in month t is defined as the sum of the dollar value of holdings by all ETFs investing in the stock, divided by the stock's capitalization at the end of the month,

$$ETF\ Ownership_{i,t} = \frac{\sum_{j=1}^{J} w_{i,j,t} AUM_{j,t}}{Mkt\ Cap_{i,t}}, \tag{3}$$

where J is the set of ETFs that hold stock I; $w_{i,j,t}$ is the weight of the stock in the portfolio of ETF j, which is extracted from the most recent quarterly report; and $AUM_{j,t}$ is the AUM of ETF j at the end of the month. The Internet Appendix provides more details on the construction of this variable.

[13] For example, using this measure, a manager who has the sequence of trades [+100, –100, +100, –100] in a given security has an adjusted churn ratio of two, which is double that of a manager with the sequence of trades [+100, –100].

Based on equation (3), variation in ETF ownership comes from three sources, which present different types of potentially spurious correlation with stock volatility. First, the number of ETFs tracking a stock depends on the number of indexes in which the stock appears. If more established, less volatile firms are more likely to be members of an index, the correlation between ETF ownership and volatility might be negatively biased. Second, variation in ETFs' AUM occurs over time and across products. For example, investor demand for existing or new ETFs may relate to how popular a given sector or asset class is. If this popularity also affects the trading intensity and volatility of the underlying securities, then popularity could generate a positive relation between ETF ownership and volatility that confounds the causal effect we are trying to identify. Third, variation exists in weighting schemes. For example, the S&P 500 and the Russell 2000 are capitalization-weighted, but the Dow Jones is price-weighted. In addition, our sample contains 17 products that explicitly mention equal-weighting in their names. If the weights in the numerator do not grow at the same pace as market capitalization in the denominator (e.g., for equal-weighted indexes), then a spurious link could exist between ETF ownership and volatility resulting from the correlation between stock size and volatility.

To avoid the issues related to weighting schemes, we control for lagged market capitalization. To guard against potentially omitted variables, we follow three approaches. First, we include stock and month fixed effects. Second, we control for stock size and liquidity, which is measured by the inverse of the stock price, the Amihud (2002) illiquidity measure of price impact, and the bid-ask spread. Finally, we include standard predictors of returns that could also relate to volatility, such as the book-to-market ratio, past 12-month returns, and gross profitability (gross income scaled by total assets, as in Novy-Marx (2013)).

We start by reporting the results of OLS regressions of daily volatility in a given month on ETF ownership at the end of the prior month. Table IV, Panel A, presents results of separate regressions for S&P 500 stocks and Russell 3000 stocks. The goal is to assess how the effect of interest varies for stocks of different size. All of the controls correspond to the end of the prior month. We also include stock and month fixed effects in all regressions. Standard errors are double-clustered at the stock and month levels. To ease interpretation, we standardize volatility and ETF ownership by subtracting the sample mean and dividing by the sample standard deviation for the full sample.

From column (1) of Table IV, Panel A, we infer that the relationship between ETF ownership and volatility is positive and strongly statistically significant. Economically, a one-standard-deviation change in ownership is associated with 16.4% of a standard deviation change in daily volatility. We provide a more detailed discussion of economic magnitudes in Section III.C. The positive and significant link between ETF ownership and stock volatility provides initial evidence in support of the liquidity trading hypothesis and against the liquidity buffer hypothesis.

Table IV
ETF Ownership and Stock Volatility

The table reports estimates from ordinary least squares (OLS) regressions of daily volatility on ETF ownership and controls. In Panel A, columns (1) to (4), the sample consists of S&P 500 stocks, and in columns (5) to (8), the sample consists of Russell 3000 stocks. Panel B presents the same regressions but with the sample split by period. The frequency of observations is monthly, and volatility is computed using all daily returns within the month. The controls in all panels include logged market capitalization, the lagged inverse share price, the lagged Amihud (2002) ratio, the lagged average bid-ask spread, the lagged book-to-market ratio, lagged past 12-month returns, lagged gross profitability (as in Novy-Marx (2013)), lagged volatility, index fund ownership, and active fund ownership. The dependent variable and the ownership variables are standardized by subtracting the mean and dividing by the standard deviation. Standard errors are double-clustered at the stock and month levels. t-statistics are presented in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample covers the period January 2000 to December 2015.

Panel A: OLS Regressions, Full Sample

Dependent Variable:	Daily Volatility (t)							
Sample:	S&P 500				Russell 3000			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ETF ownership	0.164***	0.163***	0.158***	0.077***	0.087***	0.078***	0.079***	0.053***
	(6.446)	(6.415)	(6.436)	(6.072)	(8.445)	(7.741)	(7.873)	(7.548)
log(Mktcap (t – 1))	0.019	0.010	−0.011	−0.013	−0.077***	−0.101***	−0.114***	−0.064***
	(0.544)	(0.306)	(−0.312)	(−0.912)	(−4.848)	(−6.216)	(−7.258)	(−6.953)
1/Price (t – 1)	2.906***	2.848***	2.821***	1.052***	1.586***	1.603***	1.589***	0.809***
	(4.311)	(4.237)	(4.087)	(3.574)	(10.126)	(10.299)	(9.961)	(8.625)
Amihud (t – 1)	41.750***	44.412***	55.318***	21.548**	2.072***	2.108***	2.322***	1.165***
	(2.728)	(2.897)	(3.265)	(2.540)	(8.524)	(8.635)	(9.327)	(6.717)
Bid-ask spread (t – 1)	−0.601	−0.678	0.422	−0.340	3.194***	3.280***	3.374***	2.467***
	(−0.194)	(−0.217)	(0.135)	(−0.213)	(2.639)	(2.700)	(3.340)	(4.574)
Book-to-market (t – 1)	0.111**	0.114***	0.097**	−0.027	0.118***	0.115***	0.115***	0.006
	(2.527)	(2.614)	(2.242)	(−1.338)	(5.479)	(5.324)	(5.264)	(0.478)

(Continued)

Table IV—*Continued*

Panel A: OLS Regressions, Full Sample								
Dependent Variable:	Daily Volatility (t)							
Sample:	S&P 500				Russell 3000			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Past 12-month returns ($t-1$)	−0.079**	−0.083**	−0.103***	−0.008	0.012	0.023	0.015	0.015
	(−2.398)	(−2.489)	(−3.210)	(−0.497)	(0.751)	(1.452)	(0.926)	(1.587)
Gross profitability ($t-1$)	0.010	0.038	0.010	−0.043	−0.141***	−0.137***	−0.130***	−0.083***
	(0.098)	(0.384)	(0.104)	(−1.105)	(−3.427)	(−3.359)	(−3.094)	(−3.523)
Index fund ownership ($t-1$)		0.014	0.017	0.002		0.016***	0.017***	0.009***
		(1.241)	(1.447)	(0.491)		(3.038)	(3.157)	(3.022)
Active fund ownership ($t-1$)		0.046***	0.045***	0.016***		0.048***	0.049***	0.026***
		(3.594)	(3.517)	(2.983)		(6.570)	(6.668)	(5.945)
Hedge fund ownership ($t-1$)		−0.013	−0.013	−0.026***		−0.027***	−0.025***	−0.023***
		(−1.006)	(−0.989)	(−4.220)		(−5.598)	(−5.262)	(−7.985)
Volatility ($t-1$)				0.291***				0.207***
				(17.597)				(20.102)
Volatility ($t-2$)				0.176***				0.154***
				(9.689)				(22.784)
Volatility ($t-3$)				0.204***				0.177***
				(14.983)				(28.001)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	81,243	81,243	77,675	77,675	418,748	418,748	392,684	392,684
R^2	0.643	0.644	0.644	0.740	0.607	0.608	0.613	0.670

(*Continued*)

Table IV—*Continued*

Panel B: OLS Regressions, Sample Split by Period

Dependent Variable:	Daily Volatility (t)											
Period:	2000 to 2006				2007 to 2008				2009 to 2015			
Sample:	S&P 500		Russell 3000		S&P 500		Russell 3000		S&P 500		Russell 3000	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
ETF ownership	0.082***	0.061***	0.044***	0.033***	0.276***	0.184***	0.079***	0.071***	0.110***	0.063***	0.077***	0.057***
	(3.728)	(4.557)	(5.193)	(5.434)	(5.192)	(3.667)	(4.790)	(4.890)	(4.774)	(5.701)	(6.544)	(7.063)
log(Mktcap ($t-1$))	0.220***	0.123***	0.079***	0.025	−0.513***	−0.378***	−0.576***	−0.467***	−0.149***	−0.102***	−0.163***	−0.133***
	(5.202)	(5.408)	(2.977)	(1.439)	(−4.925)	(−3.568)	(−8.455)	(−7.393)	(−2.871)	(−3.404)	(−6.807)	(−6.897)
1/Price ($t-1$)	3.072***	1.813***	2.004***	1.260***	2.253***	1.171**	1.048***	0.761***	3.896***	2.186***	1.492***	0.826***
	(3.754)	(3.525)	(7.812)	(5.810)	(4.236)	(2.734)	(4.583)	(3.930)	(4.417)	(3.798)	(7.183)	(5.755)
Amihud ($t-1$)	38.208***	25.647**	1.224***	0.992***	34.758	19.100	3.779***	2.848***	249.919***	209.336***	2.761***	1.681***
	(2.965)	(2.565)	(5.916)	(5.842)	(0.833)	(0.530)	(5.611)	(3.608)	(4.631)	(3.507)	(7.334)	(5.436)
Bid-ask spread ($t-1$)	0.410	0.912	2.511***	1.621***	11.121	2.433	10.244***	6.391**	43.292**	29.276***	14.190***	10.889***
	(0.197)	(0.596)	(3.350)	(4.187)	(1.700)	(0.522)	(3.313)	(2.686)	(2.574)	(3.217)	(6.365)	(5.834)
Book-to-market ($t-1$)	−0.036	−0.086**	−0.027	−0.059**	0.269	0.040	0.093	−0.036	0.046	−0.037	0.179***	0.068***
	(−0.615)	(−2.384)	(−0.747)	(−2.226)	(1.486)	(0.231)	(1.190)	(−0.501)	(0.806)	(−0.913)	(5.199)	(2.858)
Past 12-month returns ($t-1$)	−0.112***	−0.056**	−0.029	−0.007	−0.036	0.031	0.189***	0.171***	0.007	0.047	0.095***	0.074***
	(−3.186)	(−2.631)	(−1.642)	(−0.585)	(−0.613)	(0.529)	(7.071)	(6.354)	(0.169)	(1.603)	(5.547)	(5.573)
Gross profitability ($t-1$)	0.290**	0.074	−0.142**	−0.104**	−0.486**	−0.284	−0.064	−0.117	−0.276**	−0.063	−0.078	−0.054
	(2.185)	(1.054)	(−2.527)	(−2.434)	(−2.091)	(−1.296)	(−0.654)	(−1.251)	(−2.175)	(−0.841)	(−1.520)	(−1.277)
Index fund ownership ($t-1$)	−0.014	0.001	−0.000	0.001	0.017	0.010	0.029**	0.028**	−0.025*	−0.013*	0.011	0.007
	(−0.935)	(0.117)	(−0.068)	(0.205)	(0.678)	(0.421)	(2.360)	(2.441)	(−1.783)	(−1.680)	(1.502)	(1.189)
Active fund ownership ($t-1$)	0.047***	0.027**	0.058***	0.038***	0.097***	0.066**	0.066***	0.046**	0.010	0.021**	0.046***	0.040***
	(2.657)	(2.469)	(4.499)	(4.142)	(3.106)	(2.502)	(4.364)	(2.788)	(0.497)	(2.053)	(5.006)	(6.157)
Hedge fund ownership ($t-1$)	−0.043**	−0.038***	−0.063***	−0.045***	−0.073*	−0.055	−0.051***	−0.040**	−0.035**	−0.025**	−0.027***	−0.025***
	(−2.214)	(−2.826)	(−9.175)	(−8.850)	(−1.868)	(−1.453)	(−3.148)	(−2.209)	(−2.243)	(−2.565)	(−3.244)	(−4.084)
Volatility ($t-1$)		0.227***		0.160***		0.201***		0.141***		0.126***		0.111***
		(15.165)		(16.477)		(3.624)		(4.116)		(7.373)		(11.971)
Volatility ($t-2$)		0.113***		0.107***		0.135*		0.076***		0.092***		0.098***
		(9.920)		(16.664)		(2.068)		(2.993)		(5.869)		(14.677)
Volatility ($t-3$)		0.155***		0.129***		0.081*		0.076***		0.178***		0.147***
		(10.438)		(13.769)		(1.828)		(4.721)		(12.213)		(18.416)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,731	29,478	160,298	132,219	10,237	8,366	55,549	45,477	36,349	34,338	197,867	184,458
R^2	0.656	0.709	0.629	0.661	0.780	0.794	0.732	0.739	0.657	0.641	0.609	0.576

Next, we test whether ETF ownership captures a different effect from the ownership of other institutional investors. Among these, open-end mutual funds are the most similar to ETFs as they also tend to receive flows on a daily basis. Because hedge funds are also likely to trade at a high frequency, similar to ETF arbitrageurs with which they partly overlap, we also control for hedge fund ownership. In column (2) of Table IV, Panel A, we include lagged ownership of active and index mutual funds, as well as that of hedge funds, measured in the same way as ETF ownership (and standardized). The coefficients on both mutual fund ownership variables are positive (significant only for active funds), whereas hedge fund ownership is insignificant. The point estimates on both mutual fund ownership variables are significantly smaller in magnitude than the slope on ETF ownership, which does not change. Thus, ETF ownership appears to have an independent and stronger tie to volatility.

In column (3) of Panel A, we replicate the analysis of column (2) using the subsample for which three lags of the dependent variable are available. We find that the slope on ETF ownership is not materially impacted. In column (4), we include three lags of volatility to address the concern that the persistence in volatility could introduce reverse causality. The coefficient on ETF ownership is again significant, with a magnitude of 7.7% of a standard deviation.

Extending the universe to smaller stocks (columns (5) to (8)), the relationship between ETF ownership and volatility is somewhat weaker, at 5.3% to 8.7% of a standard deviation. The lower sensitivity of volatility to ETF ownership in a sample that is dominated by small stocks is consistent with the channel behind the liquidity trading hypothesis. In the secondary market, arbitrageurs can minimize transaction costs by concentrating on larger stocks in the ETF baskets when constructing the replicating portfolio. Such behavior, called "optimized replication," can explain why smaller stocks inherit fewer shocks from the ETF market.

To confirm that the relation of interest is a pervasive phenomenon throughout the sample period, in Table IV, Panel B, we conduct the analysis separately for the 2000 to 2006 (early period), 2007 to 2008 (crisis period), and 2009 to 2015 (late period) subsamples. In this case, the variables of interest are standardized using their standard deviation in the relevant sample. The crisis period aside, the estimates tend to be somewhat larger in the late period, especially in the specifications that do not control for lagged volatility. Moreover, the effect on Russell 3000 stocks, which are on average smaller, is more important in the late than in the early period. This finding arguably results from ETFs covering smaller and smaller corners of the market. Overall, the evidence is in line with the idea that, as ETFs become a more important asset class, their impact on the underlying securities becomes stronger. The most striking result is that during the crisis, the effect of interest is about three times as large as in the other periods. Thus, consistent with the liquidity trading hypothesis, market liquidity appears to influence the effect of ETFs on the underlying securities. The crisis led to an illiquid market, and hence ETF arbitrage had a larger impact on stock prices.

In Table IA.IV in the Internet Appendix, we study whether the presence of ETFs leads to increased tail risk for the underlying stocks. We find that stocks with higher ETF ownership display larger negative skewness during times of market stress, as measured by the VIX index as well as by aggregate selloffs of ETFs by institutional investors. These results suggest that not only do ETFs increase stock volatility on a daily basis, but they can also destabilize stock prices in periods of market turmoil, when they experience a negative order imbalance.

B. Identification Using a Quasi-Natural Experiment

Identification based on cross-sectional and time series variation in ETF ownership, which underlies the OLS results in Table IV, can be problematic if the controls and fixed effects fail to capture characteristics that codetermine ETF ownership and volatility. Accordingly, we next present results using a more robust identification approach.

Chang, Hong, and Liskovich (2015) propose an identification strategy that exploits the exogenous variation in membership in the Russell 1000 and Russell 2000 indexes. In our implementation, we closely follow Appel, Gormley, and Keim (2015), who cast this experiment within an IV framework.[14]

The Russell 1000 index comprises the top 1,000 stocks by market capitalization, while the Russell 2000 includes the next 2,000 stocks. Russell Inc. reconstitutes the indexes on the last Friday of June each year, based only on end-of-May stock capitalization. Discretion is not involved in index assignment, and index composition remains constant for the rest of the year. For stocks in close proximity to the cutoff, changes in index membership are random events after we control for the assignment variable, namely, market capitalization, because they result from random variation in stock prices at the end of May.

Chang, Hong, and Liskovich (2015) show that, although the amount of passive assets benchmarked to the Russell 1000 is 2 to 3.5 times larger than that tracking the Russell 2000, the weights of the top stocks in the Russell 2000 are about 10 times larger than those of the bottom stocks in the Russell 1000. Consequently, a significantly larger amount of passive money tracks the top Russell 2000 stocks than the bottom Russell 1000 stocks.

Figure 3, Panel A, provides evidence consistent with the latter claim in the context of ETFs. The figure plots average ETF ownership as a function of market capitalization rankings for the Russell 3000 universe, in bins of 10 stocks, for 500 stocks to the right and 500 to the left of the cutoff (the 1,000th position). We note that stocks immediately after the cutoff appear to display discontinuously higher ownership than stocks immediately to the left. Some evidence of discontinuity is also present for index funds (Figure 3, Panel B), although no such evidence exists in the case of active funds (Figure 3, Panel C)

[14] Other papers that exploit the Russell reconstitution include Cao, Gustafson, and Velthuis (2014), Crane, Michenaud, and Weston (2014), Lu (2014), Mullins (2014), and Fich, Harford, and Tran (2015).

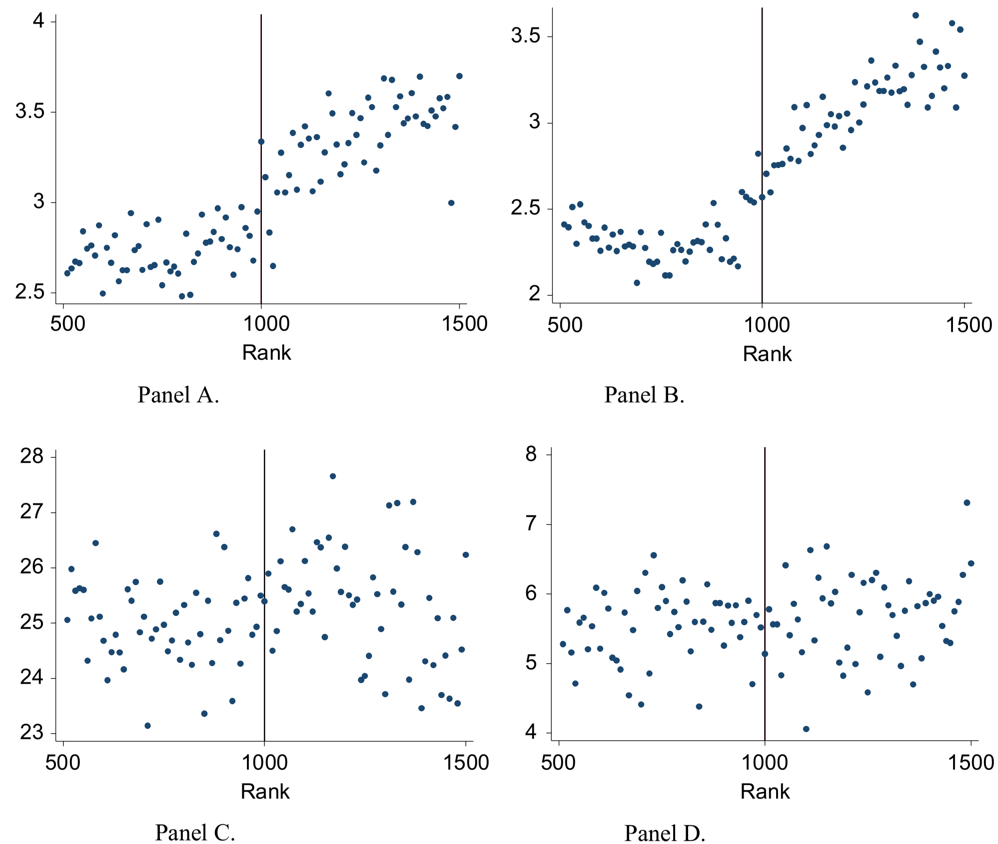


Figure 3. Fund ownership around the Russell cutoff. The figure reports average ownership (in %) by ETFs (Panel A), index funds (Panel B), active funds (Panel C), and hedge funds (Panel D) for stocks ranked by market capitalization and included in the Russell 3000. The average is computed first by ranking over time, and then ranking across bins of 10 stocks. The vertical line denotes the 1,000th rank. The sample covers the period January 2000 to May 2007. (Color figure can be viewed at wileyonlinelibrary.com)

and hedge funds (Figure 3, Panel D). The index fund evidence motivates us to control for index fund ownership in our tests and to carry out additional tests to ensure that ETF ownership, and not index fund ownership, is driving our findings.

We carry out two-stage least squares estimation. In each stage, we run our regressions on two groups of stocks: those that in May, before index reconstitution, are in the Russell 1000 and those that are in the Russell 2000. The sample composition remains constant for all months between June, the first month after index reconstitution, and May of the next year. The first stage consists of a regression of ETF ownership on an indicator variable for whether the stock switches index membership in June. For the Russell 1000 sample, the indicator variable flags stocks that switch to the Russell 2000, and vice versa for the Russell 2000 sample; the dummy captures a switch to the Russell 1000. In regression form, the first stage is given by

$$ETF\ Ownership_{it} = \alpha + \beta * I(Switched\ )_{it} + Controls + \ Fixed\ Effects + \varepsilon_{it}. \quad (4)$$

In the second stage, for the same two groups of stocks, we regress volatility on the fitted value of ETF ownership from the first stage. This regression is given by

$$Volatility_{it} = \alpha + \beta * ETF\ \widehat{Ownership}_{it} + Controls + \ Fixed\ Effects + u_{it}. \quad (5)$$

In addition to the control for market capitalization, that is, the assignment variable, we include the same set of controls as in the OLS regressions. Time fixed effects are included in the regression, but we do not include stock fixed effects because identification in this experiment is inherently cross-sectional, that is, identification results from comparing switchers to nonswitchers in a given time period. Standard errors are double-clustered at the stock and month levels. We standardize the ownership variables and volatility in the relevant samples to ease interpretation. We experiment with different specifications of the ranking variable: first-, second-, and third-degree polynomials.[15]

The identifying assumption that needs to be satisfied is the exclusion restriction—the requirement that the event affects the outcome variable only through the treatment variable. In our context, this restriction is the condition that a switch in index membership affects volatility only through ETF ownership. Later in this section, we discuss the reasons this assumption might fail, and we show that this concern does not appear to be relevant in our context.

We conduct our analysis at the monthly frequency because each month we have a different observation on the dependent variable (i.e., daily volatility). The first index reconstitution in our sample occurs in May 2000. Until the June 2006 reconstitution, the cutoff for reclassification was simply the 1,000th position in terms of market capitalization. Thus, in our main tests, we include end-of-month data between June 2000 and May 2007. We consider several

[15] Results for the second- and third-degree polynomials are reported in Table IA.VII in the Internet Appendix.

bandwidths: 100, 200, 300, 400, and 500 stocks on each side of the cutoff. The validity of the Russell experiment is questionable after 2006, when Russell Inc. adopted a banding rule whereby stocks switch from their current index only if they move beyond a 5% range around the market capitalization percentile of the 1,000th stock. As expected, switches are more frequent before the introduction of the banding rule (Table IA.V in the Internet Appendix). In Table IA.VII in the Internet Appendix, we show that our results continue to hold in the longer sample period, but are noisier in the late subsample (Table IA.VIII in the Internet Appendix). Given that during the migration of stocks across indexes, spurious effects on volatility may occur, in Table IA.VI in the Internet Appendix, we report results excluding the months of May and June. The conclusions remain unaffected.

Table V, Panel A, presents results of the first-stage regressions. We separately consider stocks that belong to the Russell 1000 before index reconstitution (columns (1) to (5)) and stocks that belong to the Russell 2000 before index reconstitution (columns (6) to (10)). The results of this test show that switching indexes has a strong effect on ETF ownership. The slope on the switch indicator in column (1) suggests that ETF ownership in the 12 months after reconstitution increases for those stocks that switch to the Russell 2000 by about 19.6% of a standard deviation. Across columns (1) through (5), the average effect is approximately 35.7%. Column (6) focuses on stocks that start out in the Russell 2000 in May prior to reconstitution, with the same bandwidth (100 stocks to the left and 100 stocks to the right of the cutoff between the Russell indexes). For the stocks that switch to the Russell 1000 after reconstitution, ETF ownership decreases by about 18% of a standard deviation. Across columns (6) to (10), the average estimate is about –33.9%. The strong statistical significance in the first stage reassures us of the validity of the instrument.

Table V, Panel B, reports the second-stage estimates of the effect of ETF ownership on volatility in the next month. Analogous to the layout in Panel A, the instruments are indicators for a switch to either index, and the sample is restricted to members of either index pre-reconstitution. The effect of ETF ownership on volatility is significant across all samples and bandwidths. The coefficients range between 16.9% and 71.7%, or about 31% of a standard deviation on average.

The larger IV estimates from Table V suggest that the endogeneity of ETF ownership induces a negative bias in the OLS estimates in Table IV. A negative bias can occur if, for example, higher ETF ownership signals companies that belong to multiple indexes, which have less volatile stocks because they are more established companies. In Section III.C, we provide a more detailed discussion of the economic magnitude of these estimates.

Finally, we come back to assessing the validity of the exclusion restriction. This assumption is violated if a correlated omitted variable exists that varies with index switches and affects volatility. For example, a violation occurs if, after appearing among the top stocks in the Russell 2000, a firm becomes more visible to investors. Prices can then react more quickly to fundamental information and returns become more volatile, as Andrei and Hasler (2015)

Table V

Quasi-Natural Experiment Based on the Russell Index Reconstitution

The table reports estimates from a quasi-natural experiment that relies on the reconstitution of the Russell 1000 and Russell 2000 indexes. The sample is at the stock-month level. Panel A presents the first-stage results where the dependent variable is ETF ownership. The explanatory variables (instruments) are a dummy for inclusion in the Russell 2000, for stocks in the Russell 1000 before index reconstitution (columns (1) to (5)), and a dummy for inclusion in the Russell 1000, for stocks in the Russell 2000 before index reconstitution (columns (6) to (10)). Stocks are ranked by market capitalization in May of each year. Columns (1) to (5) and (6) to (10) present bandwidths ranging from 100 to 500 stocks. The same stocks enter the sample in June after index reconstitution and remain in the sample until the following May, except if delisted. The controls in all panels are logged market capitalization, lagged inverse share price, lagged Amihud (2002) ratio, lagged average bid-ask spread, lagged book-to-market ratio, lagged past 12-month returns, lagged gross profitability (as in Novy-Marx (2013)), lagged volatility, index fund ownership, active fund ownership, hedge fund ownership, and a linear specification of the ranking variable (not reported). In Panels B, C, and D, the dependent variable is daily stock volatility (computed using daily returns within a month). The main explanatory variable is instrumented ETF ownership. The instruments are the switching indicators. In Panel C, the main explanatory variable is an interaction between the dummy variables for index inclusion and average ETF ownership (for Russell 2000 stocks). Panel C also includes controls for average ETF ownership as well as average ownership for index funds, active funds, and hedge funds (for Russell 2000 stocks). In Panel D, the main explanatory variable is an interaction between the dummy variable for Russell 1000 inclusion and the stock-level ratio of ETF ownership to other fund (index, active, and hedge fund) ownership in May before reconstitution. The dependent variable and the ownership variables are standardized by subtracting the mean and dividing by the standard deviation. Standard errors are double-clustered at the stock and month levels. t-statistics are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample covers the period June 2000 to May 2007.

Panel A: First-Stage Regressions, First-Degree Polynomial										
Dependent Variable:	ETF Ownership									
Instrument:	Switch to the Russell 2000					Switch to the Russell 1000				
Bandwidth:	± 100 (1)	± 200 (2)	± 300 (3)	± 400 (4)	± 500 (5)	± 100 (6)	± 200 (7)	± 300 (8)	± 400 (9)	± 500 (10)
Switch indicator	0.196***	0.294***	0.407***	0.456***	0.433***	−0.180***	−0.405***	−0.397***	−0.356***	−0.359***
	(6.214)	(7.122)	(7.887)	(8.986)	(9.099)	(−3.210)	(−8.935)	(−10.052)	(−8.246)	(−8.581)
log(Mktcap $(t-1)$)	−0.193***	−0.284***	−0.231***	−0.191***	−0.195***	−0.184***	−0.090***	−0.094***	−0.097***	−0.100***
	(−3.259)	(−5.868)	(−4.763)	(−4.894)	(−5.366)	(−5.925)	(−5.078)	(−5.343)	(−5.946)	(−5.373)
1/Price $(t-1)$	−0.680**	−0.449	−0.797***	−0.652***	−0.746***	−0.184	−0.084	0.039	−0.566***	−0.692***
	(−2.488)	(−1.611)	(−3.568)	(−3.812)	(−4.668)	(−0.863)	(−0.598)	(0.294)	(−6.409)	(−9.377)

(Continued)

Table V—*Continued*

Panel A: First-Stage Regressions, First-Degree Polynomial										
Dependent Variable:	ETF Ownership									
Instrument:	Switch to the Russell 2000					Switch to the Russell 1000				
Bandwidth:	± 100 (1)	± 200 (2)	± 300 (3)	± 400 (4)	± 500 (5)	± 100 (6)	± 200 (7)	± 300 (8)	± 400 (9)	± 500 (10)
Amihud ($t-1$)	−6.708***	−10.189***	−11.042***	−10.321***	−10.819***	−18.391***	−18.242***	−18.757***	−15.080***	−10.444***
	(−6.112)	(−8.164)	(−7.266)	(−7.338)	(−7.485)	(−11.623)	(−13.283)	(−11.000)	(−8.550)	(−9.445)
Bid-ask spread ($t-1$)	−2.787	−1.098	−0.755	−2.438**	−4.318***	6.201***	3.892**	2.989***	2.989***	1.949**
	(−1.354)	(−1.088)	(−0.785)	(−2.383)	(−3.833)	(2.853)	(2.203)	(2.742)	(3.058)	(2.005)
Book-to-market ($t-1$)	0.037	−0.008	0.023	0.022**	0.017**	0.283***	0.336***	0.318***	0.351***	0.353***
	(1.403)	(−0.416)	(1.628)	(2.029)	(2.009)	(7.261)	(13.789)	(18.003)	(19.998)	(19.685)
Past 12-month returns ($t-1$)	0.045	0.039*	0.024	0.052**	0.016	−0.064***	−0.019*	−0.020**	−0.013	−0.017*
	(1.264)	(1.868)	(1.171)	(2.386)	(0.758)	(−2.928)	(−1.963)	(−2.155)	(−1.400)	(−1.692)
Gross profitability ($t-1$)	−0.160***	−0.171***	−0.162***	−0.175***	−0.128***	−0.071	−0.101***	−0.042**	−0.007	−0.004
	(−4.142)	(−4.685)	(−5.130)	(−5.889)	(−3.965)	(−1.239)	(−3.836)	(−2.185)	(−0.430)	(−0.259)
Volatility ($t-1$)	−0.034***	−0.042***	−0.033***	−0.028***	−0.027***	−0.133***	−0.124***	−0.132***	−0.127***	−0.124***
	(−4.816)	(−5.069)	(−4.144)	(−3.816)	(−3.280)	(−13.026)	(−16.633)	(−14.908)	(−15.266)	(−13.747)
Index fund ownership ($t-1$)	0.116***	0.163***	0.165***	0.179***	0.197***	0.138***	0.159***	0.165***	0.163***	0.167***
	(12.537)	(16.642)	(15.637)	(19.056)	(20.545)	(7.803)	(8.066)	(9.682)	(10.292)	(10.655)
Active fund ownership ($t-1$)	−0.093***	−0.049***	−0.022***	−0.011	−0.020***	0.095***	0.066***	0.057***	0.056***	0.068***
	(−9.811)	(−7.210)	(−3.236)	(−1.569)	(−3.006)	(5.916)	(8.050)	(8.366)	(8.046)	(10.713)
Hedge fund ownership ($t-1$)	−0.118***	−0.105***	−0.102***	−0.100***	−0.104***	−0.093***	−0.106***	−0.080***	−0.070***	−0.072***
	(−11.806)	(−20.185)	(−23.143)	(−22.697)	(−20.998)	(−9.375)	(−12.085)	(−16.223)	(−19.097)	(−21.804)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,857	10,103	16,101	22,303	28,599	6,495	12,735	18,293	23,647	29,060

(*Continued*)

Table V—*Continued*

Panel B: Second-Stage Regressions, First-Degree Polynomial										
Dependent Variable:	Daily Volatility (t)									
Instrument:	Switch to the Russell 2000					Switch to the Russell 1000				
Bandwidth:	± 100 (1)	± 200 (2)	± 300 (3)	± 400 (4)	± 500 (5)	± 100 (6)	± 200 (7)	± 300 (8)	± 400 (9)	± 500 (10)
ETF ownership	0.512**	0.384***	0.268***	0.176**	0.249***	0.717**	0.249***	0.198***	0.222***	0.169***
	(2.224)	(2.649)	(3.269)	(2.631)	(3.656)	(2.409)	(3.572)	(3.252)	(3.486)	(3.394)
log(Mktcap ($t-1$))	−0.346***	−0.404***	−0.478***	−0.512***	−0.533***	−0.606***	−0.731***	−0.766***	−0.761***	−0.728***
	(−3.995)	(−5.576)	(−7.079)	(−7.366)	(−7.322)	(−7.121)	(−9.075)	(−9.783)	(−9.504)	(−9.266)
1/Price ($t-1$)	2.758***	1.815***	1.550***	1.326***	1.461***	1.759***	1.395***	1.322***	1.411***	1.435***
	(5.928)	(6.649)	(8.570)	(8.651)	(9.395)	(6.464)	(6.511)	(8.198)	(8.382)	(9.383)
Amihud ($t-1$)	1.241	5.238**	4.392***	5.737***	6.139***	13.736**	6.568***	7.135***	7.249***	6.084***
	(0.493)	(2.413)	(2.876)	(3.501)	(3.845)	(2.251)	(3.293)	(5.038)	(6.163)	(7.869)
Bid-ask spread ($t-1$)	−0.484	0.696	−2.117	−2.833	−3.004	−12.445***	−9.365***	−8.741***	−9.558***	−10.449***
	(−0.163)	(0.319)	(−1.293)	(−1.654)	(−1.571)	(−3.084)	(−2.662)	(−3.349)	(−3.411)	(−3.641)
Book-to-market ($t-1$)	0.063**	0.082***	0.095***	0.088***	0.109***	−0.291***	−0.248***	−0.245***	−0.235***	−0.184***
	(2.113)	(3.117)	(4.458)	(4.376)	(5.189)	(−3.162)	(−4.827)	(−5.323)	(−5.208)	(−4.293)
Past 12-month returns ($t-1$)	−0.014	−0.029	0.027	0.008	0.038	0.309***	0.156***	0.152***	0.153***	0.140***
	(−0.227)	(−0.735)	(0.749)	(0.214)	(1.101)	(5.908)	(3.318)	(4.150)	(4.738)	(5.024)
Gross profitability ($t-1$)	0.004	0.049	0.064**	0.035	0.069***	0.173**	0.109***	0.080***	0.072***	0.033
	(0.051)	(1.084)	(2.167)	(1.265)	(2.852)	(2.421)	(3.443)	(3.176)	(3.339)	(1.500)
Volatility ($t-1$)	0.229***	0.263***	0.260***	0.267***	0.270***	0.376***	0.290***	0.287***	0.297***	0.294***
	(14.041)	(19.465)	(20.091)	(18.905)	(17.987)	(9.317)	(17.415)	(18.014)	(18.037)	(19.731)
Index fund ownership ($t-1$)	−0.002	−0.022	−0.013	0.001	−0.020	−0.059	−0.003	0.005	0.001	0.010
	(−0.074)	(−0.795)	(−0.807)	(0.073)	(−1.354)	(−1.461)	(−0.238)	(0.444)	(0.043)	(0.974)
Active fund ownership ($t-1$)	0.105***	0.083***	0.082***	0.088***	0.092***	0.024	0.093***	0.098***	0.092***	0.088***
	(3.878)	(5.868)	(9.452)	(9.880)	(11.305)	(0.957)	(10.297)	(12.991)	(12.698)	(12.378)
Hedge fund ownership ($t-1$)	0.088***	0.076***	0.054***	0.048***	0.050***	0.075**	0.013	0.011	0.011	0.011
	(2.778)	(4.236)	(5.291)	(5.997)	(6.117)	(2.308)	(1.061)	(1.228)	(1.332)	(1.487)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,857	10,103	16,101	22,303	28,599	6,495	12,735	18,293	23,647	29,060

(*Continued*)

Table V—*Continued*

Panel C: Identification from Time Series Variation in Aggregate Fund Ownership

Dependent Variable:	Daily Stock Volatility									
Instrument:	Switch to the Russell 2000					Switch to the Russell 1000				
Bandwidth:	± 100	± 200	± 300	± 400	± 500	± 100	± 200	± 300	± 400	± 500
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ETF ownership × Switch	0.231	0.401**	0.351**	0.219	0.216	−0.287*	−0.431***	−0.496***	−0.592***	−0.671***
	(1.289)	(2.450)	(2.430)	(1.574)	(1.390)	(−1.866)	(−3.849)	(−4.169)	(−4.068)	(−3.769)
Index funds ownership × Switch	−0.087***	−0.060**	−0.053**	−0.046**	−0.049**	0.067***	0.016	−0.011	−0.006	−0.008
	(−3.271)	(−2.493)	(−2.461)	(−2.248)	(−2.196)	(3.201)	(0.876)	(−0.707)	(−0.358)	(−0.399)
Active funds ownership × Switch	0.237***	0.041	0.046	0.053	0.056	−0.092	0.020	0.083	0.082	0.125*
	(3.136)	(0.657)	(0.763)	(0.856)	(0.925)	(−1.554)	(0.358)	(1.337)	(1.404)	(1.770)
Hedge funds ownership × Switch	0.176	0.079	−0.059	−0.041	−0.086	−0.121	−0.053	−0.093	−0.003	0.040
	(1.237)	(0.671)	(−0.566)	(−0.384)	(−0.758)	(−0.972)	(−0.710)	(−1.320)	(−0.033)	(0.483)
Switch indicator	11.776**	12.236***	7.772**	6.098*	4.730	−9.630***	−10.797***	−11.743***	−11.380***	−11.272***
	(2.448)	(3.404)	(2.431)	(1.943)	(1.282)	(−2.716)	(−4.242)	(−4.811)	(−4.108)	(−3.645)
log(Mktcap (t − 1))	−0.570***	−0.590***	−0.612***	−0.619***	−0.652***	−0.849***	−0.846***	−0.877***	−0.860***	−0.837***
	(−6.259)	(−8.189)	(−9.086)	(−9.037)	(−9.856)	(−9.943)	(−9.605)	(−11.168)	(−11.611)	(−11.912)
1/Price (t − 1)	2.117***	1.640***	1.373***	1.252***	1.440***	1.649***	1.492***	1.374***	1.324***	1.358***
	(7.012)	(9.095)	(10.241)	(8.823)	(8.869)	(8.182)	(6.777)	(6.884)	(7.106)	(7.813)
Amihud (t − 1)	−1.524	0.713	1.136	3.570**	3.223**	−1.118	0.228	2.469***	3.405***	4.304***
	(−0.897)	(0.518)	(1.282)	(2.560)	(2.508)	(−0.742)	(0.161)	(2.643)	(4.281)	(7.037)
Bid-ask spread (t − 1)	−0.658	0.442	−1.969	−2.372	−3.742**	−8.000***	−9.079***	−8.493***	−9.282***	−10.094***
	(−0.203)	(0.242)	(−1.290)	(−1.384)	(−2.210)	(−2.775)	(−3.007)	(−3.538)	(−3.718)	(−3.976)
Book-to-market (t − 1)	0.079**	0.081***	0.099***	0.091***	0.113***	−0.066	−0.171***	−0.188***	−0.150***	−0.115***
	(2.186)	(2.873)	(4.045)	(4.015)	(5.346)	(−1.144)	(−3.108)	(−3.684)	(−3.385)	(−2.750)
Past 12-month returns (t − 1)	0.062	0.019	0.055	0.045	0.065*	0.274***	0.160***	0.161***	0.169***	0.166***
	(1.024)	(0.528)	(1.542)	(1.203)	(1.880)	(5.727)	(3.160)	(4.152)	(4.701)	(4.935)
Gross profitability (t − 1)	−0.086*	−0.021	0.023	0.010	0.044**	0.117***	0.083***	0.072***	0.074***	0.040**
	(−1.748)	(−0.555)	(0.889)	(0.409)	(2.082)	(3.033)	(3.121)	(3.507)	(3.809)	(2.008)
Volatility (t − 1)	2.214***	1.596***	1.339***	1.371***	1.255***	1.258**	1.217***	1.343***	1.243***	1.295***
	(4.114)	(3.653)	(4.186)	(4.542)	(4.535)	(2.458)	(3.244)	(3.377)	(3.303)	(3.966)
Index fund ownership (t − 1)	0.793***	0.852***	0.992***	1.053***	1.068***	1.012***	1.267***	1.280***	1.256***	1.190***
	(5.733)	(7.937)	(10.225)	(10.143)	(11.490)	(6.437)	(11.668)	(13.062)	(13.154)	(13.254)
Active fund ownership (t − 1)	0.676***	0.817***	0.612***	0.722***	0.606***	0.172	−0.173	0.040	0.071	0.144
	(3.480)	(6.219)	(5.629)	(6.892)	(6.468)	(1.031)	(−1.429)	(0.360)	(0.685)	(1.523)
Hedge fund ownership (t − 1)	0.221***	0.247***	0.251***	0.263***	0.262***	0.275***	0.255***	0.261***	0.268***	0.271***
	(17.959)	(21.347)	(21.154)	(20.220)	(19.184)	(16.202)	(15.773)	(18.007)	(18.555)	(19.530)
Time trend, interacted with switch	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,858	10,107	16,107	22,315	28,615	6,499	12,744	18,304	23,663	29,077
R^2	0.634	0.609	0.604	0.594	0.594	0.495	0.473	0.476	0.484	0.489

(*Continued*)

Table V—*Continued*

Panel D: Ratio of ETF Ownership to Ownership by Other Fund Types

Dependent Variable:	Daily Stock Volatility														
Instrument:	Switch to the Russell 1000														
Ratio:	ETF/Index Fund					ETF/Active Fund					ETF/Hedge Fund				
Bandwidth:	± 100 (1)	± 200 (2)	± 300 (3)	± 400 (4)	± 500 (5)	± 100 (6)	± 200 (7)	± 300 (8)	± 400 (9)	± 500 (10)	± 100 (11)	± 200 (12)	± 300 (13)	± 400 (14)	± 500 (15)
Ratio × Switch	−0.033	−0.091**	−0.094**	−0.109***	−0.106***	−0.352	−0.268	−0.391**	−0.372***	−0.345***	−0.040***	−0.034***	−0.018**	−0.020***	−0.014*
	(−0.526)	(−2.143)	(−2.418)	(−3.025)	(−3.047)	(−1.485)	(−1.441)	(−2.551)	(−2.772)	(−2.857)	(−2.748)	(−3.251)	(−2.417)	(−2.674)	(−1.936)
Switch indicator	−0.081**	−0.042	−0.038	−0.022	−0.010	−0.089**	−0.064*	−0.052**	−0.035	−0.024	−0.061*	−0.067**	−0.078***	−0.059***	−0.048***
	(−2.159)	(−1.482)	(−1.580)	(−0.960)	(−0.509)	(−2.284)	(−1.955)	(−2.033)	(−1.547)	(−1.249)	(−1.891)	(−2.471)	(−3.347)	(−2.874)	(−2.925)
Ratio	−0.203***	−0.202***	−0.211***	−0.206***	−0.208***	−0.689***	−0.946***	−0.870***	−0.880***	−0.902***	−0.013	−0.020**	−0.029***	−0.029***	−0.036***
	(−3.397)	(−5.489)	(−7.512)	(−9.611)	(−9.998)	(−3.043)	(−8.164)	(−9.967)	(−10.959)	(−11.838)	(−1.059)	(−2.376)	(−4.904)	(−5.143)	(−6.674)
log(Mktcap $(t-1)$)	−0.597***	−0.577***	−0.590***	−0.567***	−0.553***	−0.587***	−0.567***	−0.567***	−0.531***	−0.526***	−0.608***	−0.581***	−0.585***	−0.569***	−0.555***
	(−8.027)	(−7.372)	(−8.590)	(−8.970)	(−8.978)	(−7.966)	(−7.257)	(−8.267)	(−8.463)	(−8.739)	(−8.493)	(−7.355)	(−8.470)	(−8.969)	(−9.151)
1/Price $(t-1)$	1.459***	1.599***	1.509***	1.395***	1.461***	1.787***	1.486***	1.456***	1.399***	1.505***	1.653***	1.778***	1.523***	1.324***	1.360***
	(6.483)	(9.247)	(11.075)	(10.267)	(12.839)	(7.141)	(5.305)	(6.607)	(6.956)	(8.933)	(6.952)	(10.698)	(11.255)	(9.977)	(11.778)
Amihud $(t-1)$	−1.319***	−1.802***	−1.874***	−1.479***	−0.584	−1.850	0.867	−0.488	−0.483	0.617	−3.615**	−4.469***	−3.340***	−2.023**	1.087
	(−2.704)	(−4.049)	(−4.139)	(−3.790)	(−1.377)	(−0.852)	(0.661)	(−0.451)	(−0.731)	(0.954)	(−2.083)	(−3.219)	(−2.969)	(−2.303)	(1.347)
Bid-ask spread $(t-1)$	1.040	1.791	0.164	0.049	−0.678	1.334	0.337	−0.561	0.536	−1.138	−0.850	−0.619	−2.713	−2.053	−3.577***
	(0.416)	(0.724)	(0.078)	(0.031)	(−0.522)	(0.538)	(0.137)	(−0.271)	(0.358)	(−0.928)	(−0.287)	(−0.220)	(−1.228)	(−1.235)	(−2.807)
Book-to-market $(t-1)$	−0.030	−0.113**	−0.129***	−0.078**	−0.057	−0.038	−0.081	−0.093**	−0.050	−0.049	−0.058	−0.125**	−0.130***	−0.111***	−0.089**
	(−0.662)	(−2.254)	(−3.280)	(−2.024)	(−1.553)	(−0.834)	(−1.613)	(−2.451)	(−1.340)	(−1.262)	(−1.333)	(−2.565)	(−3.397)	(−3.193)	(−2.406)
Past 12-month returns $(t-1)$	0.240***	0.137***	0.147***	0.154***	0.151***	0.246***	0.136***	0.146***	0.157***	0.153***	0.247***	0.135***	0.144***	0.153***	0.148***
	(6.969)	(2.851)	(3.459)	(3.877)	(4.279)	(7.335)	(2.844)	(3.404)	(3.862)	(4.265)	(7.377)	(2.753)	(3.364)	(3.830)	(4.254)
Gross profitability $(t-1)$	0.148***	0.128***	0.100***	0.126***	0.097***	0.094*	0.094***	0.084***	0.102***	0.052**	0.139***	0.131***	0.114***	0.134***	0.090***
	(3.262)	(4.154)	(4.087)	(6.414)	(4.576)	(1.907)	(2.923)	(3.381)	(4.973)	(2.395)	(2.937)	(4.219)	(4.712)	(6.920)	(4.301)
Volatility $(t-1)$	0.298***	0.264***	0.264***	0.271***	0.274***	0.293***	0.250***	0.256***	0.261***	0.265***	0.299***	0.268***	0.268***	0.272***	0.275***
	(19.127)	(16.553)	(18.653)	(20.554)	(21.726)	(19.193)	(16.097)	(18.566)	(20.066)	(20.100)	(20.070)	(16.962)	(19.600)	(21.771)	(22.366)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,609	10,850	15,520	19,957	24,355	5,459	10,574	15,067	19,496	23,874	5,532	10,635	15,194	19,478	23,609
R^2	0.437	0.396	0.401	0.402	0.414	0.433	0.391	0.397	0.396	0.404	0.435	0.396	0.399	0.401	0.409

show. Appel, Gormley, and Keim (2015) find that analyst coverage is largely unaffected after inclusion in the Russell 2000. Similarly, Crane, Michenaud, and Weston (2014) show that a switch to the Russell 2000 does not lead to increased media coverage. These results are important for our study because they suggest that an increase in information production following a switch is not likely to cause the increase in volatility associated with the increase in ETF ownership. Moreover, below we show that ETFs increase mean reversion in stock prices, suggesting that the increase in volatility is due not only to improved price discovery. Therefore, the evidence seems to rule out this specific violation of the exclusion restriction.

Further corroboration of the validity of the exclusion restriction comes from combining the cross-sectional identification from the index-switching experiment with time series variation in ETF ownership. In particular, if the IV estimates are truly measuring the causal effect of ETF ownership, we should observe a stronger impact of index switching on volatility when aggregate ETF ownership is larger. Following this argument, we regress stock-level volatility on the interaction between the index-switching indicator and the (equally weighted) average ETF ownership across Russell 2000 stocks. If the exclusion restriction is satisfied, then switching to the Russell 2000 (Russell 1000) should have a more positive (negative) effect on volatility when ETF ownership is overall higher.[16] The regressions include additional interactions to control for aggregate ownership by other mutual funds (index and active) and hedge funds as well as for a time trend, given that ETF ownership increases over time. We also include the noninteracted variables and the usual stock-level controls and month fixed effects. The estimates in Table V, Panel C, broadly support the validity of the exclusion restriction. In columns (1) to (5), the addition to the Russell 2000 has a larger impact on volatility when average ETF ownership is higher. In two cases, the effect is statistically significant. In all five specifications focusing on the switch to the Russell 1000 (columns (6) to (10)), the decrease in volatility is significantly larger in months when ETF ownership is higher.[17]

One may still wonder to what extent ETF ownership captures an independent effect from that of ownership by other institutions. In particular, the concern here is the possibility that index-switching stocks experience an overall migration of institutional owners, which causes the change in volatility.

[16] Note that we expect a switch to the Russell 2000 (Russell 1000) to increase (decrease) volatility because such a switch increases (decreases) ETF ownership, based on the evidence in Panel A of Table V.

[17] Note that, in the sample consisting of 100 stocks around the cutoff, the effect of active fund ownership is comparable in sign and magnitude to that of ETF ownership. Following index reconstitution, other institutions might rebalance their portfolios. For the stocks closer to the cutoff, which are more likely to engage in rebalancing, this trading activity can affect volatility. We are confident, however, that this effect is separate from the effect of ETF ownership that we document. ETF ownership remains significant as an independent driver of volatility in specifications that also include active fund ownership (e.g., Table V, Panels B and C). Additional reassuring evidence comes from Panel D of Table V.

This concern is relevant given the evidence in Figure 3 suggesting that some discontinuity in ownership is also present for index funds around the cutoff for index reconstitution. To address this issue, in Table V, Panel D, we conduct a formal test of the importance of ETF ownership relative to the ownership of other funds in driving the volatility effect. We study whether stocks switching to the Russell 1000, for which ETF ownership goes down, experience a larger decrease in volatility if in May, before reconstitution, they have a higher ratio of ETF ownership to ownership by other fund types. Reassuringly, the tests in Panel D show that the ownership ratio (with index fund, active fund, or hedge fund ownership in the denominator) magnifies the decrease in volatility for stocks exiting the Russell 2000. For a move to the Russell 2000, which entails an increase in ETF ownership, we do not have a clear-cut prediction on the sign of the interaction between the switch indicator and the ratio of ETF ownership to other funds' ownership. Hence, carrying out the test only for switches to the Russell 1000 makes sense.

Taken together, the evidence in Table V increases confidence in a causal interpretation of the positive relation between ETF ownership and stock-level volatility. This causal evidence provides support for the liquidity trading hypothesis and against the liquidity buffer hypothesis.

C. Discussion of the Economic Magnitude of OLS and IV Estimates

In Table IA.IX in the Internet Appendix, we provide a thorough analysis of the economic magnitude of the effect of ETF ownership on stock-level volatility. Here, we summarize the main conclusions.

For the OLS analysis, we assess the economic magnitude by measuring the shift in volatility relative to the median stock in the sample following a one-standard-deviation change in ETF ownership. The median stock shifts to a place between the 58th and 64th percentiles of the volatility distribution for S&P 500 stocks, where the upper bound coincides with the less accurate specification that does not include lagged volatility. For the Russell 3000 universe, the effects are more contained. The median stocks move to a position between the 55th and 57th percentiles of the distribution.

In interpreting the economic magnitude of the IV estimates, one issue is especially relevant. The estimated IV coefficients reflect a local average treatment effect (LATE; Angrist and Imbens (1995)). In particular, the IV estimate, which results from a natural experiment inducing a switch between a control and a treatment group, measures the effect of treatment only on the units that switch groups because of the outcome of the natural experiment. These units would not otherwise receive treatment. Applied to our context, the IV estimates capture the effect of ETF ownership on the stocks that enter the ETF basket only because of the index switch. These stocks drastically change status from being intensively utilized by arbitrageurs in their replication of the index (when they are at the top of the Russell 2000) to being neglected by arbitrageurs (when they are at the bottom of the Russell 1000), and vice versa for a switch in the other direction. Arguably, this drastic change of status results in a greater impact of a given amount of ETF ownership than for the average stock. This argument can explain why the LATE is larger than the OLS effect.

Also important, the index-switching stocks experience a change in ETF ownership of about 50 bps, on average (Figure 3, Panel A). The amount of "treatment" for these stocks thus corresponds to the actual change in ETF ownership to which they are exposed because of the index switch. A more natural approach in this context is therefore to compute the economic magnitude by multiplying the IV estimates by the change in ETF ownership that results from the index switch. When we do so, the median stock's volatility shifts between the 55th and 65th percentiles (on average, to the 60th percentile). These magnitudes are comparable to those from OLS regressions.

Although economically significant, these effects do not seem out of proportion relative to the average magnitude of ETF ownership in the sample, which is not large.[18] On the other hand, ETF trading volume is an important fraction of total trading volume in the stock market (on average, 32% for the full sample). According to the liquidity trading hypothesis, trading activity in ETFs impounds price shocks, which are then passed on to the prices of the underlying securities via arbitrage. Hence, under this hypothesis, an economically important effect of ETFs on volatility would be consistent with the large volume of ETF trading.[19]

IV. Exploring the Channel: ETF Flows and Arbitrage

In this section, we explore the channel through which ETFs affect volatility. We also attempt to empirically separate the liquidity trading hypothesis from the price discovery hypothesis.

A. ETFs Attract a High-Turnover Clientele

As detailed in Section II, for ETFs to impound a new layer of liquidity shocks, absent the ETFs, liquidity traders must not have directly traded the underlying stocks or gotten indirect access to them through other vehicles, such as futures. In Section II.B, we provide suggestive evidence that ETFs attract higher turnover investors than common stocks. These investors are likely to express their liquidity demand at a higher frequency. If the liquidity trading hypothesis is correct, then this demand should propagate to the underlying stocks through arbitrage activity, exposing the stock prices to a new layer of liquidity shocks. The question, therefore, is whether the stocks in ETF baskets are exposed to this high-turnover clientele through ETF ownership.

To address this question, we study whether the stock-level churn ratio of the investors that trade in a given stock increases with ETF ownership. Using

[18] In 2017, average ETF ownership is about 2.6% in the full sample for S&P 500 stocks (2.8% for Russell 3000 stocks) and about 7% for the S&P 500 sample (7.5% for Russell 3000 stocks).

[19] In Table IA.X in the Internet Appendix, using Ancerno trade-level data, we confirm that institutional investors generate significantly higher turnover when trading stocks than when trading ETFs.

Ancerno trade-level data, for each manager-stock-quarter we compute the adjusted churn ratio as in equation (2). We then aggregate this variable at the quarter-stock level by taking the volume-weighted average across managers. Finally, we regress the stock-quarter-level churn ratio on ETF ownership at the end of the previous quarter. Both the dependent and the main explanatory variables are standardized.

Consistent with the conjecture that ETFs attract high-turnover investors to the underlying securities, in Panel A of Table VI we find a positive and significant relation between ETF ownership and the adjusted churn ratio. In column (2), a one-standard-deviation increase in ETF ownership for an S&P 500 stock is associated with 12% of a standard deviation increase in the investor churn ratio. In line with the results in Table IV, the effect is smaller for the Russell 3000.

The concern about endogeneity of ETF ownership applies here as well. Hence, we also cast the test within the IV framework that relies on the switch across Russell indexes. In Panel B of Table VI, we find broad support for a positive link between ETF ownership and the investor churn ratio. The estimates are mostly significant for additions to the Russell 2000, whereas significance is weaker for deletions from the Russell 2000.

Overall, we interpret this evidence as corroborating the view that a new clientele of high-turnover investors is attracted to the stocks in the ETF baskets. Under a causal interpretation of these regressions, these investors would not trade the stocks if they were not in the ETF portfolios.

B. The Price Impact of ETF Flows

A comparison of Figures 1 and 2 suggests a way to disentangle the two hypotheses. According to the liquidity trading hypothesis (Figure 1), following a demand shock in the ETF market, the prices of the underlying securities should move in the same direction as the initial shock, but then prices should revert to the initial level as the effect of the shock vanishes. In contrast, the price discovery hypothesis (Figure 2) posits that stock prices should stay at the new level after the demand shock. This difference motivates us to study the behavior of stock prices following changes in the demand for ETFs.

In this analysis, we identify demand shocks in the ETF market by measuring daily flows in ETFs. As explained above, ETF flows (redemptions and creations) are the result of AP activity. Stock-level flows are defined as the weighted average of the daily flows in the ETFs that own the stock. The weights are the fraction of ownership in the stock held by each ETF. Daily dollar ETF flows are a fraction of average daily trading volume over the prior month:

$$Flows_{i,t} = \frac{\sum_{j=1}^{J} w_{i,j,t} * \$Flows_{j,t}}{\$Volume_{i,t}}. \tag{6}$$

Focusing on the S&P 500, from the summary statistics in Table II, Panel C, we know that on average (absolute) flows are \$2 million per stock-day (1.5%

Table VI

ETF Ownership and Churn Ratio of Stock Traders

The table presents evidence about the relation between ETF ownership and the churn ratio of the underlying stocks, using Ancerno data. Panel A reports estimates from ordinary least squares (OLS) regressions where the dependent variable is the adjusted churn ratio at the stock-quarter level, computed as the volume-weighted average of the manager-stock-level adjusted churn ratio, which is defined as $Adjusted\ Churn\ Ratio_{k,i,q} = (1 - \frac{|\$Buy\ Trades_{k,i,q} - \$Sell\ Trades_{k,i,q}|}{\$Buy\ Trades_{k,i,q} - \$Sell\ Trades_{k,i,q}}) * \frac{N_{k,i,q}+1}{2}$. The main explanatory variable is ETF ownership. Both variables are standardized. Panel B reports second-stage, IV estimates from a quasi-natural experiment relying on the reconstitution of the Russell 1000 and Russell 2000 indexes. The frequency of the data is quarterly at the stock level. The explanatory variable is ETF ownership instrumented by a dummy for inclusion in the Russell 2000, for stocks in the Russell 1000 before index reconstitution (columns (1) to (5)), and a dummy for inclusion in the Russell 1000, for stocks in the Russell 2000 before index reconstitution (columns (6) to (10)). Stocks are ranked in terms of market capitalization in May of each year. Different ranges of this rank around the cutoff are used for inclusion in the sample: 100 stocks on each side (columns (1) and (6)), 200 stocks on each side (columns (2) and (7)), 300 stocks on each side (columns (3) and (8)), 400 stocks on each side (columns (4) and (9)), and 500 stocks on each side (columns (5) and (10)). The same stocks enter the sample in June after index reconstitution and remain in the sample until May of the next year, except if delistings occur. In all panels, the dependent variable and the ownership variables are standardized by subtracting the mean and dividing by the standard deviation. The controls in Panels A and B include logged market capitalization, the lagged inverse share price, the lagged Amihud (2002) ratio, the lagged average bid-ask spread, the lagged book-to-market ratio, the lagged past 12-month returns, lagged gross profitability (as in Novy-Marx (2013)), the lagged dependent variable in some of the columns, index fund ownership, active fund ownership, and hedge fund ownership. Stock and quarter fixed effects are included. Standard errors are double-clustered at the stock and quarter levels. t-statistics are presented in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample covers the period January 2000 to December 2014.

Panel A: OLS Specification				
Dependent Variable:	Adjusted Churn Ratio (t)			
	(1)	(2)	(3)	(4)
ETF ownership ($t-1$)	0.165***	0.123***	0.084***	0.059***
	(3.511)	(3.330)	(4.950)	(4.359)
log(Mktcap ($t-1$))	0.555***	0.420***	0.361***	0.261***
	(13.870)	(13.023)	(16.231)	(12.898)
1/Price ($t-1$)	0.320	0.215	0.405***	0.295***
	(0.611)	(0.557)	(3.182)	(2.746)

(Continued)

Table VI—*Continued*

Panel A: OLS Specification				
Dependent Variable:	Adjusted Churn Ratio (t)			
	(1)	(2)	(3)	(4)
Amihud ($t-1$)	28.446*	26.453**	0.323	0.189
	(1.825)	(2.024)	(1.350)	(0.936)
Bid-ask spread ($t-1$)	−1.529	−1.095	−1.753	−1.153
	(−0.498)	(−0.400)	(−1.333)	(−1.085)
Book-to-market ($t-1$)	−0.033	−0.013	0.005	0.010
	(−0.655)	(−0.319)	(0.247)	(0.589)
Past 12-month returns ($t-1$)	0.117***	0.111***	0.003	0.017
	(2.664)	(2.895)	(0.205)	(1.341)
Gross profitability ($t-1$)	0.077	0.045	−0.024	−0.021
	(0.505)	(0.382)	(−0.551)	(−0.648)
Index fund ownership ($t-1$)	1.425	1.108	−1.444**	−0.951*
	(0.576)	(0.548)	(−2.045)	(−1.725)
Active fund ownership ($t-1$)	1.361***	0.963***	0.909***	0.581***
	(5.010)	(4.628)	(7.748)	(6.466)
Hedge fund ownership ($t-1$)	−0.600	−0.406	−0.226***	−0.184***
	(−1.637)	(−1.325)	(−2.667)	(−2.879)
Adjusted churn ratio ($t-1$)		0.170***		0.202***
		(18.923)		(15.215)
Quarter fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	Yes	Yes	Yes	Yes
Observations	25,641	25,641	130,267	130,267
R^2	0.469	0.500	0.489	0.529

(*Continued*)

Table VI—*Continued*

Panel B: IV Specification

Dependent Variable:	Adjusted Churn Ratio (t)									
Instrument:	Switch to the Russell 2000					Switch to the Russell 1000				
Bandwidth:	± 100 (1)	± 200 (2)	± 300 (3)	± 400 (4)	± 500 (5)	± 100 (6)	± 200 (7)	± 300 (8)	± 400 (9)	± 500 (10)
Switch indicator	0.741***	0.462***	0.200**	0.051	0.137**	0.321	−0.049	−0.080	−0.025	−0.029
	(3.172)	(3.514)	(2.444)	(0.814)	(2.139)	(1.619)	(−0.785)	(−1.381)	(−0.420)	(−0.497)
log(Mktcap ($t-1$))	0.600***	0.582***	0.484***	0.438***	0.425***	0.286***	0.345***	0.254***	0.294***	0.309***
	(7.095)	(9.580)	(9.725)	(11.642)	(12.825)	(5.245)	(7.868)	(6.475)	(8.569)	(9.834)
1/Price ($t-1$)	1.832***	1.842***	1.558***	1.207***	1.316***	2.360***	1.522***	1.326***	1.432***	1.390***
	(5.270)	(5.424)	(7.152)	(7.551)	(7.701)	(9.152)	(7.769)	(7.564)	(8.199)	(9.642)
Amihud ($t-1$)	−0.215	−1.628	−2.585***	−3.355***	−1.337*	1.257	−3.382	−3.439*	−3.769**	−2.287**
	(−0.116)	(−1.152)	(−2.924)	(−4.536)	(−1.810)	(0.252)	(−1.386)	(−1.845)	(−2.370)	(−2.117)
Bid-ask spread ($t-1$)	−11.618***	−6.860***	−11.585***	−9.646***	−9.865***	−16.447***	−14.999***	−11.984***	−13.128***	−14.539***
	(−3.467)	(−2.883)	(−5.616)	(−5.516)	(−5.801)	(−6.111)	(−8.700)	(−8.748)	(−9.024)	(−9.776)
Book-to-market ($t-1$)	0.094	0.082	0.064**	0.038	0.041**	−0.171**	−0.095**	−0.095**	−0.098**	−0.063**
	(1.203)	(1.605)	(2.028)	(1.559)	(2.560)	(−2.126)	(−2.281)	(−2.543)	(−2.523)	(−2.209)
Past 12-month returns ($t-1$)	−0.141**	−0.001	0.030	0.036	0.060	0.199***	0.127***	0.144***	0.142***	0.134***
	(−2.290)	(−0.020)	(0.691)	(0.882)	(1.621)	(4.132)	(4.570)	(6.062)	(6.782)	(6.519)
Gross profitability ($t-1$)	0.073	0.094*	0.037	−0.015	−0.002	0.135**	−0.043*	−0.054***	−0.040**	−0.033*
	(1.175)	(1.795)	(0.954)	(−0.507)	(−0.084)	(2.501)	(−1.844)	(−3.061)	(−2.109)	(−1.718)
Lagged adjusted churn ratio	0.194***	0.209***	0.211***	0.218***	0.219***	0.292***	0.264***	0.242***	0.236***	0.239***
	(9.353)	(14.526)	(14.598)	(19.861)	(19.900)	(14.108)	(22.414)	(27.167)	(25.865)	(28.048)
Index fund ownership ($t-1$)	0.055*	0.063***	0.084***	0.106***	0.089***	0.026	0.108***	0.106***	0.085***	0.080***
	(1.769)	(3.039)	(6.322)	(11.103)	(7.469)	(0.871)	(7.707)	(7.766)	(6.732)	(6.725)
Active fund ownership ($t-1$)	0.175***	0.119***	0.131***	0.130***	0.135***	0.156***	0.172***	0.160***	0.148***	0.148***
	(7.490)	(9.857)	(14.279)	(13.914)	(15.288)	(11.823)	(13.826)	(15.730)	(15.222)	(15.105)
Hedge fund ownership ($t-1$)	0.136***	0.073***	0.042***	0.027**	0.046***	0.045*	0.008	0.015	0.038***	0.045***
	(4.030)	(3.549)	(2.976)	(2.354)	(4.534)	(1.681)	(0.616)	(1.353)	(3.892)	(5.037)
Quarter fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,840	10,077	16,083	22,286	28,574	6,482	12,726	18,286	23,641	29,054

of daily volume). Hence, flows far exceed the size of the National Best Bid and Offer ($517), which suggests that the demand shock from flows needs to climb up or down the order book, and likely causes prices to move.

On the day flows occur, we expect a price move in the same direction as the flows, irrespective of whether the motive for trade is new information or liquidity demand. To the extent that at least part of the originating shock is due to liquidity demand, a reversal should occur in the subsequent days. To capture this behavior, we regress returns over different horizons (5, 10, 20, and 40 days) on stock-level flows measured on the first day of each of these horizons. We include the usual stock-level controls and time fixed effects, in addition to order imbalance, which is calculated as the dollar value of buy-minus-sell trades from TAQ divided by market capitalization. Order imbalance is a natural control in this context because daily flows in ETFs could merely be a proxy for aggregate demand in the underlying securities, which would induce negative autocorrelation in returns (Chordia and Subrahmanyam (2004)). The standard errors are clustered at the stock-day level.

In Table VII, Panel A, DGTW-adjusted returns (Daniel et al. (1997)) are expressed as percentages, whereas net flows are standardized. From column (1), we note that on a given day, ETF flows and returns move in the same direction. The contemporaneous price move is 12.2 bps for a one-standard-deviation change in net flows for S&P 500 stocks. The high significance is not surprising because flows and returns are measured on the same day (i.e., this regression is not predictive). In addition, we note that the magnitude of the change in prices exceeds the half-spread, which is about 8.5 bps for the sample of large stocks. This magnitude rules out the possibility that flows cause a simple bid-ask bounce.

More relevant to separating the hypotheses of interest, ETF flows predict a reversal of the underlying stocks' prices in the 40 days starting on day t (columns (2) to (5)). This evidence is consistent with the conjecture that the demand shocks in the ETF market add a mean-reverting component to stock prices. As in the prior tables, the absolute effects are smaller in the extended universe of Russell 3000 stocks, but we still observe a complete reversal over 40 days.

In sum, the evidence of a full reversal of the price impact of ETF flows is consistent with the liquidity trading hypothesis, and rules out the conjecture that ETFs are the vehicle of choice for expressing fundamental demand, as posited by the price discovery hypothesis.

In the Internet Appendix, we also provide results from a test that aims to show that nonfundamental demand for ETFs affects stock-level volatility, consistent with the liquidity trading hypothesis. The identifying assumption is that total daily ETF flows contain a part that reflects net influxes of money into the ETFs that hold that stock and a part that is due to the reallocation of capital across the ETFs holding the stock. The former part, which is captured by net stock-level flows, may reflect fundamental information that concerns the securities in the ETF basket. The latter part depends on investors' decision to reshuffle money across the ETFs holding a given stock, which has

Table VII

ETF Flows, Stock Returns, and Price Impact

Panel A DGTW-adjusted cumulative returns (Daniel et al. (1997)) over different horizons starting on day t on stock-level ETF flows on day t. The analysis is at the stock-day level. ETF flows are scaled by market capitalization and standardized. The controls in all panels include logged market capitalization, the lagged inverse share price, the lagged Amihud (2002) ratio, lagged average bid-ask spread, lagged book-to-market ratio, lagged past 12-month returns, lagged gross profitability (as in Novy-Marx (2013)), stock-level order imbalance, and lagged returns. Stock and quarter fixed effects are included in Panel A, and quarter fixed effects are included in Panel B. Standard errors are double-clustered at the stock and time levels. Panel B reports estimates from ordinary least squares (OLS) regressions of price impact on turnover and turnover squared. The sample includes trade-level observations from Ancerno data. Price impact is defined as the percentage difference between the execution price and the price at time of placement. Turnover is measured as dollar volume of a trade as a fraction of daily dollar volume in CRSP. Day and stock fixed effects are included. Standard errors are double-clustered at the stock and time levels. The sample covers the period January 2000 to December 2014. t-statistics are presented in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Panel A: ETF Flows and Stock Returns

Sample:	S&P 500					Russell 3000				
Dependent Variable:	Ret(t) (1)	Ret(t, $t+4$) (2)	Ret(t, $t+9$) (3)	Ret(t, $t+19$) (4)	Ret(t, $t+39$) (5)	Ret(t) (6)	Ret(t, $t+4$) (7)	Ret(t, $t+9$) (8)	Ret(t, $t+19$) (9)	Ret(t, $t+39$) (10)
ETF flows (t)	0.122***	0.111***	0.062***	0.046*	0.015	0.040***	0.034***	0.027***	0.016*	−0.000
	(13.433)	(7.469)	(3.233)	(1.848)	(0.416)	(15.069)	(6.649)	(3.995)	(1.669)	(−0.006)
log(Mktcap ($t-1$))	0.048***	0.039***	0.033**	0.040	0.050	0.008***	0.022***	0.044***	0.090***	0.181***
	(14.352)	(4.750)	(2.003)	(1.216)	(0.736)	(6.956)	(7.371)	(7.942)	(8.579)	(8.896)
1/Price ($t-1$)	−0.870***	0.209	1.471**	4.131***	9.562***	−0.316***	0.100	0.613***	1.819***	4.754***
	(−4.084)	(0.835)	(2.555)	(3.365)	(3.386)	(−10.093)	(1.302)	(4.369)	(6.737)	(8.791)
Amihud ($t-1$)	37.332***	1.842	−31.973	−72.013	−159.667	0.135***	−0.813***	−1.868***	−3.664***	−6.662***
	(6.715)	(0.139)	(−1.208)	(−1.332)	(−1.482)	(2.777)	(−8.169)	(−10.661)	(−11.816)	(−11.964)
Bid-ask spread ($t-1$)	−2.450**	2.411	6.631*	19.021***	29.763**	0.781*	0.135	0.245	1.653	4.534
	(−2.505)	(1.132)	(1.759)	(2.671)	(2.220)	(1.853)	(0.157)	(0.186)	(0.716)	(1.048)
Book-to-market ($t-1$)	0.020*	0.014	0.011	−0.001	−0.020	0.017***	0.039***	0.062***	0.107**	0.177**
	(1.702)	(0.565)	(0.232)	(−0.013)	(−0.106)	(3.511)	(2.951)	(2.622)	(2.333)	(1.986)
Past 12-month returns ($t-1$)	−0.014	−0.016	−0.036	0.067	0.308	−0.020***	−0.007	−0.006	0.017	0.043
	(−1.591)	(−0.567)	(−0.699)	(0.673)	(1.490)	(−5.325)	(−0.657)	(−0.306)	(0.452)	(0.581)
Gross profitability ($t-1$)	0.035**	0.053	0.045	0.064	0.152	0.049***	0.222***	0.403***	0.690***	1.210***
	(2.052)	(1.254)	(0.627)	(0.481)	(0.571)	(5.425)	(9.468)	(10.277)	(9.918)	(9.200)
Order imbalance ($t-1$)	0.003***	0.003***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(20.893)	(18.073)	(13.541)	(9.855)	(4.915)	(75.636)	(61.639)	(53.132)	(41.312)	(27.932)

(Continued)

Table VII—*Continued*

Panel A: ETF Flows and Stock Returns

Sample:	S&P 500					Russell 3000				
Dependent Variable:	Ret(t) (1)	Ret(t, $t+4$) (2)	Ret(t, $t+9$) (3)	Ret(t, $t+19$) (4)	Ret(t, $t+39$) (5)	Ret(t) (6)	Ret(t, $t+4$) (7)	Ret(t, $t+9$) (8)	Ret(t, $t+19$) (9)	Ret(t, $t+39$) (10)
Ret($t-1$)	−2.587***					−3.364***				
	(−8.348)					(−19.595)				
Ret($t-1$, $t-5$)		−2.199***					−1.571***			
		(−7.195)					(−10.302)			
Ret($t-1$, $t-10$)			−2.266***					−1.861***		
			(−7.337)					(−12.330)		
Ret($t-1$, $t-20$)				−2.496***					−1.873***	
				(−7.010)					(−11.240)	
Ret($t-1$, $t-40$)					−3.008***					−1.609***
					(−6.340)					(−7.837)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,492,313	1,492,313	1,492,313	1,492,313	1,492,313	7,759,143	7,759,143	7,759,143	7,759,143	7,759,143
R^2	0.033	0.013	0.011	0.011	0.013	0.039	0.009	0.007	0.006	0.006

Panel B: Price Impact on Turnover

Dependent Variable:	Price Impact			
Sample:	S&P 500		Russell 3000	
	(1)	(2)	(3)	(4)
Turnover (%)	3.609***	3.621***	3.426***	3.465***
	(34.627)	(36.255)	(47.787)	(41.592)
Turnover (%) ^2	−0.148***	−0.148***	−0.131***	−0.131***
	(−29.961)	(−30.910)	(−39.559)	(−35.681)
Constant	2.598***		3.199***	
	(15.909)		(15.126)	
Day fixed effects	No	Yes	No	Yes
Stock fixed effects	No	Yes	No	Yes
Observations	27,960,051	27,960,051	53,167,782	53,167,760
R^2	0.001	0.009	0.002	0.011

nonfundamental explanations such as the fact that some ETFs revise their fees downward. Based on this decomposition, in Table IA.XII in the Internet Appendix we show that nonfundamental flows are the main driver of stock volatility, whereas fundamental flows play a smaller role.[20]

C. Making Sense of the Size and Persistence of the Price Impact

Next, we show that the magnitude of the price impact of flows in Table VII, Panel A, is consistent with microstructure evidence. To do so, we follow prior literature and compute a market-impact function from trade-level data (e.g., Hasbrouck (1991), Hausman, Lo, and MacKinlay (1992), Keim and Madhavan (1996), Frazzini, Israel, and Moskowitz (2012), Landier, Simon, and Thesmar (2015)). Hence, we use the trade-level data in Ancerno, with observations at the stock-day-side-broker-manager level (where "side" is either buy or sell). Following the literature, we construct a measure of price impact as the percentage difference between the execution price and a pretrade benchmark, specifically, the price at the time of order placement (e.g., Anand et al. (2012)). From the summary statistics in Table II, Panel D, we know that, in the sample of S&P 500 (Russell 3000) stocks, the average price impact is 3.76 (5.35) bps, whereas the average turnover as a fraction of daily volume is 0.45% (0.94%).

Table VII, Panel B, reports estimates from a regression of the trade-level price impact on turnover, measured as the fraction of the average daily volume over the prior month, and turnover squared. The quadratic function is motivated by prior literature that finds a concave relation (e.g., Keim and Madhavan (1996)). For each universe of stocks, we provide two specifications: with day and stock fixed effects and without these fixed effects. Standard errors are double-clustered at the stock-day level. The estimates are similar irrespective of the fixed effects, and confirm evidence in prior literature about the convexity of the relation between price impact and turnover.

Using these estimates, we can give a microstructure foundation to the price impact of flows. In Table VII, Panel A, we note that a one-standard-deviation change in flows, which is 2.7% for S&P 500 stocks (see Table II, Panel C), correlates with a 12 bps change in prices on the same day. Based on the price impact estimates in Table VII, Panel B, column (1), we find that a 2.7% change in turnover leads to a price impact of about 11 bps ($= 2.6 + 3.6 \times 2.7 - 0.15 \times 2.7^2$), which is consistent with the price impact of flows from Panel A of Table VII.[21,22]

[20] Also consistent with the idea that ETFs propagate nonfundamental demand shocks, in Table IA.XIV in the Internet Appendix, we find that the relation between ETF ownership and volatility is stronger during times of high sentiment, as measured by the Baker and Wurgler (2006) index and the Michigan Survey of Consumer Sentiment index.

[21] Similar conclusions emerge from Figure IA.1 in the Internet Appendix.

[22] The evidence in Panel A of Table VII of less pronounced price pressure in the universe of Russell 3000 stocks does not conform with the intuition that, because these stocks are less liquid and exposed to larger flows as a fraction of daily volume (see Table II, Panel C), they should react more strongly to ETF flows. However, the smaller price impact that we report in Panel A of

The evidence in Table VII, Panel A, suggests that, following ETF flows, the half-life in the convergence of prices to the initial level is about 10 days. This finding stands in contrast to microstructure evidence that points to high resiliency of the limit order book.[23]

Two considerations help us reconcile the evidence. First, in the sample period, the settlement of the shares of U.S. domestic equity ETFs works on a $t + 3$ basis, and in some circumstances it can be extended to $t + 6$ (e.g., if the AP is also a market maker for ETFs, which happens in the large majority of cases; see ICI Research Perspective (2014)). This institutional framework implies that APs have time to buy (in the case of creation) or short (in the case of redemption) the underlying securities through $t + 6$. Consequently, when we measure flows on day t, we should expect price pressure on the underlying securities to continue over the following days as well. This fact can explain the slow reversal of prices to the initial level. If this explanation is indeed behind the persistence of the price impact, then flows on day t should predict order imbalance on the following days. This conjecture finds confirmation in Panel A of Table IA.XI in the Internet Appendix, which shows that day t flows significantly predict order imbalance in the same direction for up to at least seven days in the future.

Second, ETF flows are themselves persistent. Because flows on a given day are followed by flows in the same direction on the subsequent days, price pressure on the underlying stocks can continue after day t. In Panel B of Table IA.XI in the Internet Appendix, we find significant persistence in ETF flows. For example, for the full sample of ETFs, the first-order daily autocorrelation is about 9% and the higher order autocorrelations remain significant for at least 20 days.

We caution that these arguments provide only suggestive explanations for the persistence of the effect of ETF flows. Further research should delve deeper into these channels and explore other potential explanations.

D. Indirect Evidence on the Arbitrage Channel

To assess the relevance of arbitrage activity in driving the impact of ETFs on volatility, we search for an interaction of this effect with measures of limits to arbitrage. We first proxy for the intensity of arbitrage activity using ETF mispricing. We then conjecture that the proxy for expected arbitrage activity should have a weaker effect on prices for stocks that are harder to arbitrage.

At the daily frequency, stock-level mispricing is computed by summing the absolute dollar mispricing (i.e., the difference between the ETF price and NAV, as a fraction of the ETF price, multiplied by the dollar holdings in the stock)

Table VII could be the result of the larger measurement error in the flows to small stocks because of the higher volatility in these stocks' returns.

[23] For example, the empirical work in Biais, Hillion, and Spatt (1995), Dufour and Engle, (2000), Degryse et al. (2005), and Large (2007) suggests that liquidity is replenished in a few minutes if not a few seconds. More recently, Hendershott and Menkveld (2014) estimate a half-life of convergence following price pressure of 0.92 days.

across all ETFs holding stock i. We express this quantity as a fraction of a stock's capitalization:

$$abs\left(Mispricing_{i,t}\right) = \frac{\sum_{j=1}^{J} w_{i,j,t} * AUM_{j,t} * \left|Mispricing_{j,t}\right|}{Mkt\ Cap_{i,t}}. \qquad (7)$$

In the computation of mispricing, the ETF price and NAV are measured using daily closing prices. Thus, mispricing on a given day is a good predictor of next-day mispricing.[24]

We raise the caveat that ETF mispricing could instead proxy for a *lack of* arbitrage activity. That is, more mispricing could be present when arbitrageurs refrain from entering the market. For instance, Gagnon and Karolyi (2010) show that the price deviation between ADRs and the corresponding home market shares are positively related to measures of holding costs, foremost among which is idiosyncratic volatility. This channel could affect our inference if arbitrageurs abstain from trading high-volatility securities. To address this concern, we control for lagged volatility so that we focus on the impact of mispricing on innovations in volatility. This helps attenuate the endogeneity concern because arbitrage trades are not likely to condition on innovations in volatility in the next period. Further, such endogeneity would lead to the opposite sign of the coefficient on the interaction variables relative to what we find.

We use two proxies for limits to arbitrage. First, because ETF arbitrage involves a round-trip transaction in the stock, a large stock-level bid-ask spread reduces the profitability of arbitrage trades and therefore the incidence of arbitrage trading in a given stock. Second, when an arbitrage transaction involves shorting the stock, higher stock lending fees discourage arbitrageurs. Also, a high share lending fee can reflect a shortage of shares for lending, meaning that some arbitrageurs may not be able to carry out the trade (Cohen, Diether, and Malloy (2007)).

Given the high-frequency fluctuations in arbitrage activity, we carry out our tests at the daily frequency, which allows us to measure the variables of interest in a more timely way. Thus, the dependent variable for these tests is intraday volatility, which we estimate from second-by-second returns within a day. The main explanatory variable is the stock-level measure of absolute ETF mispricing in the prior day, defined in equation (7).

In Table VIII, the sample consists of S&P 500 stocks, where, according to our results above, the effect of ETFs is stronger. In column (1), we test whether absolute mispricing at the close of day $t-1$, which proxies for arbitrage activity on day t, has an incremental effect on volatility for a given level of ETF ownership. In addition to the usual controls, we include the mispricing on day $t-2$ and the lagged dependent variable. The goal is to capture the effect of the innovation in mispricing on the innovation in volatility, given that mispricing

[24] In Table IA.XIII in the Internet Appendix, we provide evidence that mispricing predicts next-day volume and order imbalance in the right direction at the stock level (i.e., positive mispricing predicts positive order). Hence, the results confirm that mispricing is a signal on which arbitrageurs condition their trading strategies.

Table VIII
Limits to Arbitrage

The table reports estimates from OLS regressions of intraday volatility on absolute stock-level mispricing in the prior period, interacted with measures of arbitrage costs. The frequency is daily, and observations are at the stock level. The sample includes S&P 500 stocks. Arbitrage costs are captured in columns (2) to (4) by the bid-ask spread on the prior day, and in columns (5) to (7) by the average share lending fee in the month. For both measures of arbitrage costs, we construct dummy variables denoting whether the stock is in the top half of the distribution of that measure in the relevant period. In columns (3) and (6), we restrict the sample to observations for which the stock-level mispricing is positive. In columns (4) and (7), we restrict the sample to observations for which the stock-level mispricing is negative. The controls include logged market capitalization, the lagged inverse share price ratio, the lagged Amihud (2002) ratio, the lagged average bid-ask spread, the lagged book-to-market ratio, lagged past 12-month returns, lagged gross profitability (as in Novy-Marx (2013)), lagged returns, the lagged dependent variable, and the absolute mispricing in period $t-2$. Variable descriptions are provided in Table IA.1 in the Internet Appendix. Standard errors are double-clustered at the stock and day levels. t-statistics are presented in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample covers the period January 2000 to December 2015.

Dependent Variable:	Intraday Stock Volatility						
	All (1)	All (2)	Misp > 0 (3)	Misp < 0 (4)	All (5)	Misp > 0 (6)	Misp < 0 (7)
abs(Mispricing) ($t-1$)	0.023***	0.043***	0.076***	0.035***	0.020***	0.048***	0.018***
	(6.897)	(8.025)	(10.300)	(6.988)	(4.642)	(7.530)	(4.302)
× I(High bid-ask spread)		−0.053***	−0.055***	−0.043***			
		(−6.554)	(−4.257)	(−6.392)			
× I(High lending fee)					−0.009*	0.003	−0.010**
					(−1.956)	(0.444)	(−2.548)
High bid-ask spread		0.042***	0.038***	0.045***			
		(7.186)	(6.318)	(7.130)			
High lending fee					−0.005	−0.003	−0.005
					(−1.565)	(−0.957)	(−1.433)

(Continued)

Table VIII—*Continued*

Dependent Variable:	Intraday Stock Volatility						
	All (1)	All (2)	Misp > 0 (3)	Misp < 0 (4)	All (5)	Misp > 0 (6)	Misp < 0 (7)
ETF ownership ($t-1$)	0.022***	0.022***	−0.004	0.031***	0.021***	−0.006	0.031***
	(3.611)	(4.217)	(−0.692)	(5.531)	(4.042)	(−1.227)	(5.417)
log(Mktcap ($t-1$))	0.086***	0.042***	0.042***	0.041***	0.034***	0.035***	0.033***
	(4.060)	(5.940)	(6.017)	(5.720)	(4.945)	(5.110)	(4.719)
1/Price ($t-1$)	5.767***	4.479***	4.216***	4.718***	4.568***	4.300***	4.805***
	(7.882)	(11.778)	(11.131)	(12.275)	(11.996)	(11.333)	(12.505)
Amihud ($t-1$)	−19.124	−8.636	−3.039	−13.726	−15.911	−9.579	−21.039
	(−0.795)	(−0.399)	(−0.139)	(−0.634)	(−0.712)	(−0.427)	(−0.942)
Bid-ask spread ($t-1$)	5.128	4.059	4.005	7.552	9.859	9.416	14.033**
	(0.796)	(0.636)	(0.670)	(1.085)	(1.515)	(1.551)	(1.989)
Book-to-market ($t-1$)	0.110***	0.078***	0.075***	0.081***	0.078***	0.075***	0.081***
	(3.231)	(4.102)	(3.997)	(4.220)	(4.136)	(4.046)	(4.244)
Past 12-month returns ($t-1$)	0.009	0.027***	0.020**	0.036***	0.027***	0.020**	0.036***
	(0.805)	(2.703)	(2.078)	(3.269)	(2.668)	(2.056)	(3.243)
Gross profitability ($t-1$)	−0.018	0.035	0.038	0.038	0.033	0.037	0.035
	(−0.230)	(1.443)	(1.579)	(1.525)	(1.342)	(1.542)	(1.398)
Return ($t-1$)	−0.196***	−0.246***	−0.707***	0.249***	−0.235***	−0.697***	0.264***
	(−3.984)	(−4.956)	(−9.759)	(3.104)	(−4.704)	(−9.612)	(3.280)
Dependent variable ($t-1$)	0.408***	0.561***	0.575***	0.541***	0.565***	0.578***	0.545***
	(20.800)	(24.501)	(23.836)	(24.546)	(24.739)	(24.044)	(24.734)
abs(Mispricing) ($t-2$)	0.024***	0.021***	0.059***	0.004	0.018***	0.059***	0.000
	(7.081)	(5.305)	(9.870)	(1.060)	(4.597)	(9.901)	(0.061)
Day fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock fixed effects	Yes	No	No	No	No	No	No
Observations	1,022,548	1,022,548	509,240	513,308	1,022,548	509,240	513,308
R^2	0.549	0.500	0.505	0.498	0.499	0.504	0.497

on day $t-1$ could itself depend on volatility (i.e., ETFs holding stocks that are more volatile are more likely to be mispriced, as discussed above). We also include the return on the stock on day $t-1$ to capture variation in mispricing that is exogenous to movements in the stock price itself. We find that the effect of absolute mispricing is positive and significant. This evidence lends further support to the view that arbitrage activity is a channel through which ETF ownership affects volatility.

We next report results of specifications that include interactions between absolute mispricing and the proxies for arbitrage costs. For each measure of limits to arbitrage, we define a dummy variable for stocks that are in the top half of the distribution of the variable in the prior period.[25] We leave out stock fixed effects because here we wish to achieve identification from the cross-sectional variation in the proxies. From Table VIII, column (2), we find that the effect of arbitrage on volatility, as proxied by absolute mispricing, is significantly weaker for stocks with a high bid-ask spread, consistent with a role for limits of arbitrage. Next, we separate the sample by the sign of net mispricing. A priori, we do not expect the sign of mispricing to matter for the interaction with the bid-ask spread because the arbitrage trade involves a round-trip transaction in the underlying stock in any case. The results in columns (3) and (4) support this conjecture.

In column (5), share lending fees attenuate the effect of arbitrage on volatility, but this effect is only marginally significant. More importantly, this effect should differ based on the sign of net mispricing, as the arbitrage trade should involve a short sale of the underlying stocks only when mispricing is negative (i.e., the ETF price is below the NAV). The estimates in columns (6) and (7) square with this prediction and provide strong evidence for the role of arbitrage activity in generating the effect of interest.

Importantly, the signs on the interactions with the proxies for arbitrage costs rule out concerns about the endogeneity of mispricing. If mispricing were due to volatility discouraging arbitrageurs from trading, we would expect this effect to be even stronger for illiquid stocks or for stocks that are hard to short, since these characteristics correlate positively with volatility. That is, we would expect the signs on the interactions with arbitrage costs to be positive (as opposed to negative) and significant.[26]

In Table IA.XVI in the Internet Appendix, we show that the strength of the relationship between ETF ownership and volatility increases with the greater availability of arbitrage capital, which we measure using the trading activity of institutional investors and hedge funds, in particular.

Overall, the results in this subsection point to a role of arbitrage in the link between ETF ownership and stock volatility. We recognize, however, that

[25] Information on share lending fees is sparse, especially in the initial part of the sample. We therefore use the average fee in the month.

[26] In Table IA.XVII in the Internet Appendix, we report the analysis for the Russell 3000. As expected, the effects in this sample are weaker or nonexistent. These results confirm the view that arbitrageurs tend to focus on the larger stocks in the ETF baskets when doing optimized replication.

the evidence in this subsection is only suggestive. First, the evidence is indirect. For example, institutional investors and hedge funds may trade ETFs for reasons other than arbitrage. Moreover, a lot of arbitrage is done by high-frequency trading firms, whose trading activity we do not capture. Second, estimation in Table VIII relies on OLS regressions because of the requirement of sufficiently large cross-sectional variation in stock characteristics, which the index-switching experiment does not satisfy due to its local nature. Hence, future research should study more closely the role of arbitrage in propagating liquidity shocks between ETFs and their underlying stocks.

V. Asset Pricing Implications

In this section, we explore two dimensions of the effect of ETFs on stock prices: (1) the effect of ETFs on price efficiency, and (2) the effect of ETF ownership on expected returns.

A. Implications for Pricing Efficiency

The liquidity trading and price discovery hypotheses have different implications for price efficiency. According to the former, transitory demand shocks in the ETF market migrate to the underlying security prices in the form of a mean-reverting component. As a result, prices become noisier. For the latter hypothesis, prices of stocks with higher coverage by ETFs adjust to fundamentals more promptly. The fundamental demand in the ETF market propagates to the underlying basket, impounding a permanent shock. In this scenario, ETF ownership makes prices closer to a random walk. We can therefore test for the effect of ETFs on the transitory component of stock prices.

Lo and MacKinlay (1988) and O'Hara and Ye (2011), among others, use variance ratios to measure the transitory component of stock prices. The variance ratio is defined as the variance of k-period returns divided by k times the variance of the single-period returns in the same window. When prices follow a random walk, the variance ratio equals one. If the autocorrelation of returns at the chosen frequency differs from zero, the variance ratio diverges from one.

Given that ETF arbitrageurs operate on a daily basis to exploit ETF mispricing, we test for their effect on the autocorrelation of daily returns. Accordingly, we define

$$abs\left(VR_{i,t}\right) = \left|\frac{\mathrm{Var}\left(r_{5,i,t}\right)}{5 \cdot \mathrm{Var}\left(r_{1,i,t}\right)} - 1\right|, \tag{8}$$

where $r_{5,i,t}$ is the five-day return and $r_{1,i,t}$ is the one-day return. If ETFs cause stock prices to deviate from a random walk, the statistic in equation (8) should increase with ETF ownership.

The liquidity trading hypothesis makes an even stronger prediction about the effect of ETFs on the variance ratio. The conjecture is that ETFs impound a mean-reverting component into prices, which in turn should make returns more negatively autocorrelated. Due to this negative autocorrelation, the numerator of the variance ratio shrinks relative to the denominator. Hence, if we redefine our test statistic as

$$VR_{i,t} = \frac{\mathrm{Var}\left(r_{5,i,t}\right)}{5 \cdot \mathrm{Var}\left(r_{1,i,t}\right)}, \tag{9}$$

we should obtain a negative relation between the quantity in equation (9) and ETF ownership.

We conduct our tests using stock-quarter-level observations. More specifically, we estimate the variance ratios using nonoverlapping observations during a quarter, and regress them on ETF ownership at the beginning of the quarter. The quarterly frequency responds to the need for a sufficiently long period over which we can estimate the five-day variance ratio.

Table IX reports results of tests using the usual IV framework. Both dependent variables, as well as the ownership variables, are standardized. Panel A shows that the absolute value of the variance ratio increases with ETF ownership, consistent with lower price efficiency for stocks owned by ETFs. Panel B suggests that ETF ownership correlates with a lower variance ratio, which supports the conjecture that ETF arbitrage induces negative autocorrelation in daily returns.

Overall, this analysis suggests that the increase in stock volatility we identify in Section III is likely the result of the fact that ETFs impound a mean-reverting component in stock prices. The evidence therefore provides further support for the liquidity trading hypothesis and suggests that the price discovery hypothesis is less plausible.

B. Implications for Expected Returns

A potential implication of the finding that ETFs affect stock volatility is that ETF ownership introduces a new source of risk to the investors of the underlying stocks. This risk may not be diversifiable, given that ETFs invest across the entire universe of stocks. Thus, a natural question that arises is whether the expected returns of the stocks in the ETF baskets pay a premium to compensate investors for this risk. This question is even more relevant in light of the finding that the impact of ETFs on volatility is greater during times of market turmoil (Table IV, Panel B) and when aggregate volatility is high (Tables IA.XIV and IA.XV in the Internet Appendix), and that in bad times, the skewness of ETF-owned stocks becomes more negative (Table IA.IV in the Internet Appendix). The ETF-induced shocks tend to cancel out over time, as indicated by the evidence of mean reversion in returns. However, short-horizon investors may still consider this volatility a source of risk for which they require a reward.

We follow the standard approach in asset pricing to test whether some characteristic correlates with a premium in returns. Each month, we form portfolios based on the level of ETF ownership in the prior month. We then allocate stocks

Table IX

Variance Ratio

The dependent variable in this table is Abs(Variance Ratio), which is the absolute variance ratio over five days and is defined as $abs(VR_{i,t}) = |\frac{\mathrm{Var}(r_{5,i,t})}{5 \cdot \mathrm{Var}(r_{1,i,t})} - 1|$. Variance Ratio is the variance ratio over five days and is defined as $VR_{i,t} = \frac{\mathrm{Var}(r_{5,i,t})}{5 \cdot \mathrm{Var}(r_{1,i,t})}$. In Panels A (for the absolute value of *VR*) and B (for the level of *VR*), we report instrumental variable (IV) estimates from a quasi-natural experiment relying on the reconstitution of the Russell 1000 and Russell 2000 indexes. The frequency of the data is monthly at the stock level. The explanatory variable is ETF ownership instrumented by a dummy for inclusion in the Russell 2000, for stocks in the Russell 1000 before index reconstitution (columns (1) to (5)), and a dummy for inclusion in the Russell 1000, for stocks in the Russell 2000 before index reconstitution (columns (6) to (10)). Stocks are ranked in terms of market capitalization in May of each year. Different ranges of this rank around the cutoff are used for inclusion in the sample: 100 stocks on each side (columns (1) and (6)), 200 stocks on each side (columns (2) and (7)), 300 stocks on each side (columns (3) and (8)), 400 stocks on each side (columns (4) and (9)), and 500 stocks on each side (columns (5) and (10)). The same stocks enter the sample in June after index reconstitution and remain in the sample until May of the next year, except if delistings occur. In all panels, the dependent variable and the ownership variables are standardized by subtracting the mean and dividing by the standard deviation. The controls include the lag of logged market capitalization, the lagged inverse share price, the lagged Amihud (2002) ratio, the lagged average bid-ask spread, the lagged book-to-market ratio, lagged past 12-month returns, lagged gross profitability (as in Novy-Marx (2013)), index fund ownership, active fund ownership, and hedge fund ownership. Month fixed effects are included. Standard errors are double-clustered at the stock and quarter levels. *t*-statistics are presented in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample covers the period January 2000 to December 2015.

Panel A: Second-Stage IV Regression for Abs(Variance Ratio)

Dependent Variable:	Abs(Variance Ratio)									
Instrument:	Switch to the Russell 2000					Switch to the Russell 1000				
Bandwidth:	± 100 (1)	± 200 (2)	± 300 (3)	± 400 (4)	± 500 (5)	± 100 (6)	± 200 (7)	± 300 (8)	± 400 (9)	± 500 (10)
ETF ownership	0.305	0.207	0.243*	0.235*	0.227*	0.524	0.494**	0.511***	0.563**	0.550**
	(1.182)	(1.320)	(1.810)	(2.033)	(1.992)	(1.383)	(2.494)	(2.917)	(2.711)	(2.767)
log(Mktcap ($t-1$))	0.058	0.015	0.036	−0.022	−0.040	0.174	0.048	0.044	0.042	−0.009
	(0.414)	(0.161)	(0.546)	(−0.447)	(−1.189)	(1.347)	(0.603)	(0.786)	(1.029)	(−0.209)
1/Price ($t-1$)	−0.573	−0.719**	−0.626**	−0.801***	−0.831***	0.095	0.164	−0.176	−0.016	0.160
	(−0.854)	(−2.387)	(−2.429)	(−3.139)	(−3.446)	(0.214)	(0.533)	(−0.784)	(−0.052)	(0.520)

(Continued)

Table IX—*Continued*

Panel A: Second-Stage IV Regression for Abs(Variance Ratio)										
Dependent Variable:	Abs(Variance Ratio)									
Instrument:	Switch to the Russell 2000					Switch to the Russell 1000				
Bandwidth:	± 100 (1)	± 200 (2)	± 300 (3)	± 400 (4)	± 500 (5)	± 100 (6)	± 200 (7)	± 300 (8)	± 400 (9)	± 500 (10)
Amihud ($t-1$)	3.342	3.802	5.148**	4.697**	5.254**	10.753	11.557**	13.500***	14.604***	10.343***
	(1.315)	(1.611)	(2.078)	(2.173)	(2.326)	(1.418)	(2.295)	(3.400)	(4.017)	(3.309)
Bid-ask spread ($t-1$)	−4.951	−0.072	1.400	2.757*	1.713	−0.480	3.577	2.104	1.302	1.325
	(−1.063)	(−0.032)	(0.741)	(1.771)	(1.105)	(−0.118)	(1.063)	(0.998)	(0.683)	(0.513)
Book-to-market ($t-1$)	−0.025	−0.018	−0.025	−0.022	−0.005	−0.081	−0.082	−0.085	−0.128	−0.196**
	(−0.610)	(−0.629)	(−1.454)	(−1.072)	(−0.369)	(−0.537)	(−0.917)	(−1.164)	(−1.449)	(−2.767)
Past 12-month returns ($t-1$)	−0.154	−0.092	−0.057	−0.034	0.008	0.060	−0.002	−0.007	−0.017	0.002
	(−1.143)	(−1.635)	(−1.263)	(−0.769)	(0.169)	(0.895)	(−0.052)	(−0.165)	(−0.447)	(0.070)
Gross profitability ($t-1$)	0.087	0.133	0.099	0.113	0.094	0.040	0.137*	0.046	0.028	−0.002
	(0.640)	(1.560)	(1.444)	(1.683)	(1.624)	(0.401)	(1.838)	(0.688)	(0.500)	(−0.041)
Index fund ownership ($t-1$)	−0.079*	−0.076**	−0.065*	−0.066**	−0.069**	−0.069	−0.069**	−0.088***	−0.092***	−0.098***
	(−1.765)	(−2.101)	(−2.043)	(−2.368)	(−2.340)	(−1.437)	(−2.327)	(−3.461)	(−3.595)	(−3.751)
Active fund ownership ($t-1$)	−0.043	−0.054**	−0.061***	−0.060***	−0.054***	−0.152***	−0.127***	−0.115***	−0.124***	−0.139***
	(−1.031)	(−2.058)	(−3.307)	(−3.953)	(−4.122)	(−4.217)	(−5.411)	(−5.834)	(−6.237)	(−6.847)
Hedge fund ownership ($t-1$)	−0.008	−0.016	0.009	0.012	0.023	0.066	0.057**	0.031	0.038*	0.029
	(−0.212)	(−0.585)	(0.371)	(0.559)	(1.134)	(1.595)	(2.090)	(1.668)	(1.893)	(1.493)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,535	3,204	5,115	7,097	9,116	2,064	4,038	5,801	7,498	9,221

(*Continued*)

Table IX—*Continued*

Panel B: Second-Stage IV Regressions for Variance Ratio

Dependent Variable:	Variance Ratio									
Instrument:	Switch to the Russell 2000					Switch to the Russell 1000				
Bandwidth:	± 100 (1)	± 200 (2)	± 300 (3)	± 400 (4)	± 500 (5)	± 100 (6)	± 200 (7)	± 300 (8)	± 400 (9)	± 500 (10)
ETF ownership	−1.191***	−0.867***	−0.743***	−0.676***	−0.678***	−1.231*	−0.975***	−1.051***	−1.160***	−1.202***
	(−3.003)	(−3.322)	(−2.977)	(−2.903)	(−2.833)	(−1.834)	(−3.286)	(−3.309)	(−3.127)	(−3.096)
log(Mktcap ($t-1$))	−0.409	−0.404**	−0.218**	−0.071	−0.067	−0.448*	−0.225	−0.211	−0.169	−0.117
	(−1.570)	(−2.559)	(−2.363)	(−1.280)	(−1.656)	(−1.998)	(−1.588)	(−1.700)	(−1.577)	(−1.173)
1/Price ($t-1$)	−0.649	0.191	0.101	0.383	0.313	−0.851	−0.534	−0.247	−0.864	−0.967*
	(−0.829)	(0.436)	(0.262)	(1.470)	(1.213)	(−1.047)	(−1.192)	(−0.563)	(−1.605)	(−1.844)
Amihud ($t-1$)	−8.841**	−8.987***	−10.598***	−9.537***	−8.800***	−24.204*	−20.393***	−23.756***	−25.530***	−17.594***
	(−2.229)	(−3.237)	(−3.453)	(−3.451)	(−3.177)	(−1.739)	(−3.509)	(−4.111)	(−4.323)	(−2.836)
Bid-ask spread ($t-1$)	10.485	5.122	5.812	3.444	0.711	8.025	3.549	0.951	2.288	1.040
	(1.280)	(1.216)	(1.549)	(1.195)	(0.299)	(1.158)	(1.176)	(0.344)	(0.805)	(0.329)
Book-to-market ($t-1$)	−0.002	0.015	0.018	0.018	0.042**	0.379	0.376**	0.304**	0.378**	0.447***
	(−0.027)	(0.327)	(0.650)	(0.871)	(2.591)	(1.569)	(2.647)	(2.352)	(2.527)	(2.935)
Past 12-month returns ($t-1$)	−0.016	0.011	−0.043	−0.043	−0.055	−0.046	−0.019	−0.027	−0.025	−0.037
	(−0.112)	(0.139)	(−0.628)	(−0.692)	(−0.980)	(−0.468)	(−0.399)	(−0.530)	(−0.548)	(−0.789)
Gross profitability ($t-1$)	−0.248*	−0.154*	−0.088	−0.123*	−0.079	−0.106	−0.100	−0.048	−0.027	0.003
	(−1.858)	(−1.811)	(−1.266)	(−2.013)	(−1.193)	(−0.537)	(−0.915)	(−0.522)	(−0.307)	(0.032)
Index fund ownership ($t-1$)	0.165***	0.172***	0.140***	0.139***	0.157***	0.162**	0.157***	0.172***	0.179***	0.197***
	(2.810)	(3.570)	(3.043)	(3.180)	(3.226)	(2.198)	(4.466)	(4.322)	(3.810)	(3.834)
Active fund ownership ($t-1$)	0.011	0.050**	0.048**	0.059***	0.055***	0.150**	0.119***	0.141***	0.148***	0.189***
	(0.199)	(2.077)	(2.595)	(3.358)	(2.984)	(2.438)	(3.371)	(4.100)	(4.242)	(4.804)
Hedge fund ownership ($t-1$)	−0.107*	−0.038	−0.049	−0.033	−0.043	−0.124*	−0.086**	−0.059*	−0.066*	−0.053
	(−1.956)	(−1.083)	(−1.434)	(−1.095)	(−1.520)	(−1.800)	(−2.447)	(−2.041)	(−1.900)	(−1.483)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Linear polynomials of rank	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,535	3,204	5,115	7,097	9,116	2,064	4,038	5,801	7,498	9,221

to five quintiles and equally weight the portfolios. We therefore obtain time series of portfolio returns ranging from February 2000 to December 2015.[27]

Table X, Panel A, reports the average excess returns over the sample period for the five ETF-ownership quintile portfolios. The average returns increase almost uniformly with ETF ownership. Following the literature, we form a portfolio that is long in the top quintile and short in the bottom quintile of ETF ownership. We use this long-short portfolio to test whether ETF ownership commands a premium.

In Table X, Panel B, we report the raw returns and alphas from different factor models for the long-short portfolio, using the full sample. In terms of raw returns, the high-minus-low ETF ownership portfolio earns about 35 bps per month, which is statistically different from zero. The alpha with respect to the market model is of a somewhat smaller magnitude (30 bps) but still significant. Using the Fama and French (1993) three factors increases the significance and the magnitude of the alpha, given that the loading on HML is negative. The momentum factor does not affect the alpha in a four-factor model (Carhart (1997)). Next, we gradually introduce the two new Fama and French (2015) factors (RMW, CMA), which make the alpha somewhat larger. Finally, we test whether liquidity risk can explain the premium we observe. After including the Pástor and Stambaugh (2003) liquidity risk factor (PS_VWF), the magnitude and significance of alpha do not change.

One way to verify whether the effect we measure is indeed related to ETF ownership, as opposed to some omitted risk factor, is to test whether the alphas are larger in recent years, when ETFs have become more important. In Panels A and B of Table IA.XVIII in the Internet Appendix, we estimate the factor models by splitting the sample at the end of 2007. Consistent with the effect we capture being related to the presence of ETFs in the market, the alphas are larger and more statistically significant in the recent sample. In this period, the magnitude ranges from 44 bps to 56 bps.

The results are consistent with the conjecture that ETF ownership impounds some undiversifiable source of risk into the underlying securities, which is priced ex ante. An alternative interpretation of these results hinges on the price pressure deriving from ETF flows. Over the sample period, ETF ownership has grown consistently. Based on the results in Table VII, we know that ETF flows exert price pressure on the basket securities. Hence, according to this alternative view, the price pressure from ETF flows, which are larger for stocks with more ETF ownership, would lead ex post to the abnormal returns for high ETF ownership portfolios.

To disentangle the two explanations, we run Fama and MacBeth (1973) regressions of abnormal returns on the explanatory variable of interest, that is, lagged ETF ownership, and controls for other potential explanations, including ETF flows. The dependent variable is the monthly four-factor model-adjusted (Carhart (1997)) stock return expressed as a percentage. In Table X, Panel C,

[27] We also experimented with value-weighted portfolios and found qualitatively similar results but lower magnitudes and statistical significance.

Table X

ETF Ownership Portfolios

The table reports statistics and regression analysis for portfolios based on ETF ownership. Each month, five equally weighted portfolios are formed on the basis of the five quintiles of the distribution of ETF ownership in the previous month. Panel A presents the raw returns. Panel B presents factor model regressions for the high-ETF-ownership minus the low-ETF-ownership portfolio. The factors are the five Fama and French (2015) factors (MKTRF, HML, SMB, RMW, CMA), momentum (UMD), and the Pástor and Stambaugh (2003) traded liquidity factor (PS_VWF). Panel C reports estimates from Fama and MacBeth (1973) regressions of four-factor model-adjusted (Carhart (1997)) monthly stock returns on ETF ownership in the previous month, contemporaneous ETF stock-level flows, index, active, and hedge fund ownership, and control variables. The dependent variable is expressed as a percentage, and both the ownership variables and flows are standardized. The controls include logged market capitalization, the lagged inverse share price ratio, the lagged Amihud (2002) ratio, the lagged average bid-ask spread, the lagged book-to-market ratio, lagged past 12-month returns, and lagged gross profitability (as in Novy-Marx (2013)). t-statistics are presented in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. The sample covers the period January 2000 to December 2015.

Panel A: Raw Excess Returns for the Quintile Portfolios

	Quintiles Based on ETF Ownership				
	Low	(2)	(3)	(4)	High
Raw excess returns	0.389	0.618*	0.696*	0.739**	0.737*
	(1.051)	(1.736)	(1.941)	(2.040)	(1.828)
Number of months	191	191	191	191	191

Panel B: High-Minus-Low Portfolio, Full Sample

Dependent Variable:	Ret(High-Minus-Low ETF Ownership)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alpha	0.347**	0.303**	0.334**	0.355**	0.364**	0.350**	0.383**	0.383**
	(2.296)	(2.074)	(2.294)	(2.429)	(2.477)	(2.287)	(2.482)	(2.461)
MKTRF		0.130***	0.126***	0.136***	0.125***	0.132***	0.119***	0.119***
		(4.012)	(3.915)	(4.126)	(3.437)	(3.204)	(2.847)	(2.829)
HML			−0.089**	−0.095**	−0.101**	−0.111**	−0.051	−0.051
			(−1.996)	(−2.127)	(−2.223)	(−2.046)	(−0.749)	(−0.738)
SMB				−0.061	−0.054	−0.044	−0.036	−0.036
				(−1.316)	(−1.132)	(−0.778)	(−0.645)	(−0.620)
UMD					−0.021	−0.023	−0.014	−0.014
					(−0.747)	(−0.787)	(−0.475)	(−0.474)
RMW						0.025	0.016	0.016
						(0.340)	(0.220)	(0.219)
CMA							−0.136	−0.136
							(−1.460)	(−1.455)
PS_VWF								−0.001
								(−0.021)
Number of months	191	191	191	191	191	191	191	191
R^2	0.000	0.078	0.098	0.106	0.109	0.109	0.119	0.119

(*Continued*)

Table X—*Continued*

Panel C: Fama and MacBeth (1973) Regressions					
Dependent Variable:	Four-Factor Model-Adjusted Returns (t)				
	(1)	(2)	(3)	(4)	(5)
ETF ownership ($t-1$)	0.144**	0.161**	0.195***		0.199***
	(2.382)	(2.536)	(2.862)		(2.642)
ETF flows (t)				0.099**	0.080
				(2.009)	(1.452)
Hedge fund ownership ($t-1$)		0.062	0.024	0.026	0.021
		(1.198)	(0.444)	(0.474)	(0.376)
Index fund ownership ($t-1$)		−0.055	−0.065	−0.025	−0.067
		(−1.145)	(−1.363)	(−0.575)	(−1.398)
Active fund ownership ($t-1$)		0.075	0.049	0.053	0.040
		(1.231)	(0.973)	(1.048)	(0.819)
log(Mktcap ($t-1$))			0.109*	0.029	0.078
			(1.895)	(0.450)	(1.310)
1/Price ($t-1$)			6.458***	6.440***	6.206***
			(5.197)	(5.209)	(5.011)
Amihud ($t-1$)			3.074	−10.612	6.454
			(0.182)	(−0.607)	(0.364)
Bid-ask spread ($t-1$)			−18.181	−21.054	−17.810
			(−1.281)	(−1.479)	(−1.254)
Book-to-market ($t-1$)			−0.068	−0.095	−0.067
			(−0.311)	(−0.440)	(−0.304)
Past 12-month returns ($t-1$)			−0.430	−0.408	−0.417
			(−1.495)	(−1.442)	(−1.453)
Gross profitability ($t-1$)			0.709**	0.659**	0.702**
			(2.332)	(2.181)	(2.321)
Observations	350,237	350,237	350,237	350,237	350,237
R^2	0.003	0.011	0.037	0.036	0.039
Number of months	191	191	191	191	191

the first specification reveals a positive and significant relation between returns and standardized ETF ownership. A one-standard-deviation increase in ETF ownership correlates with a 14 bps increase in returns. The next specification includes ownership by other types of institutions, whose effect is not significant, although it is positive in the case of active mutual funds and hedge funds. We also add controls for liquidity and stock characteristics typically associated with returns. We note a significant profitability effect, consistent with Novy-Marx (2013), whereas the other asset pricing effects are insignificant, probably due to the four-factor adjustment of the dependent variable. In column (4), ETF flows display a significant correlation with returns when ETF ownership is not included. However, a horse race between the two variables reveals that ETF ownership is the dominant determinant of stock returns (column (5)). This finding rules out the possibility that the alphas we identify are due to the price pressure of ETF flows.

Overall, these results support the argument that ETFs significantly modify the underlying stocks' return distribution. As ETFs increase the volatility of stock returns, they appear to introduce a nondiversifiable source of risk, at least in the short term, for which investors require a premium in expected returns.

VI. Conclusion

The success of ETFs is due in large part to the fact that these investment vehicles offer an unprecedented source of diversification at low cost and high liquidity. This aspect of ETFs is undeniably beneficial for investors. However, due to their ease of trade, ETFs seem to attract a new breed of high-frequency investors, whose demand shocks can pass on to the underlying securities via the arbitrage activity continuously taking place between ETFs and their baskets. This mechanism can lead to higher volatility for the underlying securities. The increase in volatility would not be a desirable effect of ETFs if it were merely a reflection of increased noise trading.

In this paper, we start by showing that ETFs are indeed the preferred habitat of investors with relatively higher turnover, consistent with the view that they attract high-frequency demand. One of the main results of the paper is that stocks with more ownership by ETFs display higher volatility than otherwise similar securities. A quasi-natural experiment based on the reconstitution of the Russell indexes suggests a causal interpretation for this finding.

We next show that the demand shocks in the ETF market impound a mean-reverting component in asset prices, which plays out at the daily frequency. This result suggests that the increase in stock return volatility is not likely to be imputable to an improvement in price discovery brought about by ETFs. Rather, it is likely a reflection of the transmission of nonfundamental demand shocks from the ETF market to the prices of the underlying stocks via arbitrage. Consistent with this view, we show that proxies for the intensity of arbitrage activity between ETFs and their baskets magnify the effect of ETFs on volatility.

Finally, the paper addresses the asset pricing implications of ETFs for stocks. If the increase in stock volatility brought about by ETFs is partly nondiversifiable, it may represent systematic risk for some investors, especially for those with a short trading horizon. As such, ETF ownership may warrant a risk premium. Consistent with this conjecture, we show that portfolios of stocks with high ETF ownership display positive alphas relative to a variety of asset pricing models. These alphas are about 50 bps in the more recent sample. We confirm this finding in a regression setting and rule out the possibility that it is an ex post reflection of the growing demand for ETFs.

Future research should delve deeper into the channels linking ETF ownership to stock volatility. At this point, evidence on the role of arbitrage trades is admittedly indirect. Similarly, further investigation is necessary on the role of ETFs in attracting a different clientele of traders. Consistent with a clientele effect, we find that the composition of traders shifts toward higher turnover investors in stocks with more ETF ownership. Whether this modification in client base is the result of ETF arbitrage remains to be established.

Initial submission: September 25, 2015; Accepted: November 28, 2017
Editors: Bruno Biais, Michael R. Roberts, and Kenneth J. Singleton

REFERENCES

Agarwal, Vikas, Paul Hanouna, Rabih Moussawi, and Christof Stahel, 2017, Do ETFs increase the commonality in liquidity of underlying stocks? Working paper, Villanova University.

Amery, Paul, 2015, ETF giants bet on futures for flows, *Financial Times*, February 1. Available at https://www.ft.com/content/7979678a-a32d-11e4-9c06-00144feab7de.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Amihud, Yakov, and Haim Mendelson, 1987, Trading mechanisms and stock returns: An empirical investigation, *Journal of Finance* 42, 533–553.

Anand, Amber, Paul Irvine, Andy Puckett, and Kumar Venkataraman, 2012, Performance of institutional trading desks: An analysis of persistence in trading costs, *Review of Financial Studies* 25, 557–598.

Anand, Amber, Paul Irvine, Andy Puckett, and Kumar Venkataraman, 2013, Institutional trading and stock resiliency: Evidence from the 2007–2009 financial crisis, *Journal of Financial Economics* 108, 773–797.

Andrei, Daniel, and Michael Hasler, 2015, Investor attention and stock market volatility, *Review of Financial Studies* 28, 33–72.

Angrist, Joshua, and Guido Imbens, 1995, Identification and estimation of local average treatment effects, Working paper, National Bureau of Economic Research Cambridge.

Anton, Miguel, and Christopher Polk, 2014, Connected stocks, *Journal of Finance* 69, 1099–1128.

Appel, Ian, Todd A. Gormley, and Donald B. Keim, 2015, Passive investors, not passive owners, Working paper, University of Pennsylvania.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Barberis, Nicholas, Andrei Shleifer, and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283–317.

Basak, Suleyman, and Anna Pavlova, 2013, Asset prices and institutional investors, *American Economic Review* 103, 1728–1758.

Basak, Suleyman, and Anna Pavlova, 2016, A model of financialization of commodities, *Journal of Finance* 71, 1511–1556.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2017, Exchange traded funds, *Annual Review of Financial Economics* 9, 165–185.

Bessembinder, Hendrik, and Paul J. Seguin, 1992, Futures-trading activity and stock price volatility, *Journal of Finance* 47, 2015–2034.

Bhattacharya, Ayan, and Maureen O'Hara, 2017, Can ETFs increase market fragility? Effect of information linkages in ETF markets, Working paper, Cornell University.

Biais, Bruno, Pierre Hillion, and Chester Spatt, 1995, An empirical analysis of the limit order book and the order flow in the Paris Bourse, *Journal of Finance* 50, 1655–1689.

BlackRock, 2014, ETP landscape, *Industry Highlights*, December. Available at https://www.blackrock.com/latamiberia/insights/etp-landscape.

Cao, Charles, Matthew Gustafson, and Raisa Velthuis, 2014, Capital market consequences of index membership: Evidence from small firms, Working paper, Pennsylvania State University.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Cella, Cristina, Andrew Ellul, and Mariassunta Giannetti, 2013, Investors' horizons and the amplification of market shocks, *Review of Financial Studies* 26, 1607–1648.

Chang, Eric C., Joseph W. Cheng, and J. Michael Pinegar, 1999, Does futures trading increase stock market volatility? The case of the Nikkei stock index futures markets, *Journal of Banking & Finance* 23, 727–753.

Chang, Yen-Cheng, Harrison Hong, and Inessa Liskovich, 2015, Regression discontinuity and the price effects of stock market indexing, *Review of Financial Studies* 28, 212–246.

Chinco, Alexander, and Vyacheslav Fos, 2016, The sound of many funds rebalancing, Working paper, Boston College.

Chordia, Tarun, and Avanidhar Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, *Journal of Financial Economics* 72, 485–518.

Cohen, Lauren, Karl B. Diether, and Christopher J. Malloy, 2007, Supply and demand shifts in the shorting market, *Journal of Finance* 62, 2061–2096.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Crane, Alan D., Sebastien Michenaud, and James Weston, 2014, The effect of institutional ownership on payout policy: A regression discontinuity design approach, Working paper, Rice University.

Da, Zhi, and Sophie Shive, 2015, When the bellwether dances to noise: Evidence from exchange-traded funds, Working paper, Notre Dame University.

Daniel, Kent, Mark Grinblatt, Sheriden Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.

Danthine, Jean-Pierre, 1978, Information, futures prices, and stabilizing speculation, *Journal of Economic Theory* 17, 79–98.

Degryse, Hans, Frank De Jong, Maarten Van Ravenswaaij, and Gunther Wuyts, 2005, Aggressive orders and the resiliency of a limit order market, *Review of Finance* 9, 201–242.

Dufour, Alfonso, and Robert F. Engle, 2000, Time and the price impact of a trade, *Journal of Finance* 55, 2467–2498.

Ellis, Katrina, Roni Michaely, and Maureen O'Hara, 2000, The accuracy of trade classification rules: Evidence from NASDAQ, *Journal of Financial and Quantitative Analysis* 35, 529–551.

Evans, Richard, Rabih Moussawi, Michael Pagano, and John Sedunov, 2017, ETF short interest and failures-to-deliver: Naked short-selling or operational shorting? Working paper, Villanova University.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Fich, Eliezer M., Jarrad Harford, and Anh L. Tran, 2015, Motivated monitors: The importance of institutional investors' portfolio weights, *Journal of Financial Economics* 118, 21–48.

Frazzini, Andrea, Ronen Israel, and Tobias J. Moskowitz, 2012, Trading costs of asset pricing anomalies, Fama-Miller Working paper 14-05, Chicago Booth School of Business.

Gagnon, Louis, and Andrew Karolyi, 2010, Multi-market trading and arbitrage, *Journal of Financial Economics* 97, 53–80.

Glosten, Lawrence R., Suresh Nallareddy, and Yuan Zou, 2016, ETF trading and informational efficiency of underlying securities, Research paper no. 16–71, Columbia Business School.

Greenwood, Robin, 2005, Short- and long-term demand curves for stocks: Theory and evidence on the dynamics of arbitrage, *Journal of Financial Economics* 75, 607–649.

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

Grossman, Sanford J., and Merton H. Miller, 1988, Liquidity and market structure, *Journal of Finance* 43, 617–633.

Grossman, Sanford Jay, 1989, An analysis of the implications for stock and futures price volatility of program trading and dynamic hedging strategies, NBER Working paper 2357.

Hasbrouck, Joel, 1991, Measuring the information content of stock trades, *Journal of Finance* 46, 179–207.

Hausman, Jerry A., Andrew W. Lo, and A. Craig MacKinlay, 1992, An ordered probit analysis of transaction stock prices, *Journal of Financial Economics* 31, 319–379.

Hendershott, Terrence, and Albert J. Menkveld, 2014, Price pressures, *Journal of Financial Economics* 114, 405–423.

Holden, Craig W., and Stacey Jacobsen, 2014, Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions, *Journal of Finance* 69, 1747–1785.

Hombert, Johan, and David Thesmar, 2014, Overcoming limits of arbitrage: Theory and evidence, *Journal of Financial Economics* 111, 26–44.

Hong, Harrison, Jeffrey Kubik, and Tal Fishman, 2012, Do arbitrageurs amplify economic shocks? *Journal of Financial Economics* 103, 454–470.

ICI Research Perspective, 2014, Understanding exchange-traded funds: How ETFs work, Investment Company Institute, September. Available at https://www.ici.org/pdf/per20-05.pdf.

Israeli, Doron, Charles M. C. Lee, and Suhas Sridharan, 2015, Is there a dark side to exchange traded funds (ETFs)? An information perspective, Working paper, Stanford University.

Keim, Donald B., and Ananth Madhavan, 1996, The upstairs market for large-block transactions: Analysis and measurement of price effects, *Review of Financial Studies* 9, 1–36.

Landier, Augustin, Guillaume Simon, and David Thesmar, 2015, The capacity of trading strategies, Working paper, HEC Paris.

Large, Jeremy, 2007, Measuring the resiliency of an electronic limit order book, *Journal of Financial Markets* 10, 1–25.

Lee, Charles M. C., and Mark J. Ready, 1991, Inferring trade direction from intraday data, *Journal of Finance* 46, 733–746.

Lee, Charles M. C., Andrei Shleifer, and Richard H. Thaler, 1991, Investor sentiment and the closed-end fund puzzle, *Journal of Finance* 46, 75–109.

Leippold, Markus, Lujing Su, and Alexander Ziegler, 2015, Do index futures and ETFs affect stock return correlations? Working paper, University of Zurich.

Lo, Andrew W., and Archie Craig MacKinlay, 1988, Stock market prices do not follow random walks: Evidence from a simple specification test, *Review of Financial Studies* 1, 41–66.

Lou, Dong, 2012, A flow-based explanation of returns predictability, *Review of Financial Studies* 25, 3457–3489.

Lou, Dong, and Christopher Polk, 2014, Comomentum: Inferring arbitrage activity from return correlations, Working paper, London School of Economics.

Lu, Ruichang, 2014, How does institutional ownership affect bank loan pricing: Evidence from a regression discontinuity design, Working paper, National University of Singapore.

Madhavan, Ananth, and Aleksander Sobczyk, 2016, Price dynamics and liquidity of exchange-traded funds, *Journal of Investment Management* 14, 1–17.

Malamud, Semyon, 2015, A dynamic equilibrium model of ETFs, Working paper, Swiss Finance Institute.

Mullins, William, 2014, The governance impact of index funds: Evidence from regression discontinuity, Working paper, Massachusetts Institute of Technology.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

O'Hara, Maureen, and Mao Ye, 2011, Is market fragmentation harming market quality? *Journal of Financial Economics* 100, 459–474.

Pástor, Ľuboš, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Petajisto, Antti, 2013, Inefficiencies in the pricing of ETFs, Working paper, New York University.

Pontiff, Jeffrey, 1996, Costly arbitrage: Evidence from closed-end funds, *Quarterly Journal of Economics* 111, 1135–1151.

Puckett, Andy, and Xuemin Sterling Yan, 2011, The interim trading skills of institutional investors, *Journal of Finance* 66, 601–633.

Roll, Richard, Eduardo Schwartz, and Avanidhar Subrahmanyam, 2007, Liquidity and the law of one price: The case of the futures-cash basis, *Journal of Finance* 62, 2201–2234.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down? *Journal of Finance* 41, 579–590.

Stambaugh, Robert F., 2014, Investment noise and trends, *Journal of Finance* 69, 1415–1453.

Stein, Jeremy C., 1987, Informational externalities and welfare-reducing speculation, *Journal of Political Economy* 95, 1123–1145.

Subrahmanyam, Avanidhar, 1991, A theory of trading in stock index futures, *Review of Financial Studies* 4, 17–51.

Turnovsky, Stephen J., 1983, The determination of spot and futures prices with storable commodities, *Econometrica* 51, 1363–1387.

Vayanos, Dimitri, and Paul Woolley, 2013, An institutional theory of momentum and reversal, *Review of Financial Studies* 26, 1087–1145.

Wurgler, Jeffrey, 2011, On the economic consequences of index-linked investing, in Gerald Rosenfeld, Jay W. Lorsch, and Rakesh Khurana, eds.: *Challenges to Business in the Twenty-First Century* (American Academy of Arts & Sciences, Cambridge, MA).

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.