Supplemental Appendix

"Anticipatory Traders and Trading Speed"

Raymond P. H. Fishe
Richard Haynes
Esen Onur

This supplemental appendix provides additional tables and information about the sample data used in the study. The core data are from an order book (RAPID) sample of the WTI Crude Oil futures contract (Dec 2011 expiration) covering 48 days of trading beginning on September 18, 2011. A follow-on sample was obtained for 35 days beginning on August 1, 2014 for the October 2014 expiration. Additional information about trades was available from the Transaction Capture Report (TCR) database that provides information on both sides of a trade. These samples allowed testing of 7,554 participant accounts in 2011 and 8,087 accounts in 2014.

Analysis of the two samples showed significant attrition in customer accounts. Initially, we found only about 20% account matches between samples using the RAPID account identifiers. Using information from the TCR database these matches increased to 35%. The difference in matching frequencies is mainly due to the nature of account reporting in the two systems, with both databases together allowing additional identifiers to aid in the matching.

However, the exact size of the attrition is unknown because many new accounts may belong to participants who are in the 2011 sample, but are now using a different clearing firm or executing broker for their trades. There are several reasons that this may have occurred. First, the bankruptcy of MF Global on October 31, 2011 caused about 800 accounts in our sample to seek new clearing arrangements. A new clearing arrangement is very likely to result in a new account identifier. Second, the regulatory environment and economics of Futures Commission Merchants (FCMs) changed after the adoption of the Dodd-Frank Act.[1] There were 154 FCMs operating prior to the 2008-09 financial crises. At the beginning of 2015, the number was down to 74. These changes alone may explain the attrition we observe in sample accounts. Third, the futures business has historically experienced turnover as Boyd and Kurov (2015) document. In their study it is the introduction of electronic trading that decreased the ranks of active participants. Similarly, more recent technological advances in algorithmic trading may have impacted our sample comparisons as trading costs continued to decrease (Fishe and Smith 2017). Fourth, participants do not always trade the same commodities or expirations, so there is a subset that will not match because they are choosing to trade other commodities or expirations. Lastly, the zero-sum nature of derivatives suggests that some participants will not do well if a subset earns positive profits. Depending on the nature of those with losses—do they hedge an underlying exposure or take directional risk—they may exhaust their capital and depart the business, adding to the attrition rate over time.

The following figures and tables offer additional information about our results. They are provided with a descriptive note to aid the reader. Some of the tables are extensions of tables in the main text.

[1] See "Ranks of commodities brokers dwindle as U.S. futures industry evolves," *Reuters*, July 2, 2015.

Table A1

Monthly Price Path Characteristics

Summary statistics are shown for all local price paths in the December 2011 WTI crude oil contract. The paths are computed from intraday prices on outright trades or spread trades with one side being outright. The sample covers all trading days from September 12 to November 18, 2011. The table summarizes information by month and path direction for all price paths, price paths with Type E participants, and price paths with Type R participants. The table shows average and median path returns, average path duration in seconds, average path volume in contracts, and the average number of trades. Panel A shows this information for all participants, Panel B selects paths in which a Type E participant has trades, and Panel C selects paths in which a Type R participant trades. Type E and Type R participants are selected using the 5% FDR critical rate. The average number of unique participants is shown for all paths, but limited to the average count of Type E and Type R participants in Panels B and C, respectively.

					Average			
Month	Path Direction	Path Count	Average Path Ret. (%)	Median Path Ret. (%)	Path Duration (sec)	Path Volume	Number of Trades	Number of Unique Participants
Panel A: Statistics for All Paths								
September	down	611	-0.346	-0.291	481.2	296.6	195.5	27.4
	up	608	0.337	0.279	390.6	275.8	182.5	26.2
October	down	4,085	-0.145	-0.114	96.0	598.4	467.3	90.9
	up	4,090	0.148	0.113	89.6	534.5	416.8	83.6
November	down	3,207	-0.114	-0.094	70.9	584.7	470.7	97.8
	up	3,218	0.115	0.094	65.3	532.1	425.3	90.6
Panel B: Statistics for Paths with Type E Participants								
September	down	196	-0.466	-0.390	859.0	504.5	324.7	1.5
	up	192	0.441	0.369	646.8	459.5	296.2	1.5
October	down	1,372	-0.185	-0.144	144.9	914.0	705.7	1.5
	up	1,259	0.201	0.154	145.2	886.5	686.7	1.5
November	down	887	-0.146	-0.123	113.9	945.7	754.2	1.3
	up	839	0.155	0.137	113.3	956.3	752.5	1.3
Panel C: Statistics for Paths with Type R Participants								
September	down	359	-0.423	-0.368	627.3	400.7	259.8	2.5
	up	352	0.409	0.360	484.7	370.7	242.6	2.4
October	down	3,724	-0.149	-0.117	99.3	639.9	500.1	7.4
	up	3,665	0.154	0.118	93.6	579.4	452.0	6.9
November	down	3,057	-0.116	-0.096	72.7	604.8	486.7	7.4
	up	2,979	0.119	0.099	68.3	564.4	451.1	7.1

Table A2

Buy-Sell Frequencies on Local Price Paths

The table shows the fraction of all trades on the buyside by Type E, Type R, and all other participants in four volume segments along the local price paths. These Type E and Type R participants are those found using the FDR 5% control rate. The four segments correspond to the first 10 pecent of path volume, the next 40 percent to the volume midpoint, then the next 40 percent to the 90 percent level, and finally the last 10 percent of path volume. Panel A shows results for upward trending price paths and Panel B shows the same results for downward trending price paths. In the first 10 percent of volume, Type E traders are expected to disproportionately buy in upward trending paths and sell in downward trending paths. In the last 10 percent of volume, Type R traders are expected to sell in upward trending paths and buy in downward trending paths, anticipating the change in local price direction in the next path. The results confirm these hypotheses, but also suggest that buying is moderately high (>50% in Panel A) for Type E participants in the remaining volume segments, so they have a general buyside bias. For Type R, buying is highest in the last 10% of volume (90% in Panel B), but seem to increase to this level in the prior segments, suggesting a gradual strategy of searching for the reversal point.

Volume Bins	Type E Participants	Type R Participants	All Other Participants
Panel A: Price Paths Trending Upward			
First 10%	0.77	0.59	0.49
10% to 50%	0.63	0.25	0.50
50% to 90%	0.58	0.17	0.51
Last 10%	0.65	0.08	0.50
Panel B: Price Paths Trending Downward			
First 10%	0.20	0.39	0.51
10% to 50%	0.44	0.71	0.49
50% to 90%	0.43	0.77	0.49
Last 10%	0.52	0.90	0.49

Table A3

Do Anticipatory Traders Bring New Information to the Market?

To test whether Type E or Type R anticipatory traders bring new price information to the market, the price change in each bin relative to the absolute price change in the path is computed for specific bins, and for all bins combined. In the main paper, bins are defined by ten equal duration calendar time periods across each path to avoid the endogeneity problem of volume deciles. The results here are computed using volume deciles for comparison. Averages of these data are reported for different cases based on whether Type E or Type R participants are trading in a given time bin in a price path. The first two columns report averages across all bins separating the averages for relative price changes by "up" and "down" trending paths, respectively. The remaining columns also average the data by upward and downward paths, but only use specific time bins in the calculation. The figures shown in parentheses beneath these averages are p-values for the t-test of comparing two means (unequal variances).

	Means of Sample Data - Relative Price Changes by Bin							
	Up	Down	Up Price Paths			Down Price Paths		
	All Bins		Bin#10	Bin#1	Bin#2	Bin#10	Bin#1	Bin#2
Path Characteristics	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
All paths using bins defined by Volume deciles								
No Type E in Current Path Volume Bin#1	0.095	-0.096		0.262	0.116		-0.245	-0.108
Type E in Current Path Volume Bin#1	0.100	-0.097		0.284	0.114		-0.250	-0.188
p-value for t-Test of Diff.	(0.658)	(0.948)		(0.542)	(0.952)		(0.734)	(0.274)
No Type E_2s in Current Path Volume Bin#1	0.095	-0.096		0.261	0.118		-0.243	-0.109
Type E_2s in Current Path Volume Bin#1	0.098	-0.097		0.272	0.098		-0.263	-0.120
p-value for t-Test of Diff.	(0.515)	(0.922)		(0.471)	(0.232)		(0.072)	(0.636)
No Type R in Previous Path Volume Bin#10	0.095	-0.956	0.149	0.267		-0.128	-0.255	
Type R in Previous Path Volume Bin#10	0.095	-0.967	0.157	0.256		-0.135	-0.232	
p-value for t-Test of Diff.	(0.951)	(0.608)	(0.240)	(0.134)		(0.324)	(0.001)	
No Type R_2s in Previous Path Volume Bin#10	0.095	-0.096	0.147	0.268		-(0.125)	-(0.253)	
Type R_2s in Previous Path Volume Bin#10	0.095	-0.097	0.157	0.257		-(0.136)	-(0.237)	
p-value for t-Test of Diff.	(0.977)	(0.666)	(0.173)	(0.117)		(0.148)	(0.028)	

Table A4
Characteristics of Anticipatory Traders

An alternative to the regressions shown in the text is the inverse regression. This method infers characteristics of the Type E and Type R traders relative to the entire sample of participants. It is agruably more appropriate because the Type E and Type R traders are known at this point in the analysis, so the results are conditional on this fact. In the Type E (Type R) results, the dependent variable equals 1 if the trader is identified by the FDR methods as significant in the first 10% (last 10%) of the local price paths. The independent variables are averages by participant across all trades in the sample. Execution speed is measured as the difference between the time the Nymex confirms a trade execution and the time it receives the initial order from the participant (in seconds with fractional milliseconds). The speed variable is the natural logarithm of one plus the execution speed. The coefficients reported are transformed from initial least squares estimates following Fishe and Smith (2012). These coefficients measure the difference between the characteristic of the Type E (or Type R) trader and what would be expected given all of the other characteristics in the sample. The sample is defined to be those participants tested using the FDR method, so all traders have 30 or more transactions. Sample sizes are reduced when participants are not observed in the first or last 10% of path volume. The p-values of the underlying coefficient are shown in parentheses beneath each estimate.

Characteristics	Sample Average over Participants	Type E Participants	Type R Participants
Average Speed during the First 10% of Path Volume	1.246 sec	-0.264 (0.001)	
Average Speed during the Last 10% of Path Volume	1.348 sec		0.069 (0.161)
Average Speed over All Trades	1.255 sec	0.014 (0.818)	0.182 (0.000)
Average Trade Size	1.275	0.285 (0.000)	0.051 (0.120)
Percentage of Aggressive Trades	55.8%	-0.052 (0.026)	-0.072 (0.000)
Percentage of Trades on the Buyside	50.2%	-0.007 (0.647)	0.014 (0.250)
Binary Indicator for Proprietary Trader	9.8%	0.033 (0.214)	0.069 (0.001)
Binary Indicator for Algorithmic Trader	18.3%	0.004 (0.911)	0.055 (0.037)
R-Squared		2.1%	4.2%
Sample Size		6,886	7,208

Table A5
Effects of Anticipatory Trading on the Order Cancellations of Other Market Participants

This table shows cancellation regressions controlling for the rate of net buying behavior of Type E traders in the first 10% of path volume and Type R traders in the last 10% of path volume. The dependent variables with type E participants are the buyside order cancellation fraction (buy cancel volume divided by total buy and sell cancel volume) for other market participants between 10% and 50% and 10% and 100% of current path volume; for Type R participants they are the buyside order cancellation fraction for other market participants for the first 50% and first 90% of next path's volume . The independent variables are the fraction of "Anticipatory" participants net buying volume to total volume in the specified "Percentile" of current path volume and the same fraction computed for all other non-anticipatory participants, where "Anticipatory" participants are either Type E or Type R and "Percentile" is either the first or last 10% of volume in the path. Control variables in each rergression are a dummy variable indicating the direction of prices on the path (up/down), the lag of total path volume, and the lag of the dependent variable. The p-values are computed using heteroscedistic-consistent standard errors and are shown in parentheses below each coefficient; for control variables the coefficient sign and p-values are shown. The adjusted R-squared is shown at the bottom of the table.

	Dependent Variables:											
	Buyside Order Cancellation Rates between 10% and 50% of the *current* path volume			Buyside Order Cancellation Rates between 10% and 100% of the *current* path volume			Buyside Order Cancellation Rates for the first 50% of the *next* path volume			Buyside Order Cancellation Rates for the first 90% of the *next* path volume		
	Anticipatory Participant = TYPE E			Anticipatory Participant = TYPE E			Anticipatory Participant = TYPE R			Anticipatory Participant = TYPE R		
	Percentile = First 10%			Percentile = First 10%			Percentile = Last 10%			Percentile = Last 10%		
Independent Variable	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Fraction of "Anticipatory" participants net buying volume to total volume in specified "Percentile" of *current* path volume	0.030		0.028	0.014		0.013	-0.059		-0.053	-0.036		-0.033
	(0.048)		(0.063)	(0.236)		(0.266)	(<0.001)		(<0.001)	(0.002)		(0.006)
Fraction of all other, non-anticipatory participants net buying volume to total volume in the specified "Percentile" of *current* path volume		-0.015	-0.015		-0.008	-0.007		0.015	0.014		0.009	0.009
		(<0.001)	(<0.001)		(0.006)	(0.007)		(<0.001)	(<0.001)		(<0.001)	(<0.001)
<u>Control Variables (Sign; p-value):</u>												
Up/Down Price Direction on Path	neg (<0.001)	neg (<0.001)	neg (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (0.007)	pos (0.004)	pos (0.107)
Lag of Path Volume	neg (0.254)	neg (0.250)	neg (0.251)	neg (0.051)	neg (0.050)	neg (0.052)	neg (0.154)	neg (0.154)	neg (0.161)	neg (0.120)	neg (0.120)	neg (0.123)
Lag of Dependent Variable	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)
Intercept	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)
Number of Paths	15,583	15,583	15,583	15,642	15,642	15,642	15,765	15,765	15,765	15,765	15,765	15,765
R-Squared	1.4%	1.6%	1.6%	3.9%	4.0%	4.1%	4.1%	4.2%	4.3%	3.6%	3.6%	3.7%
p-value for Regression	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)

Table A6

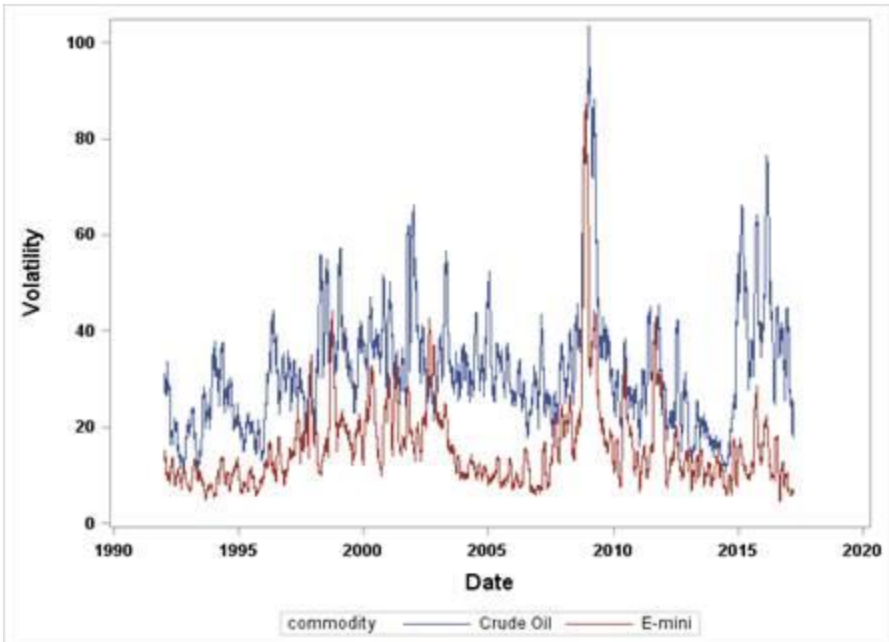
Effects of Anticipatory Trading on the Order Modifications of Other Market Participants

This table shows modification regressions controlling for the rate of net buying behavior of Type E traders in the first 10% of path volume and Type R traders in the last 10% of path volume. The dependent variables with type E participants are the buyside order modification fraction (buy modify volume divided by total buy and sell modify volume) for other market participants between 10% and 50% and 10% and 100% of current path volume; for Type R participants they are the buyside order modification fraction for other market participants for the first 50% and first 90% of next path's volume . The independent variables are the fraction of "Anticipatory" participants net buying volume to total volume in the specified "Percentile" of current path volume and the same fraction computed for all other, non-anticipatory participants, where "Anticipatory" participants are either Type E or Type R and "Percentile" is either the first or last 10% of volume in the path. Control variables in each rergression are a dummy variable indicating the direction of prices on the path (up/down), the lag of total path volume, and the lag of the dependent variable. The p-values are computed using heteroscedistic-consistent standard errors and are shown in parentheses below each coefficient; for control variables the coefficient sign and p-values are shown. The adjusted R-squared is shown at the bottom of the table.

	Dependent Variables:											
	Buyside Order Modification Rates between 10% and 50% of the *current* path volume			Buyside Order Modification Rates between 10% and 100% of the *current* path volume			Buyside Order Modification Rates for the first 50% of the *next* path's volume			Buyside Order Modification Rates for the first 90% of the *next* path's volume		
	Anticipatory Participant = TYPE E			Anticipatory Participant = TYPE E			Anticipatory Participant = TYPE R			Anticipatory Participant = TYPE R		
	Percentile = First 10%			Percentile = First 10%			Percentile = Last 10%			Percentile = Last 10%		
Independent Variable	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Fraction of "Anticipatory" participants net buying volume to total volume in specified "Percentile" of *current* path volume	0.072		0.073	0.064		0.064	0.021		0.018	0.016		0.013
	(0.001)		(<0.001)	(<0.001)		(<0.001)	(0.098)		(0.152)	(0.138)		(0.214)
Fraction of all other, non-anticipatory participants net buying volume to total volume in the specified "Percentile" of *current* path volume		0.003	0.181		0.000	0.000		-0.007	-0.007		-0.007	-0.007
		(0.516)	(0.447)		(0.888)	(0.989)		(0.005)	(0.007)		(0.002)	(0.003)
<u>Control Variables (sign; p-value):</u>												
Up/Down Price Direction on Path	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	neg (<0.001)	neg (<0.001)	neg (<0.001)	neg (<0.001)	neg (<0.001)	neg (<0.001)
Lag of Path Volume	neg (0.433)	neg (0.490)	neg (0.518)	neg (0.472)	neg (0.447)	neg (0.472)	neg (0.137)	neg (0.137)	neg (0.135)	neg (0.143)	neg (0.143)	neg (0.141)
Lag of Dependent Variable	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)
Intercept	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)	pos (<0.001)
Number of Paths	15,576	15,576	15,556	15,638	15,638	15,638	15,765	15,765	15,765	15,765	15,765	15,765
R-Squared	8.6%	8.5%	8.6%	16.9%	16.8%	17.0%	11.8%	11.8%	11.8%	18.8%	18.8%	18.8%
p-value for Regression	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)

The two figures below show the historical behavior of inter-day and intra-day volatility in crude oil and e-Mini futures contracts. They tend to mirror each other over time, so that fluctuations in equities are similar to fluctuation in the price of crude oil.

Inter-Day Volatilities



Intra-day Volatilities