Journal of Financial Economics 141 (2021) 1078–1095

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Intraday arbitrage between ETFs and their underlying portfolios[☆]

Travis Box[a,*], Ryan Davis[b], Richard Evans[c], Andrew Lynch[d]

[a] Wilbur O. and Ann Powers College of Business, Clemson University, 225 Walter T. Cox Blvd, Clemson, SC 29634, USA
[b] Collat School of Business, University of Alabama at Birmingham
[c] Darden School of Business, University of Virginia
[d] Sam M. Walton College of Business, University of Arkansas

ARTICLE INFO

Article history:
Received 25 January 2019
Revised 22 June 2020
Accepted 11 July 2020
Available online 4 May 2021

JEL classification:
G12
G14

Keywords:
Exchange-traded funds
ETFs
Limits to arbitrage
Liquidity

ABSTRACT

Prior research suggests that nonfundamental exchange-traded fund (ETF) price shocks are transmitted to their portfolios through an arbitrage mechanism. We test this proposition by examining minute-by-minute returns and order imbalances but find little evidence that ETF trading impacts underlying returns. Specifically, panel vector autoregression shows that ETF returns do not lead portfolio prices. Instead, arbitrage opportunities arise from order imbalances and price movements in the underlying securities and are subsequently eliminated by ETF quote adjustments, rather than arbitrage trading. We extend our analysis to a daily frequency but still find little relation between ETF trading and constituent security prices.

1. Introduction

There is a growing literature examining the potential impact of exchange-traded fund (ETF) trading on the securities underlying their portfolios. While several early studies suggest that ETF trading improves price discovery for the underlying securities (see, e.g., Hasbrouck, 2003; Yu, 2005; Chen and Strother, 2008; Fang and Sanger, 2012; and Ivanov *et al.*, 2013), more recent literature, both empirical and theoretical, suggests that: (i) the superior liquidity of ETFs attracts noise traders; (ii) order flow shocks, perhaps from these noise traders, create arbitrage opportunities; and (iii) authorized participants (or other sophisticated traders) correct mispricing through arbitrage, transmitting noise to underlying security prices.[1] While many of these studies analyze the relation between ETF ownership and various measures of constituent market quality, the transmission of noise through an arbitrage mechanism

[☆] We appreciate the helpful comments of seminar participants and organizers at Clemson University, University of Arkansas, University of Alabama, Auburn University, University of North Texas, University of Mississippi, Missouri State University, the 2019 Financial Management Association Meeting, the 2019 Northern Finance Association Meeting, the 2019 Southwestern Finance Association Meeting, and the Eastern Finance Association Meeting. We are especially grateful to the Mississippi Center for Supercomputing Research and its director, Ben Pharr, for computational assistance.

[*] Corresponding author

E-mail addresses: tsbox@clemson.edu (T. Box), davisrl@uab.edu (R. Davis), evansr@darden.virginia.edu (R. Evans), alynch@walton.uark.edu (A. Lynch).

[1] See Charupat and Mui (2011); Malamud (2016); and Broman and Shum (2018). ETF ownership is associated with higher trading costs and less informative prices (Israeli et al., 2017), excess comovement (Da and Shive, 2018), and excess volatility (Ben-David et al., 2018) in the underlying portfolio.

https://doi.org/10.1016/j.jfineco.2021.04.023
0304-405X/© 2021 Elsevier B.V. All rights reserved.

is more often presumed than tested directly.[2] In this paper, we assess the extent to which data supports each of these assertions by examining whether ETF trading affects the intraday prices of underlying securities across a sample of 423 passively managed US equity ETFs from 2006 to 2015.

Ultimately, we find little support for the notion that ETF activity transmits noise to the underlying portfolio. Furthermore, we find strong evidence that ETF prices are more likely to follow constituent returns, even during episodes of extreme price divergence or fund and portfolio return shocks (i.e., stochastic jumps). Rather than transmitting nonfundamental shocks to underlying security prices, our results suggest that ETFs may actually shield their portfolios from demand shocks by supporting liquidity provision in underlying securities.

Our analysis begins with simple intraday correlations between ETF and underlying returns and order flows. The hypothesis that ETF activity affects underlying prices requires at least some statistical relation between ETF order flow and constituent returns. While we find a strong correlation between underlying order flow and ETF returns, we find no such relation between ETF trading and portfolio returns at 1-, 5-, and 10 min intervals, and only a weak relation across daily intervals. Furthermore, we observe strong correlations between constituent returns and their own order flow, whereas the statistical relation between ETF trading and price changes is effectively zero. Given that ETF order flow does not seem to affect fund share prices, it should not be surprising that underlying security prices are also unaffected by ETF order flow.

It is possible that our simple bivariate analysis masks the true relation between ETF and constituent prices and trading by failing to account for contemporaneous and intertemporal covariation in these time series. Thus, we proceed with a panel vector autoregressive (PVAR) framework wherein all four relations (ETF and underlying order imbalances and returns) can be modeled simultaneously across our entire sample. Impulse response functions (IRFs) from this analysis describe the expected impact of a shock to one variable on future realizations of the other three. If ETF activity affects constituent securities, we expect to find shocks in fund order imbalances and/or returns predicting order imbalances and/or returns in the underlying securities. However, we observe no statistically significant response in constituent prices following shocks to ETF order flow. While we find a small response in portfolio returns from a shock to ETF prices, we also report that the response of ETF returns to underlying price shocks is approximately six times larger, and that the difference between the two is statistically significant.

While IRFs describe one variable's response to a hypothetical exogenous shock in another, these functions provide no information about how much realized variation in one time series can be explained by another. Thus, we also perform a forecast error variance decomposition (FEVD) of our PVAR specifications to determine whether independent shocks to each variable impact future realizations of the other three. Over the 1-, 5-, or 10 min interval window, we find that orthogonalized shocks to ETF returns explain at most 0.26% of underlying return variance, whereas independent shocks to fund order flow explain only 0.0007%. Even after limiting our analysis to underlying portfolios most exposed to nonfundamental ETF shocks, those with low market capitalization or high spreads relative to their ETFs, we still find underlying returns largely unresponsive to either fund returns or order imbalances.

While our PVAR analysis covers the entire cross-section of intraday data, it is possible that ETF trading only impacts constituent securities whenever prices diverge dramatically. To address this possibility, we revisit our 1 min frequency intraday analysis but focus only on episodes where mispricing (measured by best bid and ask prices for the ETF and portfolio) is severe enough for arbitrage trading to become potentially profitable. On average, we observe selling pressure and negative returns in the constituents prior to the ETF becoming overvalued (where ETF bid is greater than underlying ask), with buying pressure and positive portfolio returns before the fund becomes undervalued (where ETF ask is less than underlying bid). In either case, these potential arbitrage opportunities are preceded by shocks to underlying security prices, which are impounded by directional trading in the portfolio. Therefore, profitable arbitrage opportunities are not instigated by short-term price pressure in the ETF's shares, on average. Additionally, our results indicate that the constituent price shocks preceding these mispricing events do not reverse during the subsequent 30 min window, suggesting that they are not transitory in nature.

During these apparent arbitrage opportunities, we still observe swift convergences between bid and ask prices for ETFs and their underlying portfolios. As discussed in Brogaard *et al.* (2019), price discovery can occur both through order flows and updated limit orders. However, our analysis suggests order flows do not affect the convergence of prices. After the ETF becomes overvalued (undervalued), we see additional buying (selling) pressure in the fund and net selling (buying) pressure in the portfolio. Thus, price discrepancies between ETF and underlying securities are corrected in the limit order book. Furthermore, marketable orders appear to be moving prices for the constituent securities and the ETF farther away from parity. If arbitrage activity were the primary motivation for directional trading, we should expect to see order flow pushing prices closer together.

In addition to our analysis of mispricing everts, we also examine cumulative returns and order flow around stochastic jumps, as identified by the measure of Lee and Mykland (2008), in ETF or portfolio prices. In cases where the underlying experiences a stochastic jump, ETF and portfolio prices mirror each other almost perfectly. In contrast, ETF return jumps are usually preceded by price changes in the underlying securities. Overall, the arbitrage

[2] While previous work has not tested this arbitrage mechanism directly, several potential limitations of the ETF arbitrage process have been suggested. For instance, market frictions allow some mispricings to persist longer than a potential arbitrageurs' planned trading horizon (Ben-David et al., 2018). Furthermore, investors may not enforce efficient pricing due to execution risks (Malamud, 2016), transaction costs (Broman and Shum, 2018), difficulty borrowing shares, as well as varying arbitrage speeds or difficulties in observing intrinsic values, especially during times of market stress (Madhavan and Sobczyk, 2016).

and stochastic jump analyses are consistent with liquidity providers hedging their exposure to informed trading in the underlying shares, but we find little evidence that ETF shocks are transmitting noise to their underlying portfolios.

If liquidity providers are hedging their exposure to underlying price shocks, a sequence of mispricing events might lead to imbalances between fund and constituent shares in a market maker's inventory. ETFs have a second source of liquidity, however, that mitigates much of the risk associated with market maker inventory accumulation. Specifically, designated market makers, known as Authorized Participants (APs), can exchange the underlying portfolio for ETF shares directly with the fund sponsor.[3] Because ETF shares can be created and redeemed at the fund's closing net asset value (NAV), APs are not required to speculate whether mispricings will eventually converge. When inventories become large enough, liquidity providers simply sell their accrual of constituent securities or fund shares in the ETF's primary market, then use the resulting creation or redemption to offset any accumulated short position.[4] With this in mind, we repeat our PVAR analysis at a daily frequency, while including measures of daily mispricing, creation/redemption activity and closing discount/premium as additional time series. Similar to our intraday PVAR analysis, we find little evidence that ETF trading in the secondary market impacts constituent prices. Likewise, we observe that portfolio returns are also unaffected by the fund's primary market activity.

While our analysis focuses on direct tests of the relation between ETF and underlying returns and order flows, there are two empirical details that are also worth highlighting. First, as discussed above, the argument for ETF noise transmission presumes that ETFs offer greater liquidity than their underlying securities. Across all of our 94 million fund-minutes, the weighted-average quoted spread (hereafter, quoted spread) of the underlying securities only exceeds that of the ETF in roughly 45% of our observations.[5] This means that for our domestic equity focused sample, ETFs do not offer superior liquidity relative to their underlying securities on average. Second, it is also assumed in the noise transmission argument that ETF order flow shocks generate arbitrage opportunities, the resolution of which underlies transmission. However, only 1.66% of our 94 million fund-minute observations represent tradeable arbitrage opportunities.

Overall, while our findings preserve the possibility that nonfundamental shocks to ETF prices may extend to the constituent securities in some way, they also suggest that arbitrage trading is not a primary channel through which these disruptions are transmitted.

2. Data and variable definitions

From the CRSP Survivor-Bias-Free US Mutual Fund Database, we collect monthly holdings for 423 passively managed US equity funds between 2006 and 2015.[6] Together, these ETFs account for nearly \$1.16 trillion of the \$1.23 trillion total net assets invested in US domestic equity ETFs during 2015. Due to the prevalence of errors in self-reported holdings, we verify the accuracy of each Mutual Fund Database position by corroborating the holding's value using the CRSP Daily Stock File. The details of this verification process are provided in Appendix A.

Based on best bid and ask prices collected from the Trade and Quote Database (TAQ), we are able to calculate quotes during 94,605,900 fund-minutes for both the ETFs (ETF_t^{Bid} and ETF_t^{Ask}) and their underlying portfolios (Und_t^{Bid} and Und_t^{Ask}). In appendices B and C, we describe the construction of bid and ask prices for the ETFs and their underlying portfolios. At each time t, we also calculate percentage returns ($RetRaw_t$) based on changes in the midpoint between bid and ask quotes for both the ETF and its underlying portfolio. For most of the securities in our analysis, the distribution of these intraday midpoint returns is highly leptokurtic. Thus, to minimize the influence of extreme price changes, we examine intraday returns (Ret_t) after performing the following adjustment:

$$Ret_t = \begin{cases} \ln(1 + RetRaw_t), & if\ RetRaw_t \geq 0 \\ -\ln(1 + |RetRaw_t|), & if\ RetRaw_t < 0 \end{cases}. \quad (1)$$

By computing the midpoint return's natural log independently for positive and negative values, we preserve the distribution's natural symmetry around zero.[7]

Next, we calculate directional order flow ($VolDiffRaw_t$), scaled by average total volume and reported in basis points, during each period t:

$$VolDiffRaw_t = \frac{BuyVol_t - SellVol_t}{\frac{1}{T} \times \sum_{t=1}^{T} (BuyVol_t + SellVol_t)} \times 10,000, \quad (2)$$

where $BuyVol_t$ is the dollar value of all trades occurring at prices above the prevailing midpoint, $SellVol_t$ represents

[3] Malamud (2016) argues that reductions in primary market trading costs may strengthen the shock propagation channel that allows ETFs to affect stock returns. Petajisto (2017) suggests that primary market activity may not always provide an affordable way to correct price deviations between a fund and its NAV.

[4] Another form of potential ETF arbitrage involves a pairs trading strategy, whereby investors exploit price discrepancies between correlated equivalent assets. Box et al. (2019, 2020) document competition between ETFs that hold nearly identical portfolios. Marshall et al. (2013) study price discrepancies between two ETFs that track the S&P 500 and, ultimately, find evidence of intraday arbitrage. Likewise, Broman (2016) reports that unexpected shocks to ETF trading activity are correlated across funds with similar holdings.

[5] Underlying bid and ask quotes are measured as the dollar-value-weighted mean bid and ask quotes across all of a fund's constituent securities. The quoted spread for the underlying portfolio is the difference between the two.

[6] The sample is limited to the 423 passively managed US equity funds for which we can accurately identify daily holdings, as described in Appendix A, and match those holdings to TAQ data. All analyses include all funds in this sample.

[7] Using logged returns, instead of the symmetric log transformation described in Eq. (1), exaggerates the magnitude of negative outliers, inducing negative skewness into both ETF and underlying returns. Nonetheless, our analyses are robust to the method of log transformation. To further ensure that our approach for mitigating outliers is not driving the results, we repeat our primary PVAR analysis using entirely untransformed variables. The resulting impulse response functions (IRF) and forecast error variance decomposition (FEVD) are included in Appendix D.

Table 1
Summary statistics for midpoint return and order imbalance variables.
Ret_t^{Und} represents the midpoint return of the underlying securities during period t defined in Eq. (1). Likewise, Ret_t^{ETF}, represents the midpoint return of the ETF. $VolDiff_t^{Und}$ and $VolDiff_t^{ETF}$, defined in Eq. (3), measure the difference between buy and sell volume for the underlying securities and ETF shares, respectively. Correlations are estimated across the entire sample.

	Summary Statistics								Correlations		
	Mean	P1	P10	P25	P50	P75	P90	P99	Ret_t^{Und}	Ret_t^{ETF}	$VolDiff_t^{Und}$
Panel A: 1-Minute Window											
Ret_t^{Und}	0.00	-0.17	-0.05	-0.02	0.00	0.02	0.05	0.18			
Ret_t^{ETF}	0.00	-0.18	-0.05	-0.02	0.00	0.02	0.05	0.18	81.6%		
$VolDiff_t^{Und}$	-0.01	-8.66	-7.49	-6.66	0.00	6.66	7.49	8.68	9.5%	8.4%	
$VolDiff_t^{ETF}$	0.24	-11.01	-7.41	0.00	0.00	0.00	8.01	11.28	-1.2%	-2.0%	2.2%
Panel B: 5-Minute Window											
Ret_t^{Und}	0.00	-0.38	-0.13	-0.05	0.00	0.05	0.12	0.38			
Ret_t^{ETF}	0.00	-0.38	-0.13	-0.05	0.00	0.05	0.12	0.38	91.8%		
$VolDiff_t^{Und}$	-0.09	-9.72	-8.57	-7.77	-3.45	7.73	8.55	9.74	19.9%	18.8%	
$VolDiff_t^{ETF}$	0.63	-12.64	-9.89	0.00	0.00	8.20	10.36	12.83	-0.3%	-0.3%	2.0%
Panel C: 10-Minute Window											
Ret_t^{Und}	0.00	-0.49	-0.17	-0.07	0.00	0.07	0.17	0.49			
Ret_t^{ETF}	0.00	-0.50	-0.17	-0.07	0.00	0.07	0.17	0.50	94.3%		
$VolDiff_t^{Und}$	-0.10	-10.19	-9.05	-8.26	-4.04	8.20	9.03	10.22	23.2%	22.2%	
$VolDiff_t^{ETF}$	0.87	-13.25	-10.75	-8.14	0.00	9.56	11.20	13.41	0.2%	0.4%	2.3%

the dollar value of trades occurring below the quoted midpoint, and T is the total number of periods within each trading day. As with midpoint return, we use a symmetric log transformation of $VolDiffRaw_t$ to minimize the effect of outliers on our analyses:

$$VolDiff_t = \begin{cases} \ln(1 + VolDiffRaw_t), & if\ VolDiffRaw_t \geq 0 \\ -\ln(1 + |VolDiffRaw_t|), & if\ VolDiffRaw_t < 0 \end{cases}. \tag{3}$$

Table 1 provides summary statistics for each of the four variables of interest, Ret_t^{Und}, Ret_t^{ETF}, $VolDiff_t^{Und}$, and $VolDiff_t^{ETF}$, calculated over 1-, 5-, and 10-minute windows within each trading day. Regardless of frequency, constituent midpoint returns and directional order flow demonstrate minimal skewness and both have average values centered near zero. For ETFs, however, Table 1 reports that the average $VolDiff_t^{ETF}$ is positive across all three frequencies. Sellers, of any asset, should raise prices in response to strong demand and lower prices whenever demand is weak. If order imbalance captures the net demand for a security, it is not obvious why liquidity providers would choose to offer ETFs at prices low enough to encourage a disparity between buy and sell volume. Persistent order imbalances, especially over such a lengthy sample period, raise the possibility that ETF prices might not respond to buying and selling pressure in the same way as their underlying securities.

For a cursory look at how price changes and order imbalances might be associated, Table 1 also reports the estimated correlation coefficients between each of our four variables. Not surprisingly, we find a strong connection between ETF and underlying portfolio returns across all frequencies, with correlations becoming even stronger as measurement frequency falls. We also see that the correlation between Ret_t^{Und} and $VolDiff_t^{Und}$ is positive, implying that directional order flow usually trends in the same direction as price changes for the underlying securities. Once again, this association becomes stronger as the measurement window lengthens. For ETFs, however, Table 1 provides little evidence that directional order flow is positively associated with their own midpoint returns. Furthermore, the estimated correlation between Ret_t^{ETF} and $VolDiff_t^{ETF}$ is even slightly negative at the 1- and 5 min frequencies, whereas ETF returns are positively related to order imbalances within their underlying portfolios. Thus, for ETFs, price changes coincide with buying and selling pressure in their constituent securities, but not with trading in their own shares.

In previous studies, it is presumed that order imbalance in ETF shares, possibly initiated by short-term noise traders, pushes fund prices away from fundamental values, and that these fluctuations are impounded into underlying securities through an arbitrage mechanism. From Table 1, it seems unlikely that ETF order imbalances are contemporaneously influencing their own returns, much less those of their underlying portfolios. For the passively managed domestic equity strategies populating our sample, persistent order imbalances, and negligible associations between trading and returns, suggest that underlying portfolio values determine ETF bid and offer prices, not buying and selling pressures in the fund's shares.

3. Intraday analysis

Even though bivariate analysis does not reveal a contemporaneous association between ETF trading and portfolio returns, a more robust empirical approach might reveal a link between ETF order flow and constituent prices. Multiple confounding empirical characteristics may obscure this link. For instance, shocks to the supply and demand of underlying shares, regardless of whether those shocks are related to ETF activity, might not impact security prices contemporaneously. Also, momentum in either order imbalance or returns could mask a fundamental connection between these variables. Lastly, if order imbalances and price changes for ETFs and their constituents are jointly determined, measuring responses to independent shocks in one of these time series would require an approach that accounts for all potential interdependencies.

One approach to addressing the interdependencies between these variables is panel vector autoregression (PVAR). More commonly used vector autoregression (VAR) models account for intertemporal dependence between groups of macroeconomic variables (see, e.g., Hamilton, 1985; Koijen *et al.*, 2017; and Duffee, 2018). Unlike structural models with simultaneous equations, VAR models require little knowledge of the economic forces that impact each time series. Variables are usually treated as endogenous within a VAR system, but the impact of shocks to individual time series can still be modeled through an impulse response function (IRF) and forecast error variance decomposition (FEVD). Holtz-Eakin *et al.* (1988) extend the estimation of vector autoregressive systems to panel data, accounting for heterogeneity in levels, variances, and the time series correlation patterns of each panel cross-section. To do this, they apply an instrumental variables approach to the quasi-differenced autoregressive equations (Love and Zicchino, 2006; Abrigo and Love, 2016). In our setting, PVAR allows us to simultaneously quantify the reaction of ETF and portfolio returns and order flow to independent innovations in these same variables using a large cross-section of fund-days.[8]

The following system of linear equations represents a PVAR specification of order p:

$$\boldsymbol{Y}_{fdt} = \boldsymbol{Y}_{fdt-1}\boldsymbol{A}_1 + \boldsymbol{Y}_{fdt-2}\boldsymbol{A}_2 + \ldots + \boldsymbol{Y}_{fdt-p}\boldsymbol{A}_p + \boldsymbol{\alpha}_{dt} + \boldsymbol{\gamma}_{fd} + \boldsymbol{\varepsilon}_{fdt}, \tag{4}$$

where $\boldsymbol{Y}_{fdt}$ is a (1×4) vector of dependent variables, Ret^{Und}_{fdt}, Ret^{ETF}_{fdt}, $VolDiff^{Und}_{fdt}$ and $VolDiff^{ETF}_{fdt}$ for each fund (f), date (d), and time $(t \in \{1, 2, \ldots, T\})$. The vector $\boldsymbol{\alpha}_{dt}$ contains variable-specific fixed effects for each date-time combination, and $\boldsymbol{\gamma}_{fd}$ is vector of variable-specific panel effects for each fund-date. The parameters to be estimated are included in the 4×4 matrices $\boldsymbol{A}_1$, $\boldsymbol{A}_2$,..., $\boldsymbol{A}_p$, and the vector $\boldsymbol{\varepsilon}_{fdt}$ consists of idiosyncratic error terms for each dependent variable, such that $E[\boldsymbol{\varepsilon}_{fdt}] = 0$, $E[\boldsymbol{\varepsilon}'_{fdt}\boldsymbol{\varepsilon}_{fdt}] = \boldsymbol{\Sigma}$ and $E[\boldsymbol{\varepsilon}'_{fdt}\boldsymbol{\varepsilon}_{fds}] = 0$ for all $t > s$. To correct for estimate biases due to panel effects with lagged dependent variables, we utilize an instrument set made up of all future observations.[9]

The consistency of generalized method of moments (GMM) estimation relies on optimal model and moment selection. Andrews and Lu (2001) propose criteria resembling the widely used Bayesian information criterion to select each specification's order p and how many lags q of the dependent variables form the basis of the moment conditions.[10] Once the optimal model has been chosen, IRFs, derived from the parameter estimates of $\boldsymbol{A}_1$, $\boldsymbol{A}_2$,..., $\boldsymbol{A}_p$, describe each dependent variable's evolution following some exogenous shock. These IRFs have no causal interpretation, however, because the idiosyncratic errors, $\boldsymbol{\varepsilon}_{fdt}$, are likely to be correlated contemporaneously.

To determine whether independent shocks to one variable, such as $VolDiff^{ETF}_{fdt}$, impact future realizations of another, like Ret^{Und}_{fdt}, we rely on FEVD analysis for most of our qualitative interpretations. Here, idiosyncratic errors are orthogonalized to isolate the contribution of exogenous shocks in one variable to the forecast-error variances of all the others. Sims (1980) proposes a Cholesky decomposition of $\boldsymbol{\Sigma}$ whereby an instantaneous causal ordering of each time series is specified by their arrangement within $\boldsymbol{Y}_{fdt}$.[11] For the sequence $[Ret^{Und}_{fdt}, Ret^{ETF}_{fdt}, VolDiff^{Und}_{fdt}, VolDiff^{ETF}_{fdt}]$, shocks to fund order flow are orthogonalized relative to the immediate impacts of each preceding variable. Thus, FEVD analysis can tell us how much independent shocks to ETF order flow imbalance, $VolDiff^{ETF}_{fdt}$, at $t = 0$ influence realizations of $VolDiff^{Und}_{fdt}$, Ret^{ETF}_{fdt}, and, most importantly, Ret^{Und}_{fdt} during subsequent periods.[12]

3.1. Full sample panel vector autoregression results

To ease the computational burden of estimating Eq. (4), the variables Ret^{Und}_{fdt}, Ret^{ETF}_{fdt}, $VolDiff^{Und}_{fdt}$, and $VolDiff^{ETF}_{fdt}$ are centered across date and time combinations so that variable-specific fixed effects $\boldsymbol{\alpha}_{dt}$ can be omitted from the specification.[13] Each time series is also standardized by their fund-specific intraday volatilities realized during the

[8] Panel vector autoregressions are used in a similar context by Hilscher et al. (2015), Hollifield et al. (2017), and Lee et al. (2018).

[9] While the bias approaches zero as the total number of time periods increases, Judson and Owen (1999) find significant bias even when $T = 30$. Based on the assumption that idiosyncratic errors are serially uncorrelated, consistent GMM estimators have been proposed whereby parameter estimates are based on a first-difference transformation of the dependent variables. Thus, lagged levels and differences of $\boldsymbol{Y}_{fdt}$ become eligible instruments for the transformed dependent variables. With first-difference transformation, a second-order PVAR requires that $T_i \geq 5$ realizations are available for each subject. Here, $\Delta\boldsymbol{Y}_{fdt-1}$ is modeled as a function of $\Delta\boldsymbol{Y}_{fdt-1} = \boldsymbol{Y}_{fdt-1} - \boldsymbol{Y}_{fdt-2}$ and $\Delta\boldsymbol{Y}_{fdt-2} = \boldsymbol{Y}_{fdt-2} - \boldsymbol{Y}_{fdt-3}$, so the levels $\boldsymbol{Y}_{fdt-1}$, $\boldsymbol{Y}_{fdt-2}$ and $\boldsymbol{Y}_{fdt-3}$ are ineligible as instruments. Thus, the levels $\boldsymbol{Y}_{fdt-4}$ and $\boldsymbol{Y}_{fdt-5}$ must be available for the model to be just-identified. Instead of first-differencing, Arellano and Bover (1995) propose a forward orthogonal deviation that subtracts the average of all future observations and preserves all but the most recent realizations for the instrument set. Further, to improve efficiency, especially in unbalanced panels, Holtz-Eakin et al. (1988) recommend substituting missing observations with zero.

[10] The Andrews and Lu (2001) model selection criteria suggest that the specification's order p should equal 5, 9, and 16 for the 1-, 5-, and 10-minute window samples, respectively. In all cases, q is suggested to equal $p + 3$.

[11] The ordering is such that the first time series in $\boldsymbol{Y}_{fdt}$ may have an immediate impact on all other variables. The second time series may have an instantaneous impact on the remaining components of $\boldsymbol{Y}_{fdt}$, excluding the first variable, and so on. This type of causality is often referred to as Wold causality.

[12] This ordering ensures that examinations of ETF variables measure shocks that are independent of underlying returns and order flows, a necessary condition for assessing how ETFs impact underlying securities Brown et al., 2019). We also estimate Eq. ((4) with several different orderings, but qualitative inferences regarding the impact of fund order flow on subsequent constituent prices are unchanged by these alternative specifications.

[13] Centering variables across date and time combinations removes the systematic component of intraday returns and order flow during each period.

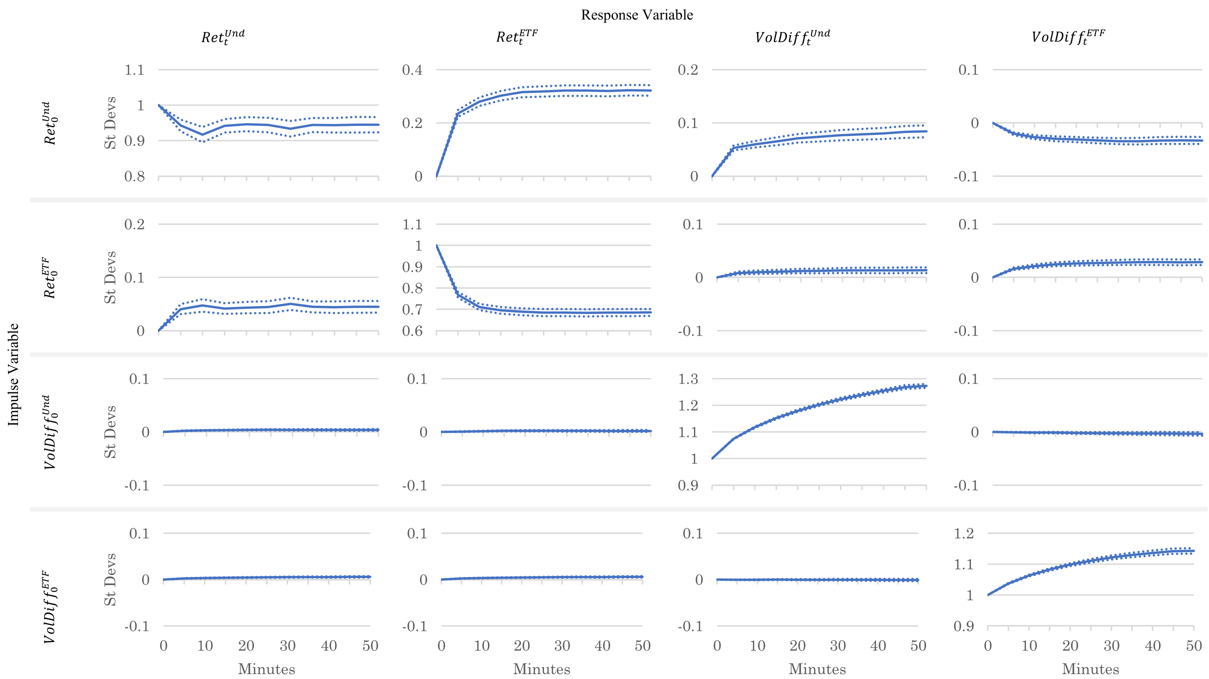

Fig. 1. 5-minute cumulative impulse response functions
This figure depicts the cumulative impulse response functions (IRF) derived from the estimated parameters from Eq. (4). These IRFs describe each dependent variable's evolution following a one standard deviation shock in the associated impulse variable. The four variables included are Ret_t^{Und} and Ret_t^{ETF}, as defined in Eq. (1), and $VolDiff_t^{Und}$ and $VolDiff_t^{ETF}$, as defined in Eq. (3), measured over 5-minute intervals. Confidence intervals, 97.5% and 2.5%, are denoted by dotted lines.

prior 200 days, with a minimum of 50 days required for inclusion in the analysis. After estimating the parameters from Eq. (4), we derive ten-period IRFs following a unit shock to the idiosyncratic error terms of each dependent variable.

Fig. 1 depicts cumulative IRFs based on our parameter estimates, which describe a variable's accumulated reaction in each of the proceeding ten periods. Response variables are denoted along the top of each column, and impulse variables are indicated to the left of each row. The 97.5% and 2.5% confidence intervals are designated by dotted lines above and below the average impulse response. By standardizing Ret_{fdt}^{Und}, Ret_{fdt}^{ETF}, $VolDiff_{fdt}^{Und}$, and $VolDiff_{fdt}^{ETF}$ prior to the estimation of Eq. (4), the cumulative IRFs in Fig. 1 describe each variable's expected reaction following a one-standard deviation shock to an impulse variable during $t = 0$. While only 5 min results are tabulated, cumulative IRFs from the 1- and 10 min samples are qualitatively similar.

Despite their inability to generate causal inferences, cumulative IRFs can help us understand how variables within a dynamic system interact with each other on average. The IRFs along the diagonal of Fig. 1 represent the cumulative response of each variable (Ret_{fdt}^{Und}, Ret_{fdt}^{ETF}, $VolDiff_{fdt}^{Und}$, and $VolDiff_{fdt}^{ETF}$) from a one standard deviation shock to itself. For the underlying, we observe a small reversal (5% dissipation to 0.95 at 50 min after the shock) compared with a strong continuation in underlying directional order flow. For the ETF, we observe a larger reversal (32% dissipation to 0.68) and a more modest continuation in the ETF directional order flow (to 114%). Across the top row, a positive shock to underlying prices is followed by a large positive ETF return response (0.32 standard deviations).[14] The directional order flow responses, however, present a more interesting pattern. The positive underlying return shock is followed by buying in the underlying shares (0.08 standard deviation response) and selling in the ETF (-0.03 standard deviation response). Here, fund trading trends in the opposite direction as price changes suggesting, once again, that ETFs might respond differently to price pressures than their underlying portfolios.

In the second row of Fig. 1, even though we observe ETF price shocks generating positive responses in underlying returns and order flow, the cumulative reactions never amount to more than a small fraction of one standard deviation. For comparison, a one standard deviation shock to Ret_t^{ETF} only generates an accumulated 0.06 standard

[14] While IRFs depict hypothetical exogenous shocks to idiosyncratic error terms, realized shocks are often correlated between certain variables. Consider a scenario where idiosyncratic shocks to Ret_{fdt}^{Und} and Ret_{fdt}^{ETF} are only ~70% correlated during a typical 5-minute period but, eventually, the prices of ETFs and their portfolios must converge. Without continuations or reversals, ~30% of the shock in one of these time series would manifest itself in the other during subsequent periods. Despite full convergence between their prices, the cumulative IRF would only suggest a ~0.3-unit response in one variable *resulting from* a unit shock to the other.

Table 2
Intraday panel vector autoregression forecast-error variance decomposition.
This table presents the fraction of forecasted error variance explained by exogenous shocks to impulse variables after ten periods. The four variables included are Ret_t^{Und} and Ret_t^{ETF}, as defined in Eq. (1), and $VolDiff_t^{Und}$ and $VolDiff_t^{ETF}$, as defined in Eq. (3), measured over 1-, 5-, and 10-minute intervals. Shocks are orthogonalized from top to bottom in the order presented.

		Panel A: 1 Min Window			
		Response Variable			
		Ret_{10}^{Und}	Ret_{10}^{ETF}	$VolDiff_{10}^{Und}$	$VolDiff_{10}^{ETF}$
Impulse Variable	Ret_0^{Und}	99.73%	18.67%	0.19%	0.00%
	Ret_0^{ETF}	0.26%	81.32%	0.01%	0.05%
	$VolDiff_0^{Und}$	0.01%	0.00%	99.80%	0.00%
	$VolDiff_0^{ETF}$	0.00%	0.01%	0.00%	99.95%
	Observations	87,421,327			
		Panel B: 5 Min Window			
		Response Variable			
		Ret_{10}^{Und}	Ret_{10}^{ETF}	$VolDiff_{10}^{Und}$	$VolDiff_{10}^{ETF}$
Impulse Variable	Ret_0^{Und}	99.83%	37.14%	0.41%	0.01%
	Ret_0^{ETF}	0.16%	62.86%	0.01%	0.01%
	$VolDiff_0^{Und}$	0.00%	0.00%	99.58%	0.00%
	$VolDiff_0^{ETF}$	0.00%	0.00%	0.00%	99.99%
	Observations	15,309,936			
		Panel C: 10 Min Window			
		Response Variable			
		Ret_{10}^{Und}	Ret_{10}^{ETF}	$VolDiff_{10}^{Und}$	$VolDiff_{10}^{ETF}$
Impulse Variable	Ret_0^{Und}	99.89%	43.20%	0.51%	0.01%
	Ret_0^{ETF}	0.11%	56.80%	0.00%	0.00%
	$VolDiff_0^{Und}$	0.00%	0.00%	99.48%	0.00%
	$VolDiff_0^{ETF}$	0.00%	0.00%	0.00%	99.99%
	Observations	6821,930			

deviation response in underlying returns after 50 min, whereas a unit shock to Ret_t^{Und} generates a 0.34 standard deviation reaction in ETF returns over the same horizon. Furthermore, confidence intervals around each IRF show that this difference is statistically significant. For the other off-diagonal figures in the two bottom rows, we see almost no evidence that demand shocks to the underlying or ETF shares are correlated with their subsequent returns. Similar to our correlation findings, the IRF results are not consistent with ETF trading negatively impacting the underlying portfolio through an arbitrage mechanism.

To better understand the causal links between ETF trading and portfolio returns, we now turn to a FEVD analysis of our parameters estimated from Eq. (4). The contributions of an impulse variable to the forecast-error variance of each response variable, after ten periods, are tabulated in Table 2. Panels A, B, and C report FEVD estimates for the 1-, 5-, and 10 min windows, respectively.

Regardless of the frequency with which the variables are measured, we find almost no evidence that independent shocks to ETF returns (Ret_t^{ETF}) or ETF order imbalance ($VolDiff_t^{ETF}$) impact future price changes or trading in the underlying securities. While fund returns may contribute a tiny fraction to their portfolio return's total forecast-error variance, 0.26% in the 1 min sample, orthogonalized shocks to ETF order flow have no relation to future innovations in Ret_t^{Und} or $VolDiff_t^{Und}$. Conversely, the results in Table 2 also suggest that future fund returns are influenced by price changes in their underlying securities, or at least by underlying and fund price changes that occur simultaneously.[15] Roughly 18.67% of Ret_t^{ETF}'s forecast-error variance can be explained by shocks to constituent prices in the 1-minute sample, rising to 43.20% in the 10 min sample. Altogether, our results imply that, while future ETF prices respond to, potentially simultaneous, shocks in their underlying portfolios, orthogonalized ETF shocks do not affect their portfolios reciprocally during later periods.[16]

3.2. ETF liquidity

Across a broad sample of US equity ETFs, our PVAR analysis demonstrates that independent shocks to fund prices and order imbalances have little effect on subsequent constituent returns and trading. It is possible, however, that the relation depends on a liquidity differen-

[15] Idiosyncratic shocks to underlying returns are not orthogonalized relative to Ret_t^{ETF}, $VolDiff_t^{Und}$, and $VolDiff_t^{ETF}$ because of the variable ordering within $\boldsymbol{Y}_{fdt}$. Statistically, it is not possible to determine whether temporally correlated shocks to the fund and portfolio are attributable to one or the other. As basket securities, however, ETF returns should be impacted instantaneously by shocks to constituent prices. This presumption is also supported by our graphical analysis.

[16] The lack of association between Ret_t^{Und} and $VolDiff_t^{Und}$ in the FEVD analysis may seem curious given their positive correlation reported in Table 1. FEVD only measures the contribution of independent shocks in one time series to the forecast-error variance of another. While innovations in Ret_t^{Und} and $VolDiff_t^{Und}$ are obviously related contemporaneously, the FEVD results show that orthogonalized shocks to order flow do not impact future realizations of returns.

Table 3

Intraday 5-minute panel vector autoregression forecast-error variance decomposition across size categories.

This table presents the fraction of forecasted error variance explained by exogenous shocks to impulse variables after ten periods. The four variables included are Ret_t^{Und} and Ret_t^{ETF}, as defined in Eq. (1), and $VolDiff_t^{Und}$ and $VolDiff_t^{ETF}$, as defined in Eq. (3), measured over 5-minute intervals. Shocks are orthogonalized from top to bottom in the order presented. The PVAR specification is estimated separately on four subsamples divided on the market capitalization of the ETF and underlying securities. Funds and constituents are included in the "Large" group if their mean daily market capitalization falls in the top tercile of our sample during a particular month, "Small" if in the bottom two terciles. Underlying securities are similarly included in the Large group if the daily average of the value-weighted market capitalization of the underlying securities falls in the top tercile of our sample during a particular month, Small if in the bottom two terciles.

		Small underlying		Large underlying	
Small ETF		Response Variable			
		Ret_{10}^{Und}	Ret_{10}^{ETF}	$VolDiff_{10}^{Und}$	$VolDiff_{10}^{ETF}$
Impulse Variable	Ret_0^{Und}	99.93%	39.37%	0.40%	0.00%
	Ret_0^{ETF}	0.07%	60.63%	0.00%	0.01%
	$VolDiff_0^{Und}$	0.00%	0.00%	99.60%	0.00%
	$VolDiff_0^{ETF}$	0.00%	0.00%	0.00%	99.99%
		Response Variable			
		Ret_{10}^{Und}	Ret_{10}^{ETF}	$VolDiff_{10}^{Und}$	$VolDiff_{10}^{ETF}$
Impulse Variable	Ret_0^{Und}	99.85%	19.76%	0.54%	0.00%
	Ret_0^{ETF}	0.15%	80.24%	0.00%	0.01%
	$VolDiff_0^{Und}$	0.00%	0.00%	99.46%	0.00%
	$VolDiff_0^{ETF}$	0.00%	0.00%	0.00%	99.99%
Large ETF		Response Variable			
		Ret_{10}^{Und}	Ret_{10}^{ETF}	$VolDiff_{10}^{Und}$	$VolDiff_{10}^{ETF}$
Impulse Variable	Ret_0^{Und}	99.68%	51.93%	0.35%	0.03%
	Ret_0^{ETF}	0.32%	48.07%	0.00%	0.00%
	$VolDiff_0^{Und}$	0.00%	0.00%	99.64%	0.01%
	$VolDiff_0^{ETF}$	0.00%	0.00%	0.00%	99.96%
		Response Variable			
		Ret_{10}^{Und}	Ret_{10}^{ETF}	$VolDiff_{10}^{Und}$	$VolDiff_{10}^{ETF}$
Impulse Variable	Ret_0^{Und}	99.75%	46.37%	0.46%	0.01%
	Ret_0^{ETF}	0.25%	53.63%	0.03%	0.00%
	$VolDiff_0^{Und}$	0.00%	0.00%	99.51%	0.00%
	$VolDiff_0^{ETF}$	0.00%	0.00%	0.00%	99.98%

tial between ETFs and their underlying securities. For instance, Charupat and Mui (2011), Malamud (2016), Israeli *et al.* (2017), and Broman and Shum (2018) suggest that short-term traders, those that are most likely to introduce noise into security prices, are encouraged by the superior liquidity of ETFs. Thus, to identify cases where a fund's liquidity might be superior, relative to its underlying holdings, we divide our sample into four distinct subsets: Small Underlying and Small ETF, Small Underlying and Large ETF, Large Underlying and Small ETF, and Large Underlying and Large ETF. Funds are included in the "Large" group if their mean daily market capitalization falls in the top tercile of our sample during a particular month, and in the "Small" group if in the bottom two terciles. Underlying securities are similarly included in the Large group if the daily average of value-weighted market capitalization for the underlying securities falls in the top tercile of our sample during a particular month, and the Small group for average capitalizations in the bottom two terciles.[17]

After generating unique parameter estimates for Eq. (4) from each of the four subsamples, we tabulate results from a FEVD analysis in Table 3. We expect that disparities in liquidity will be greatest whenever ETFs are large relative to their underlying holdings.[18] Even though the impact of independent fund return shocks on underlying returns, Ret_t^{Und}, is greatest for the Small Underlying and Large ETF subset, these impulses contribute only 0.32% to the portfolio return's total forecast-error variance. Thus, even in a subset where the fund's market capitalization is most likely to exceed that of its average constituent, independent ETF price changes have almost no influence on subsequent portfolio returns.

Where we find the most meaningful variation across the four subsamples is in the share of ETF returns' total forecast-error variance that is attributable to the underlying, or potentially simultaneous, price changes. For funds that are large relative to the average market capitalization of their holdings, Table 3 demonstrates that future innovations in ETF prices are more strongly influenced, 51.93% of forecast-error variance, by shocks to constituent returns. Conversely, shocks to Ret_t^{Und} contribute only 19.76% to the forecast-error variance of Ret_t^{ETF} in the Large

[17] Underlying market capitalization is the dollar-weighted average of all stocks held by the ETF, based on their CRSP price and shares outstanding reported at the end of the previous year. The top tercile of ETFs (underlying portfolios) ranked by market capitalization accounts for 95.6% (78.6%) of aggregate value on average.

[18] Even amongst the largest ETFs, the market capitalization of the fund is usually dwarfed by the average size of their holdings. For instance, the largest ETF, the SPDR S&P 500, had an average market value of $263 billion in 2019 according to *The Wall Street Journal*. By comparison, the fund's largest holding, Microsoft, is worth more than $1 trillion.

Table 4
Intraday 5-minute panel vector autoregression forecast-error variance decomposition across spread difference quintiles.

This table presents the fraction of forecasted error variance explained by exogenous shocks to impulse variables after ten periods. The four variables included are Ret_t^{Und} and Ret_t^{ETF}, as defined in Eq. (1), and $VolDiff_t^{Und}$ and $VolDiff_t^{ETF}$, as defined in Eq. (3), measured over 5-minute intervals. Shocks are orthogonalized from top to bottom in the order presented. The PVAR specification is estimated separately on five subsamples divided daily on the mean minute-by-minute difference in spread between the fund and its underlying portfolio over the prior 200 days.

Impulse Variable	Response Variable	ETF Spread Difference Quintile				
		ETF Spread Low Relative to Underlying ↔ ETF Spread High Relative to Underlying				
		1	2	3	4	5
Ret_0^{Und}	Ret_{10}^{Und}	99.90%	99.90%	99.74%	99.95%	99.98%
	Ret_{10}^{ETF}	59.92%	64.46%	48.68%	43.51%	32.50%
	$VolDiff_{10}^{Und}$	0.66%	0.49%	0.48%	0.59%	0.53%
	$VolDiff_{10}^{ETF}$	0.02%	0.00%	0.00%	0.01%	0.01%
Ret_0^{ETF}	Ret_{10}^{Und}	0.10%	0.10%	0.25%	0.05%	0.01%
	Ret_{10}^{ETF}	40.07%	35.53%	51.31%	56.48%	67.50%
	$VolDiff_{10}^{Und}$	0.00%	0.00%	0.01%	0.00%	0.00%
	$VolDiff_{10}^{ETF}$	0.01%	0.01%	0.00%	0.00%	0.00%
$VolDiff_0^{Und}$	Ret_{10}^{Und}	0.00%	0.00%	0.01%	0.00%	0.00%
	Ret_{10}^{ETF}	0.00%	0.00%	0.00%	0.00%	0.00%
	$VolDiff_{10}^{Und}$	99.33%	99.51%	99.51%	99.40%	99.46%
	$VolDiff_{10}^{ETF}$	0.00%	0.00%	0.00%	0.00%	0.00%
$VolDiff_0^{ETF}$	Ret_{10}^{Und}	0.00%	0.00%	0.00%	0.00%	0.00%
	Ret_{10}^{ETF}	0.00%	0.00%	0.00%	0.00%	0.00%
	$VolDiff_{10}^{Und}$	0.00%	0.00%	0.00%	0.00%	0.00%
	$VolDiff_{10}^{ETF}$	99.97%	99.99%	99.99%	99.99%	99.99%

Underlying and Small ETF subsample. Instead of conducting more noise into constituent security prices, the superior liquidity of relatively large ETFs only seems to facilitate a more efficient response to innovations in portfolio values.

To examine how differences in liquidity between the funds and their underlying securities may affect our results, we also divide our 5 min sample into quintiles based on the ETF's and constituents' relative bid-ask spreads. Quintile breakpoints are based on the difference between average minute-by-minute intraday quoted spreads for the fund and its market-weighted portfolio during the previous 200 trading days. The first quintile consists of ETFs with bid-ask spreads that are low relative to their holdings, whereas the fifth quintile contains funds with quotes that are wide compared to the constituents.

The results in Table 4 show that independent shocks to ETF returns, Ret_t^{ETF}, and ETF order imbalance, $VolDiff_t^{ETF}$, still have almost no impact on future price changes or trading in the underlying securities, regardless of any disparities between fund and constituent quoted spreads. Furthermore, the small fraction of the portfolio return's total forecast-error variance that can be attributed to ETF returns does not seem to vary monotonically with relative liquidity. Here, 0.25% of the forecast-error variance of underlying returns is contributed by independent fund price changes in the middle subset, falling to 0.10% and 0.01% in the lowest and highest spread difference quintiles, respectively.

Once again, the most interesting differences across the subsamples are found in the response of Ret_t^{ETF} from orthogonalized shocks to Ret_t^{Und}. While variation across the quintiles is not perfectly monotonic, a larger proportion of the ETF return's total forecast-error variance is contributed by portfolio price changes when the fund's quoted spread is low relative to its constituents. Altogether, the results from Table 3 and Table 4 suggest that fund liquidity is not promoting the transfer of shocks from equity ETFs into their holdings.[19] Even if a fund's liquidity encouraged short-term trading, it might also dampen the price impact of nonfundamental demand shocks. Empirically, improvements in an ETF's relative liquidity only seem to improve the efficiency of the fund's reaction to underlying price changes without promoting a transfer of noise into security prices.

3.3. ETF mispricing

The results from our FEVD analysis strongly suggest that independent shocks to ETF prices or order imbalances have little economic impact on the future returns and trading of their portfolio securities. However, the arbitrage mechanism, as described in prior studies, requires at least some divergence between fund and underlying prices. Perhaps the effect of ETF trading on constituent shares is only observable when mispricing is large enough to incentivize arbitrageurs to enter opposing positions. To address this possibility, we examine the relation between ETF and

[19] Instead of trading directly in the fund's holdings, arbitrageurs could take an opposing position in some other highly correlated security. Similar to our strategy for classifying funds according to their relative bid-ask spreads, we also group ETFs by their degree of comovement with other funds. Just as in the previous analyses, we are unable to isolate a subset of ETFs where shocks to the demand or pricing of a fund affect its holdings after the shock. Funds without a strongly correlated alternative to trading directly in the underlying securities respond less efficiently to innovations in portfolio values. Yet, the absence of suitable alternatives to direct arbitrage still does not seem to encourage the transmission of price shocks from the ETF to its holdings. See Appendix E for more on our analysis of correlated alternatives.

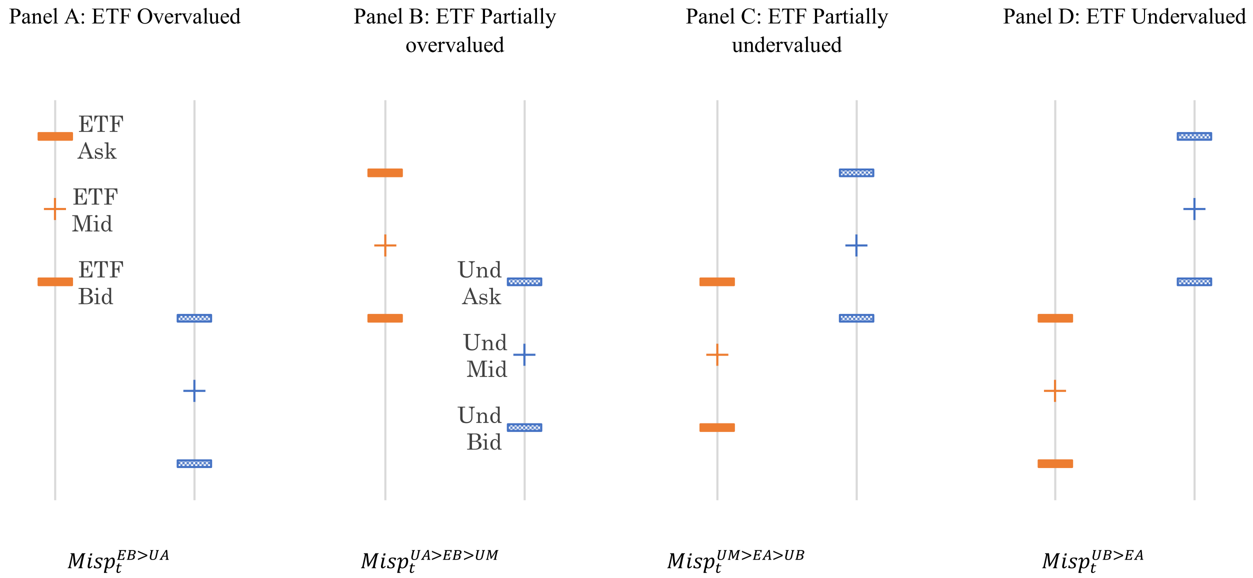

Fig. 2. Mispricing events
This figure depicts four mispricing associations between an ETF and its underlying portfolio.

underlying returns and order flow during episodes of significant price divergence.

3.3.1. Relative bid-ask spread inferred mispricing

Fig. 2 describes four mispricing associations between an ETF and its underlying portfolio. In all four cases, the ETF is trading at a premium or discount as measured by the percentage difference in midpoint prices. However, only Panel A and Panel D describe a tradeable arbitrage opportunity. In Panel A, for instance, the ETF's bid price is above the ask price for the underlying portfolio. Therefore, the simultaneous purchase of the underlying portfolio, and sale of the ETF, would result in realized profits if prices later converged. In Panel B, the midpoint price of the ETF is still higher than that of its underlying portfolio, but the simultaneous purchase of the underlying portfolio, and sale of the ETF at the prevailing ask and bid, respectively, would not result in a profitable trade.[20]

We next compare all ETF and underlying bid and ask prices at the end of every minute during our sample period. All fund-minute observations are designated as $Misp_t^{EB>UA}$, $Misp_t^{UA>EB>UM}$, $Misp_t^{UM>EA>UB}$ or $Misp_t^{UB>EA}$, in accordance with Fig. 2, or classified as $NoMisp_t$ when none of these conditions is satisfied. Table 5 reports the percentage of fund-minute observations that are mispriced across our entire sample. Out of 94,605,900 total fund-minute observations, 813,815 and 749,441 are classified as overvalued and undervalued, respectively. Thus, tradeable arbitrage opportunities exist between an ETF and its underlying portfolio approximately 1.66% of the time. The classifications $Misp_t^{UA>EB>UM}$ and $Misp_t^{UM>EA>UB}$ account for 3.69% and 3.63% of our observations, respectively, leaving 91.02% of fund-minutes designated as $NoMisp_t$.

While smaller ETFs are likely to be less liquid, and ETFs that hold smaller securities are probably more difficult to value, the results in Table 5 indicate that the proportion of overvalued and undervalued fund-minutes is roughly equal across all size subsamples. To understand why, quoted spreads for the ETF and underlying portfolio are also reported in the in the far-right columns of the table. As expected, the Large ETF and Large Underlying classifications have lower spreads. However, these narrower spreads mean that relatively minor discrepancies between the ETF and its portfolio can lead to tradeable arbitrage opportunities. Thus, the margin for error shrinks with fund and security size. Furthermore, contrary to the common presumption that superior ETF liquidity could attract short-term traders and introduce noise into constituent prices (Charupat and Mui, 2011; Malamud, 2016; Israeli *et al.*, 2017; and Broman and Shum, 2018), quoted spreads for US equity ETFs are usually wider than those of their underlying holdings. Only when ETFs are large relative to the size of their portfolios do transaction costs favor the fund.

3.3.2. Cumulative midpoint returns

To better understand the price behavior of ETFs and their underlying securities around arbitrage opportunities, we identify all of the bid-ask spread inferred mispricing events in our sample and cumulate the ETF and underlying portfolio midpoint return after each minute beginning at $t-10$ and extending through $t+30$.[21] Next, we

[20] While this does not rule out the possibility that a skilled arbitrageur could profit from the submission of non-marketable orders in the partially overvalued and undervalued scenarios depicted in Panels B and C respectively, the four different designations shown here help to categorize the degree and direction of potential mispricing between the ETF and underlying.

[21] We remove any observations that proceed another event by ten minutes or less. Extreme mispricing events, $Misp_t^{EB>UA}$ and $Misp_t^{UB>EA}$, are only excluded if there is another extreme mispricing event of the same direc-

Table 5
Frequency of mispricing events across size categories.
The following table reports the frequency of mispricing events, as well as the average quoted spreads from the end of each minute t, across fund and ETF size categories. All fund-minute observations are designated as $Misp_t^{EB>UA}$, $Misp_t^{UA>EB>UM}$, $Misp_t^{UM>EA>UB}$ or $Misp_t^{UB>EA}$, in accordance with Fig. 2, or classified as $NoMisp_t$ when none of these conditions is satisfied.

		Frequency					Average Quoted Spread in Basis Points	
	N	$Arb_t^{EB>UA}$	$Arb_t^{UA>EB>UM}$	$NoArb_t$	$Arb_t^{UM>EA>UB}$	$Arb_t^{UB>EA}$	Und	ETF
All Observations	94,305,900	0.86%	3.69%	91.02%	3.63%	0.79%	12.4	19.3
Small ETF	58,886,490	0.85%	1.70%	95.04%	1.65%	0.76%	13.7	26.0
Large ETF	35,419,410	0.88%	7.00%	84.34%	6.93%	0.85%	10.4	8.2
Small Underlying	58,121,310	0.83%	4.24%	89.95%	4.19%	0.78%	16.3	21.4
Large Underlying	36,184,590	0.91%	2.80%	92.75%	2.73%	0.81%	6.2	16.0
Small Und/Small ETF	41,385,630	0.80%	2.09%	94.33%	2.04%	0.74%	16.8	26.5
Small Und/Large ETF	16,735,680	0.91%	9.56%	79.12%	9.53%	0.89%	15.2	8.8
Large Und/Small ETF	17,500,860	0.97%	0.76%	96.72%	0.73%	0.81%	6.4	24.9
Large Und/Large ETF	18,683,730	0.85%	4.71%	89.02%	4.60%	0.82%	6.0	7.6

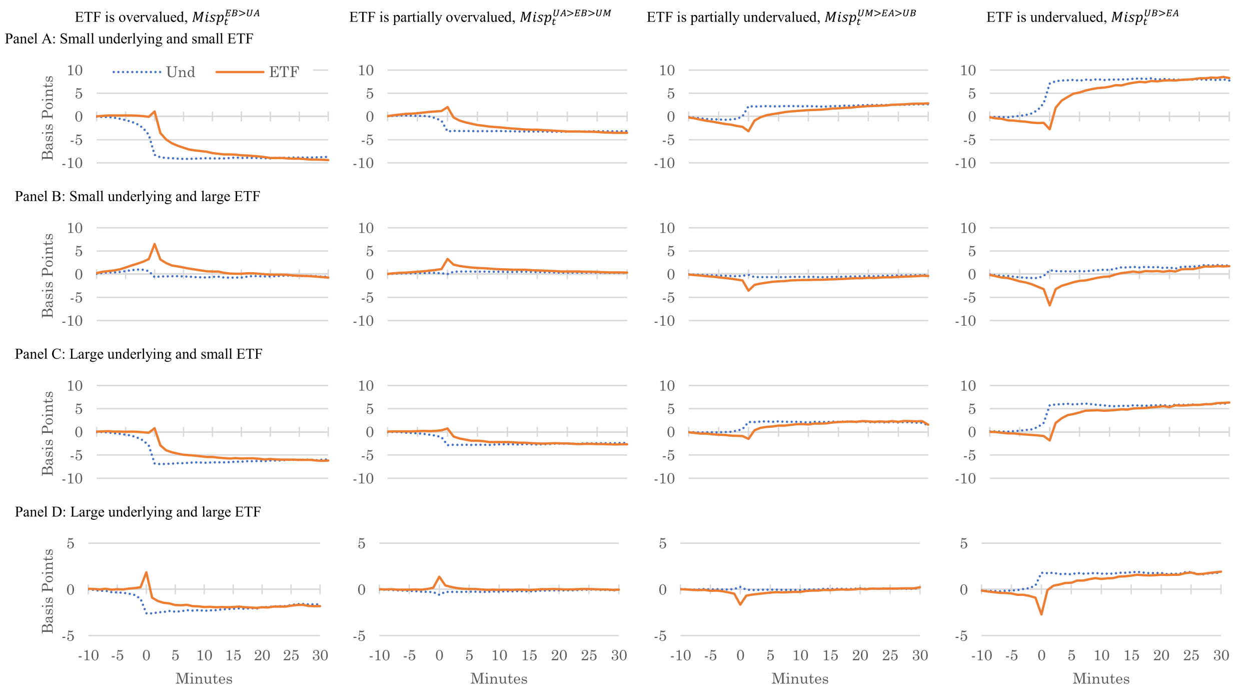

Fig. 3. Average midpoint returns around arbitrage opportunities
This figure shows the average midpoint returns for ETFs and constituents in a 40-minute window around arbitrage. Each panel presents average midpoint returns based on the size of the ETF and the underlying. The results are partitioned using the four mispricing conditions described in Fig. (2).

calculate the average cumulative midpoint return for each event-window-minute across all observations within a particular size category and then winsorize the variable at the 1st and 99th percentiles. When these minute-by-minute averages are depicted in Fig. 3, several results become clear. Except for Panel B, where the ETF is large relative to its holdings, the figure suggests mispricing events center around the arrival of information relevant to underlying security values. The portfolio's cumulative returns, represented by the dotted line, experience persistent declines (increases) during event windows where ETFs become overvalued (undervalued). Following a price shock to constituent prices during the first ten minutes, the portfolio stays at this new level for the remaining thirty minutes of the event window. Thus, these substantial adjustments to portfolio prices are not, on average, transient

tion in the prior ten minutes. Partial mispricing events, $Misp_t^{UA>EB>UM}$ and $Misp_t^{UM>EA>UB}$, are excluded if there is another partial or extreme mispricing event of the same direction in the prior ten minutes.

deviations from fundamental values. While this effect is strong for extreme mispricing ($Misp_t^{EB>UA}$ and $Misp_t^{UB>EA}$), it also holds in cases of partial mispricing ($Misp_t^{UA>EB>UM}$ and $Misp_t^{UM>EA>UB}$), especially for Small ETFs.

The sole exception to this behavior, when the ETF is large relative to the value of constituents (i.e., Panel B), fluctuations in the price of the underlying do not persist through the end of the event window. For these cases, which represent 17.75% of the sample's observations, arbitrage opportunities are less likely to be associated with material information about the underlying holdings. According to Table 5, this is the only size classification where transaction costs are lower in the ETF than in the underlying securities. Therefore, as suggested by Charupat and Mui (2011), Malamud (2016), Israeli *et al.* (2017), and Broman and Shum (2018), the superior liquidity of these ETFs might attract short-term traders that could introduce noise into the prices for constituent securities.

Fig. 3 is informative regarding this proposition as it allows us to compare the timing of price adjustments between an ETF and its constituent securities. First, we see that the 40-minute event window provides enough time to observe total convergence between the fund's price, represented by the solid line, and that of its portfolio. In three of the four size classifications, we also observe that the value of the underlying securities begins moving several minutes before the arbitrage opportunity materializes. Meanwhile, the ETF price does not begin adjusting to new information about the constituent securities until $t + 1$. As before, we observe different behavior when the ETF is large relative to the underlying portfolio. Here, the fund's price shifts several minutes before the arbitrage event. Furthermore, the prices of constituent securities appear to follow those of the ETF, at least temporarily, before the fund and portfolio shift back towards their $t - 10$ values. Thus, for the Small Underlying and Large ETF case, Fig. 3 provides some evidence that short-term fluctuations in ETF prices could introduce noise into their underlying portfolios. Yet, for all other size classifications, we observe persistent shocks to the constituent securities with ETF price adjustments lagging by several minutes.

One remaining empirical detail visible in each of the cases described by Fig. 3 is the peculiar short-term reversal in ETF midpoint returns between minutes $t - 1$ and $t + 1$. To better understand this result, we also provide the sample-average cumulative bid and ask returns around arbitrage opportunities in Fig. 4. The top and bottom of each vertical line represent the average path of the best offer and bid price, respectively, and the horizontal dash in the center represents the average path of the midpoint.

For the underlying securities, the bid and ask prices adjust smoothly and symmetrically to the arrival of new information during the minutes preceding an arbitrage opportunity. Conversely, the adjustment process is less balanced for the ETF's quotes. In Panel A, negative information lowers the prices of constituent securities and, at $t + 1$, the ETF's best offer price begins to fall as expected. For bid prices, however, we observe a sharp increase during minute t that temporarily pushes up midpoint prices. Here, the market appears to temporarily offer liquidity to ETF sellers at favorable prices before adjusting quotes downward to reflect new information about the constituent securities. We observe a similar pattern in Panel B, except that ETF bid prices adjust smoothly while offer prices fluctuate. In either case, these asymmetrical shifts in quoted prices could arise from conditioning our graphical analysis on arbitrage opportunities. Had the quotes adjusted smoothly and symmetrically, instead, the arbitrage opportunity would never materialize. Regardless, the overall behavior shown in Fig. 4 is still consistent with hedging on the part of liquidity providers, whereby aggressive price quotes for the ETF might serve to offset liquidity provision in the underlying securities.

3.3.3. On-balance volume

Next, we explore the effects of directional trading around arbitrage opportunities. While the results in the previous section suggest mispricing events are the result of an information shock to the underlying portfolio, studying the flow of marketable orders for the ETF and its underlying securities allows us to determine whether buying or selling pressures are also consistent with informed trading in the portfolio. In Fig. 5, we report on-balance volume for the ETF and its constituents during the same 40 min event windows depicted in Fig. 3. To calculate on-balance volume, we cumulate $VolDiffRaw_t$, defined in Eq. (2), after each minute from $t - 10$ through $t + 30$ and then winsorize at the 1st and 99th percentiles. Then, we take the average event-window-minute on-balance volume across all observations within a particular size category. Once again, we observe a consistent pattern for all three size classifications where the size of the ETF is not large relative to its constituents. Specifically, we find persistent buying or selling pressure in the underlying portfolio during the minutes immediately surrounding the arbitrage event, followed by relatively balanced trading during the remainder of the event window. Therefore, the underlying price shocks highlighted in Fig. 3 appear to be impounded by informed directional trading in the constituent securities. Conversely, on-balance volume for the ETF flows in the opposite direction as midpoint prices throughout the event window. Here, marketable ETF buy orders arrive more frequently as midpoint prices fall, whereas the intensity of ETF sell orders increases with rising prices.

For these situations where the underlying portfolio experiences a persistent price shock, marketable order flow can be classified as informed or uninformed depending on whether the direction of trading is consistent with the ETF or portfolio's return. Negative price shocks are accompanied by informed marketable sell orders for the constituents executed below the midpoint. Thus, for liquidity providers, negative price shocks lead to buildup of inventory in the underlying securities. As exposure to the portfolio rises, liquidity providers might choose to offset their position by selling shares of the ETF. From Fig. 3 and Fig. 5, lowering the offer price for ETF shares encourages uninformed marketable orders and reduces the number of fund shares in the liquidity providers' inventory. Thus, a series of mispricing events associated with negative price shocks would lead to an accumulation of constituent securities held by liquidity providers. Similarly, repeated positive shocks would lead to a buildup in ETF shares. Should

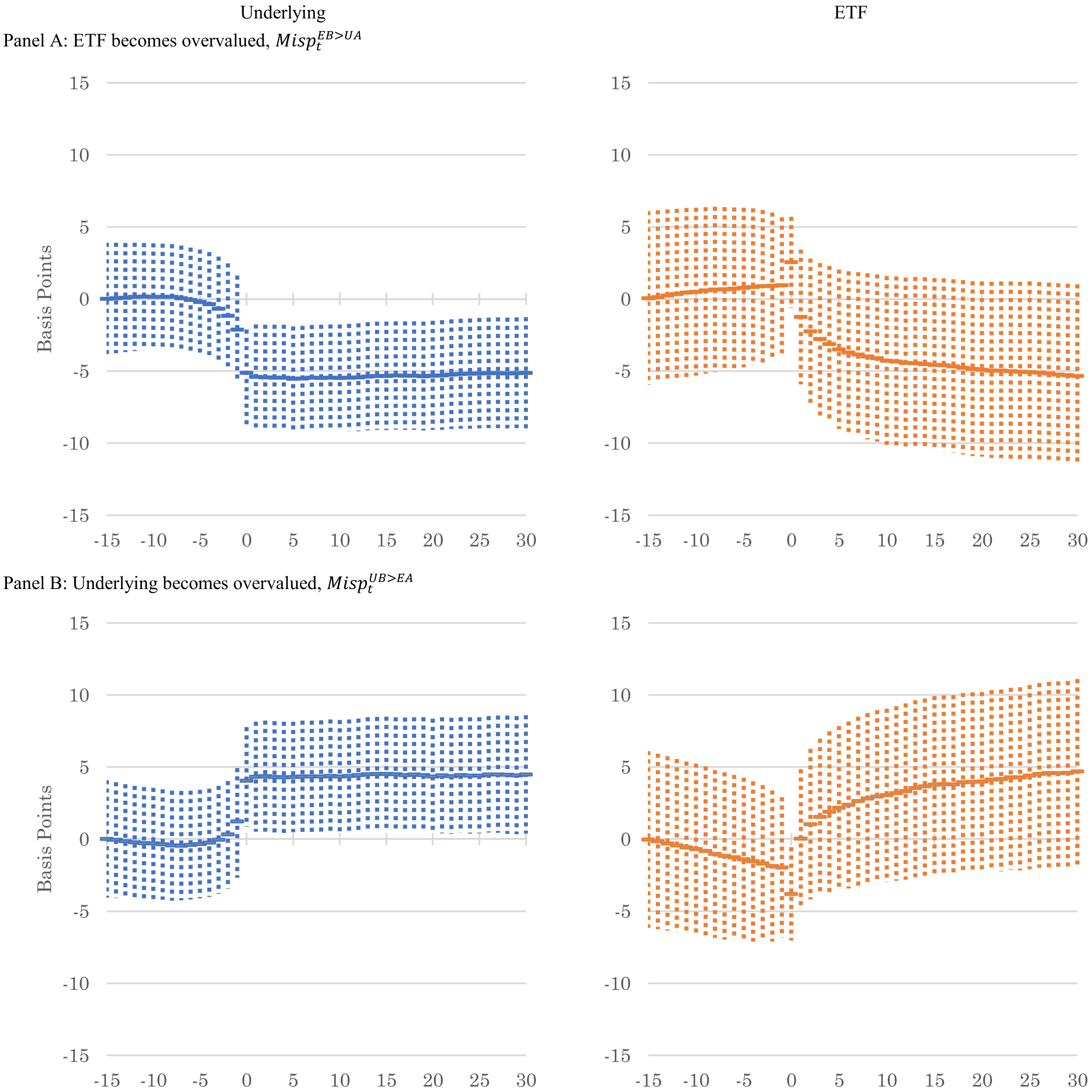

Fig. 4. Cumulative bid and ask returns around arbitrage opportunities
This figure presents cumulative bid and ask returns for ETFs and constituents in the 45-minute window around arbitrage, classified using the two clear arbitrage opportunities presented in Fig. 2 (i.e., when the ETF bid exceeds the underlying ask or when the underlying bid exceeds the ETF ask).

their inventories become large enough, liquidity providers could sell their accrual of constituent securities or fund shares in the ETF's primary market.

For the Small Underlying and Large ETF subsample, we observe less informed trading in the underlying portfolio and almost no evidence of liquidity provider hedging in the ETF. Here, transient shocks to the price of the ETF coincide, rather weakly, with fluctuations in ETF order flow. Yet, on-balance ETF volume returns to zero by the end of the window. Thus, it is difficult to determine why ETF prices adjust to create arbitrage opportunities in these situations. Furthermore, cumulative order imbalance for the constituents is negative when $Misp_t^{EB>UA}$, but roughly zero when $Misp_t^{UB>EA}$. Therefore, it is difficult to draw conclusions about whether trading in portfolios of small securities is influenced by short-term price fluctuations in larger ETFs.

3.3.4. Stochastic jumps

While the arbitrage mechanism proposed in prior studies requires some divergence between fund and underlying values, our graphical results suggest that these mispricing events are usually preceded by persistent shocks to con-

Fig. 5. Average on-balance volume around arbitrage opportunities
This figure shows the average on-balance volume for ETFs and constituents in a 40-minute window around arbitrage. Each panel presents average on-balance volume based on the size of the ETF and the underlying. The left (right) axis corresponds to the underlying (ETF). The results are partitioned using the four mispricing conditions described in Fig. 2.

stituent prices. At the same time, the very existence of arbitrage opportunities may signal unusual market conditions (e.g., limits-to-arbitrage, such as a lack of available arbitrage capital), which limit the generalizability of the results. As an alternative, we examine stochastic jumps in the returns of both ETFs and their constituent securities. Based on the jump identification measure developed by Lee and Mykland (2008), we use the prior fifteen days of 1 min return volatility to test the null of a diffusion process for each fund and portfolio.[22] Across our entire sample we find that 0.27% of 1 min underlying returns contain a jump, compared with 0.39% of ETF observations.[23]

With these jump events centered at $t = 0$, average cumulative midpoint returns are depicted in Panel A of Fig. 6. During the minutes surrounding underlying portfolio jumps, pictured on the left side of Fig. 6, average fund and constituent cumulative returns overlap almost perfectly whether the shock is positive or negative. Conversely, ETF jumps, depicted on the right side, are anticipated by small price changes in the underlying securities on average. While cumulative fund returns are slightly larger in magnitude than those of the portfolio following ETF jumps, price changes persist for the remaining thirty minutes of the event window in all four cases. Thus, fund and constituent price shocks are not, on average, transient deviations from fundamental values in our sample.

As with our graphical analysis conditioned on mispricing events, the paths of average on-balance volume depicted in Fig. 6 are consistent with liquidity providers hedging their exposure to informed trading in the underlying shares. Similar to our examination of arbitrage opportunities, our analysis of stochastic jumps provides little evidence that ETF shocks are transmitting noise into their portfolios.[24]

4. Daily analysis

To this point, our analyses have been predicated on the search for an intraday arbitrage mechanism whereby nonfundamental ETF demand shocks are transmitted to constituent security prices. It is possible, however, that any trading process propagating shocks from funds into their holdings occurs at lower frequencies than we have analyzed thus far. Previous studies have primarily examined such transmissions across days or months, not minutes or hours. Expanding our analysis beyond intraday trading also allows us to incorporate other potential sources of nonfundamental disruption, such as ETF share creations and redemptions or end-of-trading discounts and premiums.

[22] See Appendix F for a discussion of stochastic jumps and the Lee and Mykland (2008) jump detection measure.

[23] Given the volatility inherent in the opening auction, we exclude the first 15 minutes of each trading day both as a candidate for a jump and in the estimation of prior volatility.

[24] We repeat this analysis using big returns instead of stochastic jumps. We define a big return as one that is larger in absolute value than three times the prior 15 days' 1-minute return standard deviation. We find 1.80% (1.55%) of underlying portfolio (ETF) returns to be big. Our results are qualitatively similar to those reported in Fig. 6.

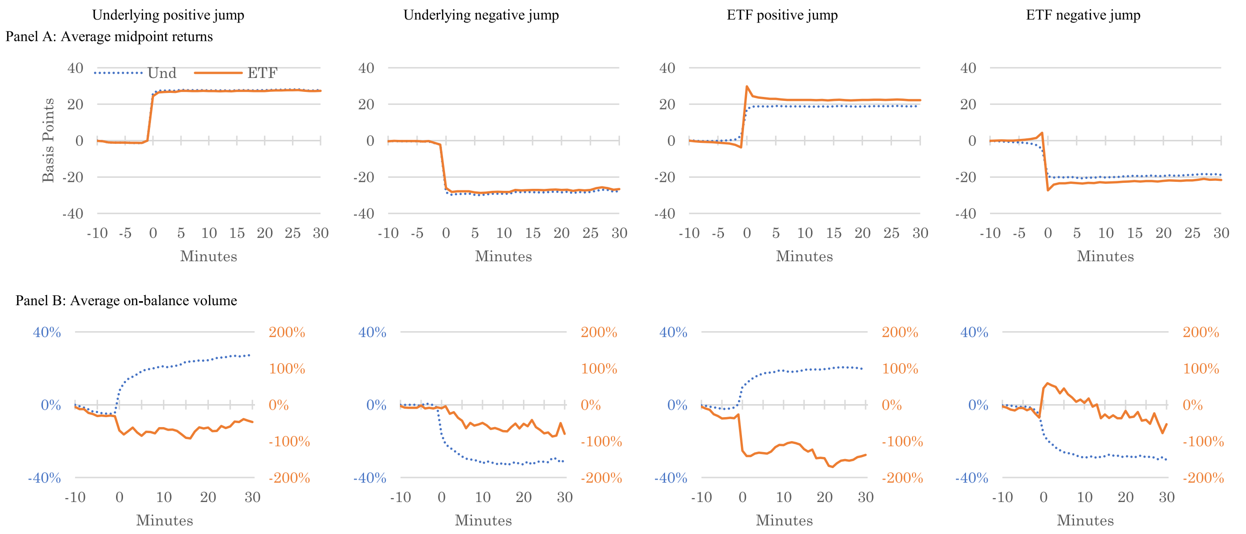

Fig. 6. Average midpoint returns and on-balance volume around stochastic jumps in ETF returns
This figure shows the average midpoint returns and on-balance volume for ETFs and constituents in a 40-minute window around stochastic jumps in either the fund or the underlying. Panel A shows midpoint returns. Panel B shows on-balance volume. The left (right) axis corresponds to the underlying (ETF). Stochastic jumps are identified for both funds and their underlying portfolios using the Lee and Mykland (2008) jump detection measure using the prior 15 days of minute-by-minute returns. See Appendix F for a discussion of the measure and its calculation.

For most traders, arbitrage profits could be realized by entering into, then later reversing, opposing positions through the secondary market for fund and constituent shares. What makes ETFs unique is that a subset of potential arbitrageurs, APs, can choose to unwind their positions in the primary market by creating or redeeming shares directly with the fund sponsor. To measure this form of primary market activity, we collect total shares outstanding, $ShrOut_d$, from Bloomberg for each fund at the end of every trading day, d, in our sample. We analyze daily percentage changes in shares outstanding, $\Delta ShrOut_d$, after performing a symmetric log transformation similar to what is described in Eqs. (1) and (3).

Regardless of whether arbitrageurs would seek to unwind their positions in primary or secondary markets, some divergence in ETF and constituent prices must occur to incentivize their activity. In Fig. 2, we show four situations where intraday mispricing between a fund and its holdings might provide opportunities to earn arbitrage profits. The following summation, across all intraday periods t, captures the persistence of these opportunities during the trading day d:

$$\begin{aligned} MispSumRaw_d = \sum_t \Big[& \left(2 \times Misp_t^{EB>UA}\right) \\ & + \left(1 \times Misp_t^{UA>EB>UM}\right) \\ & + \left(-1 \times Misp_t^{UM>EA>UB}\right) \\ & + \left(-2 \times Misp_t^{UB>EA}\right) \Big]. \end{aligned} \quad (5)$$

Values for $MispSumRaw_d$ that differ from zero suggest the prevalence of mispricing events throughout the day. Positive (negative) values indicate the persistence of ETF overpricing (underpricing).

Lastly, arbitrageurs can participate in secondary markets by posting orders for ETF and underlying shares on several different stock exchanges and dark pools throughout the day. However, shares can also be bought and sold in centralized closing auctions hosted by each security's primary listing exchange during the conclusion of trading. The constituent prices determined by these auctions are often used by fund sponsors to calculate an ETF's reported NAV. The deviation between this value and the fund's own closing price is described by the widely disseminated daily discount or premium. Most previous studies examining ETF and portfolio dynamics rely on these daily reported values to quantify mispricing and infer intraday arbitrage opportunities. Therefore, we also examine, after performing another symmetric log transformation, each fund's daily basis point discount or premium, $DiscPrem_d$, based on ETF closing prices and their reported NAVs collected from CRSP.

4.1. Panel vector autoregression analysis

We begin our lower-frequency analysis of ETF and constituent trading dynamics by expanding the PVAR framework described in Eq. (4) to include the summation of intraday mispricing, the daily change in shares outstanding, and reported basis point discounts or premiums. Ret_d^{Und} is the percentage change in ETF closing NAV from day $d-1$ through day d. Ret_d^{ETF} is the CRSP daily ETF return for day d. The calculation of daily order imbalance, $VolDiff_d$, is similar to what is described by Eqs. (1), (2), and (3), except that $VolDiffRaw_d$ is calculated in percentage terms instead of basis points and cumulated throughout the day.

Summary statistics for each time series included in $\boldsymbol{Y}_{fd}$ are reported in Table 6. Even though the summation of intraday fund mispricing is slightly positive on average, at 0.137, the median $MispSum_d$ is exactly zero. Likewise, we observe no strong tendency in the deviation between an ETF's closing price and its reported NAV, as the 1st and

Table 6
Summary statistics for daily analysis.

Ret_d^{Und} represents the percentage change in ETF NAV from day $d-1$ through day d. Likewise, Ret_d^{ETF}, represents the daily ETF return reported in CRSP on day d. $VolDiff_d^{Und}$ and $VolDiff_d^{ETF}$ measure the total daily difference between buy and sell volume for the underlying securities and ETF shares, respectively. All four variables are reported in percentage terms and symmetrically logged. $\Delta ShrOut_d$ and $MispSum_d$ are symmetrically logged percentage change in ETF shares outstanding and logged daily sum of mispricing events, respectively. $DiscPrem_d$ is the symmetrically logged percent deviation in ETF price from NAV as reported by CRSP at the close of day d, reported in symmetrically logged basis points. The sample includes all funds included in our intraday sample, as described in Section 2.

	N	Mean	P1	P10	P25	P50	P75	P90	P99
Ret_d^{Und}	241,341	0.026	-1.91	-1.03	-0.52	0.08	0.58	0.99	1.85
Ret_d^{ETF}	241,679	0.030	-1.86	-1.01	-0.51	0.08	0.58	0.98	1.81
$VolDiff_d^{Und}$	241,498	0.020	-2.18	-1.50	-1.00	0.02	1.02	1.55	2.37
$VolDiff_d^{ETF}$	236,671	0.682	-4.61	-3.98	-2.84	2.04	3.70	4.30	4.61
$MispSum_d$	241,688	0.137	-4.45	-2.94	-1.61	0.00	1.95	3.04	4.45
$\Delta ShrOut_d$	219,943	0.033	-1.926	0.000	0.000	0.000	0.000	0.350	2.097
$DiscPrem_d$	241,688	0.027	-4.606	-3.020	-2.063	0.079	2.061	2.916	4.536

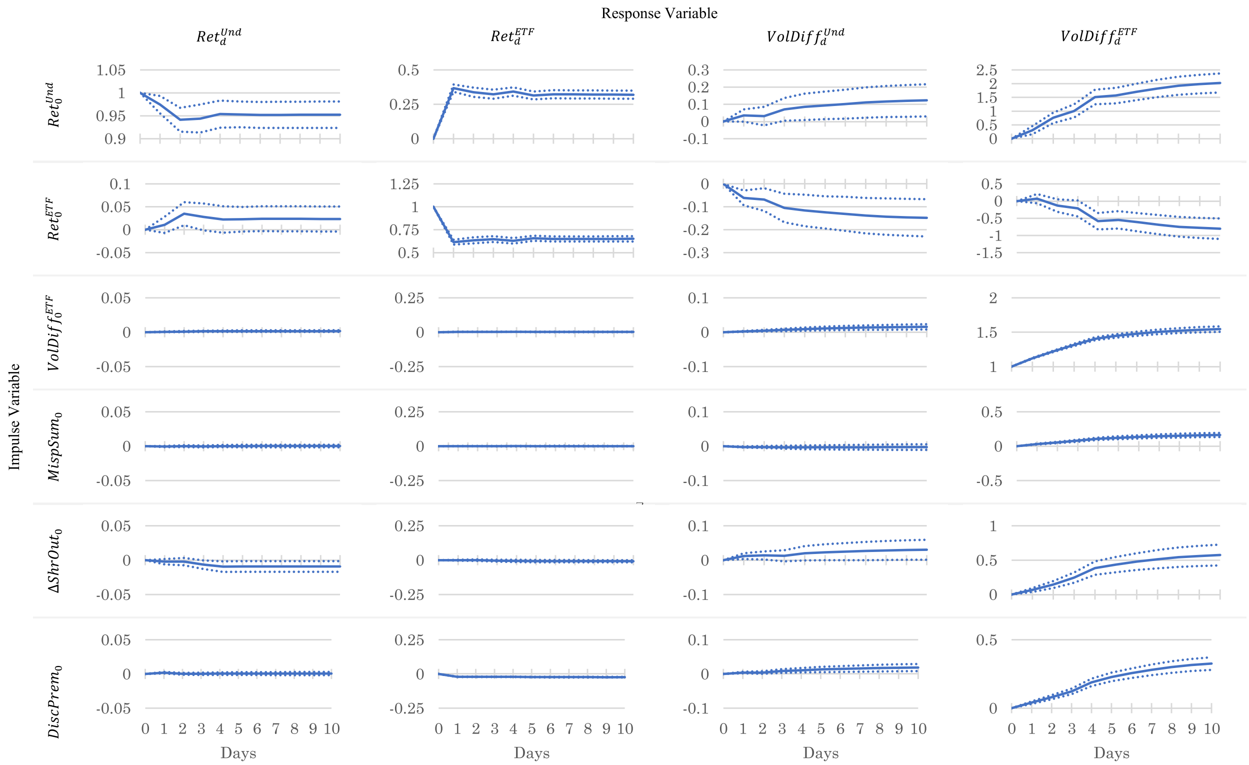

Fig 7. Daily cumulative impulse response functions
This figure depicts the cumulative impulse response functions (IRF) after expanding the PVAR framework introduced in Eq. (4) to a daily setting. These IRFs describe each dependent variable's evolution following a one standard deviation shock in the associated impulse variable. The seven variables included are $\boldsymbol{Ret_d^{Und}}$, $\boldsymbol{Ret_d^{ETF}}$, $\boldsymbol{VolDiff_d^{Und}}$, and $\boldsymbol{VolDiff_d^{ETF}}$, as in Fig. 1, and $\boldsymbol{MispSum_d}$, $\boldsymbol{\Delta ShrOut_d}$, and $\boldsymbol{DiscPrem_d}$, as defined in Table 6. Confidence intervals, 97.5% and 2.5%, are denoted by dotted lines.

99th percentile $DiscPrem_d$ values are -4.6 and 4.5 bps, respectively. There appears, however, to be some asymmetry in $\Delta ShrOut_d$, with the frequency of share creations exceeding that of redemptions by at least a small margin.

Fig. 7 depicts ten-day cumulative IRFs derived from a PVAR specification, where $\boldsymbol{Y}_{fd}$ is a (1×7) vector of dependent variables. Out of 49 possible impulse and response combinations, we focus on those relevant to the transmission of shocks from ETFs to their underlying portfolios with the IRF of underlying return shocks included in the first row as a point of comparison. The IRFs in Fig. 7 describe the ten-day reactions of Ret_{fd}^{Und}, Ret_{fd}^{ETF}, $VolDiff_{fd}^{Und}$, and $VolDiff_{fd}^{ETF}$ after a unit shock to Ret_{fd}^{Und}, Ret_{fd}^{ETF}, $VolDiff_{fd}^{ETF}$, $MispSum_{fd}$, $\Delta ShrOut_{fd}$, or $DiscPrem_{fd}$ when $d = 0$.

Along the top row of Fig. 7, we see a modest reversal in constituent returns following a unit shock to Ret_{fd}^{Und} and an increase in Ret_{fd}^{ETF} between 32% and 37% during the subsequent days. A positive shock to Ret_{fd}^{ETF} also results in a re-

Table 7
Daily panel vector autoregression forecast-error variance decomposition.
This table presents the fraction of forecasted error variance explained by exogenous shocks to impulse variables after ten days. The seven variables included are Ret_d^{Und}, Ret_d^{ETF}, $VolDiff_d^{Und}$, and $VolDiff_d^{ETF}$, as in Fig. 1 and $MispSum_d$, $\Delta ShrOut_d$, and $DiscPrem_d$, as defined in Table 6. Shocks are orthogonalized from top to bottom in the order presented.

		Response Variable						
		Ret_d^{Und}	Ret_d^{ETF}	$VolDiff_d^{Und}$	$VolDiff_d^{ETF}$	$MispSum_d$	$\Delta ShrOut_d$	$DiscPrem$
Impulse	Ret_0^{Und}	99.93%	77.82%	1.62%	0.71%	0.15%	0.31%	0.32%
Variable	Ret_0^{ETF}	0.02%	21.33%	0.00%	0.90%	0.14%	0.17%	30.96%
	$VolDiff_0^{Und}$	0.02%	0.01%	98.32%	0.11%	0.02%	0.04%	0.05%
	$VolDiff_0^{ETF}$	0.00%	0.01%	0.03%	97.47%	1.64%	1.27%	3.10%
	$MispSum_0$	0.00%	0.00%	0.00%	0.19%	97.70%	0.10%	0.07%
	$\Delta ShrOut_0$	0.01%	0.00%	0.01%	0.17%	0.11%	97.74%	0.22%
	$DiscPrem_0$	0.01%	0.83%	0.02%	0.44%	0.24%	0.36%	65.29%

versal in the ETF return but, consistent with our intraday results, we see almost no response in Ret_{fd}^{Und}. While point estimates for the response range between 3% and 6%, these values are only statistically significantly different from 0 on day two of the ten-day window.

Examining the other four variables for evidence of noise transmission from ETFs to their portfolios, we see that shocks to ETF order flow, intraday mispricing, primary market activity, and daily discounts or premiums are only correlated with future trading, not returns, in the underlying and ETF. Thus, any demand shocks associated with these four variables do not appear to cause nonfundamental disruptions in market prices.

For a more deliberate evaluation of these causal links, we turn to another FEVD analysis of our estimated parameters. Within $\boldsymbol{Y}_{fd}$, the sequencing of each time series is such that shocks to $MispSum_{fd}$, $\Delta ShrOut_{fd}$, and $DiscPrem_{fd}$ are orthogonalized relative to the instantaneous impacts of Ret_{fd}^{Und}, Ret_{fd}^{ETF}, $VolDiff_{fd}^{Und}$, and $VolDiff_{fd}^{ETF}$. The contributions of an impulse variable to the forecast-error variance of each response variable, after ten days, are given in Table 7.

Consistent with our intraday FEVD analysis, we find almost no evidence that independent shocks to ETF returns or order imbalances impact future price changes or trading in their underlying securities. Furthermore, as presaged by Fig. 7's IRFs, orthogonalized shocks to intraday mispricing, shares outstanding and daily discounts or premiums have no association with future innovations in Ret_d^{Und} or $VolDiff_d^{Und}$. Finally, while our results imply that future ETF prices respond to potentially simultaneous shocks in their underlying portfolios, independent fund shocks do not affect their constituents reciprocally during later periods. Altogether, the five potential sources of nonfundamental disruption we consider in Table 7 appear to have little effect on subsequent portfolio returns and trading.

5. Conclusion

Recent studies suggest that ETFs may attract short-term traders that introduce noise in their underlying securities prices through the arbitrage mechanism. As traders take opposing positions in the ETF and underlying shares, price pressures resulting from ETF demand may extend to the constituent securities. Using a sample of 423 passively managed US equity funds between 2006 and 2015, we directly examine this proposition by examining the minute-by-minute relation between the returns and order imbalances of ETFs and their constituent securities. Intraday impulse response functions and forecast error variance decompositions generated from a panel vector autoregression suggest ETF returns and order imbalance have little to no impact on underlying returns. Conversely, we find ETF returns follow underlying returns.

Identifying intraday arbitrage opportunities between ETFs and the constituents, we find little evidence that arbitrage opportunities precede trading in the underlying. Instead, arbitrage opportunities are initiated by shocks to the underlying and subsequently corrected through updates in the best bid and offer quotes. Thus, while we observe quote adjustments in response to price discrepancies, we find limited evidence of arbitrage trading. Additionally, our results indicate that bid-ask spreads remain steady during arbitrage opportunities. Not only are we unable to document arbitrage in the face of mispricing, we also find no evidence that the convergence of prices removes liquidity from the market.

After expanding our analysis beyond intraday trading to consider primary activity and other possible sources of nonfundamental disruption, we find little evidence that trading in the ETF propagates into the underlying. Daily PVAR results show constituent security returns are unaffected by shocks to ETF demand, intraday mispricing, primary market activity and daily discounts or premiums. Another FEVD analysis of our estimated parameters reveals almost no evidence that independent shocks to ETF returns or order imbalances impact future price changes or trading in their underlying securities.

In total, our results stand in sharp contrast to recent studies suggesting that ETF shares serve as a shock propagation channel that allows temporary demand shocks to leave an enduring impact on constituent security prices. In fact, our results suggest that in the presence of an information event in the underlying, ETF trading may help shield the portfolio from demand shocks by offsetting liquidity provision in the underlying securities.

References

Abrigo, M., Love, I., 2016. Estimation of panel vector autoregression in Stata. The Stata J. 16, 778–804.

Andrews, D., Lu, B., 2001. Consistent model and moment selection procedures for GMM estimation with application to dynamic panel data models. J. Econ. 101, 123–164.

Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimation of error-components models. J. Econ. 68, 29–51.

Ben-David, I., Franzoni, F., Moussawi, R., 2018. Do ETFs increase volatility? J. Finance 73, 2471–2535.

Box, T., Davis, R., Fuller, P., 2019. ETF competition and market quality. Financial Manage. 48, 873–916.

Box, T., Davis, R., Fuller, P., 2020. The dynamics of ETF fees. Financial Anal. J. 76, 11–18.

Brogaard, J., Hendershott, T., Riordan, R., 2019. Price discovery without trading: Evidence from limit orders. J. Finance 74, 1621–1658.

Broman, M., 2016. Liquidity, style investing and excess comovement of exchange-traded fund returns. J. Financial Markets 30, 27–53.

Broman, M., Shum, P., 2018. Relative liquidity, fund flows and short--term demand: Evidence from exchange-traded funds. Financial Rev. 53, 87–115.

Brown, D., Davies, S., Ringgenberg, M., 2019. ETF Arbitrage, Non-Fundamental Demand, and Return Predictability. University of Arizona Unpublished working paper.

Charupat, N., Miu, P., 2011. The pricing and performance of leveraged exchange-traded funds. J. Bank. Finance 35, 966–977.

Chen, G., Strother, T., 2008. On the Contribution of Index Exchange Traded Funds to Price Discovery in the Presence of Price Limits Without Short Selling. University of Otago Unpublished working paper.

Da, Z., Shive, S., 2018. Exchange traded funds and asset return correlations. Eu. Financial Manage. 24, 136–168.

Duffee, G., 2018. Expected inflation and other determinants of treasury yields. J. Finance 73, 2139–2180.

Fang, Y., Sanger, G., 2012. Index Price Discovery in the Cash Market. Louisiana State University, Baton Rouge Unpublished working paper.

Hamilton, J., 1985. Uncovering financial market expectations of inflation. J. Polit. Econ. 93, 1224–1241.

Hasbrouck, J., 2003. Intraday price formation in US equity index markets. J. Finance 58, 2375–2399.

Hilscher, J., Pollet, J., Wilson, M., 2015. Are credit default swaps a sideshow? Evidence that information flows from equity to CDS markets. J. Financial Quant. Anal. 50, 543–567.

Hollifield, B., Neklyudov, A., Spatt, C., 2017. Bid-ask spreads, trading networks, and the pricing of securitizations. Rev. Financial Studies 30, 3048–3085.

Holtz-Eakin, D., Newey, W., Rosen, H., 1988. Estimating vector autoregressions with panel data. Econometrica 56, 1371–1395.

Israeli, D., Lee, C., Sridharan, S., 2017. Is there a dark side to exchange traded funds (ETFs)? An information perspective. Rev.Account. Studies 22, 1048–1083.

Ivanov, S., Jones, F., Zaima, J., 2013. Analysis of DJIA, S&P 500, S&P 400, NASDAQ 100 and Russell 2000 ETFs and their influence on price discovery. Glob. Finance J. 24, 171–187.

Judson, R., Owen, A., 1999. Estimating dynamic panel data models: a guide for macroeconomists. Econ. Lett. 65, 9–15.

Koijen, R., Lustig, H., Van Nieuwerburgh, S., 2017. The cross-section and time series of stock and bond returns. J. Monetary Econ. 88, 50–69.

Lee, J., Naranjo, A., Velioglu, G., 2018. When do CDS spreads lead? rating events, private entities, and firm-specific information flows. J. Financial Econ. 130, 556–578.

Lee, S., Mykland, P., 2008. Jumps in financial markets: a new nonparametric test and jump dynamics. Rev. Financial Stud. 21, 2535–2563.

Love, I., Zicchino, L., 2006. Financial development and dynamic investment behavior: Evidence from panel VAR. Q. Rev. Econ. Finance 46, 190–210.

Madhavan, A., Sobczyk, A., 2016. Price dynamics and liquidity of exchange-traded funds. J. Invest. Manage. 14, 1–17.

Malamud, S., 2016. A Dynamic Equilibrium Model of ETFs. Swiss Federal Institute of Technology, Lausanne Unpublished working paper.

Marshall, B., Nguyen, N., Visaltanachoti, N., 2013. ETF arbitrage: intraday evidence. J. Bank. Finance 37, 3486–3498.

Petajisto, A., 2017. Inefficiencies in the pricing of exchange-traded funds. Financial Anal. J. 73, 24–54.

Sims, C., 1980. Macroeconomics and reality. Econometrica 48, 1–48.

Yu, L., 2005. Basket Securities, Price Formation, and Informational Efficiency. University of Notre Dame Unpublished working paper.