Online Appendix to Stock-Specific Price Discovery From ETFs

Thomas Ernst*

MIT Sloan School of Management

[Click Here for Latest Version of the Paper]

July 2, 2020

Abstract

Conventional wisdom warns that exchange traded funds (ETFs) harm stock price discovery, either by "stealing" single-stock liquidity or forcing stock prices to co-move. Contra this belief, I develop a theoretical model and present empirical evidence which demonstrate that investors with stock-specific information trade both single stocks and ETFs. Single-stock investors can access ETF liquidity by means of this tandem trading, and stock prices can flexibly adjust to ETF price movements. Using high-resolution data on SPDR and the Sector SPDR ETFs, I exploit exchange latencies in order to show that investors place simultaneous, same-direction trades in both a stock and ETF. Consistent with my model predictions, effects are strongest when an individual stock has a large weight in the ETF and a large stock-specific informational asymmetry. I conclude that ETFs can provide single-stock price discovery.

In this online appendix, I consider several robustness checks or alternative specifications to the results of the paper.

Keywords: Exchange Traded Fund, ETF, Liquidity, Asymmetric Information, Market Microstructure, Trading Costs, Comovement, Cross Market Activity, High-Frequency Data, Microsecond TAQ Data

JEL Classification: G12, G14

*First Draft: October 2017. I am very grateful to my advisors: Haoxiang Zhu, Chester Spatt, Leonid Kogan, and Jiang Wang. Additional helpful comments were provided by Andrey Malenko, Simon Gervais, Shimon Kogan, Antoinette Schoar, Jonathan Parker, Lawrence Schmidt, David Thesmar, Hui Chen, Daniel Greenwald, Deborah Lucas, Dobrislav Dobrev, Andrew Lo, Robert Merton, Christopher Palmer, Adrien Verdelhan, Eben Lazarus, Austin Gerig, Peter Dixon, and Eddy Hu, as well as seminar participants at MIT and the SEC DERA Conference.

I. Overview

In this online appendix, I consider several additional tests and robustness checks of my results. Section II examines stock-ETF-stock and stock-ETF-ETF triple trades. Section III considers alternative definitions or specifications of the results. Section IV is an event study around the re-ordering of SPDRs when the Real Estate and Communications Sector SPDR ETFs were created. Section V examines alternative methods for controlling for the level of trading.

II. Trade Timing

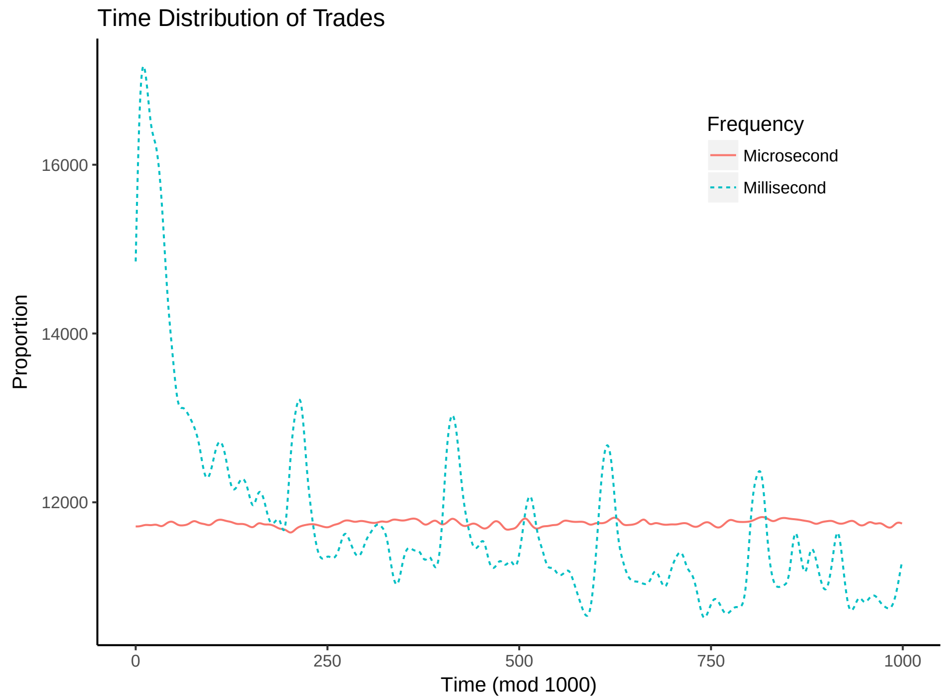
In this section, I verify that the results are not due to patterns in the timing of trading activity. Trading activity is not uniform throughout the day. **?**, for example, note that there is a large spike in trading activity in the first few milliseconds of each second, suggesting algorithmic trading at a one-second frequency. Higher levels of algorithmic trading would lead to higher levels of trading in the first few microseconds of each second.

I plot the distribution of trades at the millisecond and microsecond level in Figure **??**. At the microsecond level, trade distribution is uniform. To confirm that my results are not driven by some form of periodic algorithmic trading, I re-estimate Regression 1 on a restricted subsample. From the sample of all trades, I eliminate trades occurring in the first 200 milliseconds of each second. This leaves trades occurring only between $(200, 999)$ of each millisecond to avoid any spikes in activity around 0 microseconds. Results for this regression are presented in Table **??** and are similar to the previous estimation of Regression 1.

Table I: **Robustness Check of Regression 1- Time Restricted Sub-Sample**
This table reports estimates of Regression 1 on a time restricted subsample. To control for changes in algorithmic trading, I exclude all trades which occur in the first 200 milliseconds of each second. I measure stock–ETF simultaneous trades on the trades which remain. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Simultaneous Trades	
	(1)	(2)
Weight	16.064***	15.110***
	(5.882)	(5.468)
Earnings Date	−1.099	
	(0.909)	
Weight∗Earnings Date	2.400***	
	(0.640)	
Abs Return		1.367**
		(0.583)
Weight∗Abs Return		3.738**
		(1.754)
Abs ETF Return	37.335***	36.548***
	(14.203)	(13.908)
Weight∗Abs ETF Return	4.089	0.898
	(3.179)	(3.749)
Observations	873,244	873,226
R^2	0.166	0.169
Adjusted R^2	0.166	0.169
Residual Std. Error	116.811	116.553
Standard Errors Clustered by ETF.	*p<0.1; **p<0.05; ***p<0.01	

Figure 1. Time Distribution of Trades. I plot clock-time periodicities of trades for a subset of the data: XLE and the top 10 underlying stocks for September, 2016. Consider a timestamp in the form HH:MM:SS.mmm$\mu\mu\mu$. The three blue digits are the milliseconds, while the three red digits are the microseconds. There is a clear spike in trades in the first few milliseconds of each second (dashed blue line). **?** argue this could be algorithmic trading activity. The distribution of trades at the microsecond level (solid red line), however, is uniform.



III. Alternative Restrictions and Definitions

This section explores variations of the main results in the paper subject to alternative definitions or restrictions.

In the main regression of the paper, restated below, I estimate two versions: one with a control of the Absolute ETF Return with a weight–Absolute ETF Return interaction, and one without. The absolute value of the ETF return allows ruling out the possibility that the increase in simultaneous trades is purely driven by changes in the ETF return, rather than the stock-specific return.

REGRESSION 1: *For stock i, ETF j, and day t:*

$$\begin{aligned} Simultaneous\ Trades_{ijt} &= \alpha_0 + \alpha_1 Earnings\ Date_{it} + \alpha_2 Weight_{ij} \\ &\quad + \alpha_3 Weight_{ij} * Earnings\ Date_{it} + \alpha_4 Controls_{ijt} + \epsilon_{ijt} \end{aligned} \tag{1}$$

$$\begin{aligned} Simultaneous\ Trades_{ijt} &= \alpha_0 + \alpha_1 Abs\ Return_{it} + \alpha_2 Weight_{ij} \\ &\quad + \alpha_3 Weight_{ij} * Abs\ Return_{it} + \alpha_4 Controls_{itj} + \epsilon_{ijt} \end{aligned} \tag{2}$$

Using the ETF return as a control, however, is problematic because the ETF return is correlated with the individual stock return. The central hypothesis of the paper that investors trade the ETF based on stock-specific information depends on the fact that the ETF return will be correlated with the stock-specific information. As an alternative, I re-estimate Regression 1 using an alternative control for the ETF return. For this alternative control, for each stock i, I calculate the ETF Return Ex_i as the absolute absolute value of the total return on all the other stocks in the ETF.

Results are presented in Table IV. Estimates are similar in both cases, consistent with the effect of the stock return on simultaneous stock-ETF trades existing over and above that of any simultaneous trades driven by ETF returns alone.

As an additional control ETF or market-wide changes in trading, I re-estimate Regression 1 excluding the dates of macro announcements. I take the dates given by Baker, Bloom, Davis, and Sammon (2015) and exclude them from the analysis. Results are presented in Table **??** and are similar.

In Regression 1, for the estimate of earnings date, I use the trading session after earnings announcement, to reflect potential larger disagreement over the interpretation of stock-specific information. Alternatively, I could use the trading session before earnings announcement as the earnings date session, reflecting differences of private information or speculation over the forthcoming earnings. This change in estimation is reported in Table V. After controlling for ETF return, results for large stocks are similar between the two definitions. For small stocks, pre-earnings dates predict slightly more trades while post-earnings dates predict slightly fewer simultaneous trades.

In Regression 1, simultaneous trades are measured with stock-ETF trades which occur within 20 microseconds of each other. Given the latencies at the major exchanges, as reported in Table VI, this is the maximum possible distance in time for which it can be certain that two trades are

Table II: Estimation of Regression 1

This table reports estimates of Regression 1, which estimates the effect of changes in stock-specific information on simultaneous trades under an alternative control. This alternative control of Abs ETF Ret Ex_i is the absolute value of the total return of stocks in the ETF except the return of stock i. Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Simultaneous Trades					
	(1)	(2)	(3)	(4)	(5)	(6)
Weight	48.015***	49.816***	51.383***	60.867***	53.408***	54.191**
	(16.075)	(18.547)	(19.594)	(20.408)	(19.972)	(21.120)
Abs Ret	21.432***	4.435**	3.721**			
	(8.247)	(2.048)	(1.645)			
Abs ETF Ret Ex_i		101.606**			104.265**	
		(42.721)			(43.778)	
Abs ETF Ret			104.848**			106.948**
			(43.799)			(44.355)
Weight∗ Abs Ret	14.732***	10.288**	10.809*			
	(5.715)	(4.792)	(6.033)			
Weight∗ Abs ETF Ret Ex_i		3.515			12.189	
		(10.331)			(9.927)	
Weight∗Abs ETF Ret			−0.579			8.652
			(11.832)			(9.632)
Earnings Date				−12.717***	−9.086***	−8.568***
				(2.679)	(1.418)	(1.486)
Weight∗ Earnings Date				6.188***	4.802***	3.827**
				(1.659)	(1.696)	(1.761)
Observations	873,178	873,178	873,178	873,196	873,178	873,196
R^2	0.110	0.133	0.133	0.095	0.131	0.131
Adjusted R^2	0.110	0.133	0.133	0.095	0.131	0.131
Residual Std. Error	406.833	401.397	401.431	410.309	401.397	402.051
Degrees Freedom	873164	873162	873162	873182	873162	873180
Note:						*p<0.1; **p<0.05; ***p<0.01

Table III: **Estimation of Regression 1 with Earnings Date Definitions**

This table reports estimates of Regression 1 using two definitions of earnings dates. Earnings Before is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session (either in the morning or previous evening). Earnings After is an indicator which takes the value 1 for stocks which announce earnings after the day's trading session (either in the evening or following morning). Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Simultaneous Trades			
	(1)	(2)	(3)	(4)
Weight	60.867***	54.191**	60.908***	54.195**
	(20.408)	(21.120)	(20.393)	(21.108)
earningsAfter	−12.717***	−8.568***		
	(2.679)	(1.486)		
absETFret		106.948**		106.965**
		(44.355)		(44.365)
Weight:earningsAfter	6.188***	3.827**		
	(1.659)	(1.761)		
Weight:absETFret		8.652		8.657
		(9.632)		(9.633)
earningsBefore			8.866***	12.259***
			(2.020)	(1.716)
Weight:earningsBefore			3.533*	3.304*
			(2.043)	(1.758)
Observations	873,196	873,196	873,196	873,196
R^2	0.095	0.131	0.095	0.131
Adjusted R^2	0.095	0.131	0.095	0.131
Residual Std. Error	410.309	402.051	410.308	402.047
Degrees of Freedom	873182	873180	873182	873180
Note:				*p<0.1; **p<0.05; ***p<0.01

not responding to each other. Over longer horizons, however, there is a danger that two trades are not placed by the same individual. While these trades may be placed by different individuals, they are still close enough in time that they must be based on the same information; empirical analysis of the trades also shows that they are placed in the same direction. So if the trades are placed by different traders, these traders still have the same signal and same interpretation of that signal. Nonetheless, it is worth investigating the extent to which the results are robust to alternative definitions of simultaneous trades. As an alternative, I consider simultaneous trades which are timestamped to the exact same microsecond.

Results for this restriction to exactly the same timestamped trades are presented in Table VII. Coefficient estimates are lower, but this is due to the fact that same-microsecond trades are much less common. In both cases, the weight-return interaction is positive and significant without the control for ETF return, but not significant with the control. Over more restrictive intervals, one drawback is that trades which are placed at the same time will not appear as simultaneous, due to the microsecond or two it takes for exchange servers to process the order in the electronic limit order book.

Table IV: **Gateway to Limit Order Book Latency from Major Exchanges**. This table gives the latencies reported by the major exchanges. All times are in microseconds, and reflect the total round-trip time from the gateway to limit order book of an exchange. All traders, including co-located high-frequency traders, must make this trip. Note that IEX and NYSE American have significantly longer round-trip times due to the inclusion of a 350 microsecond speed-bump.

	Min	Average
CBOE/BATS	31	56
Nasdaq	25	sub-40
NYSE	21	27
NYSE Arca	26	32
NYSE American	724	732
IEX	700+	700+

To estimate the differences between large and small stocks, I split my sample into large, medium, and small stocks based on their ETF weight. Small stocks are those with a sector weight of 0-2%, medium stocks with a sector weight of 2-5%, and large stocks have a sector weight above 7%. These restrictions can then be estimated with one regression, for which the fixed effect and standard errors are estimated using the full sample, or as separate regressions, where fixed effects and standard

Table V: Estimation of Regression 1 with Same Microsecond Simultaneous Trades
This table reports estimates of Regression 1 using an alternative definition of simultaneous trades. These simultaneous trades are restricted to trades in which the stock and ETF have the exact same timestamp, down to the microsecond. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Exact Simultaneous			
	(1)	(2)	(3)	(4)
Weight	4.723***	4.314**	3.769***	4.094**
	(1.705)	(1.791)	(1.349)	(1.658)
earningsAfter	−1.242***	−0.906***		
	(0.372)	(0.239)		
absETFret		8.956**		8.833**
		(3.809)		(3.780)
Weight:earningsAfter	0.367**	0.187		
	(0.156)	(0.196)		
Weight:absETFret		0.531		−0.175
		(0.755)		(0.974)
absret			1.739**	0.233*
			(0.681)	(0.134)
Weight:absret			1.098**	0.827
			(0.459)	(0.514)
Observations	873,728	873,728	873,710	873,710
R^2	0.068	0.095	0.078	0.097
Adjusted R^2	0.068	0.095	0.078	0.097
Residual Std. Error	39.125	38.553	38.907	38.517
Degrees Freedom	873714	873712	873696	873694
Note:		*p<0.1; **p<0.05; ***p<0.01		

errors are estimated separately within each subset of the data. In the paper, I estimate the following regressions:

REGRESSION 2: *For stock i, ETF j, and day t:*

$$Simultaneous\ Trades_{ijt} = \alpha_0 + Size * \alpha_1 Earnings\ Date_{it} + \alpha_2 Controls_{ijt} + \epsilon_{ijt} \tag{3}$$

$$Simultaneous\ Trades_{ijt} = \alpha_0 + Size * \alpha_1 Abs\ Return_{it} + \alpha_2 Controls_{itj} + \epsilon_{ijt} \tag{4}$$

REGRESSION 3: *For stock i, ETF j, and day t:*

$$Simultaneous\ Trades_{ijt} = \alpha_0 + Size * \alpha_1 Largest\ X\ Abs\ Return_{it} + \alpha_2 Controls_{itj} + \epsilon_{ijt} \tag{5}$$

Sizes of return categories in Regression 3 are given in Table VIII. In the paper, I estimate Regression 2 in one regression for brevity. In Table IX, I report the analogous estimate where effects are estimated separately for each group. I estimate Regression 3 in separate regressions for clarity. In Table X, I report the analogous estimate where effects are estimated jointly as one group. Results are qualitatively similar in all cases.

Table VI: Extreme Returns by Absolute Value Categories

Stock Category:	Small	Medium	Large
1%	6.387	5.762	4.688
5%	3.857	3.640	3.017
10%	2.991	2.859	2.353
20%	2.194	2.114	1.728

Table VII Estimation of Regression 2: Size Comparison

This table reports estimates of Regression 2, which estimates the effect of changes in stock-specific information across different ETF weights. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. I categorize small stocks as those with a weight less than 2%, medium stocks with weight between 2% and 5%, and large stocks with a weight greater than 5%. The sample is the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Simultaneous Trades					
	Small	Small	Medium	Medium	Large	Large
Earnings Date	-2.537		12.690^{**}		38.076^{***}	
	(2.969)		(5.557)		(11.789)	
Abs Return		12.375^{***}		40.523^{***}		117.845^{***}
		(2.899)		(11.143)		(33.675)
Observations	310,127	310,119	96,276	96,274	37,097	37,097
R^2	0.162	0.179	0.255	0.285	0.248	0.284
Adjusted R^2	0.162	0.179	0.255	0.285	0.248	0.284
Residual Std. Error	121.224	119.974	224.999	220.404	496.940	484.695
Degrees of Freedom	310116	310108	96265	96263	37086	37086
Note:						*p<0.1; **p<0.05; ***p<0.01

Table VIII : Estimation of Regression 3 - Return Comparison

This table reports estimates of Regression 3. For each stock, largest $X\%$ Abs Return is an indicator which takes the value one on days for which the intraday return is among the most positive $X\%$ or most negative $X\%$ of returns for that stock. I include an equivalently defined daily indicator on the ETF return for whether the ETF return is among the most positive $X\%$ or most negative $X\%$. Small stocks are stocks with less than 2% ETF weight, medium stocks are 2% to 5%, and large stocks have greater than 5% weight. The sample is the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Simultaneous Trades			
	(1)	(2)	(3)	(4)
Large ETF 1%	115.402***			
	(23.172)			
Large Stock 1%	67.569***			
	(15.234)			
Large ETF 5%		89.526***		
		(17.292)		
Large Stock 5%		42.648***		
		(10.146)		
Large ETF 10%			63.200***	
			(11.909)	
Large Stock 10%			38.547***	
			(8.264)	
Large ETF 20%				49.609***
				(8.621)
Large Stock 20%				32.286***
				(6.781)
Observations	443,500	443,500	443,500	443,500
R^2	0.126	0.141	0.139	0.134
Adjusted R^2	0.126	0.141	0.139	0.134
Residual Std. Error (df = 443488)	233.478	231.527	231.763	232.388
Note:				*p<0.1; **p<0.05; ***p<0.01

IV. Reclassification Event Study

The Sector SPDR ETFs are based on Standard and Poor's Global Industry Classification Standard groupings.[1] These industry codes, and thus the Sector SPDRs themselves, have gone through two major reclassifications in the last five years. In September 2016, the Real Estate Sector SPDR (XLRE) was created from stocks previously categorized as Financials. In October 2018, the Communications Sector SPDR (XLC) was created from stocks previously categorized as Technologies or Consumer Discretionary.

These reclassifications present an opportunity for an event study. For example, the real estate sector stocks which were previously part of the financials ETF would see their ETF weight in the financials sector go to zero, as they were spun off into a real estate ETF. During this time, however, both their weight in the S&P 500 Index and relationship with the overall S&P 500 index remained unchanged.

APPENDIX REGRESSION 1: *For stock i, ETF j, and day t:*

$$\begin{aligned}
\mathit{Simultaneous\ Trade}_{ijt} = \alpha_0 &+ \alpha_1 \mathit{Abs\ Return}_{it} + \alpha_2 \mathit{SPY\ Weight}_{ij} && (6)\\
&+ \alpha_3 \mathit{SPY\ Weight}_{ij} * \mathit{Abs\ Return}_{it} + \alpha_4 \mathit{SPY\ Weight\ Post\ Spinoff}_{ij} && (7)\\
&+ \alpha_5 \mathit{Abs\ Return\ Post\ Spinoff}_{it} + \alpha_6 \mathit{Abs\ Return\ Pre\ Spinoff}_{it} && (8)\\
&+ \alpha_7 \mathit{SPY\ Weight\ Pre\ Spinoff}_{ij} && (9)\\
&+ \alpha_8 \mathit{SPY\ Weight\ Post\ Spinoff}_{ij} * \mathit{Abs\ Return\ Post\ Spinoff}_{it} && (10)\\
&+ \alpha_9 \mathit{SPY\ Weight\ Pre\ Spinoff}_{ij} * \mathit{Abs\ Return\ Pre\ Spinoff}_{it} && (11)\\
&+ \alpha_{10} \mathit{Controls}_{itj} + \epsilon_{ijt} && (12)
\end{aligned}$$

Controls include indicators for the post-spinoff time period, a fixed effect for each group of stocks, the ETF return, and the level of simultaneous trades in SPY.

Following the spin-off, the stock-ETF simultaneous trades should be lower. Thus α_6 should be negative, as the weight-return interaction should be lower for stocks which are spun-off relative to the stocks which remain in their original ETF. Results are presented in Table XI. Before the spinoff,

[1] The GICS groups are: Financials, Communications, Energy, Health Care, Consumer Discretionary, Consumer Staples, Industrials, Materials, Real Estate, Technology, and Utilities

the weight-return interaction is positive for the stocks which are to be spun-off: $\alpha_3+\alpha_9 = 16-12 = 4$. Thus before the spin-off, there is a positive weight-return interaction in the stocks which will move to become part of the new ETFs. After the spin-off, the interaction term becomes negatve: $\alpha_3+\alpha_9 = 16-23 = -9$. The weight term alone has a similar change. Before the spinoff, the larger weight stocks which are to be spun-off have more simultaneous trades: $\alpha_2 + \alpha_7 = 111 - 32 = 79$. After the spinoff, this relationship becomes negative: $\alpha_2 + \alpha_4 = 111 - 177 = -66$. Thus after the SPDR reorganization, the large-weight stocks which moved no longer see more simultaneous trades or a positive weight-return interaction term with the ETFs they left.

Table IX: Estimation of Appendix Regression 3

This table reports estimates of Appendix Regression 3, which estimates the how the stock-ETF relationship changes with the reorganization of the sector SPDR ETFs. SPY Weight is the stock weight in the S&P 500 ETF, SPY. Abs Return is the absolute value of the intraday return, measured as a percentage. Before the spinoff date, SPY Weight Pre Spinoff is the weight only for the stocks which are going to leave XLK, XLY, or XLK, and zero otherwise. After the spinoff date, SPY Weight Post Spinoff is the weight only for the stocks which are going to leave XLF, XLY, or XLK, and zero otherwise. The sample is XLF and the stocks of XLF and XLRE from September to December 2016, and XLY and XLK, along with the stocks of XLK, XLY, and XLC from September to December 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Simultaneous Trades
Cspin	14.465^{***}
	(2.635)
Rspin	3.079^{**}
	(1.396)
SPY Weight	111.230^{***}
	(4.908)
Abs Return	4.934^{***}
	(1.273)
SPY Weight$*$ Abs Return	16.061^{***}
	(2.538)
SPY Weight Pre Spinoff	-32.362^{***}
	(8.237)
SPY Weight Post Spinoff	-177.973^{***}
	(13.648)
Abs Return Pre Spinoff	-5.763^{***}
	(1.579)
Abs Return Post Spinoff	-12.319^{***}
	(2.508)
SPY Weight$*$Abs Return Pre Spinoff	-12.490^{**}
	(5.137)
SPY Weight$*$ Abs Return Post Spinoff	-23.669^{***}
	(7.786)
Abs ETF Return	X
SPY Return	X
SPY Simultaneous	X
Observations	15,825
R^2	0.942
Adjusted R^2	0.942
Residual Std. Error	95.353 (df = 15806)
Note:	*p<0.1; **p<0.05; ***p<0.01

V. Order Control

In this section, I consider alternative controls for the random chance or baseline trades. This baseline estimate measures how many trades occur in both markets at a point close in time, but not exactly simultaneously. For each stock trade in my sample, I calculate how many ETF trades occur exactly X microseconds before or after the stock trade. I calculate the average number of trades as X ranges from 1000 to 1200 microseconds; this boundary is far enough away to avoid picking up high-frequency trading response trades, but close enough to pick up patterns in trading at the millisecond level. I then scale this average up by 20 and subtract this baseline level of trades from each daily calculation of simultaneous trades. The level of simultaneous trades between stock i and ETF j on day t can be written as:

$$\text{Simultaneous Trades}_{ijt} = \text{Raw Simultaneous}_{ijt} - \frac{20}{200}\text{Baseline}$$

This baseline-corrected measure of simultaneous trades accounts for chance simultaneous trades which varies with changes in daily trading volume. Figure **??** plots a sample observation of cross market activity, along with the raw simultaneous and baseline regions.

As a first check, I re-estimate Regression 1 without the baseline correction. Results are presented in Table XII. Results are qualitatively similar, with slightly larger values as to be expected with the larger overall level of trading. I also re-estimate Regression 2 without the baseline correction, and present results in Table XIII.

As an alternative control, I re-estimate Regression 1 without the baseline correction, but with a control for the total number of orders. Results are presented in Table XIV. The stock weight-absolute return interaction is similar, as is the estimate for earnings date as measured by the before-earnings date. When earnings dates are measured by the day after announcement, the interaction term is no longer significant. In Table XV, I re-estimate Regression 1 with a control for volume. In the volume regression, weight as well as the weight-return interaction term are positive and strongly significant. Earnings dates, however, are not significant.

I also consider the use of the baseline trade estimates as a linear predictor, rather than subtracting the baseline level of trades from raw simultaneous trades. Results are presented in Table

Table X: **Estimation of Regression 1 without Baseline Correction**
This table reports estimates of Regression 1, which estimates the effect of changes in stock-specific information on simultaneous trades between stocks and ETFs which contain them. In this regression, I do not remove the baseline level of trading. I consider two different measures of stock-specific information: earnings dates and absolute value of the return. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Raw Simultaneous Trades			
	(1)	(2)	(3)	(4)
Weight	62.026***	55.383**	48.998***	52.540***
	(20.835)	(21.644)	(16.445)	(20.096)
Earnings Date	−12.052***	−7.783***		
	(2.449)	(1.448)		
Abs ETF Return		110.550**		108.182**
		(45.913)		(45.265)
Weight∗Earnings Date	6.295***	3.886**		
	(1.686)	(1.792)		
Weight∗Abs ETF Return		8.609		−0.801
		(9.846)		(12.115)
Abs Return			22.407***	4.113**
			(8.619)	(1.747)
Weight∗Abs Return			14.965**	11.013*
			(5.817)	(6.152)
Observations	873,196	873,196	873,178	873,178
R^2	0.095	0.131	0.110	0.134
Adjusted R^2	0.095	0.131	0.110	0.134
Residual Std. Error	419.519	410.958	415.921	410.315
Degrees Freedom	873182	873180	873164	873162
Note:			*p<0.1; **p<0.05; ***p<0.01	

Table XI Estimation of Regression 2: Without Baseline

This table reports estimates of Regression 2, which estimates the effect of changes in stock-specific information across different ETF weights. In this regression, I do not remove the baseline level of trades. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. I categorize small stocks as those with a weight less than 2%, medium stocks with weight between 2% and 5%, and large stocks with a weight greater than 5%. The sample is the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Raw Simultaneous Trades					
	Small	Small	Medium	Medium	Large	Large
Earnings Date	−1.664		8.794*		26.752	
	(2.275)		(4.529)		(16.601)	
Abs Return		5.030***		18.121***		42.704***
		(1.539)		(3.806)		(8.246)
Abs ETF Return	41.036***	37.242***	85.081***	70.663***	200.611***	165.139***
	(5.438)	(4.578)	(23.835)	(21.384)	(62.642)	(56.715)
Observations	310,127	310,119	96,276	96,274	37,097	37,097
R^2	0.212	0.214	0.308	0.313	0.306	0.309
Adjusted R^2	0.212	0.214	0.308	0.313	0.306	0.309
Residual Std. Error	119.723	119.548	220.603	219.923	485.307	484.326
Degrees Freedom	310115	310107	96264	96262	37085	37085
Note:						*p<0.1; **p<0.05; ***p<0.01

XVI. Results are similar to to the order level control: the weight-return interaction is positive and significant. The weight-earnings interaction is positive and statistically significant for the trading session before earnings, but not significant for the trading session after earnings.

As a more flexible yet coarse estimate, I re-estimate Regression 1 using a fixed effect for each date rather than a control for the changes in orders. Results with a fixed effect for each date are presented in Table XVII while results with a fixed effect for each date and a fixed effect for each stock are presented in Table XVIII.

Table XII: Estimation of Regression 1 with Order Control
This table reports estimates of Regression 1, which estimates the effect of changes in stock-specific information on simultaneous trades between stocks and ETFs which contain them. In this regression, I do not remove the baseline level of trading. For easy to read coefficients, stock or ETF volume are measured in tens of millions of dollars. I consider two different measures of stock-specific information: earnings dates and absolute value of the return. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Raw Simultaneous Trades		
	(1)	(2)	(3)
Weight	15.430***	28.160***	29.182***
	(3.709)	(7.514)	(7.860)
Abs Ret	−42.724***		
	(11.580)		
Earnings After		−228.361***	
		(51.520)	
Earnings Before			−45.766***
			(13.305)
Stockorders	51.228***	46.620***	44.307***
	(5.974)	(3.889)	(3.767)
ETForders	−1.536***	−2.343***	−2.358***
	(0.216)	(0.218)	(0.224)
Weight∗Abs Return	9.554***		
	(2.916)		
Weight∗Earnings After		10.306	
		(8.944)	
Weight∗Earnings Before			5.157***
			(1.921)
Stockorders∗ETF orders	3.142***	3.162***	3.180***
	(0.126)	(0.104)	(0.101)
Observations	873,178	873,196	873,196
R^2	0.634	0.626	0.622
Adjusted R^2	0.634	0.626	0.622
Residual Std. Error	266.576 (df = 873161)	269.684 (df = 873179)	270.942 (df = 873179)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table XIII: Estimation of Regression 1 with Volume Control

This table reports estimates of Regression 1, which estimates the effect of changes in stock-specific information on simultaneous trades between stocks and ETFs which contain them. In this regression, I do not remove the baseline level of trading. For easy to read coefficients, stock or ETF volume are measured in tens of thousands of orders. I consider two different measures of stock-specific information: earnings dates and absolute value of the return. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Raw Simultaneous Trades		
	(1)	(2)	(3)
Weight	25.001***	32.748***	33.347***
	(8.139)	(10.539)	(10.790)
Abs Return	−12.096*		
	(6.293)		
Earnings After		−121.762***	
		(35.769)	
Earnings Before			−11.298
			(6.923)
Stock Volume	2.188**	2.537***	2.447***
	(0.887)	(0.942)	(0.943)
ETF Volume	0.141***	0.138***	0.138***
	(0.009)	(0.010)	(0.010)
Weight∗Abs Return	10.216***		
	(3.285)		
Weight∗Earnings After		−1.898	
		(8.738)	
Weight∗Earnings Before			2.086
			(1.765)
Stock Volume∗ETF Volume	0.004***	0.004***	0.004***
	(0.0004)	(0.0004)	(0.0004)
Observations	873,178	873,196	873,196
R^2	0.439	0.438	0.437
Adjusted R^2	0.439	0.438	0.437
Residual Std. Error	330.145 (df = 873161)	330.517 (df = 873179)	330.861 (df = 873179)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table XIV: Estimation of Regression 1 with Linear Baseline Control

This table reports estimates of Regression 1, which estimates the effect of changes in stock-specific information on simultaneous trades between stocks and ETFs which contain them. In this regression, I do not remove the baseline level of trading, but instead use it as a linear control in the regression. I consider two different measures of stock-specific information: earnings dates and absolute value of the return. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Raw Simultaneous Trades		
	(1)	(2)	(3)
Weight	9.938***	16.243***	16.241***
	(2.564)	(4.516)	(4.510)
Abs Return	−16.497***		
	(5.847)		
Earnings After		−39.309***	
		(13.411)	
Earnings Before			−7.629*
			(4.249)
Baseline	3.987***	3.968***	3.967***
	(0.070)	(0.082)	(0.082)
Weight∗Abs Return	5.696***		
	(1.885)		
Weight∗Earnings After		2.178	
		(2.098)	
Weight∗Earnings Before			2.632***
			(1.013)
Observations	873,178	873,196	873,196
R^2	0.885	0.883	0.883
Adjusted R^2	0.885	0.883	0.883
Residual Std. Error	149.303 (df = 873163)	150.582 (df = 873181)	150.649 (df = 873181)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table XV: Estimation of Regression 1 with Date Fixed Effect
This table reports estimates of Regression 1, which estimates the effect of changes in stock-specific information on simultaneous trades between stocks and ETFs which contain them. In this regression, I have fixed effects for each date. I consider two different measures of stock-specific information: earnings dates and absolute value of the return. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Simultaneous Trades					
	(1)	(2)	(3)	(4)	(5)	(6)
Weight	60.570***	52.052***	60.605***	52.063***	49.501***	49.406***
	(20.411)	(19.807)	(20.398)	(19.798)	(16.693)	(18.318)
Earnings After	−23.309***	−22.822***				
	(3.531)	(3.661)				
Abs ETF Return		35.294*		35.276*		38.209**
		(18.432)		(18.433)		(19.353)
Weight∗Earnings After	6.731***	5.244***				
	(1.631)	(1.495)				
Weight∗Abs ETF Return		10.983		10.990		2.597
		(8.930)		(8.930)		(10.485)
Earnings Before			5.040***	5.266***		
			(1.750)	(1.738)		
Weight∗Earnings Before			4.407**	4.197***		
			(1.765)	(1.597)		
Abs Return					0.544	−0.112
					(2.155)	(1.488)
Weight∗Abs Return					11.854**	9.783*
					(5.022)	(5.796)
Observations	873,196	873,196	873,196	873,196	873,178	873,178
R^2	0.189	0.193	0.189	0.193	0.193	0.194
Adjusted R^2	0.188	0.192	0.188	0.192	0.192	0.193
Residual Std. Error	388.477	387.657	388.482	387.661	387.520	387.247
Degrees Freedom	872323	872321	872323	872321	872305	872303
Note:						*p<0.1; **p<0.05; ***p<0.01

Table XVI: Estimation of Regression 1 with Date-Stock Fixed Effect
This table reports estimates of Regression 1, which estimates the effect of changes in stock-specific information on simultaneous trades between stocks and ETFs which contain them. In this regression, I have fixed effects for each date and for each stock. I consider two different measures of stock-specific information: earnings dates and absolute value of the return. Earnings Date is an indicator which takes the value 1 for stocks which announce earnings before the day's trading session. Abs Return is the absolute value of the intraday return, measured as a percentage. The sample is SPDR and the ten Sector SPDR ETFs and their stock constituents from August 1, 2015 to December 31, 2018. The frequency of observations is daily. I include a fixed effect for each ETF and cluster standard errors by ETF.

	Dependent variable: Simultaneous Trades					
	(1)	(2)	(3)	(4)	(5)	(6)
Weight	-37.515^{***}	-49.428^{***}	-37.489^{***}	-49.431^{***}	-46.187^{***}	-50.028^{***}
	(9.011)	(15.321)	(9.027)	(15.350)	(12.359)	(15.401)
Earnings After	-22.921^{***}	-22.277^{***}				
	(3.191)	(3.323)				
Abs ETF Return		28.880^{**}		28.860^{**}		28.239^{**}
		(13.956)		(13.959)		(13.164)
Weight∗Earnings After	6.712^{***}	4.954^{***}				
	(1.470)	(1.481)				
Weight∗Abs ETF Return		14.751^{*}		14.758^{*}		12.155
		(8.505)		(8.504)		(8.025)
Earnings Before			4.749^{***}	5.005^{***}		
			(1.633)	(1.623)		
Weight∗Earnings Before			4.905^{***}	4.709^{***}		
			(1.715)	(1.481)		
Abs Return					1.068	2.766
					(3.469)	(2.068)
Weight∗Abs Return					9.261^{**}	2.939^{***}
					(4.124)	(0.785)
Observations	873,196	873,196	873,196	873,196	873,178	873,178
R^2	0.470	0.474	0.470	0.474	0.472	0.474
Adjusted R^2	0.469	0.473	0.469	0.473	0.471	0.474
Residual Std. Error	314.301	312.944	314.306	312.948	313.576	312.870
Degrees Freedom	871703	871701	871703	871701	871685	871683
Note:						*p<0.1; **p<0.05; ***p<0.01

REFERENCES

Baker, Scott, Nicholas Bloom, Steven J Davis, and Marco Sammon, 2015, What triggers stock market jumps?, in *Work in progress presented at the January 2015 ASSA meetings*.