Order Flows and Financial Investor Impacts in Commodity Futures Markets*

Mark Ready
University of Wisconsin-Madison

Robert Ready
University of Oregon

October 10, 2021

Abstract

Using intraday data, we document a statistically strong, but temporary, association between commodity-index trader flows and commodity futures prices. Reexamining the positive returns associated with the issuance of commodity-linked notes documented by Henderson, Pearson, and Wang (2015), we find that these returns are too large to be explained by the small trades necessary to hedge these notes, and provide new evidence that they are instead the result of endogenous issuance. Our results provide novel support for commodity financialization, but highlight the importance of measuring the magnitude of financial investment, since even large financial flows have economically modest impacts on prices.

JEL Codes: G12, G13
Keywords: Commodities, Futures, Order flow, Financialization

*We would like to thank Steffen Hitzemann and Brian Henderson (discussants) and participants at the 2017 Pacific Northwest Finance Conference, the 2018 Midwest Macroeconomics Meetings, the 2019 Midwest Finance Association Meetings, and the 2019 University of Oklahoma Energy and Commodities Finance Research Conference. We would also like to thank two anonymous referees for their excellent comments.

1 Introduction

Increasing financial investment in commodity futures over the last 20 years has generated substantial interest in the impact of this investment on commodity markets. Theoretical work (e.g. Hamilton and Wu (2014), Sockin and Xiong (2015), and Goldstein and Yang (2017)) predicts that, under certain conditions, trading by financial investors can impact prices even if these investors are uninformed about market fundamentals, a hypothesis often referred to as the "financialization" of commodity markets. Another large literature tests for empirical evidence of financialization, but little consensus has emerged as to whether, or how much, financial investors affect prices in these markets. For instance, much of this work focuses on commodity-index traders (CITs), and generally finds no evidence that changes in the positions of CITs impact commodity prices.[1] Conversely, an influential paper by Henderson, Pearson, and Wang (2015) (henceforth HPW) finds commodity price increases on days with the pricing of commodity-linked notes (CLNs), and argues that these price increases are evidence that the uninformed trades necessary to hedge these notes can create significant movements in the underlying commodity prices.

In this paper, we reexamine the empirical relations between commodity prices and financial investors, but we pursue a novel approach by using intraday data on returns and order flow in commodity futures markets. Using these data we provide new evidence that the flows associated with CITs are large relative to futures markets, and unlike the previous literature, we find strong statistical evidence that these flows do in fact impact prices. However, these impacts are small relative to the size of CIT flows and are highly concentrated in the minutes immediately prior to the daily futures settlement. Moreover, these impacts partially reverse in the minutes after the settle, and reverse more fully in the subsequent week. In contrast to our results on CITs, we find that the trades necessary to hedge CLNs are roughly an order of magnitude too small to explain the price increases documented by HPW. We also present new evidence that these price increases are the result of notes with flexible issuance dates being more likely to issue on days with rising prices, rather than the result of price impacts from hedging trades. Taken together, our results

[1] For instance, Hamilton and Wu (2015) review previous work (e.g. Stoll and Whaley (2010), Irwin and Sanders (2012), and Fattouh, Kilian, and Mahadeva (2013)) and conduct their own analysis, finding that "*consistent with most of the earlier literature, (...) index-fund investing seems to have little impact on futures prices in these markets*". See Cheng and Xiong (2014) or Boyd, Harris, and Li (2018) for a review of this literature.

suggest that commodity prices are indeed impacted by large financial flows in a manner consistent with theories of financialization, but that this impact is modest and temporary, and that smaller financial flows have little or no impact on commodity prices.

The new results presented here rely crucially on our use of intraday data, which allows us to make two important methodological contributions relative to previous work. First, we use intraday trade-and-quote data for a variety of commodity futures contracts to sign trades as buy or sell, and then estimate the price impact per dollar traded in these markets. These estimates reflect the cost of trading with marketable orders, and at least some of these orders are motivated by information, so the estimates are likely larger than the impact of financial investors, who would typically be uninformed and may choose to trade less aggressively. Thus, the impact of an average trade in these markets provides a reasonable upper bound for the impact of trades from financial investors. The second methodological contribution is that our intraday data allow us to examine order flow and price changes at specific times within the trading day, and thus provide more power for our empirical tests. In particular, we find that volumes increase and price impacts substantially decrease in the minutes prior to the daily futures settlement. Consistent with theoretical models of trading (e.g. Admati and Pfleiderer (1988)), this suggests that uninformed investors are concentrating their trading in this time period. This is therefore the period where one might expect financial investors to concentrate, because they will be drawn by the lower transaction costs and because they often have an incentive to trade at or near the daily settlement price.

In our first set of tests, we use our price impact estimates and find that changes in CIT positions are large, in the sense that they could explain a substantial amount of the overall variation in commodity prices if the associated trades had the same impact as the average trade. The lack of such findings in the earlier literature suggests either that these traders are able to lessen the impact of their trades, or that empirical tests are mis-specified, perhaps due to an omitted variable driving both commodity prices and index trader positions (Cheng, Kirilenko, and Xiong (2014)). When we look within the trading day we find evidence for the former explanation. We find that weekly index flows are strongly related to futures returns, but that the size of the impact is substantially smaller than that of the average trader. The apparent impact of these trades is also highly concentrated in the minutes leading up to the daily futures settlement, so our intraday results are statistically much stronger than those that would be obtained using daily data. Additionally, we find strong evidence of partial reversal of this impact in the minutes immediately following the futures settlement and a more complete reversal in the subsequent week. These results together suggest that index traders are able to mitigate their price impact, but their flows do result in small temporary price changes, as is consistent with theories of financialization.

In our second set of tests we revisit the results of HPW. The primary finding in the original paper is that on approximately 350 days with the pricing of CLNs (identified via 424b filings) there is statistically significant average positive return of approximately 29 basis points. As a first step, we replicate this result by collecting the relevant filings and find a highly significant average return of 28 basis points on these days. In order to evaluate whether or not this effect is consistent with price impacts from hedging trades, we extend the analysis of HPW and use the specific features of each CLN to estimate the size of the associated hedge for each note. We then use this size to estimate the predicted price impact for each note, and find that these notes are quite small in magnitude. For HPW's sample, we find an average hedging trade size of approximately $11 million with a predicted price impact of slightly less than 6 basis points. When we focus on the approximately 90% of these notes issued in the most liquid markets, this difference is even more stark. In these markets we find a strongly significant average return of approximately 24 basis points associated with the pricing of the notes, with an average predicted price impact of only 2.6 basis points. In other words, for these notes the positive return is approximately an order of magnitude larger than the impact from a typical trade of this size, which again is a likely upper bound for the impact of an uninformed hedging trade.

Looking within the day, we also find that the positive returns occur mostly early in the trading day rather than in the minutes just prior to the pricing of the notes. This is also puzzling because the issuers would likely hedge as close to the pricing time as possible to avoid basis risk, and because the pricings tend to occur either at the daily settlement or during another fixing time (e.g. the London 3:00 p.m. Fix for gold), which tend to be highly liquid periods that could easily accommodate the relatively small volumes necessary to hedge the notes.[2]

What then is the source of these positive average returns? HPW's empirical strategy implicitly assumes that the pricing date of the note is fixed *ex ante* and so is not subject to a possible reverse

[2]We do find some evidence of positive returns near the pricing of the notes, but they are roughly the same magnitude as the predicted price impacts and therefore too small to explain the overall result.

causality whereby notes are more likely to be issued on days with positive return.[3] We extend their analysis by collecting additional SEC filings known as "Free-Writing Prospectuses" (FWPs) that were not collected by HPW. These FWPs are required whenever the issuer communicates important features of the offering in advance to retail investors. For approximately half the notes, we find an FWP filed prior to the pricing of the note with an indication of an expected pricing date. For these notes with a pre-specified pricing date we find no positive average return. For the remaining notes, it is either the case that 1) the FWP was filed prior to the pricing of the note but without an expected pricing date, 2) no FWP was ever filed, or 3) the FWP was filed after the note priced but before it was issued. In these cases the issuer presumably had flexibility to decide the pricing date, and strikingly, it is only for these notes that we find positive average returns. This difference in return between notes with and without a pre-specified pricing date is highly significant and robust to the inclusion of controls for the size of the note, the linked commodity, the year of issuance, and the issuer. This evidence, along with the large size of the pricing-day returns relative to the small sizes of the notes, suggests that the positive average return documented by HPW is most likely the result of an increased likelihood of issuing notes on days with positive returns, rather than being evidence of financialization.

Theories of financialization (e.g. Hamilton and Wu (2014), Acharya, Lochstoer, and Ramadorai (2013), Sockin and Xiong (2015), Baker (2021), Basak and Pavlova (2016), and Goldstein and Yang (2017)) are closely related to broader models featuring non-fundamental price impacts, such as Scholes (1972), Grossman and Stiglitz (1980), Kyle (1985), and Hendershott and Menkveld (2014). Given that many of these models are targeted at explaining higher-frequency market movements, we believe that is natural to use intraday behavior to study this question in commodity markets. This is in contrast to the previous literature, which primarily tests for the impact of financialization using daily, weekly, or monthly data. As mentioned previously, there is little agreement on the overall effect of financial traders. Several papers find evidence supporting the impacts of financialization, either in the form of price impacts or predictable returns, including Buyuksahin and Robe (2011), Tang and Xiong (2012), Singleton (2013), Cheng et al. (2014), and HPW, while

[3]HPW acknowledge this potential for reverse causality and address it by including an analysis of the determination days near the notes' maturities, which are specified far in advance and are therefore not subject to this issue. Here they find a significant negative return when the hedges would presumably be unwound. We were unable to replicate this result. See section 4.3 for detailed discussion.

others find no evidence of impacts, including Stoll and Whaley (2010), Irwin and Sanders (2011), Irwin and Sanders (2012), Silvennoinen and Thorp (2013), Fattouh et al. (2013), Alquist and Gervais (2013), Hamilton and Wu (2015), and Chari and Christiano (2017). More recently, a concurrent paper of Yan, Irwin, and Sanders (2021) examines index-fund rebalancing and finds evidence of price impacts in futures markets.[4]

While the above empirical work studies prices at daily or longer frequencies, there is a small set of papers that study intraday trading and liquidity in commodity markets. Bessembinder, Carrion, Tuttle, and Venkataraman (2016) study liquidity around the predictable rolling of index funds, and Bessembinder (2015) reviews the empirical and theoretical framework for understanding predictable roll trades.[5] Raman, Robe, and Yadav (2017) examine price impacts and liquidity over a one-year period around the advent of electronic trading in WTI oil futures markets in 2007, but do not examine price impacts of financial investors.

Other related work includes Elder, Miao, and Ramchander (2014), who study intraday price patterns in Brent and WTI futures, Marshall, Nguyen, and Visaltanachoti (2011), who study liquidity proxies in commodity prices, and Halova, Kurov, and Kucher (2014), who study price reactions to inventory announcements. However, these papers do not study financial investment.

The remainder of the paper is organized as follows. In section 2 we use our intraday trade-and-quote data to study the price impacts of typical trades in seven large commodity markets for which we have data. In section 3, we build on these results and use an extensive sample of intraday futures returns to study the impacts of CITs on futures markets. Section 4 re-examines the results of HPW regarding the impacts of CLNs. Section 5 concludes.

2 The Price Impact of Order Flows in Commodity Futures

As a first step for our analysis of financial trader impacts, we use intraday trade-and-quote data in seven major commodity markets to quantify the price impact of a given amount of buying or

[4]We view our results as complementary to those of Yan et al. (2021). They use daily returns and focus on the once-yearly reweighting of the GSCI. They find a modestly significant impact of from changing weights, that appears to reverse, though this reversal in not significant. While we do not utilize a plausibly exogenous source of variation for index flows, we examine a broader set of flows and find highly significant evidence of both impacts and reversals.

[5]In contrast to our results, these papers find little evidence of price impact from these roll trades. This is perhaps due to the fact that roll trades are generally predictable and involve the simultaneous buying and selling of multiple maturities, as opposed to the mostly unpredictable changes in the total amount of index investment across all maturities that we focus on here.

selling. We then use the results for these seven commodities, along with volume and return data available at daily frequencies to extrapolate our price impact measures to other commodities and years for which we do not have trade-and-quote data.[6]

2.1 Calculating Futures Returns and Near Month Imbalance

The seven contracts for which we have trade-and-quote data include two energy contracts: both the West Texas Intermediate (WTI) contract traded on the NYMEX (now owned by the CME) and the Brent contract traded on the ICE. We also include the CME gold, corn, soybeans, wheat, and copper contracts. In terms of open interest and volume, these contracts are generally largest in their respective commodity classes. Moreover, the gold, corn, soybeans, and wheat contracts on the CME are the dominant futures markets for each commodity. The copper contract on the CME rivals the contract traded on the London Metal Exchange, but generally has slightly lower volume. We use trade-and-quote data from the Globex electronic trading platform for all six of the U.S. commodities in our sample. For Brent futures, we use trade-and-quote data from the ICE, which was fully electronic for our entire sample period.[7]

We classify each trade in a single-month contract as a buy or sell by comparing the price to the current quote for that contract (similar to Lee and Ready (1991)), and we aggregate buying and selling volume by minute. We also measure the return over each minute using quote midpoints as of the end of each minute.[8] We exclude floor trades from our imbalance measure because they are executed manually, making it impossible to accurately align them in time with the Globex quotes, and therefore impossible to assign trade direction. We also exclude calendar spread trades from our imbalance measure, motivated in part by results from supplemental tests where we found that the imbalance in calendar spread trades has little impact on the level of front and next month futures

[6]See section IA.1 in internet appendix for a description of our data sources and their coverage.

[7]Globex trading in WTI futures runs from Sunday night at 6:00 p.m. in New York to Friday night at 5:00 p.m. with one-hour breaks at 5:00 p.m. each day. The bulk of the trading occurs during the day, so when we do analysis by minute, we limit our WTI sample to the time periods from 7:30 a.m. to 4:00 p.m. New York time. This time window captures 88% of the total WTI volume in the front and next month contracts. We use this same time window for minute-by-minute analysis of Brent, gold and copper. For corn, soybeans, and wheat we use 9:30 a.m. through the close, which was 2:15 p.m. for most of our sample, but was delayed until 3:00 p.m. in late 2012 and early 2013.

[8]There may be some concern (e.g. O'Hara (2015)) that such a procedure incorrectly classifies trades. However, our soybeans data contain an explicit aggressor flag, and we confirm in this data set that this method of classifying buys or sells is extremely accurate (well above 99%). We also report robustness using the tick-test to identify trades in the internet appendix, and find similar results. See Tables IA.5 and IA.6.

prices[9]. Finally, we do not include data for Trade-at-Settlement (TAS) trades in our analysis. These trades allow an investor to enter into a position early in the day that will be executed at the closing price and thus may be useful for non-fundamental investors.[10] The exclusion of these other sources of trading are likely to bias up our measures relative to the true predicted impacts for uninformed financial traders.[11]

The CME procedures for determining daily settlement prices begin by focusing the contract that generally has the highest volume. This is called the "Active Month" for WTI, gold and copper, and is called the "Lead Month" for corn, soybeans, and wheat. We measure returns using the quote midpoints for the Active/Lead Month contracts. We measure imbalance using the total difference between buy and sell volume for trades in all contract months from the front month through the month that is currently the Active/Lead month or is within three weeks of becoming the Active/Lead month. Note that our definition of imbalance effectively nets out any trades that are a result of a trader rolling between the nearest contract months. The Active Month in WTI is the generally the front month but switches to the next month for the two days prior to expiration. So for example, if a WTI trader uses market orders (or a calendar-spread order) to sell the front month and buy the next month (within roughly three weeks of the front month expiration), our measure will reflect zero net imbalance for those trades.[12] This is a marked contrast to work such as Bessembinder et al. (2016) or Mou (2010) that studies predictable roll trades.

2.2 Summary Data for Returns and Imbalance

Table 1 shows summary statistics for our seven futures contracts. Panel A shows summaries for all minutes in the sample, while Panel B shows summaries for only the minute prior to futures settlement. We measure returns in percent, and express both volumes and imbalances as millions of dollars of futures notional (calculated using the previous minutes closing midpoint price).

Trade volumes are large and, trade volumes, imbalances, and returns are quite volatile over

[9]See internet appendix Table IA.2.

[10]We were able to obtain TAS data for WTI and Gold contracts. In unreported analysis we find that buys and sells in TAS contracts have little impact on the price when they are executed, and therefore they are unlikely to explain the large returns early in the trading day associated with CLN pricing, and we find no evidence of abnormal TAS imbalance or volume on these days. We also find that TAS trading throughout the day does not predict price moves at the settle, and therefore is unlikely to explain our findings of price impacts near the settle from CITs.

[11]See internet appendix section IA.6

[12]See the internet appendix section IA.2 for more detail on determination of Active/Lead month and imbalance.

the period. Average one-minute volume ranges from approximately \$34 million of notional for WTI to approximately \$2.8 million of notional for Copper. Average imbalances are near zero, but they are quite volatile with standard deviations between \$12 and \$15 million per minute for gold, Brent, and the WTI, and between \$2.2 and \$7.6 million per minute for copper, corn, soybeans, and wheat. Both volume and the volatility of imbalance are much higher in the settlement minutes. For instance, in the WTI, volume goes up by a factor of six and the volatility of imbalance is more than double in the settlement minute.

2.3 Price Impacts and Volumes Across the Trading Day

Our primary impact measure is the estimate of a slope in a regression of one-minute returns on one-minute imbalance.[13] We believe that this measure is likely an upper bound for the impact of financial traders in a given periods, due to both the exclusion of floor trades and the presence of both informed and uninformed traders in our imbalance measure.[14]

$$r_{c,t} = \alpha + \beta Imbalance_{c,t} + \epsilon_{c,t} \tag{1}$$

Here returns are measured in percent, and imbalance in millions of dollars. We estimate the univariate regression separately for each commodity c and Table 2 reports the results. Columns (1) - (4) show the results for regressions using all minutes for each commodity. For each commodity, we find that imbalance in futures markets has significant explanatory power for futures prices, suggesting that trading in these markets play an important role in price discovery.[15] The R^2 values range from 34% for WTI to 13% for Brent. The slope estimates provide our measure of impact, and range from 0.0022 for gold to 0.0155 for wheat. The interpretation is that a one million dollar buy (sell) will lead to a return in gold markets of positive (negative) 0.0022%, or 0.22 basis points. In contrast, in the smaller wheat market, a one million dollar trade will lead to a return impact of 1.55 basis points.

[13] This measure in effect treats all returns and trade imbalance as unpredictable, in contrast to the VAR formulation of Hasbrouck (1991). We perform these VARs and report results in the internet appendix (section IA.3). The primary takeaways are that the price impacts of both order flow and public return news are mostly permanent at one-minute horizons, and that most of the imbalance in each minute represents an unpredictable innovation.

[14] In the internet appendix (section IA.6) we show this formally in a simple model, and also address other potential issues with this price-impact measure, including those arising from mis-classification of trades due to the use of limit orders, and potential non-linearities in the relation between returns and imbalance.

[15] Evans and Lyons (2002) find a similar result in currency markets.

Columns (5) - (8) shows the slope estimates from the same regression, but with the sample restricted to only the minute prior to futures settlement. Here we see large reductions in the impact associated with a given amount of trading, with impacts typically between two-thirds and one-half of what see in the full sample. We are likely to observe the impact of financial investors during this period because they will be drawn by the lower transaction costs and because they often have an incentive to trade at or near the daily settlement price. Indeed, the lower transaction cost may in part result from their tendency to concentrate their trading in the interval.

To help visualize how trading impacts change through the day we estimate our univariate regression for each minute of the trading day. Figure 1 shows the results for these regressions, along with average volume, for each of the seven commodities.[16] The first panel shows the minute-by-minute average volume and price impacts throughout the trading day for WTI futures. The volume rises on the open of pit trading at 9:00 a.m., and then spikes at times of various announcements, including the EIA's weekly energy outlook published each Wednesday at 10:30 a.m. The largest spike however occurs at 2:30 p.m. in New York when the daily futures settlement price is set.

The implication of this finding is that even large trades during this period are unlikely to have a large impact on the market. For instance, a $10 million hedging trade (roughly the size of our average CLN), would only have an impact of 1.2 basis points if traded with a market order in the last minute before settlement. Note that a trade of this size would be less than one third of the standard deviation of imbalance for the settlement minute and less than 10% of the average settlement minute volume (see Table 1).

This pattern is repeated for each of the seven commodities. For all of the commodities volume spikes and price impact falls around the futures settlement. The reduction in price impact is most notable for the WTI and gold, but is apparent in all seven commodities. We also see similar patterns at other times where various price indices are calculated. For instance, the 10:00 a.m. (New York time) volume spike and drop in impact in the gold market corresponds to the London p.m. Fix.[17]

[16]There are approximately 1,800 trading days in the sample, so five of the regressions have approximately 1,800 observations (Brent data start in January 2008 and soybean data start in January 2009). For WTI, Brent, gold, and copper we consider the interval from 7:30 a.m through 4:00 p.m. in New York. Corn, soybeans, and wheat have extremely low volume after their close of floor trading at 2:15 p.m., so we end the analysis here. Corn, soybeans, and wheat also had their settlements delayed to 3:00 p.m. New York time for the 11-month period from 5/22/2012 to 4/5/2013, so we omit this period for the analysis in Figure 1. These data are included in all other analyses.

[17]The high volume and volatility of imbalance at the settlement means that the impacts in these minutes are estimated with high levels of statistical accuracy, and as such the difference between impacts in this period and those in the rest of the day is highly statistically significant.

2.4 Inferring Price Impacts from Daily Data for Other Commodities

In the following sections we examine futures price impacts of changes in positions of CITs and the issuance and maturity of CLNs. In both sections our results incorporate estimates of price impacts based on the results from this section. The index trading data cover a broad rage of commodities and years, and roughly 25% of our CLNs are linked to commodities for which we do not have intraday data. In order to estimate the potential price impacts for these commodities, in this subsection we regress our observed price impacts on data that we have daily frequency for a larger set of commodities. Intuitively we find that markets with lower volumes and higher volatilities have higher estimates of price impact. To formalize this intuition, we first calculate univariate regressions following the specification in Table 2 for each calendar year and commodity in our intraday data. We then regress estimates on the average daily volume in millions of dollars for the active (or highest volume contract where an active contract is not specified) maturities for a commodity market, as well as the daily volatility of returns to the active contract. All variables are in logs. The specification is therefore

$$\text{Log}(Impact_{c,yr}) = \alpha + \beta_1 \text{Log}(AvgVolume_{c,yr}) + \beta_2 \text{Log}(DailyVolatility_{c,yr}) + \epsilon_{c,yr} \qquad (2)$$

The first column of Panel A in Table 3 shows the regression using impact over all minutes in the calendar year as the dependent variable. The second column shows the same specification but with the dependent impact measured in the settlement minute. The third column shows the results from a pooled specification which adds a dummy variable for the settlement impacts

$$\text{Log}(Impact_{c,yr,p}) = \alpha + \beta_1 \text{Log}(AvgVolume_{c,yr}) + \beta_2 \text{Log}(DailyVolatility_{c,yr}) + \beta_3 \mathbb{1}_{settle} + \epsilon_{c,yr} \quad (3)$$

Here p denotes the period the impact was measured, either across all minutes or only in the settlement minute, and $\mathbb{1}_{settle}$ is a dummy variable that takes a value of one if the period is the settlement minute.

As the table shows, even with the relatively small sample, both volume and volatility are highly significant predictors of impacts with the expected signs. Moreover, the fit of the regression is extremely strong, with R^2 greater than 80%. Figure 2 illustrates the fit from the three regression specifications. As the figure shows, the regression performs extremely well in predicting impacts both across commodities and across years.

We then use the pooled regression to estimate impacts for each commodity-year for all of the contracts with CLNs and data on CIT positions. Panel B of Table 3 shows the averages for estimates across all years from 2003 to 2018 where data are available.[18] We find that gold, due to its high volume and relatively low volatility has the lowest estimated impact. Palladium contracts on the CME have the highest estimated impact. While these estimates are likely imperfect, the strong fit shown in Figure 2 suggests that they should provide reasonable estimates for impacts in commodities where we do not have intraday data.

We note for all of the subsequent analysis, we calculate predicted price impact as

$$PredictedPriceImpact_{c,t} = Settle\widehat{Impact}_{c,yr} \times FinancialFlow_{c,t} \tag{4}$$

Here $FinancialFlow_{c,t}$ is typically either a CLN hedging trade or an index flow measured in millions of dollars. $Settle\widehat{Impact}_{c,yr}$ is the predicted impact per million dollars of imbalance traded in the settlement minute in the year containing the flow, calculated using the daily volumes and volatilities for that year along with the pooled regression results from the estimate of equation 3. Note that we use this measure in the settlement minute, as opposed to the measure of impacts from all minutes. We believe that, particularly for the smaller CLNs that could easily be trader near the settlement, this is the more appropriate measure for the likely impact of an uninformed trader. While it may be reasonable to use the full day measure for the larger CIT flows, we use the settlement minute measure for consistency.[19]

3 Impacts of Index Trading in Commodity Futures Markets

In this section we explore the impacts of CITs on futures markets using intraday data. There are two sources of index positions compiled by the CFTC. The first is the weekly Supplemental Commitments of Traders (SCOT) report, which provides futures positions of traders following

[18]We use daily data from 2003 to 2018 to extrapolate our 2008 - 2013 price impacts to other commodities and years. For analysis of index flows in 2019 and 2020 we use the 2018 estimates of price impacts.

[19]We also use the predicted value from the regression even for the commodity-years where we have intraday data (rather than the actual value), again for consistency, but the very high in-sample performance of the regression means this has a negligible effect.

indexing strategies for thirteen agricultural markets. These reports are available from 2006 to the present (soybean meal data start in 2015). The second source is the CFTC Index-Investment Data (IID) that covers a broader range of commodities, but is only available quarterly or monthly from 2007 to 2015. We use both these sources to calculate index flows as changes in the net long positions of CITs. Due to the higher frequency and longer period, we focus on the SCOT data for our primary analysis. We do however find similar results in the non-agricultural commodities using IID data. In some of our tests we also use the CFTC's Disaggregated Commitments of Traders (DCOT) reports, which provide a breakdown by trader type, but do not specifically separate out index positions[20].

Several studies have examined the link between these reports and futures prices. Our main contribution is to bring intraday data to bear. In particular, we would expect that traders who are trying to match a published index will tend to focus their trading near the daily settlement in order to have their return closely track the calculated index return. In fact, this is what we find, and this new result is one of our contributions. More importantly, if the index trading tends to concentrate near the daily settlement, we can use these narrower time periods to avoid changes in price from other events occurring over the full day, and thus get more precise estimates for the impacts of index trading.

We use the trade and quote data sets described in the previous section to examine the link between trade imbalance and index trading for corn, soybean and soft wheat futures. We also employ intraday futures price and volume data measured at 5-minute intervals to examine price effects. We obtained these additional futures data from Barchart for all of the commodities in the SCOT and IID reports and they extend through 2020. [21]

[20]See Irwin and Sanders (2012) for a detailed description of the various reports.

[21]These data do not have quote information, so the returns are based on trade prices and therefore subject to bid-ask bounce (although the five-minute returns constructed from our quote based data sets for corn, soybeans and soft wheat generally have correlations around 0.99 with the five-minute trade-based returns).To select the contract month used to measure five-minute returns in our tests, we use the Lead or Active month where available. Where it is not available we use volume to identify the most active contract by moving to a later contract month whenever the day's volume for the later contract month is higher than the day's volume for the contract month currently in use.

3.1 Summary Data for Index Flows

Panel A of Table 4 shows the summary statistics for index flows (weekly changes in the net-long notional value of positions of index traders) for corn, soybeans, and soft wheat. [22] The statistics in Panel A are limited to the 2008-2014 time where we have intraday trade and quote data for these commodities. The index flows are quite substantial in magnitude. The standard deviations of weekly flows range from $309 million for soybeans to $139 million for soft wheat. For these commodities over this period, there is also a positive correlation among the index flows, averaging out to about 0.35. The remaining columns in Panel A report the standard deviations of weekly trade imbalances from the intraday trade and quote data. These columns show imbalances summed across the week and then summed over just the periods near the settlement.

Panel B shows all of the agricultural commodities included in the weekly SCOT reports in an expanded sample that spans from 2008 to 2020. Column (2) of Panel B uses the estimated price impact per dollar of imbalance in the minute before settlement that is calculated in the previous section, and multiplies it times the size of the index flow. Column (3) reports the average pairwise correlation of each commodity's predicted index impact with the other 12 commodities. While there is a positive correlation, the average pairwise correlation of the index flows is a modest 0.16. As a comparison, Irwin and Sanders (2012) report a pairwise correlation of 0.36 for the sample prior to 2011, and this agrees with our calculations. The remaining columns in Panel B report weekly standard deviations of returns for these commodities, for the full week and then summed across the periods in the week that are near the settlement.

As one might expect, the standard deviation of weekly returns is larger than the standard deviation of predicted price impacts from index flows, both because index traders may be able to reduce the impacts of their trades, and because there are many other sources return variation. However, the standard deviations of returns near the settlement are somewhat lower than the standard deviation of predicted flow impacts. The main message from Table 4 is that the index flows are substantial enough that they might cause significant price pressure, particularly if they are concentrating their trading near the close.

To illustrate this point, consider a back-of-the-envelope calculation using the last row of Table

[22]Throughout this section when calculating dollar values of flows, we multiply the change in contracts by the price per contract at the start of the week.

4. The ratio of the standard deviation of the predicted impacts to the standard deviation of weekly returns over the full day is ($\frac{0.92}{3.53} = 0.26$), implying an r-squared of approximately $0.26^2 = 7\%$ in a univariate regression of weekly returns on weekly index flows, and a similar calculation would imply an r-squared of 66% if index trades were fully concentrated in the last 30 minutes of the trading day. This suggests that even though these measures are likely an overestimate of true impacts, it is quite plausible that index traders would move prices and that this effect would be detectable, particularly if it is concentrated within the trading day. This basic reasoning motivates our subsequent tests for the linkage between the index flows and imbalances and/or returns, particularly around the daily futures settlement. We turn to that in the next subsection.

3.2 The Relation between Index Flows and Futures Markets

To test for the relation between index flows and various outcomes in the futures market, we rely on weekly regressions where various measures of weekly index flow are the independent variable, and the futures market outcome variables are the dependent variable. For most of our analysis, we first report regressions for individual commodities estimated as

$$FuturesOutcome_{c,t} = \alpha + \beta IndexFlow_{c,s} + \epsilon_{c,t} \tag{5}$$

Here $FuturesOutcome_{c,t}$ is either futures imbalance or return measured over various parts of the trading day for commodity c in week t. $IndexFlow$ is either the net change in index position in millions of dollars, or the predicted price impact, calculated as the net change in position multiplied by the predicted price impact calculated using the regression specification in Table 3. These variables are calculated in week s for commodity c. We estimate both contemporaneous regressions ($s = t$), and regressions at one lag ($s = t - 1$). We also estimate pooled time-series and cross-sectional regressions for all commodities using the specifications

$$FuturesOutcome_{c,t} = \alpha + \beta IndexFlow_{c,s} + \Gamma_c \mathbb{1}_c + \epsilon_{c,t} \tag{6}$$

$$FuturesOutcome_{c,t} = \alpha + \beta IndexFlow_{c,s} + \Gamma_c \mathbb{1}_c + \Gamma_t \mathbb{1}_t + \epsilon_{c,t} \tag{7}$$

Here $\mathbb{1}_c$ is a commodity fixed effect and $\mathbb{1}_t$ is a week fixed effect. To control for cross-sectional correlation we cluster standard errors within week for both specifications. The cross-sectional regression is broadly analogous to a Fama and MacBeth (1973) specification (see Cochrane (2009)), and allows us to partially control for correlation in index flows across commodities. While the correlation of index trader positions across the 13 commodities is relatively low (Table 4), we view this as the cleanest specification and therefore rely on it for much of our analysis.

As a first test, Table 5 presents direct tests of the linkage between index flows and trade imbalances for corn, soft wheat, and soybeans. The table shows the results of univariate regressions of imbalances summed across the week on the net flow for the week.

The first column of Table 5 shows that on average roughly half of the index flow for the week shows up in the imbalance. We would certainly expect these coefficients to be below one if index traders make some of their trades using floor trades, or use even a modest amount of passive orders. For example, if an index trader bought four contracts over the week, where three were purchased using marketable orders (classified as buys) and one was purchased using a limit order (classified as a sell), then the imbalance would be 3-1=2 contracts, so the imbalance for this trader in that week would be one half of the net flow.

The remaining columns in Table 5 show the imbalance near the end of the day is highly significantly related to the net index flow for the week. The coefficients are smaller than in the first column, which is to be expected if not all index trades are done in these time windows. The trading volume in the 30 minutes prior to settlement is approximately 10% of total volume for the week, so if index traders spread their trades across the trading day along with the rest of the trading volume, then we would expect the coefficients in the middle column to be approximately 10% of those in the first column. The pooled regression results in the final row of Table 5 indicate that the relation between flow and imbalance in the intervals near the settlement is stronger than what would be expected if the index traders simply matched daily volume. Instead it appears these traders concentrate their trading and/or trade more aggressively near the settlement.[23]

To help visualize this effect, Figure 3 plots the results from the univariate regressions where the dependent variable is imbalance measured in expanding windows across the trading day. As the left-hand panels show, the coefficient of futures imbalance on index flows rises as more of the

[23]In the internet appendix we show that this is result is primarily driven by large trades (See Table IA.7). While this is broadly consistent with the index traders being institutional traders (e.g. Kaniel, Saar, and Titman (2008)), it is unclear how much individual trading occurs in commodity futures markets.

trading day is included, and then there is a pronounced spike in the coefficient in the minutes just prior to settle. This is most pronounced for soft wheat, but is present in all three commodities. The right hand panels repeat this test but narrowing in on expanding windows in the 30 minutes prior to settlement. Here we see the increased statistical significance and again the pronounced jump in the coefficient when the minutes just prior to settlement are included.

Having established that index trader flows are linked to imbalances, and that this linkage seems to be the strongest near the settlement, we then turn to the linkage between flows and returns. Table 6 presents the main results of this section. The coefficients in the table are from univariate regressions where the independent variable is the estimated return impact of the index trader flow over the week. Note that by multiplying the index flow by the predicted price impact variable in the return regressions, we are effectively scaling the flow so that the coefficients are more easily interpreted. If the entire index trade imbalance is executed with marketable orders within a particular time window and if the average price impact of these trades is the same as for other trades, then one would expect a coefficient of 1.0. An additional benefit of this scaling is that it makes it more sensible to run a pooled regression across all of the commodities. The larger commodity markets have larger index flows and also have more liquid markets that can absorb the flows, so without the scaling a pooled regression could be highly mis-specified.

Though the predicted price impact provides an easily interpretable variable that is uniformly scaled across commodities, it does introduce a potential estimation error from the regression in Table 3. To address this, we report p-values from our regressions using both the asymptotic measures, which take the predicted price impacts as given, and using a bootstrap procedure to account for this estimation error.[24]

The dependent variable in the regression reported in columns (2) - (6) of Table 6 is the total return over the week. There is a significant linkage between index flows and returns for nearly all of the commodities in the weekly reports. The result for the pooled regression indicate that the price impact is 46% of what would be expected if the index traders used only aggressive orders with price impacts equal to the average estimated impacts from the previous section. As with the

[24]In the bootstrap procedure we construct 1000 simulated samples. In each sample, we construct a sample to estimate predicted impacts by drawing 41 commodity-year impact observations (the variables used in the regression in Table 3), as well as a sample of 643 weeks of index flows and futures returns. As in the primary regressions, this results in an unbalanced panel of observations. We then calculate the p-value as the percentage of these simulations with a positive coefficient.

imbalance results from 5, these results are consistent with index traders using passive strategies to some degree.[25]

In columns (7) - (11) of Table 6 the dependent variable is the return over the thirty minutes up to and including the settlement minute, and columns (12) - (16) the five minutes up to and including the settlement minute, again summed over the week. Again the results are significant for nearly all of the commodities. The pooled regression results in the final row of the table indicate that on average about 40% of the weekly impact of index flows occurs in the final 30 minutes and nearly 1/3 of the impact occurs in the final 5 minutes. As with the imbalance results from Table 5, these results are consistent with index traders concentrating and/or trading more aggressively near the settlement. In addition, the results in Table 6 indicate this trading had an impact on futures prices. We also note that the bootstrapped p-values across the specifications are very similar to the asymptotic p-values, and generally are larger, particularly for the results near the settlement. We therefore report asymptotic p-values results for the remainder of the CIT analysis.

Figure 4 visualizes the response of returns across different portions of the trading day. In this figure, we plot the slope coefficient on the predicted price impact from the pooled cross-sectional regression described in equation 7 and reported in the last row of Table 6. As in Figure 3, the figure shows an increasing response to the index flows as more of the day is included, and this response accelerates near the settlement, withe a large spike in the last five minutes. In this figure we also include 15 minutes after the settlement.[26]. Here we see that the positive response to the index flows in the last five minutes of the day is followed by a partial subsequent reversal. Such a reversal is consistent with several, but not all, models of financialization. For instance, in models such as Goldstein and Yang (2017), you may see reversal associated with impacts from index flows created by incomplete information. In contrast, Baker (2021) and Hamilton and Wu (2014) predict essentially permanent effects created by changes in expected returns and limited risk-bearing capacity. In addition, microstructure inventory effects, such as those described by Stoll (1978) and

[25] We note that the significant relation between index flows and aggregated daily returns is inconsistent with prior literature. In the following sections, we show that this is mainly due to a strong relation between futures returns and index flows in the most recent portions of our sample. We also show that the result in the periods prior to the settle is not robust to the inclusion of other types of trader flows.

[26] Globex trading of agricultural futures stops at the settlement time, but the CME continues to report over-the-counter trades through ClearPort, so price data are available from the period after settlement. We do not include this period in the imbalance analysis, as there is almost zero volume in this period for corn, wheat, and soybeans prior to 2015

documented by Hendershott and Menkveld (2014) in equity markets, might explain reversals at higher frequency. Therefore we now examine the data for evidence of reversals, both in the minutes following the settlement and in the following week.

3.3 Reversals

The results presented thus far suggest index trading causes price pressure near the settlement, and the pattern shown after the settlement in 4 suggests that some of this impact may be temporary. In order to directly assess the tendency for reversals, Table 7 reports the results of regressions where the independent variable is again the predicted price impact of the index trader flows for the week and the dependent variable is a return measured over an interval where one might expect to see reversal of the price pressure from the index trading leading up to the settlement. In the first two sets of regressions in Table 7 the return is the weekly sum over the five and 30-minute intervals immediately following the settlement and independent variable is the predicted price impact from the contemporaneous index flow. In the remaining columns of Table 7 the return is measured over one, two and all days of the subsequent week. Given the returns are measured using trade prices, we would expect some price reversals due to bid-ask bounce, so all of the regressions include the return over the five minutes leading up to the settlement as an additional control variable. For the first two columns, this is the weekly sum of the five minute returns leading up to the daily settlements, and for the remaining columns it is the five minute return leading up to the last settlement for the week. [27]

As the table shows, even controlling for the returns in the five minutes leading up to the futures settlement, there is clear evidence of reversal in the minutes after the settle associated with index flows. All but one of the commodities have a negative coefficient for the 5-minutes after settle, and the pooled regressions are highly significant. Looking at the remaining columns, we also see evidence for reversal in the first part of the subsequent week, as well as for the whole week. These results, when combined with the reversal just after the settle, are roughly equal in magnitude to the price impacts we see in the last 30-minutes of the day. This again is consistent with theories of financialization that predict a temporary impact. It is also similar to the findings of Hendershott

[27] As expected, the coefficients on these control variables (not reported) are negative and significant, but they have little impact on the index flow coefficients.

and Menkveld (2014) who document temporary price pressure related to market-maker inventories in equity markets. It is interesting to note here that the report specifying the index trader position for a Tuesday is released on Friday afternoon after most futures market settle. Therefore the next week reversal is occurring primarily before this release, suggesting it is either due to inventory effects, or there is another mechanism for information on index flows to be incorporated into prices.

3.4 Subperiod Analysis

The first row of Panel A in Table 8 repeats the results for the pooled cross-sectional regressions reported at the bottom of Table 6, and then the remaining rows show the results of these pooled regressions by subperiod. The results for the periods prior to the settlement are remarkably consistent across the various subperiods. The results for the full week are less consistent and are much weaker from 2011 to 2014. his finding helps explain the difference between our full-day return results and the findings of previous papers that conduct similar tests and find no relation using daily data. Most of these papers focus on individual commodities, and use data from the first half of our sample, or in the two years prior which we do not include due to the lack of intraday data prior to 2008.[28] Since the full day regressions are much weaker in this period, it is perhaps unsurprising that no result was found in earlier papers. However, the subperiod results in Panel A of Table 8 show a strong relation between index flows and price movements prior to the futures settlement in this earlier period.[29]

Panel B repeats the subperiod analysis for the reversal regressions shown in Table 7. Here we see very little post-settlement reversal prior to 2015. This may be due to the fact that volumes were extremely small in this period. Looking at the reversals in the following week, both in the first day of the trading week and the full week, we see that, though these results are weaker, they are largely consistent across the sample.

[28] One exception is Irwin and Sanders (2012) who run cross-sectional pooled regressions using the index flows from the IID data. We confirm their primary results in section 3.6.

[29] In the internet appendix we provide evidence that the full day results shown in the later time periods in Panel A of Table 8 are driven by an increase in momentum trading by at least some traders classified as index traders. See section IA.8.

3.5 Controlling for Other Flows

The significant link between index flows and returns is consistent with price pressure from index traders, but it could also partially arise if some of these traders following positive momentum strategies. In order to shed more light on this issue, we examine the impact of index flows on returns after controlling for measures of flows from other types of traders. The DCOT reports provide the positions of four reported categories of traders: Managed Money, Producer/Merchant/Processor/User, Swap Dealers, and Other. It is the case that swap dealers positions most closely track our index investment positions, but index flows are also positively correlated with changes in the positions of Managed Money traders, while being strongly negatively correlated with the Producer category.[30] Given that Kang, Rouwenhorst, and Tang (2020) document that Managed Money traders tend to behave as momentum traders, and Robe and Roberts (2019) show that some of these traders may be classified as index traders, we feel that they are the most important control.[31]

Columns (1), (4), and (7) in Table 9 repeat the pooled sample results from the last line of 6. The remaining columns of Table 9 show the results of the pooled sample regressions where the flow from Managed Money, Producer/Merchant/Processor/User, or Swap Dealers included as a control (as in the other regressions the flows for each commodity are scaled by the predicted return impacts for that commodity). The results in column (2) show that weeks with rising return tend to be weeks where Managed Money traders are buyers, and the results in column (3) show that these weeks tend to be ones where those in the Producer category are net sellers. Including these controls in the full day regression has a large impact on the coefficient associated with index flows, and in the case of the Producer flows actually flips the sign. While these results do not necessarily prove that the relation between index investors and contemporaneous returns is being driven by an omitted variable, they certainly indicate that the full day results need to be viewed with caution.

In stark contrast to the full day results, the coefficients on index trader flows in the periods near the settlement (columns (6) - (8) and (10) - (12)) are essentially unchanged when including the other flows. On balance, these results suggest that the full week results might be explained by positive momentum strategies of some of the market participants classified as index traders, but this

[30]See Irwin and Sanders (2011) for a detailed description of how the various trader categories in the DCOT reports relate to index trader positions.

[31]See section IA.8 in the internet appendix for more evidence on the relations between index trading, Managed Money flows, and momentum strategies over the two halves of our sample period.

explanation seems much less reasonable for the results using the time windows near the settlement. Instead, it appears more likely that index traders are timing their trades near the settlement to control basis risk and reduce price impact, and that their trading results in price pressure in the futures markets.

Figure 5 visualizes these regressions across the trading day when the flows from Managed Money traders are included as a control. As the figure shows, after controlling for Managed Money flows, there is essentially no association between returns and index flows early in the trading day, but as in the prior analysis, there is the pronounced spike and partial reversal near the daily futures settlement.

3.6 Monthly Index-Investment Data for a Broader Set of Commodities

In addition to the weekly reports of index trader positions in agricultural commodity futures, the CFTC collected monthly futures positions of index traders from July 2010 through October 2015 (64 months) for the same agricultural commodities in the SCOT report, plus gold, silver, copper and four energy contracts. These data are also reported quarterly from 2007 - 2010, but we focus on the period with monthly observations to get a consistent panel. Table 10 uses a similar approach to Table 6 and reports the results of monthly regressions of returns on the predicted price impacts from index flows. The right hand set of regressions follows Table 7 and reports the relation between index flows and returns in the five minutes after futures settlement. The sample size is much smaller than in Table 6, but the monthly results for the agricultural commodities in Table 10 are generally consistent with the results in Table 6. The additional commodities in Table 10 also show a strong linkage between index flows and returns in the 30 minutes prior to settlement, and a subsequent partial reversal in the minutes after settlement which suggests index trading concentrated near the settlement in these commodities as well and that this trading impacted prices. What is striking for these data is that even though we see large impacts near the settle, the full day pooled sample results are not statistically significant, a finding consistent with earlier literature.[32]

[32] For some smaller market energy commodities, the coefficient at the end of the day is greater than one. In unreported results, we find that controlling for crude oil index flows, which are highly correlated with these flows and much larger, negates the result. This suggests that index flows to the WTI may be driving the returns to other energy contracts.

4 Impacts of Commodity-Linked Notes

Having documented that index flows are large relative to commodity futures markets, and having found that they do appear to create a price impact concentrated near the daily futures settlement, we now turn to an analysis of financial trader flows associated with commodity-linked notes.

To construct our sample of CLNs, we follow the procedure of HPW and search the universe of 424b filings for issuers of CLNs from the SEC's Edgar website to identify CLNs linked to a single commodity with face value of at least $2 million.[33] We find 597 notes, of which approximately 75% are in the commodities for which we have intraday trade-and-quote data. Our sample of notes appears to closely track the set captured by HPW in terms of number and size.[34]

HPW argue that the pricing dates of these notes are set in advance, which would imply that the positive average return that they find on the pricing day is caused by the issuance, and presumably results from the hedging trades made by the issuers. On the other hand, if pricing dates are not always set in advance, then it is possible that at least some of the notes are purposely issued on days with rising prices. To test this, we collect the Free Writing Prospectuses (FWPs) associated with the notes in the sample. These FWPs are required by the SEC when specific information about the issue is shared with retail investors. We find that slightly more than half of the notes in the sample have FWPs that are filed in advance and specify the ultimate pricing date, so in these cases it is clear that the HPW assumption is correct. However, for about 25% of the notes, the FWP does not specify a pricing date, and for about 4% of the notes the FWP is filed on or after the pricing date.[35] Finally for the remainder of the notes it appears there was no FWP filed. Figure 6 illustrates the four cases. To allow for the possibility that some notes may be issued in response to the price activity on the pricing date, we separate the sample into CLNs where the FWP was filed in advance with an indication of the expected pricing date (Prefiled Pricing Date),[36] and all

[33] HPW also include notes linked to multi-commodity indices if all the indices are in the same sector (i.e. energy). We collected these notes but do not include them because they complicate some portions of the analysis. None of these notes have more than $10 million linked to a single commodity, and we mainly focus on the comparison to the HPW results for notes with face values greater than $10 million.

[34] We find some notes that were created and then transferred to a subsidiary for later sale to investors. We do not include these notes in our sample. We also exclude exchange-traded notes following HPW.

[35] For FWPs that are filed on the pricing day we consider FWPs that are received by the SEC after 4 p.m. on the pricing day to be late, and those that are received prior to 4 p.m. on the pricing day to be prefiled. We do this to account for some time in preparing the filing. This decision only effects a small number of notes and therefore has minimal impact on any results.

[36] We found a small number of notes where the expected pricing date was different than the ultimate date the note

others (No Prefiled Pricing Date). Of course, it is possible that the pricing date was determined in advance in our No Prefiled Pricing Date subsample. If this is the case, then we should see no significant difference between the average pricing-date returns across the two categories.

When conducting their analysis, HPW exclude notes pricing during the Goldman Roll documented by Mou (2010). This period includes the 5th to 9th trading days of the month and the five previous business days. They do this to avoid potential returns coming from the price pressure associated with the roll trades. However, it is not clear that this is the correct choice. In particular, as documented by both Mou (2010) and Neuhierl and Thompson (2016), the predictable returns associated with the roll trades had disappeared by 2007. Moreover, in unreported results, we found no difference in the price impacts of order flow in the Goldman Roll period. Accordingly, we would not expect to see differences in the price impacts of hedging trades for notes issued during this period, and thus we report results for both notes issued outside of the Goldman Roll, and the full sample of notes.

4.1 Summary Data for Commodity-Linked Notes

We extend the analysis of HPW by using standard option pricing techniques to calculate the hypothetical notional value of the hedge for each note, both at initial issue and on the date when the final payoff is determined (the determination date). Many of the CLNs have complicated features (e.g. call provisions, caps, floors, knock-outs, and buffer regions). In order to accommodate the various structures, we calculate the required delta hedge size of the note via Monte Carlo valuation with 10,000 sample paths of daily returns over the life of the note. [37] We refer to the ratio of this notional hedge value to the face value of the note as delta. Combined with price impact estimates from above, the size of the hedging trade allows us to estimate the price impact from the associated hedging trades.[38]

---

priced. However these cases were appear to be due to the expected pricing date being specified on a trading holiday, and the pricing date was then the date following.

[37]We assume that returns are log-normally distributed with daily standard deviation equal to the realized daily standard deviation of the underlying over the month prior to issuance. The simulated risk-neutral drift of the underlying depends on the type of index used to calculate the note payoff. In some cases, the underlying is an excess return index, so the risk-neutral proportional expected return is zero. If the note uses a total return index, then the risk-neutral expected return equals the risk free rate. In many of the notes the underlying is a spot price, in which case we set the drift so that the expected value on the determination date is the futures price for the contract whose maturity is closest to the determination date.

[38]See the internet appendix section IA.9 for an example of this calculation for a sample CLN with a knock-out provision.

Panel A of Table 11 shows summary statistics for sample of CLNs, focusing on the pricing date. We follow HPW and combine notes in the same commodity with the same pricing date, and report face value, the size of the delta hedging trades, and the predicted price impacts of these trades. Here, in contrast to the changes in positions of index traders shown in Table 4, the notes are very small relative to the size of the futures markets. The average size of the delta hedging trades is approximately $11 million. This yields an average predicted price impact of approximately four basis points if the note was hedged near the futures settlement.

Panel B of Table 11 reports summary statistics for the CLN determination date. We start with the sample of notes issued prior to February 2014 with determination dates prior to January 2019. About 10% of the notes in the sample have multiple determination dates, and we follow HPW and exclude these from the determination date analysis. The next line removes the many notes that have zero delta at the determination date. This occurs either because the note is retired early or because the underlying commodity price is in a region in which the note has no exposure to the underlying commodity price. Then we remove the smaller notes to focus on the larger notes where HPW find their significant determination date results. Finally we restrict the sample to determination dates prior to February 2014 to obtain the notes that were available at the time of the original HPW study. While the predicted price impacts are again small, they are larger in magnitude than the pricing date impacts as the deltas tend to be higher for those notes which still have exposure to the underlying at the determination date.

4.2 Pricing Date Returns for Commodity-Linked Notes

HPW find that days with the pricing of CLNs have significant positive average returns. They attribute this to the price impact of hedging trades made in the futures market. Because the exposure to the underlying commodity starts at a particular time on the pricing date (often the daily settlement, but usually the London afternoon Fix for metals), that is where we would expect to see the hedging trades and their associated price impact.[39]

Table 12 shows the average realized pricing date returns and the predicted average impacts for different subsets of notes. As in HPW, we aggregate notes priced on the same day in the

[39] HPW also report results for returns on days after the pricing date and results using abnormal returns controlling for various systematic variables. Since their results are similar using raw returns, and are the most significant on the actual pricing date, we focus on these specifications for parsimony.

same commodity. Panel A includes all notes. Column (1) shows an average pricing date return of 28 basis points for all notes priced outside of the Goldman Roll. This is very similar the to the average return of 29 basis points reported by HPW in their Figure 2. Columns (2) and (3) show the separate results based on whether the FWP was issued in advance and specified the pricing date. Comparing these columns shows that the average positive pricing date returns are entirely due to the subsample with no prefiled pricing date. Columns (4) through (6) show the results for all notes including those with pricing dates during the Goldman Roll.[40] For this larger set, the average pricing date return is only marginally statistically significant, but the average return for the notes with no prefiled pricing date is still strongly significant. The table reports the average of the predicted impacts from the hedging trades across all of the observations, as well as a bootstrapped standard error of this average based on the estimation error from the price impact regressions.[41] Comparing the average predicted impact to the average realized return in column (1) suggests that the HPW result is much too large to be explained by price pressure. The small standard errors on this predicted price measure suggest that this difference is unlikely to be the result of estimation error.

Panel B repeats the analysis from Panel A, but restricts the sample to those commodity-years where the expected impact of a 1 million dollar trade is less than 1 basis point. This eliminates about 10% of the observations. The results in Panel B are very similar to those in Panel A, indicating the results are not driven by the observations in the least liquid markets. Here the difference between the predicted price impacts and the observed returns is even more stark. For instance, for notes issued without prefiled pricing dates in the Goldman Roll, the predicted price impact is 3.9 basis points, while the observed average returns are 43 basis points, more than an order of magnitude larger.

The results in Table 12 suggest that the pricing day results reported by HPW are driven by the observations where the pricing date was not set in advance. However, the table also shows that the No Prefiled Date category has somewhat larger predicted price impacts. In order to confirm

[40]In unreported results we found similar point estimates and no significant differences in average returns for notes with an FWP and no pricing date, no FWP, or a late FWP (Cases 2, 3, and 4 of Figure 6). We also find no evidence of positive returns on the dates with the filings of FWPs for the notes with prefiled pricing dates.

[41]In this bootstrap we create 1000 samples where we re-estimate the regression in Table 3 by drawing 41 commodity-years with replacement. We do not resample the notes, so this standard error can be interpreted as the error on the predicted price impact for this particular set of CLNs.

that the results are not driven by a correlation between the FWP-based classification and note size or predicted impact, Table 13 reports regressions where the return on the commodity-day is the independent variable and the dependent variables include a dummy variable set to one for the No Prefiled Date category along with other potential explanatory variables for the impact of hedging trades. Note that Table 12 shows a difference of 48 basis points between the Prefiled Date and No Prefiled Date categories. Table 13 shows that the coefficient on the dummy variable is consistently very close to this value across all of the regressions. Column (1) of Table 13 includes just the dummy variable, so the coefficient equals the difference between the two category averages in Table 12 by construction. Columns (2) through (4) show that the impact of the dummy variable is largely unaffected by inclusion of the size or estimated price impact of the CLN. Column (5) includes the face value along with issuer, commodity market, and year fixed effects, and shows that the result is not driven by a tendency of some issuers to fail to prefile the pricing dates. Finally, columns (6) through (10) repeat the analysis including the observations during the Goldman Roll and find similar results.

Together, these results strongly suggest that the hedging trades for these notes are not the cause of the positive return, and instead it is the demand or supply of notes with flexible pricing days responding to the changing price of the underlying commodity. The precise reason for this uncertain, but we think there are plausible explanations. One is simply marketing, as it may be easier for a broker to elicit interest from a retail investor on a day where prices are rising. It is also possible that unsophisticated retail investors misunderstand the payoffs of the note (e.g. Egan (2018)), and mistakenly believe that the positive return will improve the performance of the note. In any case, the fact that the positive average returns occur only in the set of notes where the notes have flexible dates strongly suggest a potential bias driving the pricing day results.

HPW acknowledge this bias, and address this by examining returns on the determination dates when final payoff of the CLN is set. This date is specified when the note is issued, so these results are not subject to the endogeneity concerns related to the pricing date analysis. In the next section we examine these results.

4.3 Determination Date Returns for Commodity-Linked Notes

Table 14 shows our analysis of futures returns on days with CLN determination. Following HPW, we restrict our analysis to notes which still have positive exposure to the underlying commodity on the determination date. We exclude notes where there are multiple determination dates. The table shows average determination date returns for various subsets of the sample. In particular, column (1) of Panel A is our attempt to replicate the main finding of HPW. This is the return on determination dates of notes that are outside of the Goldman Roll, have at least $10 million of face value, and have a maturity prior to February of 2014. As in HPW we combine notes with the same underlying commodity and the same determination date, but in spite of attempting to match their approach exactly, our sample size is larger (we initially had 50, and after consultation with HPW included four notes we had missed, giving us 54 as opposed to their sample of 42). More notably, while they report a significant average return of -42 basis points (t-stat of 2.50), we find an insignificant average return of only -10 basis points (t-stat of 0.49). Looking at the table, we do not find a negative average return that is significant for additional samples of the determination dates, either including the Goldman Roll or extending the sample of determination dates through 2018.

After seeing our initial results, HPW re-examined their set of 42 notes, and found that they had mistakenly included 10 notes and mistakenly excluded 24 notes.[42] They provided a refined sample of 56 notes in which they find a determination day return of -15 basis points and a t-stat of (0.76).[43] However, despite finding no significant negative return on the day after the determination date in the published paper, their new set has a return of -51 basis points (t-stat 1.78) in the two-day window starting at the determination date. It is not clear why hedgers would wait until the futures trading day after the determination date to unwind their hedge. This is especially true because 36 of the 54 commodity/days in column (1) are based on metals and have a payoff linked to the London 3:00 p.m. Fix price, which is established in the morning of the futures trading day in New York. Even if the hedges are not unwound exactly when the note is priced, it seems unlikely that hedgers would wait until the next morning to unwind, particularly since the notes tend to be quite small relative to the typical futures volume. HPW also propose a change of methodology from

[42] See Henderson, Pearson, and Wang (2019).

[43] See internet appendix section IA.10 for a comparison of our set of 54 notes and their set of 56 notes.

their published paper, namely the exclusion of determination dates coinciding with the pricing of other notes, that again changes the average returns on the determination date to -33 basis points (t-stat of 1.92), and a two-day return of -53 basis points (t-stat of 2.46). They did not provide the list of notes pricing on the indicated days, and we were unable to replicate this result, so we do not include a formal evaluation of this analysis. At the very least, the sensitivity of this result to small choices about the empirical specification suggests that this result is much less robust than the positive pricing day returns.

4.4 Intraday Patterns on CLN Pricing and Determination Dates

Our final set of tests focuses on the notes for which we have intraday data, and looks within the trading day to see if we observe any patterns similar to those we see in the analysis of commodity-index traders. In order to maximize the power of the tests we include observations from the Goldman Roll periods.

Table 15 shows the results. Panel A considers pricing days for notes with No Prefiled Pricing Date and column (1) shows the full day return. This value of 35 basis points differs from the 38 basis point value in column (6) of Table 12 because the sample in Table 15 is restricted to the commodities where we have intraday data. Column (2) of Panel A shows how much of this return accumulates before the period beginning 30 minutes prior to the pricing of the note. For Brent, WTI, and corn, the notes all price at the close of the associated futures market. However for gold, most of the notes price at the London 3:00 p.m. Fix, which is 10:00 or 11:00 a.m. in New York depending on the time of year. Unlike the patterns we see with index traders, we find that most of the return in this No Prefiled Date sample has accumulated before the note pricing. When we look at the windows 30-, 15-, and 5-minutes prior to the pricing in columns (3) - (5) of Panel A, we do see moderate positive returns concentrated just prior to the pricing of these notes, and the magnitude of these returns is roughly consistent with our impact estimates.[44]

In columns (6) - (8) we test for imbalance near the pricing time. These columns report the average abnormal imbalance during each time window with a CLN pricing, where the abnormal imbalance is calculated using a regression that controls for the average imbalance for each com-

[44] Columns (2) and (3) do not sum up to column (1) due to the early pricing of gold notes, and the number of observations is larger in columns (3) - (5) due to dates with multiple pricings at different times.

modity in that particular time window across all of the days in our sample period.[45] As columns (6) - (8) in Panel A show, we find no evidence of positive abnormal imbalance prior to 5 minutes before the pricing, but we do see a significant abnormal imbalance in the five minutes prior to the price for these notes. This coefficient is also broadly consistent in magnitude with the predicted sizes of the hedging trades.

Panel B repeats this analysis for all notes in the sample, including those with prefiled dates. Here, though there is no full day effect, we still find average positive return and imbalance near the pricing minute. Again the magnitude of the average imbalance is consistent with the predicted hedging trades. Panel C repeats this analysis again but using the full set of determination days. Here, while we don't find a significant return effect, we do find a modestly significant negative abnormal imbalance of $17.8 million in the 15 minutes prior to the determination of the final payoff.

Interestingly, in unreported results we do not find stronger imbalance effects in the minutes near pricing for larger notes, suggesting that larger notes may be more likely to be traded on the floor or earlier in the day. However, we view these intraday findings as consistent with the notes being hedged and creating a modest impact.

5 Conclusion

In this paper we construct order flow imbalances for seven major commodity futures markets. We find that order flows in these futures markets have a large explanatory power for prices. We also document substantial intraday variation in price impacts, with high volumes and low price impact around futures settlements. We use our findings on price impacts to examine the potential impacts of financial investors in this market.

Our main contribution is the examination of the impacts of changes in the positions of commodity-index investors. Prior empirical research on financialization has generally found little or no price impact from these flows. However, consistent with theoretical models of financialization, we find strong evidence that changes in index investor positions are associated with intraday order flow imbalances and price impacts in the futures markets that are concentrated in the minutes leading

[45]See section IA.11 for a detailed description of these regressions.

up to the daily futures settlement. These impacts appear to reverse partially in the minutes after the settle, and we find some evidence that they reverse further in the following week.

In our second set of tests we examine commodity-linked notes following Henderson et al. (2015). We find that the positive returns associated with the issuance of these notes documented by Henderson et al. (2015) are surprisingly large given the notes' size, occur primarily prior to the pricing of the notes, and are only present in notes without a preliminary filing specifying the expected pricing date. In contrast to Henderson et al. (2015), we find no evidence of significant negative returns on CLN determination dates. These findings suggest that the positive returns are potentially the result of CLN issuers or purchasers favoring days with increasing commodity prices, rather than evidence of impacts from associated hedging trades.

Taken together, our results highlight the usefulness of using intraday data to understand the impacts from non-fundamental traders in financial markets, even when the data on their positions is only available at lower frequencies. While our results to provide new evidence in support of commodity financialization, they are also reassuring for policy makers, as they suggest that while large flows from non-fundamental traders do have some impact on commodity futures markets, they do not result in large permanent impacts on commodity prices.

References

Acharya, V. V., L. A. Lochstoer, and T. Ramadorai. 2013. Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics* 109:441–465.

Admati, A. R., and P. Pfleiderer. 1988. A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies* 1:3–40.

Alquist, R., and O. Gervais. 2013. The role of financial speculation in driving the price of crude oil. *The Energy Journal* pp. 35–54.

Baker, S. D. 2021. The financialization of storable commodities. *Management Science* 67:471–499.

Basak, S., and A. Pavlova. 2016. A model of financialization of commodities. *The Journal of Finance* 71:1511–1556.

Bessembinder, H. 2015. Predictable ETF Order Flow and Market Quality. *Available at SSRN 2506322* .

Bessembinder, H., A. Carrion, L. Tuttle, and K. Venkataraman. 2016. Liquidity, resiliency and market quality around predictable trades: Theory and evidence. *Journal of Financial Economics* 121:142–166.

Boyd, N. E., J. H. Harris, and B. Li. 2018. An update on speculation and financialization in commodity markets. *Journal of Commodity Markets* 10:91–104.

Buyuksahin, B., and M. A. Robe. 2011. Does paper oil matter? Energy markets' financialization and equity-commodity co-movements. *Available at SSRN 1855264* .

Chari, V. V., and L. Christiano. 2017. Financialization in commodity markets. Tech. rep., National Bureau of Economic Research.

Cheng, I.-H., A. Kirilenko, and W. Xiong. 2014. Convective risk flows in commodity futures markets. *Review of Finance* 19:1733–1781.

Cheng, I.-H., and W. Xiong. 2014. Financialization of commodity markets. *Annu. Rev. Financ. Econ.* 6:419–441.

Cochrane, J. H. 2009. *Asset pricing: Revised edition*. Princeton university press.

Egan, M. 2018. Brokers vs. retail investors: Conflicting interests and dominated products. *Journal of Finance, Forthcoming* .

Elder, J., H. Miao, and S. Ramchander. 2014. Price discovery in crude oil futures. *Energy Economics* 46:S18–S27.

Evans, M. D., and R. K. Lyons. 2002. Order flow and exchange rate dynamics. *Journal of political economy* 110:170–180.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of political economy* 81:607–636.

Fattouh, B., L. Kilian, and L. Mahadeva. 2013. The role of speculation in oil markets: What have we learned so far? *The Energy Journal* pp. 7–33.

Goldstein, I., and L. Yang. 2017. Commodity financialization and information transmission. In *AFA 2016 Annual Meeting*.

Grossman, S. J., and J. E. Stiglitz. 1980. On the impossibility of informationally efficient markets. *The American economic review* 70:393–408.

Halova, M. W., A. Kurov, and O. Kucher. 2014. Noisy inventory announcements and energy prices. *Journal of Futures Markets* 34:911–933.

Hamilton, J. D., and J. C. Wu. 2014. Risk premia in crude oil futures prices. *Journal of International Money and Finance* 42:9–37.

Hamilton, J. D., and J. C. Wu. 2015. Effects of index-fund investing on commodity futures prices. *International economic review* 56:187–205.

Hasbrouck, J. 1991. The summary informativeness of stock trades: An econometric analysis. *The Review of Financial Studies* 4:571–595.

Hendershott, T., and A. J. Menkveld. 2014. Price pressures. *Journal of Financial Economics* 114:405–423.

Henderson, B. J., N. D. Pearson, and L. Wang. 2015. New evidence on the financialization of commodity markets. *The Review of Financial Studies* 28:1285–1311.

Henderson, B. J., N. D. Pearson, and L. Wang. 2019. Response to "Order Flows and Financial Investor Impacts in Commodity Futures Markets". *Available at SSRN 3475065* .

Irwin, S. H., and D. R. Sanders. 2011. Index funds, financialization, and commodity futures markets. *Applied Economic Perspectives and Policy* 33:1–31.

Irwin, S. H., and D. R. Sanders. 2012. Testing the Masters Hypothesis in commodity futures markets. *Energy economics* 34:256–269.

Kang, W., K. G. Rouwenhorst, and K. Tang. 2020. A tale of two premiums: the role of hedgers and speculators in commodity futures markets. *The Journal of Finance* 75:377–417.

Kaniel, R., G. Saar, and S. Titman. 2008. Individual investor trading and stock returns. *The Journal of finance* 63:273–310.

Kyle, A. S. 1985. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society* pp. 1315–1335.

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

Marshall, B. R., N. H. Nguyen, and N. Visaltanachoti. 2011. Commodity liquidity measurement and transaction costs. *The Review of Financial Studies* 25:599–638.

Mou, Y. 2010. Limits to arbitrage and commodity index investment: front-running the Goldman roll. *Available at SSRN 1716841* .

Neuhierl, A., and A. Thompson. 2016. Liquidity Timing in Commodity Markets and the Impact of Financialization. *Available at SSRN 2682698* .

O'Hara, M. 2015. High frequency market microstructure. *Journal of Financial Economics* 116:257–270.

Raman, V., M. A. Robe, and P. K. Yadav. 2017. The Third Dimension of Financialization: Electronification, Intraday Institutional Trading, and Commodity Market Quality. *WBS Finance Group Research Paper* .

Robe, M. A., and J. S. Roberts. 2019. Who Holds Positions in Agricultural Futures Markets. *Available at SSRN 3438627* .

Scholes, M. S. 1972. The market for securities: Substitution versus price pressure and the effects of information on share prices. *The Journal of Business* 45:179–211.

Silvennoinen, A., and S. Thorp. 2013. Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money* 24:42–65.

Singleton, K. J. 2013. Investor flows and the 2008 boom/bust in oil prices. *Management Science* 60:300–318.

Sockin, M., and W. Xiong. 2015. Informational frictions and commodity markets. *The Journal of Finance* 70:2063–2098.

Stoll, H. R. 1978. The supply of dealer services in securities markets. *The Journal of Finance* 33:1133–1151.

Stoll, H. R., and R. E. Whaley. 2010. Commodity index investing and commodity futures prices. *Journal of Applied Finance (Formerly Financial Practice and Education)* 20.

Tang, K., and W. Xiong. 2012. Index investment and the financialization of commodities. *Financial Analysts Journal* 68:54–74.

Yan, L., S. H. Irwin, and D. R. Sanders. 2021. Sunshine vs. predatory trading effects in commodity futures markets: New evidence from index rebalancing. *Journal of Commodity Markets* p. 100195.

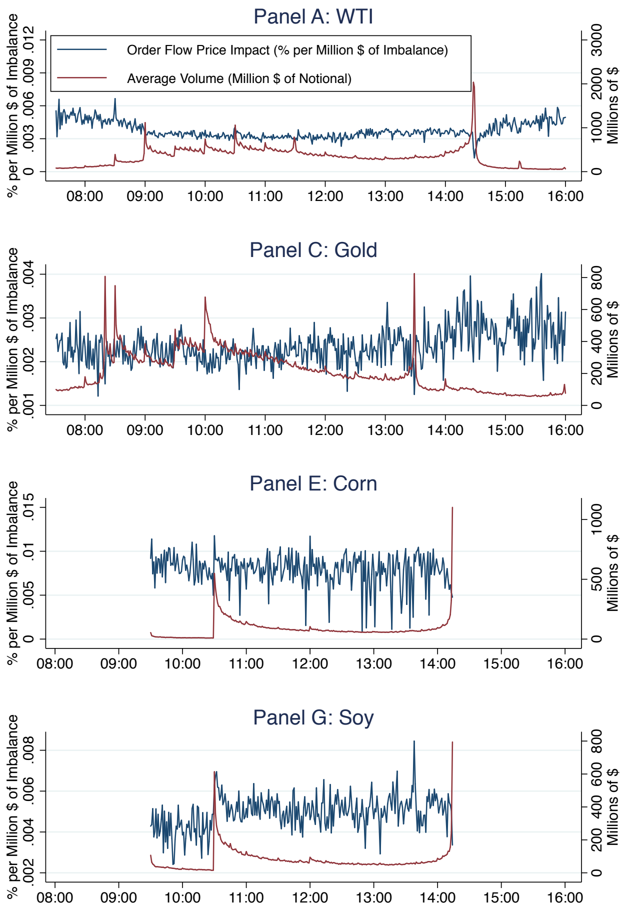


Figure 1. Volume and Price Impact of Order Flow Across the Trading Day
The figure shows the average intraday volume (in red) by minute for each commodity as well as the minute-by-minute price impact (in blue). The price impact is measured as the slope in a univariate regression of return (%) on order flow imbalance (millions of $) estimated using imbalance and returns in each minute of the day. For instance, for the 12:00 average volume we calculate the total volume from 12:00:00 to 12:00:59 for each day, and take the average of this value across all trading days. Similarly, to calculate the 12:00 imbalance, we calculate the return and imbalance from 12:00:00 to 12:00:59 for each day, and then run a univariate regression of return on imbalance for this minute across all trading days. The sample period is January 2007 through March 2014, except that Brent data begin in January 2008 and Soy data begin in January 2009. Corn, Wheat and Soy data exclude the period from 5/22/2012 to 4/5/2013 when settlement was delayed to 15:00 eastern time.

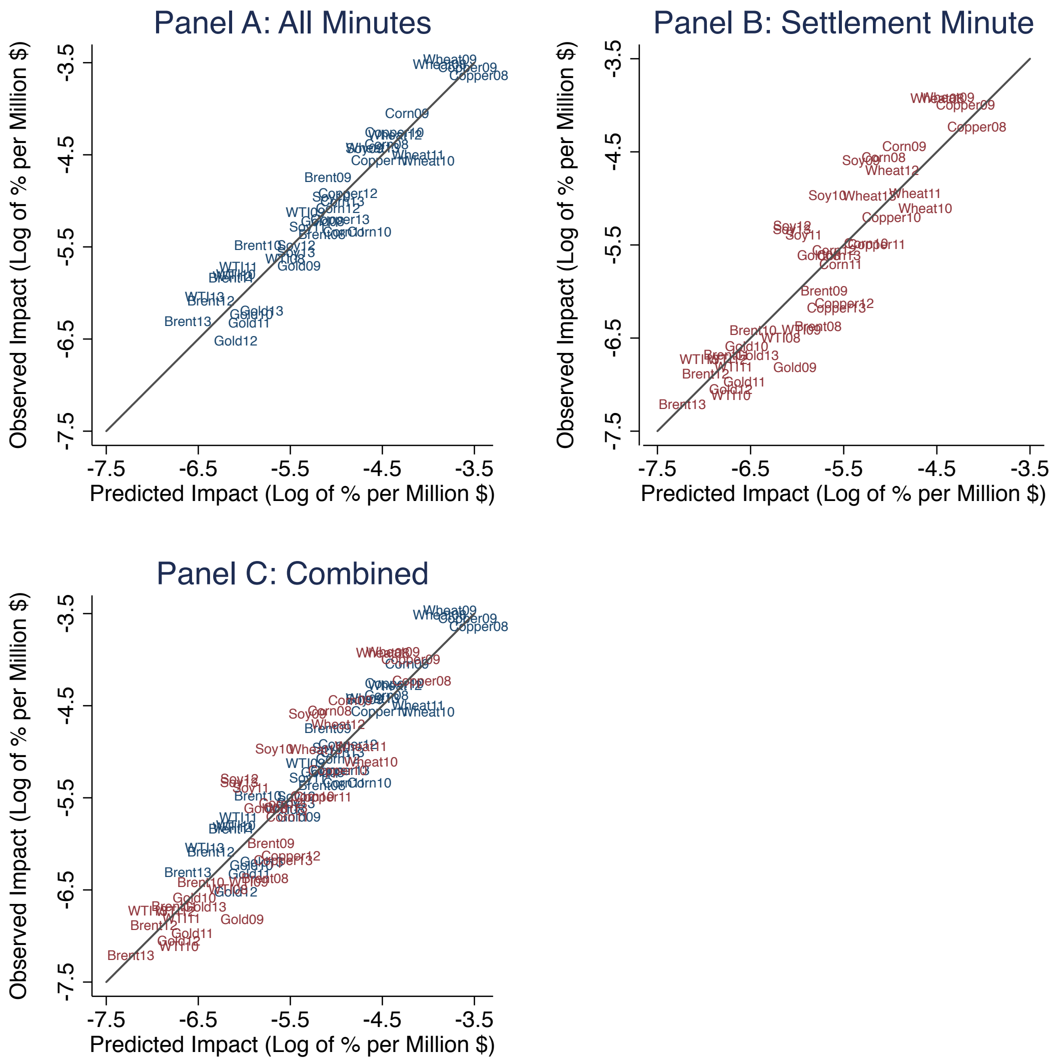


Figure 2. Regression fit for inferring order flow price impact from daily data
This figure plots the fit for regressions of the log of predicted price impacts for each commodity in a calendar year on the logs of average daily futures volume and daily futures volatility (See columns (1) - (3) of Panel A in Table 3).

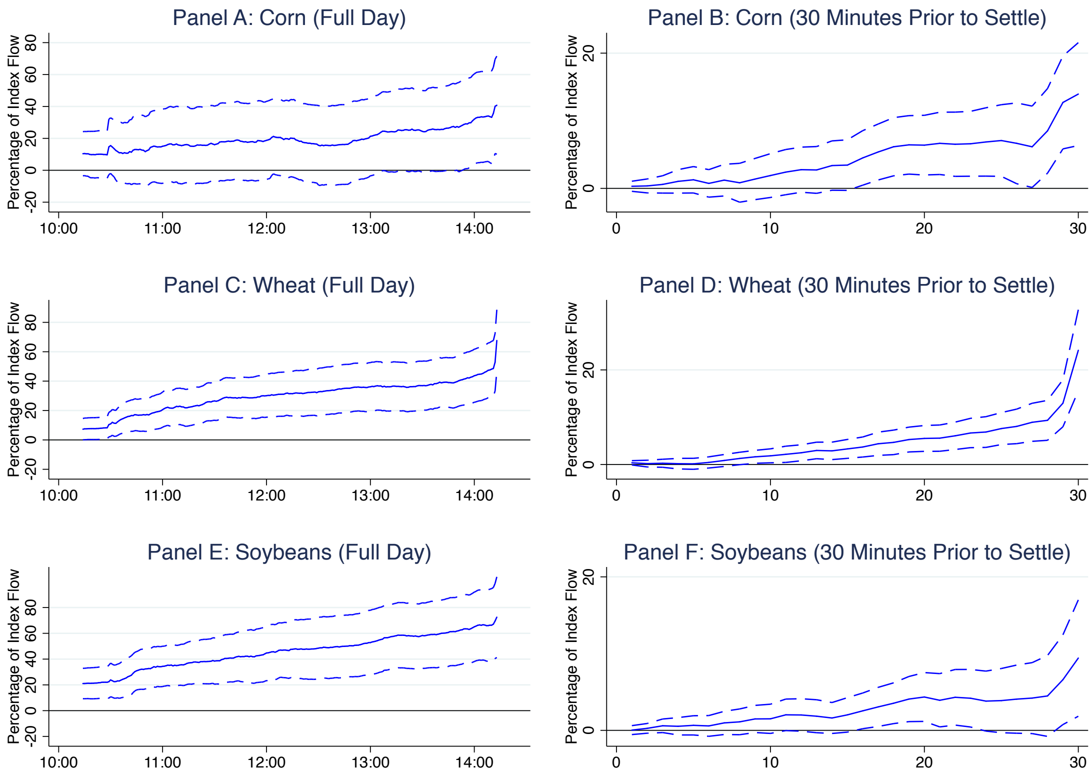


Figure 3. Regressions of Intraday Imbalance on Index Flows
The figure plots the slope coefficient and 95% confidence interval from regressions where the dependent variable is weekly cumulative futures imbalance calculated in expanding windows across the trading day, and the independent variable is weekly index flows (the specification is equivalent to the first three rows in Table 5). In Panels A, C, and E, the dependent variable is the imbalance measured from the previous day's settlement summed across the trading days in the week. In Panels B, D, and F, the dependent variable is the cumulative imbalance starting 30 minutes prior to the daily futures settlement and ending in the settlement minute, again summed across the trading week. Both imbalance and index flows are measured in millions of dollars. The sample period is January 2007 through March 2014, except that Soybean data start in January 2009. Panels A, C, and D exclude the period for corn, wheat, and soybeans in which the future settlement was delayed until 15:00 eastern time (5/22/2012 to 4/5/2013)

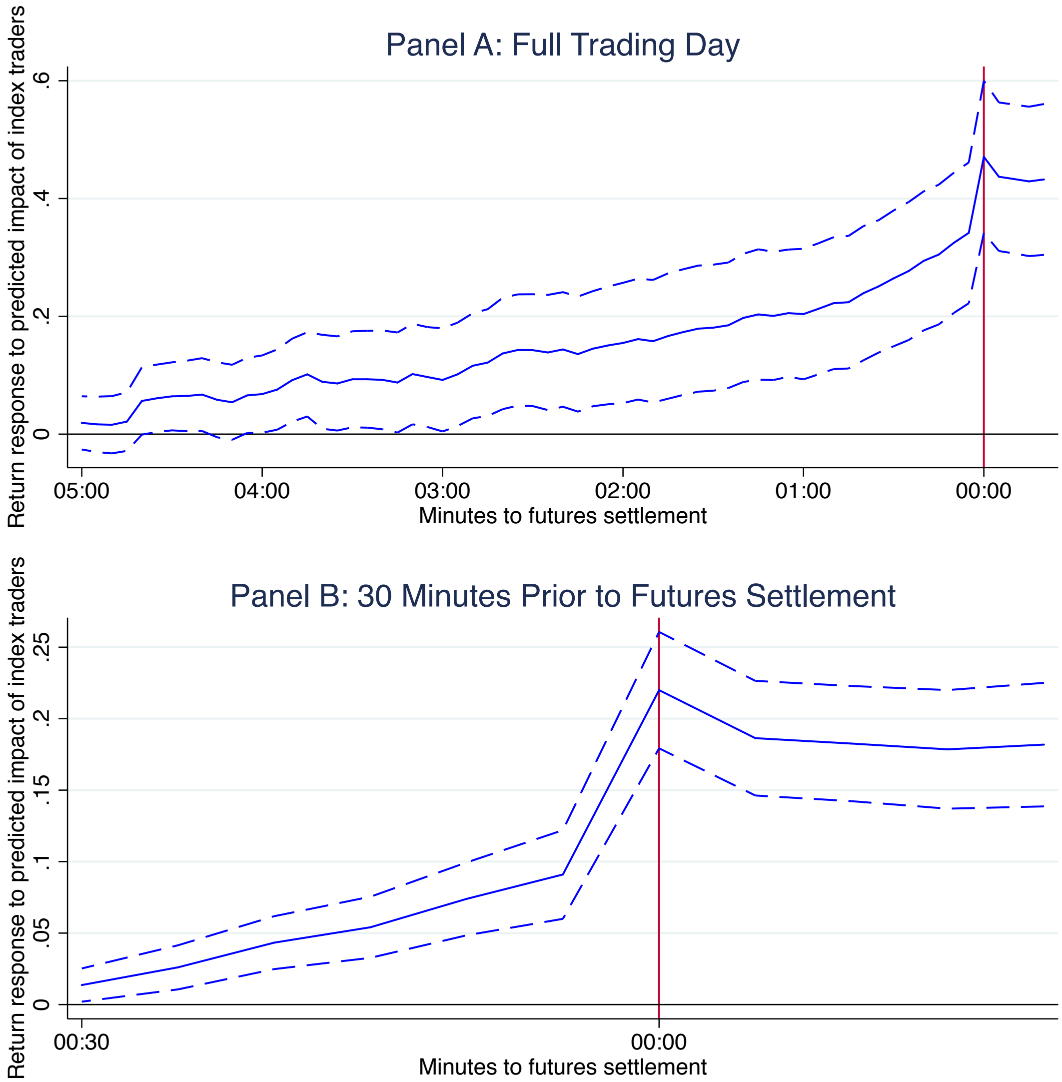


Figure 4. Regressions of Intraday Returns on Index Flows
The figure plots the slope coefficient and 95% confidence interval from weekly cross-sectional regressions using the 13 agricultural commodities included in the CFTC SCOT report. The dependent variable is weekly cumulative futures returns calculated in expanding windows across the trading day, and the relevant independent variable is the predicted price impact from the weekly index flows (the specification is equivalent to the last row in Table 6). In Panel A, the dependent variable is the cumulative return or imbalance measured from the previous day's settlement summed across the trading days in the week. In Panel B, the dependent variable is the cumulative return starting 30 minutes prior to the daily futures settlement and ending 15 minutes after the settlement minute, again summed across the trading week. Both returns and predicted impact from index flows are measured in percent. Data are from January 2008 to September 2020.

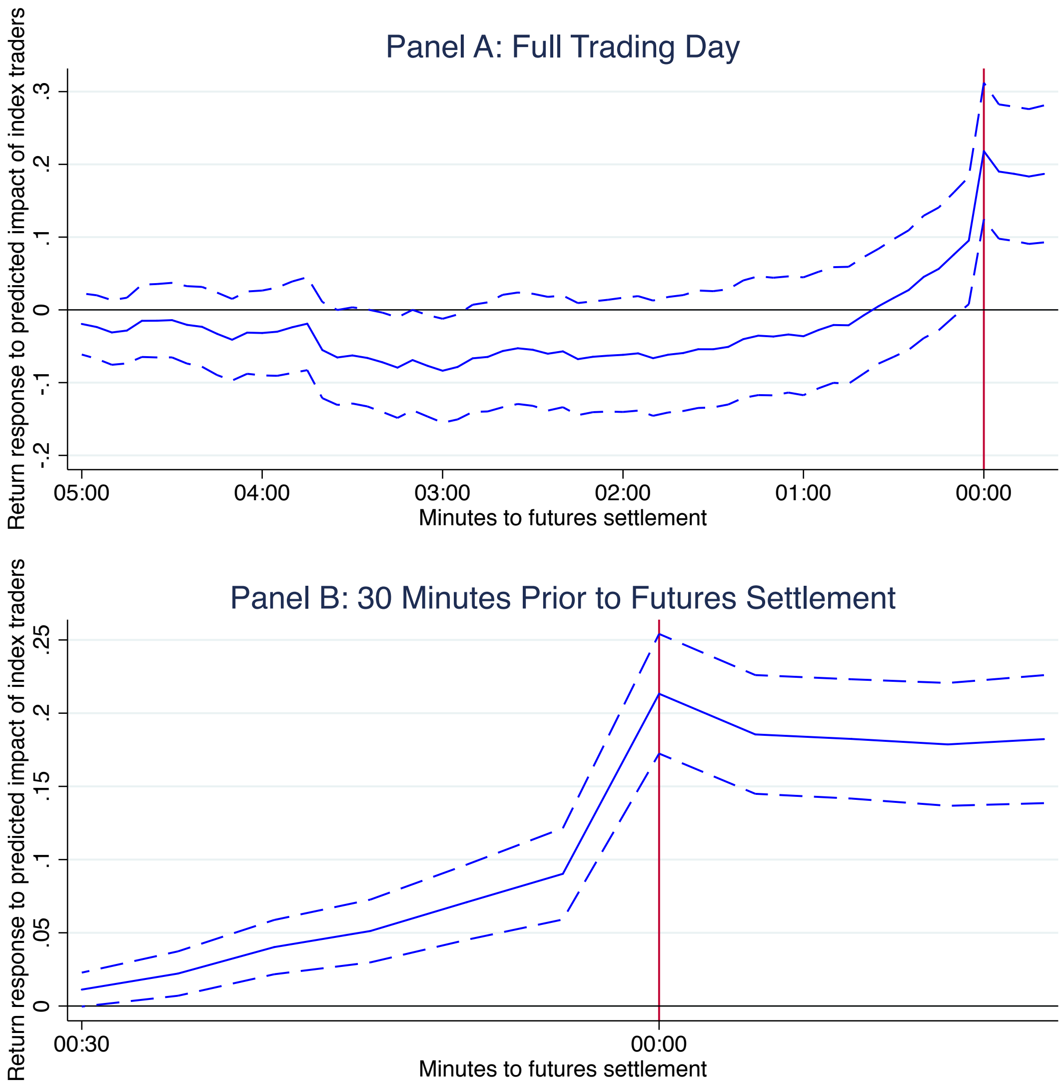


Figure 5. Regressions of Intraday Returns on Index Flows: Controlling for Managed Money Flows

The figure plots the slope coefficient and 95% confidence interval from weekly cross-sectional regressions using the 13 agricultural commodities included in the CFTC SCOT report. The dependent variable is weekly cumulative futures returns calculated in expanding windows across the trading day, and the relevant independent variable is the predicted price impact from the weekly index flows, with the predicted impact from weekly changes in the positions of Managed Money Traders included as a control (the specification is equivalent to column (3) in Table 9). In Panel A, the dependent variable is the cumulative return or imbalance measured from the previous day's settlement summed across the trading days in the week. In Panel B, the dependent variable is the cumulative return starting 30 minutes prior to the daily futures settlement and ending 15 minutes after the settlement minute, again summed across the trading week. Both returns and predicted impact from index flows are measured in percent. Data are from January 2008 to September 2020.

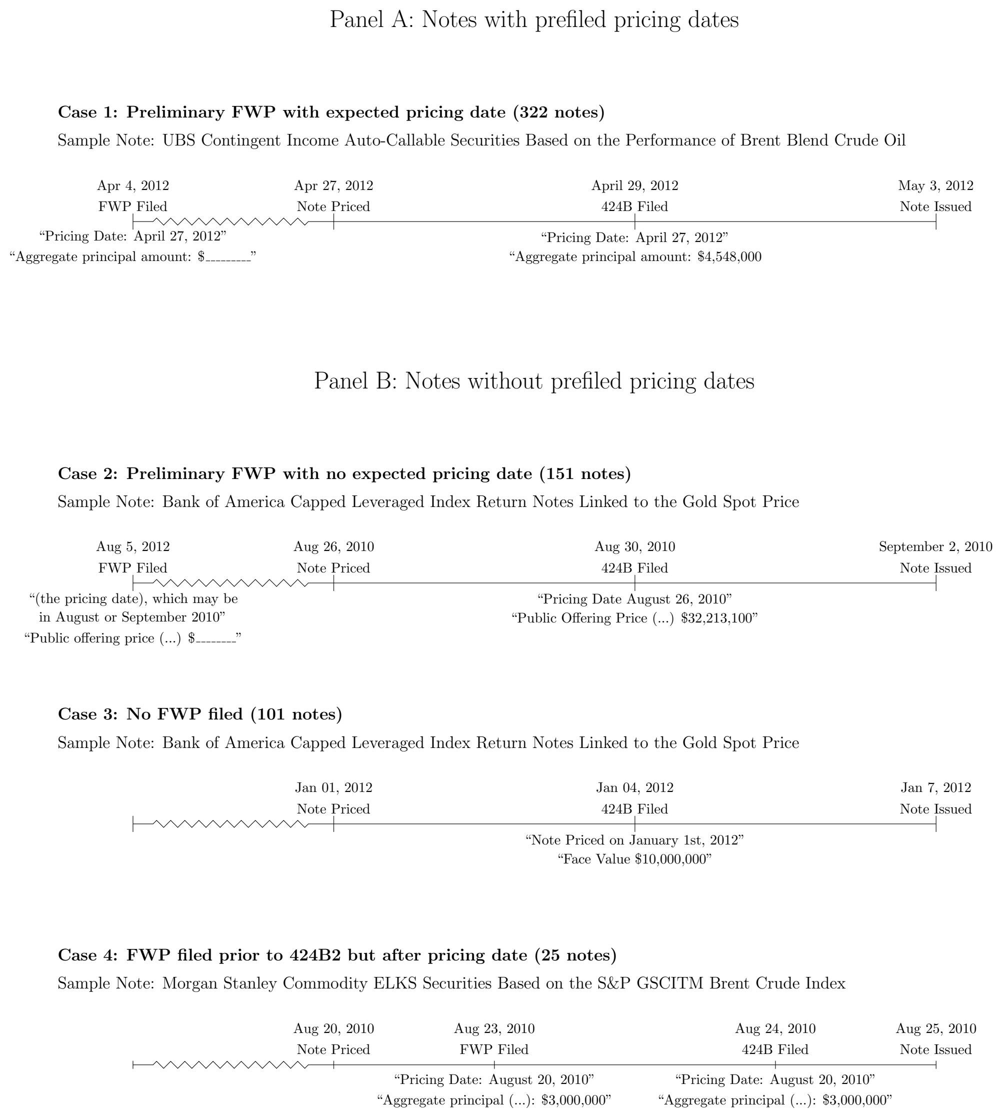


Figure 6. Pricing Dates in the Free Writing Prospectus
Panel A shows an example CLN in the Prefiled Pricing Date category, and Panel B shows three examples of CLNs in the No Prefiled Pricing Date category.

Table 1. Summary of Trade and Quote Data for Seven Commodities

The table shows means and standard deviations for minute-by-minute returns, trading volume, and signed trading volume (imbalance). Imbalance is calculated using a quote-based method similar to Lee and Ready (1991). Statistics for volume and imbalance are reported in millions of dollars of notional value. The sample is January 2007 through March 2014, except that Brent data start in January 2008 and Soybean data start in January 2009. Corn, Soybean and Wheat data are from 9:30 to the settlement time and from 7:30 to 16:00 for the remaining commodities (New York times). Minutes with no trade or quote activity are excluded.

Panel A: All Minutes

	# of Min	Return (%)		Volume (Mil $)		Imbalance (Mil $)	
		Mean	Sd Dev	Mean	Sd Dev	Mean	Sd Dev
Commodity	(1)	(2)	(3)	(4)	(5)	(6)	(7)
WTI	920,567	0.00	0.09	33.98	42.81	-0.17	14.94
Brent	791,420	0.00	0.08	14.72	23.58	-0.01	10.00
Gold	919,713	0.00	0.05	21.54	35.99	-0.15	12.00
Copper	909,542	0.00	0.07	2.79	5.43	-0.01	2.22
Corn	539,702	0.00	0.12	9.25	17.59	-0.14	7.43
Soft Wheat	509,464	0.00	0.15	4.41	9.15	-0.07	3.58
Soybeans	467,965	0.00	0.10	10.60	18.22	-0.24	7.58

Panel B: Settlement Minute

	# of Min	Return (%)		Volume (Mil $)		Imbalance (Mil $)	
		Mean	Sd Dev	Mean	Sd Dev	Mean	Sd Dev
Commodity	(1)	(2)	(3)	(4)	(5)	(6)	(7)
WTI	1,816	-0.01	0.11	193.25	88.46	-4.13	36.04
Brent	1,563	0.00	0.10	81.81	68.19	0.00	22.68
Gold	1,825	0.00	0.06	83.22	63.33	2.03	23.87
Copper	1,867	0.00	0.12	23.23	24.49	0.66	9.17
Corn	1,810	0.01	0.24	120.51	96.23	4.28	29.92
Soft Wheat	1,808	-0.04	0.36	73.17	62.36	-1.89	20.38
Soybeans	1,271	0.00	0.17	130.89	76.94	-3.57	31.25

Table 2. Price Impact Regressions

The table shows the results from separate univariate regressions where the dependent variable is one-minute returns and the independent variable is the contemporaneous one-minute imbalance. Return is measured in percentage and imbalance is measured in millions of dollars, (i.e. a coefficient of 0.01 represents a return response of 0.01%, or one basis point, per million dollars of imbalance). Columns (1)-(4) show the results for all minutes using imbalance measured with a quote-based method similar to Lee and Ready (1991). Columns (5)-(8) only include the minute prior to daily futures settlement. Standard errors are shown in parentheses. The sample is January 2007 through March 2014, except that Brent data start in January 2008 and Soybean data start in January 2009. Corn, Soybean and Wheat data are from 9:30 to the settlement time and from 7:30 to 16:00 for the remaining commodities (New York times). Minutes with no trade or quote activity are excluded.

Independent Variable:	Imbalance (Mil $)							
Dependent Variable:	Return (%) All Minutes				Return (%) Settlement Minute			
Commodity	Estimate (1)	Std. Err (2)	R-sq (3)	Obs (4)	Estimate (5)	Std. Err (6)	R-sq (7)	Obs (8)
WTI	0.0034***	(0.00001)	0.34	929,219	0.0013***	(0.00007)	0.17	1,816
Brent	0.0029***	(0.00002)	0.13	800,225	0.0010***	(0.00008)	0.05	1,563
Gold	0.0022***	(0.00002)	0.31	928,085	0.0013***	(0.00011)	0.22	1,825
Copper	0.0114***	(0.00010)	0.15	903,774	0.0042***	(0.00034)	0.11	1,867
Corn	0.0068***	(0.00019)	0.22	515,471	0.0044***	(0.00024)	0.29	1,808
Soft Wheat	0.0155***	(0.00050)	0.18	489,820	0.0091***	(0.00078)	0.26	1,806
Soybeans	0.0047***	(0.00004)	0.15	446,319	0.0032***	(0.00013)	0.37	1,271

*** p<0.01, ** p<0.05, * p<0.1

Table 3. Inferring the Price Impact of Order Flow from Daily Data

Panel A shows the results of regressions that are used to estimate the price impact of order flow from daily data. The dependent variables are the log of the price impacts estimated for a single commodity in a calendar year. The independent variables are the log of daily futures return volatility and the log of the average daily volume (in millions of $ of futures notional) across all maturities of futures for a given commodity. Price impacts are defined as the slope of a regression of minute-by-minute returns (in %) on minute-by-minute imbalance (in millions of $) as in Table 2. All variables are obtained for Brent, WTI, Gold, CME Copper, Soft Wheat, Soybeans, and Corn for each calendar year from 2008 to 2013 (Soybeans data start in 2009). In column (1), the impacts are computed using all minutes. In column (2), the impacts are computed using only the minute prior to daily futures settlement. In column (3), the two sets of impacts are pooled in a single regression with a settlement minute dummy variable. Panel B uses the regression estimates of specification (3) in Panel A and estimates impacts for a broad set of commodity contracts. The contracts are sorted from lowest impact to highest. Estimates are calculated for the period 2003 to 2018 where data are available and averages for all years are reported

Panel A: Regressions of Price Impacts on Volume and Volatility

	Log(Order Flow Impact)		
	All Minutes (1)	Settlement Minutes (2)	Combined (3)
Log(Daily Volatility)	0.854***	0.578***	0.716***
	[6.450]	[2.971]	[5.987]
Log(Average Volume)	-0.550***	-0.701***	-0.626***
	[-13.391]	[-11.622]	[-16.864]
Settlement Dummy			-0.619***
			[-7.706]
Constant	3.466***	3.158***	3.621***
	[6.237]	[3.870]	[7.194]
Observations	41	41	82
R-squared	0.887	0.826	0.860

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Panel B: Average Predicted Price Impacts by Commodity

Commodity	Daily Volatility (%)	Average Volume ($Mil/Day)	Est. Impact All Min (%/$Mil)	Est. Impact Settle Min (%/$Mil)
Gold	0.011	33,745	0.003	0.001
LME Copper	0.012	17,054	0.004	0.002
WTI Crude	0.027	54,937	0.004	0.002
Brent Crude	0.028	44,476	0.005	0.003
Aluminum	0.010	8,422	0.005	0.003
Soybeans	0.014	10,402	0.006	0.003
Corn	0.017	6,606	0.008	0.005
Natural Gas	0.028	10,592	0.009	0.005
Silver	0.020	7,824	0.010	0.005
Heating Oil	0.018	5,788	0.010	0.006
Zinc	0.014	4,161	0.011	0.006
CME Copper	0.016	5,358	0.012	0.006
Nickel	0.016	3,580	0.013	0.007
Live Cattle	0.010	2,108	0.014	0.008
Soft Wheat	0.019	3,013	0.015	0.008

Commodity	Daily Volatility (%)	Average Volume ($Mil/Day)	Est. Impact All Min (%/$Mil)	Est. Impact Settle Min (%/$Mil)
Soybean Oil	0.012	1,952	0.016	0.009
Soybean Meal	0.028	4,571	0.017	0.009
Sugar	0.019	2,195	0.018	0.010
RBOB Gasoline	0.029	12,064	0.019	0.010
Coffee	0.019	1,887	0.020	0.011
Lead	0.015	1,630	0.021	0.012
Lean Hogs	0.017	1,097	0.027	0.015
Cotton	0.017	986	0.029	0.016
Hard Wheat	0.017	813	0.032	0.017
Cocoa	0.017	821	0.033	0.018
Feeder Cattle	0.011	644	0.034	0.018
Platinum	0.013	575	0.035	0.019
Tin	0.013	522	0.037	0.020
Palladium	0.020	306	0.091	0.049

Table 4. Summary Data for Index Flows and Futures Imbalance and Returns

The table shows summary statistics for the standard deviation of commodity-index flows and predicted price impacts relative to the imbalance and return in the associated futures market. Index flows are the weekly changes in position of commodity-index traders for agricultural commodities taken from the CFTC's Supplemental Commitment of Traders reports. Panel A shows the weekly standard deviation of flows (in millions of $) for the three agricultural commodities (column (1)) for which we have trade-and-quote data (January 2007 through March 2019, except Soybeans start in January 2009), along with the average pairwise correlation of the index flows with the other two commodities (column (2)). The remaining columns of Panel A show the standard deviation of weekly imbalance in the futures markets over different subperiods of the trading day. Panel B shows the standard deviation of the predicted price impact of the weekly index flows (column (1)), which is calculated as the change in position multiplied by the estimate of price impact per million dollars of imbalance traded in the settlement minute for the applicable commodity-year combination (see Table 3). Panel B also reports the average pairwise correlation between this predicted impact and the predicted impacts of the other 12 commodities. The remaining columns of Panel B also show the weekly standard deviation of the associated futures returns for different subperiods of the trading day.

Panel A: Index Flows and Futures Imbalance (2008 - 2014)

		Index Flows (Mil $)		Standard Deviation of Futures Imbalance (Mil $)		
Commodity	# of Weeks (1)	Std. Dev (2)	Avg. Corr. w/ other Commodities (3)	Full Week (4)	30 Min Pre Settle (5)	5 Min Pre Settle (6)
Corn	382	223.0	0.30	578.0	168.5	112.6
Soybeans	270	309.4	0.35	754.1	124.4	42.2
Soft Wheat	382	138.8	0.35	280.4	97.2	71.6

Panel B: Predicted Price Impacts of Index Flows and Futures Returns (2008 - 2020)

		Predicted Impacts of Index Flows (%)		Standard Deviation of Futures Returns (%)		
Commodity	# of Weeks (1)	St. Dev (2)	Avg. Corr. w/ Other Commodities (3)	Full Week (4)	30 Min Pre Settle (5)	5 Min Pre Settle (6)
Cocoa	642	0.89	0.15	3.31	1.26	0.74
Coffee	644	0.84	0.21	4.12	1.14	0.60
Corn	640	0.95	0.21	3.45	1.07	0.73
Cotton	644	1.30	0.17	3.47	1.13	0.73
Feeder Cattle	630	0.54	0.16	2.32	0.82	0.52
Live Cattle	643	0.83	0.24	2.20	0.77	0.49
Lean Hogs	643	0.99	0.18	3.70	1.27	0.85
Soybeans	642	0.82	0.18	2.71	0.80	0.53
Soybean Meal	387	0.61	0.10	5.65	1.71	1.15
Soybean Oil	644	0.58	0.14	2.81	0.82	0.55
Sugar	637	1.24	0.12	4.07	1.18	0.61
Hard Wheat	644	0.92	0.11	3.74	1.29	0.96
Soft Wheat	643	0.96	0.17	3.93	1.31	0.97
All	8083	0.92	0.16	3.53	1.13	0.74

Table 5. Regressions of Futures Order Flow Imbalance on Commodity-Index Flows

The table shows the results from regressions of weekly futures order flow imbalance on contemporaneous commodity-index flows as reported by the CFTC in their supplemental commitments of traders reports. The first three rows show results from regressions where the independent variable is the weekly index flow (in millions of $) in a single commodity, and the dependent variable is weekly imbalance (see Table 1 for a description) in the associated futures market measured over the entire trading day (columns (2) - (4)), only in the 30 minutes prior to futures settlement (columns (5) - (7)), or only in the five minutes prior to futures settlement (columns (8) - (10)). Each set of three columns reports the estimate of the slope coefficient on the weekly index flows, the associated t-statistic, and the regression R-squared (regression constants are not reported). The last two rows show the slope coefficients on changes in index trader positions of pooled regressions using all three commodities. The first pooled regression is a time-series regression, and thus includes only commodity fixed effects. The second regression is cross-sectional, and thus includes both commodity and week fixed effects. For both pooled regressions standard errors are clustered by week. Data are from January 2007 through March 2014, except Soybeans data begin in January 2009.

Independent Variable:		Index Flows in Current Week								
Dependent Variable:		Weekly Imbalance in All Minutes			Weekly Imbalance in 30 Minutes Prior to Settle			Weekly Imbalance in 5 Minutes Prior to Settle		
Commodity	Obs (1)	Estimate (2)	T-stat (3)	R-sq (4)	Estimate (5)	T-stat (6)	R-sq (7)	Estimate (8)	T-stat (9)	R-sq (10)
Single Commodity										
Corn	382	0.383**	[2.55]	0.027	0.139***	[3.59]	0.036	0.071***	[2.70]	0.023
Soft Wheat	382	0.527***	[4.69]	0.073	0.240***	[5.53]	0.118	0.173***	[5.11]	0.113
Soybeans	270	0.741***	[3.70]	0.094	0.125***	[2.59]	0.037	0.073**	[2.13]	0.054
Pooled										
w/ commodity FE	1034	0.454***	[3.90]	0.056	0.152***	[4.78]	0.046	0.087***	[3.98]	0.021
w/ com. & week FE	1034	0.320*	[1.88]	0.571	0.139***	[2.99]	0.533	0.062**	[2.18]	0.174

Table 6. Regressions of Futures Returns on Index Flows

The table shows the results from regressions of weekly futures returns on the predicted price impacts of contemporaneous commodity-index flows (see Table 4 for a description of predicted price impacts). The first 13 rows show results from regressions where the independent variable is the weekly predicted impact (in percent) for a single commodity, and the dependent variable is weekly return in the associated futures market measured over the entire trading day (columns (2) - (6)), only in the 30 minutes prior to futures settlement (columns (7) - (11)), or only in the five minutes prior to futures settlement (columns (12) - (16)). Each set of five columns reports the estimate of the slope coefficient on the weekly index flows, the associated t-statistic and p-value calculated asymptotically taking the predicted price impact as given, the p-value from a boot strap analysis to account for potential error from the estimation of predicted impact, and the regression R-squared (regression constants are not reported). The last two rows show the slope coefficients on changes in index trader positions of pooled regressions using all 13 commodities. The first pooled regression is a time-series regression, and thus includes only commodity fixed effects. The second regression is cross-sectional, and thus includes both commodity and week fixed effects. For both pooled regressions standard errors are clustered by week (and weeks are drawn with replacement in the bootstrap). Data are January 2008 to September 2020 where available.

Independent Variable:		Predicted Index Trader Impact in Current Week														
Dependent Variable:		Weekly Return In All Minutes					Weekly Return In 30 Minutes Prior to Settle					Weekly Return In 5 Minutes Prior to Settle				
Commodity	Obs (1)	Estimate (2)	Asymptotic T-stat (3)	Asymptotic P-value (4)	Boot P-value (5)	R-sq (6)	Estimate (7)	Asymptotic T-stat (8)	Asymptotic P-value (9)	Boot P-value (10)	R-sq (11)	Estimate (12)	Asymptotic T-stat (13)	Asymptotic P-value (14)	Boot P-value (15)	R-sq (16)
Single Commodity																
Cocoa	641	1.085***	[6.31]	(0.000)	(0.000)	0.08	0.343***	[4.94]	(0.000)	(0.001)	0.06	0.179***	[4.92]	(0.000)	(0.000)	0.05
Coffee	642	1.104***	[4.41]	(0.000)	(0.000)	0.05	0.208***	[3.35]	(0.001)	(0.001)	0.02	0.091***	[2.89]	(0.004)	(0.008)	0.02
Corn	643	0.611***	[3.44]	(0.001)	(0.000)	0.03	0.212***	[4.15]	(0.000)	(0.000)	0.04	0.114***	[3.00]	(0.003)	(0.004)	0.02
Cotton	643	0.374***	[3.01]	(0.003)	(0.001)	0.02	0.236***	[6.16]	(0.000)	(0.000)	0.07	0.133***	[5.80]	(0.000)	(0.000)	0.06
Feeder Cattle	619	0.937***	[3.77]	(0.000)	(0.000)	0.05	0.405***	[5.94]	(0.000)	(0.000)	0.07	0.272***	[6.47]	(0.000)	(0.000)	0.08
Live Cattle	641	0.709***	[5.41]	(0.000)	(0.000)	0.07	0.275***	[6.61]	(0.000)	(0.000)	0.09	0.179***	[6.29]	(0.000)	(0.000)	0.09
Lean Hogs	638	0.761***	[4.31]	(0.000)	(0.000)	0.04	0.444***	[6.74]	(0.000)	(0.000)	0.12	0.305***	[7.15]	(0.000)	(0.000)	0.12
Soybeans	638	0.752***	[4.40]	(0.000)	(0.000)	0.05	0.127***	[2.82]	(0.005)	(0.000)	0.02	0.092***	[2.90]	(0.004)	(0.000)	0.02
Soybean Meal	387	0.474	[0.87]	(0.382)	(0.574)	0.00	0.391**	[2.52]	(0.012)	(0.005)	0.02	0.202*	[1.82]	(0.069)	(0.035)	0.01
Soybean Oil	636	0.909***	[3.78]	(0.000)	(0.000)	0.03	0.122*	[1.86]	(0.063)	(0.014)	0.01	0.127***	[2.97]	(0.003)	(0.002)	0.02
Sugar	636	0.020	[0.15]	(0.883)	(0.411)	0.00	0.058	[1.42]	(0.157)	(0.086)	0.00	0.059***	[2.98]	(0.003)	(0.000)	0.01
Hard Wheat	641	1.111***	[5.34]	(0.000)	(0.000)	0.08	0.230***	[3.37]	(0.001)	(0.000)	0.03	0.233***	[4.80]	(0.000)	(0.000)	0.05
Soft Wheat	643	0.592***	[2.68]	(0.007)	(0.002)	0.02	0.442***	[6.68]	(0.000)	(0.000)	0.11	0.319***	[5.92]	(0.000)	(0.000)	0.10
Pooled																
w/ commodity FE	8048	0.637***	[8.06]	(0.000)	(0.000)	0.03	0.250***	[12.53]	(0.000)	(0.000)	0.05	0.167***	[12.13]	(0.000)	(0.000)	0.05
w/ com. & week FE	8048	0.470***	[7.04]	(0.000)	(0.000)	0.27	0.208***	[10.59]	(0.000)	(0.000)	0.20	0.128***	[9.75]	(0.000)	(0.000)	0.21

*** p<0.01, ** p<0.05, * p<0.1

Table 7. Regressions of Futures Return Reversals on Index Flows

The table shows the results from regressions of weekly futures returns on the predicted price impacts of contemporaneous commodity-index flows (see Table 4 for a description of predicted price impacts). Each set of three columns reports the estimate of the slope coefficient on the weekly index flows, the associated t-statistic, and the regression R-squared (regression constants are not reported). Columns (2) - (4) show the results from regressions where the independent variable is the contemporaneous weekly index flow, and the dependent variable is the return in the five minutes after the daily futures settle summed across the week. Columns (5) - (7) are the same except the dependent variable is the return in the 30 minutes after settle. The remaining columns report regressions where the independent variable is the previous weeks index flow, and the return is the cumulative return on the first day of the trading week (Columns (8) - (10)), the first two days of the trading week (Columns (11) - (13)), or the full trading week (Columns (14) - (16)) The last two rows show the slope coefficients on changes in index trader positions of pooled regressions using all 13 commodities. The first pooled regression is a time-series regression, and thus includes only commodity fixed effects. The second regression is cross-sectional, and thus includes both commodity and week fixed effects. For both pooled regressions standard errors are clustered by week (and weeks are drawn with replacement in the bootstrap). In all specifications we include the return in the 5-minutes prior to the observation (coefficient not reported) to control for bid-ask bounce. Data are January 2008 to September 2020 where available.

Independent Variable:		Predicted Index Trader Impact in Current Week						Predicted Index Trader Impact in Previous Week								
Dependent Variable:		Weekly Return In 5 Minutes After Settle			Weekly Return In 30 Minutes After Settle			Return on First Trading Day in Week			Return Across First Two Trading Days in Week			Return Across All Trading Days in Week		
Commodity	Obs (1)	Estimate (2)	T-stat (3)	R-sq (4)	Estimate (5)	T-stat (6)	R-sq (7)	Estimate (8)	T-stat (9)	R-sq (10)	Estimate (11)	T-stat (12)	R-sq (13)	Estimate (14)	T-stat (15)	R-sq (16)
Single Commodity																
Cocoa	641	-0.012	[-0.57]	0.03	0.045	[1.13]	0.01	0.140*	[1.76]	0.01	0.042	[0.39]	0.00	0.116	[0.70]	0.00
Coffee	643	0.014	[0.24]	0.00	-0.051	[-0.99]	0.00	-0.205**	[-2.19]	0.01	-0.011	[-0.09]	0.00	-0.018	[-0.07]	0.00
Corn	643	-0.009	[-0.79]	0.00	-0.008	[-0.66]	0.01	-0.106	[-1.29]	0.01	-0.206*	[-1.81]	0.01	-0.153	[-0.77]	0.00
Cotton	643	-0.028**	[-2.34]	0.05	-0.026	[-1.40]	0.04	-0.085	[-1.50]	0.01	-0.106	[-1.18]	0.00	-0.165	[-1.06]	0.00
Feeder Cattle	629	-0.008	[-0.24]	0.05	-0.055	[-1.42]	0.04	-0.074	[-0.86]	0.01	-0.210	[-1.52]	0.01	-0.357	[-1.53]	0.02
Live Cattle	642	-0.000	[-0.02]	0.00	0.005	[0.18]	0.00	0.065	[1.11]	0.01	0.156*	[1.96]	0.01	0.201	[1.47]	0.01
Lean Hogs	642	-0.024	[-1.06]	0.09	-0.036	[-1.45]	0.06	-0.139*	[-1.81]	0.01	-0.207	[-1.64]	0.01	-0.282	[-1.56]	0.01
Soybeans	641	-0.008	[-1.01]	0.00	-0.001	[-0.09]	0.00	-0.083	[-1.11]	0.01	-0.035	[-0.39]	0.00	0.119	[0.71]	0.00
Soybean Meal	387	-0.165***	[-3.15]	0.04	-0.163***	[-3.11]	0.04	-0.086	[-0.38]	0.00	-0.072	[-0.22]	0.00	0.117	[0.21]	0.00
Soybean Oil	638	-0.038***	[-2.84]	0.03	-0.018	[-1.09]	0.00	-0.040	[-0.32]	0.00	0.013	[0.09]	0.00	0.207	[0.85]	0.00
Sugar	636	-0.016	[-1.05]	0.00	-0.036	[-0.92]	0.00	-0.082	[-1.33]	0.01	-0.097	[-0.98]	0.00	-0.195	[-1.19]	0.00
Hard Wheat	643	-0.032***	[-2.75]	0.03	-0.032***	[-2.74]	0.03	0.008	[0.10]	0.00	-0.094	[-0.72]	0.00	0.025	[0.14]	0.00
Soft Wheat	643	-0.018*	[-1.72]	0.01	-0.027**	[-2.37]	0.02	-0.121	[-1.40]	0.00	-0.256*	[-1.92]	0.01	-0.490**	[-2.43]	0.01
Pooled																
w/ commodity FE	8071	-0.021***	[-3.41]	0.02	-0.025***	[-2.73]	0.02	-0.069**	[-2.32]	0.00	-0.094**	[-2.35]	0.00	-0.109	[-1.47]	0.00
w/ com. & week FE	8071	-0.025***	[-3.53]	0.13	-0.032***	[-2.88]	0.02	-0.085***	[-3.70]	0.28	-0.125***	[-3.58]	0.28	-0.162***	[-2.84]	0.30

*** p<0.01, ** p<0.05, * p<0.1

Table 8. Regressions of Futures Returns and Reversals on Index Flows by Subperiod

The first row of Panel A repeats the cross-sectional regression results in the last row Table 6, and the first row of Panel B repeats the cross-section regression results for three of the specifications in the last row of Table 7. The remaining rows in both panels show results for the same specifications in two-year subperiods.

Panel A: Regressions of Futures Returns on Index Flows

Independent Variable:		Predicted Index Trader Impact in Current Week								
Dependent Variable:		Weekly Return in All Minutes			Weekly Return in 30 Minutes Prior to Settle			Weekly Return in 5 Minutes Prior to Settle		
Sub-period	Obs (1)	Estimate (2)	T-stat (3)	R-sq (4)	Estimate (5)	T-stat (6)	R-sq (7)	Estimate (8)	T-stat (9)	R-sq (10)
Full Sample	8048	0.470***	[7.04]	0.27	0.208***	[10.59]	0.20	0.128***	[9.75]	0.21
2008-2010	1626	0.372***	[2.91]	0.36	0.203***	[5.53]	0.36	0.106***	[3.81]	0.28
2011-2012	1260	0.242	[1.58]	0.25	0.185***	[4.21]	0.25	0.110***	[4.73]	0.20
2013-2014	1339	0.055	[0.50]	0.19	0.164***	[2.97]	0.19	0.076**	[2.24]	0.14
2015-2016	1352	0.614***	[3.42]	0.25	0.218***	[4.31]	0.25	0.208***	[5.95]	0.21
2017-2018	1365	0.740***	[6.00]	0.24	0.311***	[5.98]	0.24	0.170***	[5.50]	0.19
2019-2020	1129	0.930***	[5.88]	0.30	0.185***	[4.93]	0.30	0.145***	[5.20]	0.22

Panel B: Regressions of Futures Return Reversals on Index Flows

Independent Variable:		Predicted Index Trader Impact in Current Week			Predicted Index Trader Impact in Previous Week					
Dependent Variable:		Weekly Return in 30 Minutes after Settle			Return on First Trading Day in Week			Return Across All Trading Days in Week		
Sub-period	Obs (1)	Estimate (2)	T-stat (3)	R-sq (4)	Estimate (5)	T-stat (6)	R-sq (7)	Estimate (8)	T-stat (9)	R-sq (10)
Full Sample	8048	-0.032***	[-2.88]	0.02	-0.085***	[-3.70]	0.28	-0.162***	[-2.84]	0.30
2008-2010	1626	0.019	[0.82]	0.11	-0.089**	[-2.14]	0.41	-0.246**	[-1.99]	0.42
2011-2012	1260	-0.019	[-0.61]	0.13	-0.093*	[-1.75]	0.31	-0.145	[-1.12]	0.30
2013-2014	1339	-0.001	[-0.08]	0.12	-0.084	[-1.21]	0.16	-0.133	[-1.16]	0.20
2015-2016	1352	-0.125***	[-3.83]	0.17	-0.103	[-1.45]	0.20	0.036	[0.25]	0.25
2017-2018	1365	-0.046***	[-2.69]	0.17	-0.097	[-1.59]	0.22	-0.291*	[-1.77]	0.24
2019-2020	1129	-0.067***	[-3.40]	0.20	-0.023	[-0.40]	0.24	-0.009	[-0.07]	0.30

*** p<0.01, ** p<0.05, * p<0.1

Table 9. Regressions of Futures Returns on Predicted Price Impact from Commodity-Index Flows Controlling for Other Trader Types

The table shows the results from regressions of weekly cross-sectional futures returns on the predicted price impacts of contemporaneous commodity-index flows (see Table 4 for a description of predicted price impacts), and includes controls for predicted price impact from other types of traders reported in the CFTC's Disaggregated Commitments of Traders Reports. Columns (1), (5), and (9) report the regression results from the last row of Table 6. The remaining columns add the predicted impacts from weekly changes in the positions of Managed Money, Physical Traders, and Swap Traders. Data are 2008 to 2020 where available.

Dedependent Variable:	Weekly Return in All Minutes				Weekly Return in 30 Minutes Prior to Settle				Weekly Return in 5 Minutes Prior to Settle			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Index Traders	0.471***	0.221***	-0.175***	0.640***	0.206***	0.202***	0.191***	0.175***	0.127***	0.121***	0.113***	0.101***
	[7.08]	[4.61]	[-3.24]	[7.66]	[10.63]	[10.29]	[9.33]	[6.89]	[9.76]	[9.31]	[8.19]	[6.23]
Managed Money		0.620***				0.011*				0.015***		
		[35.95]				[1.73]				[3.76]		
Producer/Merchant/Processor/User			-0.616***				-0.014**				-0.013***	
			[-34.51]				[-2.30]				[-3.56]	
Swap Dealers				-0.209***				0.039				0.033***
				[-3.20]				[1.60]				[2.64]
Week FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Commodity FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Obs	8071	8071	8071	8071	8071	8071	8071	8071	8071	8071	8071	8071
R-sq	0.27	0.44	0.45	0.27	0.20	0.20	0.20	0.20	0.20	0.21	0.20	0.20

*** p<0.01, ** p<0.05, * p<0.1

Table 10. Regressions of Futures Returns on Index Flows from Monthly IID Data

The table shows the results from regressions of monthly futures returns on the predicted price impacts of contemporaneous commodity-index flows (see Table 4 for a description of predicted price impacts). The first rows show results from regressions where the independent variable is the monthly predicted impact (in percent) for a single commodity, and the dependent variable is monthly return in the associated futures market measured over the entire trading day (columns (2) - (4)), only in the 30 minutes prior to futures settlement (columns (5) - (7)), or only in the five minutes after futures settlement (columns (8) - (10)). Each set of three columns reports the estimate of the slope coefficient on the monthly index flows, the associated t-statistic, and the regression R-squared (regression constants are not reported). The last two rows show the slope coefficients on changes in index trader positions of pooled regressions using all 20 commodities. The first pooled regression is a time-series regression, and thus includes only commodity fixed effects. The second regression is cross-sectional, and thus includes both commodity and week fixed effects. For both pooled regressions standard errors are clustered by month. Data are 2010 to 2015 where available.

Independent Variable:		Predicted Index Trader Impact in Current Month								
Dependent Variable:		Monthly Return In All Minutes			Monthly Return In 30 Minutes Prior to Settle			Monthly Return In 5 Minutes After Settle		
Commodity	Obs (1)	Estimate (2)	T-stat (3)	R-sq (4)	Estimate (5)	T-stat (6)	R-sq (7)	Estimate (8)	T-stat (9)	R-sq (10)
Agricultural										
Cocoa	64	0.743	[1.16]	0.02	1.125***	[3.47]	0.16	0.015	[0.15]	0.00
Coffee	64	0.357	[1.11]	0.01	0.059	[0.67]	0.00	-0.088	[-0.84]	0.01
Corn	64	0.438	[0.66]	0.01	0.144	[0.79]	0.01	-0.005	[-0.85]	0.00
Cotton	64	0.413**	[2.45]	0.03	0.398***	[7.77]	0.24	-0.110***	[-2.79]	0.07
Feeder Cattle	62	-0.142	[-0.27]	0.00	0.321	[1.53]	0.04	0.022	[0.37]	0.00
Live Cattle	64	0.610	[1.58]	0.05	0.425**	[2.55]	0.10	0.005	[0.07]	0.00
Lean Hogs	64	0.336	[0.74]	0.01	0.278**	[2.64]	0.06	-0.005	[-0.09]	0.00
Soybeans	64	0.440	[0.43]	0.00	0.077	[0.25]	0.00	-0.004	[-0.61]	0.00
Soybean Meal	57	-0.284	[-0.69]	0.00	0.610***	[5.41]	0.15	0.002	[0.70]	0.00
Soybean Oil	64	0.272	[0.29]	0.00	0.093	[0.52]	0.00	-0.011	[-1.18]	0.00
Sugar	64	0.078	[0.19]	0.00	0.224*	[1.82]	0.06	0.053	[0.81]	0.02
Hard Wheat	64	1.080	[1.62]	0.05	0.176	[1.57]	0.02	-0.004	[-0.61]	0.00
Soft Wheat	64	1.134	[1.53]	0.04	0.670***	[3.08]	0.25	0.010	[1.08]	0.01
Other										
WTI Crude Oil	64	-0.566	[-1.40]	0.02	0.829***	[3.55]	0.32	-0.139***	[-4.43]	0.16
Heating Oil	64	0.915	[1.38]	0.03	1.202***	[2.70]	0.30	-0.145**	[-2.31]	0.09
Natural Gas	64	-0.571***	[-3.57]	0.06	0.292*	[1.85]	0.10	0.009	[0.15]	0.00
RBOB Gasoline	64	-0.644	[-0.36]	0.00	3.539***	[2.78]	0.31	-0.367***	[-2.73]	0.09
Copper	64	0.788**	[2.11]	0.09	0.358***	[3.93]	0.26	-0.036	[-1.24]	0.03
Gold	64	-0.294	[-0.60]	0.01	0.182**	[2.16]	0.05	-0.039	[-1.00]	0.02
Silver	64	2.042***	[2.83]	0.10	0.612***	[4.11]	0.15	-0.141**	[-2.47]	0.09
Pooled										
w/ commodity FE	1278	0.228	[1.46]	0.02	0.440***	[6.00]	0.14	-0.038*	[-1.68]	0.08
w/ com. & week FE	1278	0.087	[0.67]	0.23	0.373***	[4.19]	0.25	-0.041	[-1.61]	0.14

*** p<0.01, ** p<0.05, * p<0.1

Table 11. Summary CLN Data

Panel A shows the total face value and the calculated trade size necessary to delta hedge the notes on the pricing dates for each commodity. Panel B summarizes days with CLN determination for all commodities. Both panels also report the predicted price impact of order flow associated with the CLNs, which is calculated as the delta hedge size multiplied by the estimate of price impact per million dollars of imbalance trade in the settlement minute for the applicable commodity-year combination (see Table 3).

Panel A: Pricing Days of Commodity Linked Notes

		Face Value ($Mil)			Δ Hedge Size ($Mil)			Predicted Impact(%)		
	N	mean	min	max	mean	min	max	mean	min	max
Gold	200	17.1	2.0	157.9	12.1	0.1	108.2	0.01	0.00	0.31
Brent Crude	114	12.5	2.0	103.4	7.5	0.4	56.0	0.05	0.03	0.06
WTI Crude	80	13.8	2.0	75.9	8.8	0.5	63.0	0.01	0.00	0.08
Palladium	41	13.3	2.3	80.2	7.7	0.7	43.5	0.02	0.00	0.20
LME Copper	34	16.0	2.1	155.5	13.8	0.4	172.4	0.02	0.00	0.31
Silver	25	23.2	2.0	84.9	15.9	0.4	54.9	0.04	0.01	0.11
Corn	21	27.9	2.0	205.0	23.8	1.2	182.7	0.20	0.01	1.07
Natural Gas	8	15.1	3.1	55.4	7.6	1.4	27.1	0.70	0.06	1.34
RBOB Gasoline	4	16.9	2.5	42.3	12.5	0.4	33.8	0.04	0.00	0.10
Cotton	2	7.5	5.0	10.0	4.3	3.2	5.5	0.08	0.08	0.08
Platinum	2	35.4	7.3	63.6	43.1	4.7	81.6	0.07	0.07	0.07
Lead	1	5.0	5.0	5.0	4.7	4.7	4.7	0.19	0.19	0.19
Zinc	1	11.0	11.0	11.0	10.3	10.3	10.3	0.04	0.04	0.04
Nickel	1	23.0	23.0	23.0	21.6	21.6	21.6	0.09	0.09	0.09
Aluminum	1	17.0	17.0	17.0	15.9	15.9	15.9	0.07	0.00	0.27
Tin	1	4.0	4.0	4.0	3.7	3.7	3.7	0.19	0.00	1.93
Soybeans	1	19.6	19.6	19.6	19.2	19.2	19.2	0.10	0.10	0.10
Total	537	15.9	2	205.0	11.1	0.1	182.7	0.04	0.00	0.35

Panel B: Determination Dates of Commodity Linked Notes

		Face Value ($Mil)			Δ Hedge Size ($Mil)			Predicted Impact (%)		
	N	mean	min	max	mean	min	max	mean	min	max
All Days	534	16.0	2.0	205.0	7.4	-23.3	228.0	-0.03	-2.84	0.02
... and positive Δ	219	18.5	2.0	205.0	18.6	0.2	228.0	-0.07	-2.84	-0.00
... and $10+ Mil	104	33.2	10.0	158.0	33.2	0.9	228.0	-0.12	-2.84	-0.00
Prior to 2014/02	425	16.3	2.0	158.0	7.9	-18.6	228.0	-0.03	-2.84	0.01
... and positive Δ	160	19.3	2.0	205.0	21.3	0.2	228.0	-0.09	-2.84	-0.00
... and $10+ Mil	79	33.6	10.0	156.0	37.0	0.9	228.0	-0.15	-2.84	-0.00

Table 12. Average Returns and Predicted Impacts of Delta Hedges on CLN Pricing Dates

The table shows average futures returns and the average value of the predicted price impacts of delta hedging trades on days with CLN pricing. Panel A includes all dates with CLN pricing in the sample. Panel B includes only commodity/years where the predicted price impact is less than one basis point per million dollars traded. Price impacts are calculated as in Table 11, and standard errors below the reported averages are obtained using a bootstrap to account for estimation error from the regressions reported in Table 3.

Panel A: Average Pricing Date Returns

	Days Outside of Goldman Roll			All Days		
	All (1)	Prefiled Pricing Date (2)	No Prefiled Pricing Date (3)	All (4)	Prefiled Pricing Date (5)	No Prefiled Pricing Date (6)
Realized Pricing Date Returns						
Avg.	0.28***	-0.00	0.48***	0.13*	-0.12	0.38***
t-stat	[3.33]	[-0.02]	[4.28]	[1.71]	[-1.08]	[3.81]
Predicted Impact of Delta Hedges						
Avg.	0.058	0.024	0.082	0.051	0.029	0.074
Bootstrap SE	(0.006)	(0.003)	(0.009)	(0.006)	(0.004)	(0.008)
Obs	342	145	197	537	274	263

Panel B: Average Pricing Date Returns in Liquid Markets

	Days Outside of Goldman Roll			All Days		
	All (1)	Prefiled Pricing Date (2)	No Prefiled Pricing Date (3)	All (4)	Prefiled Pricing Date (5)	No Prefiled Pricing Date (6)
Realized Pricing Date Returns						
Avg.	0.24***	-0.02	0.43***	0.09	-0.15	0.35***
t-stat	[2.77]	[-0.14]	[3.63]	[1.19]	[-1.29]	[3.28]
Predicted Impact of Delta Hedges						
Avg.	0.026	0.008	0.039	0.023	0.009	0.038
Bootstrap SE	(0.002)	(0.001)	(0.003)	(0.002)	(0.001)	(0.003)
Obs	309	131	178	482	244	238

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 13. Regressions of Pricing Date Returns on Dummy for No Prefiled Pricing Date

The table reports results from regressions where the return on the CLN pricing day is the dependent variable and the independent variables include a dummy variable that takes a value of one if the note is in the No Prefiled Date category as well as various controls. Columns (1) - (5) show results for notes outside of the Goldman Roll period. Columns (6) - (10) show results for all pricing days. Columns (2) and (6) control for the predicted impact of the hedging trade as calculated in Table 11. Columns (3) and (6) control for a dummy variable that takes a value of one if a notes face value is greater than or equal to $10 million. Columns (4) and (9) include the face value of the note in millions of $. Columns (5) and (10) include the face value and add fixed effects for commodity, year, and CLN issuer.

	Pricing Days Outside of Goldman Roll					All Pricing Days				
	Pricing Date Return (1)	Pricing Date Return (2)	Pricing Date Return (3)	Pricing Date Return (4)	Pricing Date Return (5)	Pricing Date Return (6)	Pricing Date Return (7)	Pricing Date Return (8)	Pricing Date Return (9)	Pricing Date Return (10)
Dummy: No Prefiled Pricing Date	0.48***	0.45***	0.44**	0.46***	0.45*	0.50***	0.47***	0.46***	0.48***	0.67***
	[2.96]	[2.74]	[2.39]	[2.68]	[1.83]	[3.38]	[3.17]	[2.97]	[3.14]	[3.00]
Predicted Impact of Delta Hedges		0.53*					0.67**			
		[1.80]					[2.17]			
Dummy: FV $\geq$\$10 Mil			0.14					0.14		
			[0.76]					[0.89]		
Face Value				0.002	0.003				0.002	0.004
				[0.647]	[0.849]				[0.793]	[1.094]
Fixed Effects	N	N	N	N	Y	N	N	N	N	Y
Obs	342	342	342	342	342	537	537	537	537	537
R-sq	0.02	0.03	0.03	0.03	0.16	0.02	0.02	0.02	0.02	0.15

*** p<0.01, ** p<0.05, * p<0.1

Table 14. Average Returns and Predicted Impacts of Delta Hedges on CLN Determination Dates

The table repeats the analysis in Table 12, but uses the average return on days with CLN determination rather than CLN pricing, and with predicted impact calculated as negative one times the determination date delta times the face value. Panel A uses all commodity/days with a note that has a delta greater than zero on the determination date prior to 2/1/2014, which is the sample period from HPW. Panel B extends this to commodity/days with a note that has a delta greater than zero on the determination date prior to 1/1/2019.

Panel A: Determination Dates prior to 2014/02 (HPW Sample)

	Days Outside of Goldman Roll		All Days	
	≥$10 Million Face Value (1)	All (2)	≥$10 Million Face Value (3)	All (4)
Realized Determination Date Returns				
Avg.	-0.10	0.05	-0.01	0.05
t-stat	[-0.49]	[0.38]	[-0.06]	[0.44]
Predicted Impact of Delta Hedges				
Avg.	-0.22	-0.13	-0.19	-0.10
Bootstrap SE	(0.03)	(0.02)	(0.02)	(0.01)
Obs	54	108	75	157

Panel B: Determination Dates Prior to 2019/01

	Days Outside of Goldman Roll		All Days	
	≥$10 Million Face Value (1)	All (2)	≥$10 Million Face Value (3)	All (4)
Realized Determination Date Returns				
Avg.	-0.04	0.14	-0.10	0.03
t-stat	[-0.22]	[1.15]	[-0.59]	[0.31]
Predicted Impact of Delta Hedges				
Avg.	-0.19	-0.10	-0.16	-0.09
Bootstrap SE	(0.02)	(0.01)	(0.02)	(0.01)
Obs	67	141	92	202

*** p<0.01, ** p<0.05, * p<0.1

Table 15. Intraday Futures Returns and Imbalance around CLN Pricings and Determinations

Table shows the average returns (in %) and imbalance (in millions of $) over various intraday periods during the pricing and determination dates of CLNs. The sample is restricted to the commodities (Brent crude oil, WTI crude oil, gold, and corn) and periods where we have intraday data. Panel A examines pricing days for notes in the No Prefiled Date category, Panel B examines pricing days for all notes, and Panel C examines determination dates. Column (1) shows the average full day return ending at the close of the futures market. Column (2) shows the average return ending 30 minutes prior to the pricing or determination time of the first note on the day. (Some days have multiple gold notes, and most, but not all, of the gold notes price at 3:00 p.m. London time. For the other commodities all of our notes price at the close of the futures market). Columns (3) - (5) show average returns in the periods 30, 15, and five minutes before the pricing or determination of the note. Columns (6) - (8) show the average abnormal imbalance in the windows just prior to the pricing or determination of the note. We calculate average abnormal imbalance with a regression approach. The regression sample is constructed by including the imbalance observations for all minutes and commodities that have any note pricing during the sample. For instance, in column (7) part of our sample is the imbalance in the 30 minutes up to and including the closing minute of the WTI on all days in our trade-and-quote data, since WTI notes price at this time. We control for commodity-time fixed effects, and our estimate for abnormal imbalance is the coefficient on a dummy that takes a value of one if there is the pricing or determination of a note.

Panel A: Pricing Days without Prefiled Pricing Date

	Mean Returns (%)					Abnormal Imbalance (Mil $)		
	Full Day (1)	Before 30 min prior to pricing (2)	30 Min Prior to pricing (3)	15 min Prior to pricing (4)	5 min Prior to pricing (5)	Before 30 min prior to pricing (6)	30 Min Prior to pricing (7)	5 min Prior to pricing (8)
Average	0.35***	0.24**	0.03	0.04	0.04***	27.43	-5.12	10.73***
	[3.20]	[2.49]	[1.01]	[1.59]	[3.00]	[1.12]	[-0.69]	[3.58]
Notes	223	223	223	223	223	181	181	181

Panel B: All Pricing Days

	Mean Returns (%)					Abnormal Imbalance (Mil $)		
	Full Day (1)	Before 30 min prior to pricing (2)	30 Min Prior to pricing (3)	15 min Prior to pricing (4)	5 min Prior to pricing (5)	Before 30 min prior to pricing (6)	30 Min Prior to pricing (7)	5 min Prior to pricing (8)
Average	0.10	0.02	0.01	0.02	0.03***	3.16	-7.51	6.01***
	[1.32]	[0.33]	[0.36]	[1.07]	[2.67]	[0.19]	[-1.54]	[3.03]
Notes	509	509	509	509	509	428	428	428

Panel C: All Determination Days

	Mean Returns (%)					Abnormal Imbalance (Mil $)		
	Full Day (1)	Before 30 min prior to pricing (2)	30 Min Prior to pricing (3)	15 min Prior to pricing (4)	5 min Prior to pricing (5)	Before 30 min prior to pricing (6)	30 Min Prior to pricing (7)	5 min Prior to pricing (8)
Average	-0.03	-0.05	0.02	0.01	0.02	-7.27	-15.67**	-4.54
	[-0.31]	[-0.59]	[0.57]	[0.29]	[1.29]	[-0.31]	[-2.02]	[-1.55]
Notes	195	195	195	195	195	155	155	155

Robust t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Internet Appendix for Order Flows and Financial Investor Impacts in Commodity Futures Markets

Mark Ready
University of Wisconsin-Madison

Robert Ready
University of Oregon

October 10, 2021

IA.1 Data Sources and Coverage

We combine data from several public and private sources for this paper. Table IA.1 shows these sources and the commodities they cover. We obtain intraday trade-and-quote data from Thomson Reuters and the CME for the seven commodities described in the paper. Our soybeans data is unique in that it also includes an aggressor flag that allows us to verify our trade classification. We use this data from the beginning in 2007 (2009 for soybeans) to 2014. We also obtain intraday volume and price data at 5-minute intervals for a broader and longer sample from Barchart. This data covers 2008-2020. We obtain various daily commodity index and futures prices from Bloomberg and Barchart. We obtain FWP and 424B2 filings related to commodity-linked notes from the SEC's EDGAR database. We obtain reports on index investment from the CFTC.

IA.2 Details on Calculating Imbalance

Section 2.1 of the paper gives an overview of our procedure for calculating near month imbalance as an aggregation of all imbalance up through the active or lead month, and including any contract within three weeks of becoming the lead or active month. Here we provide details for each of our commodities. We begin by reviewing typical front and next month volumes for WTI futures and we use these volumes to illustrate our procedure.

WTI futures contracts are available for every month going out five years and for June and December delivery months going out an additional four years. Unlike stock index futures, where nearly all of the trading is in the contract with the nearest delivery dates, there is substantial trading and open interest in longer-dated WTI futures contracts. However, most of this trading in the longer-dated contracts is through calendar spread trades, wherein traders agree to simultaneously buy one maturity and sell another. Most of the trading in a single contract is concentrated in the nearer months.

We use trade and quote data from the Globex electronic trading platform for all six of the U.S. commodities in our sample. For Brent futures, we use trade and quote data from the ICE, which was fully electronic for our entire sample period. Note that the NYMEX adopted the CME Globex platform for electronic trading of the WTI contracts in June of 2006 (the CME announced its acquisition of the NYMEX in March of 2008). We are able to separately identify floor and calendar spread trades, and we exclude them from our imbalance measures.

In order to illustrate the typical pattern in trading volumes, Table IA.2 shows the WTI contract volumes (in thousands of contracts, each for 1,000 barrels of oil) for the trading days in June 2013. Table IA.2 shows that the July 2013 contract last traded on June 20, but most of the trading volume had moved to the August 2013 contract the day before that. The table also shows that calendar spread trading makes up a fairly substantial portion of the front and next month volume, and it constitutes the vast majority of trading in the remaining months. Finally, the table shows that floor trading volume is non-trivial, but is smaller than Globex volume. NYMEX suspended floor trading in WTI futures and many other futures products in July of 2015.

We exclude floor trades from the price impact analysis because they are executed manually, making it impossible to accurately align them in time with the GLOBEX quotes, and therefore impossible to assign trade direction.

We classify each Globex single-month trade as a buy or sell by comparing the price to the current quote for that contract, and we aggregate buying and selling volume by minute. We also measure the return over each minute using quote midpoints as of the end of each minute.

Globex trading in WTI futures runs from Sunday night at 6:00 p.m. to Friday night at 5:00 p.m. with one-hour breaks at 5:00 p.m. each day. The bulk of the trading occurs during the day, so when we do analysis by minute, we limit our WTI sample to the time periods from 7:30 a.m. to 4:00 p.m. New York time. This time window captures 88% of the total WTI volume in the front and next month contracts. We use this same time window for minute-by-minute analysis of Brent, gold and copper. For corn, soybeans, and wheat we use 9:30 a.m. through the close, which was 2:15 p.m. New York for most of our sample, but was delayed until 3:00 p.m. in late 2012 and early 2013.

The CME procedures for determining daily settlement prices begin by focusing the contract that generally has the highest volume. This is called the "Active Month" for WTI, gold and copper, and is called the "Lead Month" for corn, wheat and soybeans. We measure returns using the quote midpoints for the Active/Lead Month contracts. We measure imbalances using the difference between buy and sell volume for trades in all months from the front month through the month that is currently the Active/Lead Month or is within three weeks of becoming the Active/Lead Month.

WTI futures contracts are available for every calendar month out through 5 years. The Active Month in the WTI futures is the nearest month contract, except for the last two trading days prior to expiration, at which point the next month contract becomes the Active Month. Thus, referring back to Table IA.2, our return data on June 18, 2013 use the July 2013 contract and our return data on June 19, 2013 use the August 2013 contract. Our imbalance data include both the July 2013 and August 2013 through June 20, 2013, and reflect just the August 2013 contract starting June 21, 2013.

Brent futures are traded on the ICE, and are also available for every calendar month out through 5 years. The ICE does not follow the CME's Active Month approach for settlement calculations, however, similar to WTI the volume in the nearest month Brent contract falls off in the final two days prior to expiration. Accordingly, we use the same approach for Brent futures that we use for WTI and define the Active Month as the near month except for the last two trading days when we define the next month as the Active Month.

The volume patterns in the other commodities are more complex. Gold futures contracts are available for the nearest three calendar months and for all even calendar months (February, April, June, etc.) for the next two years. Although some trading occurs in odd calendar months that are close to expiration, the volume in odd expiration months is much lower than in the nearby even calendar months. In addition, volume for October tends to be lower that for the other even months. The Active Months in gold are the even months, except for October. The current Active Month is the nearest of these contracts that is not in the final calendar month of trade. For example, on February 1 the April contract becomes the Active Month. The Active Months in copper are March, May, July, September and December, and the current Active Month works the same way it does in gold. So for example, on March 1 the May contract becomes the Active Month.

Corn and wheat futures contracts are available for expirations March, May, July, September and December. Trading occurs through the business day prior to the 15th calendar day of the expiration month. For wheat, each of these months is the Lead Month until the 12th business day of the calendar month prior to expiration. For example, on the 12th business day of November, the Lead Month changes from December to March. Corn is very similar to wheat, except September is never considered the Lead Month in corn. Soybean futures expiration months are January, March, May, July, August, September, and November. As with corn and wheat, soybean futures trading occurs through the business day prior to the 15th calendar day of the expiration month, and the Lead Month rolls on the 12th business day of the calendar month prior to expiration. Lead Months in soybean futures are January, March, May, July and November, so August and September are never Lead Months.

IA.2.1 Excluding Calendar Spreads

We exclude calendar spread trades from our imbalance measure. Table IA.3 motivates this decision. The table shows imbalance measured for several contracts in the WTI including the front and next-month contract, and calendar spreads containing either of these contracts. Panel A of the table shows summary statistics for returns and the various types of imbalance. As the table shows, front month imbalance tends to be more highly volatile (indicating larger volumes as it is more often the active month) then next month, but the calendar spreads in both have even higher volume.

However, what we are most interested in what types of trades move prices. We begin with our benchmark imbalance specification which is either the front month, the next month, or a combination of the two depending on the period in the trading month.[1] Column (1) shows that a million dollars of benchmark imbalance has on average a price impact of 0.0028%, or 0.28 basis points. The R-squared from this regression suggests that this imbalance can explain 48.8% of the variation in prices. When we look at front-month only, we find that the magnitudes are larger, 0.033 basis point/million \$, but the R-sq falls to 46.4%. We get an even higher coefficient of 0.048 substantially weaker R-sq of 18.2% when we only use next month. There is a fair amount of correlation between buying in the front- and next-month contracts, and so including only one or the other biases the estimates up (see model below).

What is most important for us is that including both the front and the next month contracts together does not result in a substantial increase in explanatory power over our benchmark specification. As Column (4) shows, though both come in significant, the total R-sq only rises to 49.1%, thus supporting our decision to aggregate up the two measures to a single measure.

Columns (5) and (6) add calendar spread imbalance. Column (5) adds calendar spread buying with in the front month (with an associated sell somewhere in the future, often the next month). As the table shows, the impacts here are two orders of magnitude smaller than the impacts from

[1] Our data here start in 2010 when we have calendar spreads, so the results are slightly different those in the paper.

non-spread trades in the front month. While there is some statistical significant, including these trades does not results in a meaningful increase in R-sq (it is still 49.1%). Further including the calendar spreads that have buying in the second month adds no extra explanatory power.

The results documented here (we find similar results for other commodities) motivate our choice to both net imbalance across the front two months, and to exclude calendar spreads. Our focus here is not about index rolls and their impact on the term structure of futures (e.g. Bessembinder, Carrion, Tuttle, and Venkataraman 2016), it is instead about changes in the level of index investment impacting the level of the front of the term structure.

IA.3 Vector Autoregressions as in Hasbrouck 1991

Our primary measure of order flow impact is a univariate regression of returns on imbalance measured at the one minute horizon. This essentially treats all variation in imbalance and returns as unpredictable. To support this assumption, we follow the Vector Autoregression approach developed in Hasbrouck (1991). (Note we do not report results for soybeans in this section, but they are qualitatively similar to the other six commodities.) Specifically, assume that the (log) quote midpoint for the commodity evolves according to:

$$q_t = m_t + s_t$$

Where m_t is the "efficient price" based on all relevant information, including public announcements and order flow up to time t, and the s_t component captures transient market microstructure effects. The efficient price evolves according to:

$$m_t = m_{t-1} + w_t$$

where the increments w_t are mean zero, have variance σ_w^2, and are serially independent at all lags. The s_t process has zero unconditional mean and is jointly covariance stationary with w_t. We observe the evolution of log quote midpoints, $r_t = q_t - q_{t-1}$, and the signed order flow x_t, and following

Hasbrouck 1991 we assume these evolve according to the following VAR:

$$
\begin{aligned}
r_t &= a_1 r_{t-1} + a_2 r_{t-2} + \ldots + b_0 x_t + b_1 x_{t-1} + b_2 x_{t-2} + \ldots + v_{1,t} \\
x_t &= c r_{t-1} + c_2 r_{t-2} + \ldots + d_1 x_{t-1} + d_2 x_{t-2} + \ldots + v_{2,t}
\end{aligned}
$$

In the above VAR, $v_{1,t}$ denotes the impact of public announcements in period t and $v_{2,t}$ denotes the surprise in current period order flow, and these have variances σ_1^2 and σ_2^2, respectively. The assumption that the current period order flow does not depend on the current period public announcement allows the above VAR to be recast in the following VMA representation:

$$
\begin{aligned}
r_t &= v_{1,t} + a_1^* v_{1,t-1} + a_2^* v_{1,t-2} + \ldots + b_0^* v_{2,t} + b_1^* v_{2,t-1} + b_2^* v_{2,t-2} + \ldots \\
x_t &= c_1^* v_{1,t-1} + c_2^* v_{1,t-2} + \ldots + v_{2,t} + d_1^* v_{2,t-1} + d_2^* v_{2,t-2} + \ldots
\end{aligned}
$$

The VAR is estimated using OLS, giving the coefficients as well as estimates for σ_1^2 and σ_2^2. Then a Cholesky decomposition recovers the coefficients. This VMA representation allows for the calculation of impulse response functions. Hasbrouck 1991 shows that the fraction of the variance of the efficient price innovations w_t that is due to the innovations in the order flow is given by:

$$
R_w^2 = \frac{(\sum_{t=0}^{\infty} b_t^*)^2 \sigma_2^2}{(\sum_{t=0}^{\infty} b_t^*)^2 \sigma_2^2 + (1 + \sum_{t=1}^{\infty} a_t^*)^2 \sigma_1^2} \tag{1}
$$

When examining equity data, Hasbrouck 1991 applies the approach to trade-by-trade data, although trades within 5 seconds of each other are aggregated into a single observation. In contrast, we aggregate data into one-minute time intervals. As in Hasbrouck 1991, we set the lagged values returns and imbalances to zero at the start of each trading day. We examine three primary dimensions of liquidity based on the VAR, including: b_0 and b_0^* , which are the initial impact of order flow and the initial price impact of the unpredictable portion of order flow. Higher values for these coefficients may suggest a higher fraction of trades come from the informed, or that the information held by informed traders is more valuable, or that the market is illiquid for other reasons. $\sum b_t^*$, the permanent price impact of an innovation in order flow. We illustrate this with impulse response functions to test if the impact of an innovation in order flow is reversed in subsequent minutes. R_w^2, the fraction of the efficient price variance explained by order flow innovations (as with b_0^*, a higher value implies more information coming from trades, but this measure is relative to the amount of information that arrives through public announcements).

Table IA.4 shows the results of the regressions shown in equation (1) for the full sample. Imbalance is measured in 100s of contracts, and return is expressed in percentage to facilitate interpretation. Again, for most of the commodities, 100 contracts translates into roughly \$10 million of notional (with the exception of Corn and Wheat, where 100 contracts translates into approximately \$2.5 million of notional over the sample).

The parameter b_0 from equation the VAR is shown in the first row of each of the return columns in Table IA.4. This is the estimated response of the futures price to the order imbalance in the current minute. When the regressions from Table IA.4 are converted to the VMA representation (results not shown), we find that the values of b_0 from the VAR are very close to the values of b_0^*. This is not surprising, because as shown in the remaining rows of Table IA.4, current returns are not sensitive to past imbalances and there is only modest persistence in imbalances. The low R^2 values in the imbalance regressions indicates that most of the current minute imbalance is unpredictable.

The b_0 value of 0.033 for WTI futures shows that a minute with 100 contracts (about \$10 million notional) of buy (sell) imbalance will create a same-minute price increase (decrease) of 3.3 basis points. A roughly \$10 million dollar flow yields an impact of approximately 3 basis points for gold, similar to the WTI, but a trade of \$10 million notional value trade moves copper and corn prices approximately 10 basis points (the coefficient for corn must be multiplied by four to adjust for the lower notional value per contract). For all four of these commodities, the R^2 of these return regressions is relatively large, and results in a correspondingly high value of R_w^2 from the VMA representation, both results suggesting that order flow imbalance in these markets is playing a major role in price discovery.

To ascertain whether or not these price impacts from order flow reverse in subsequent minutes. We use the VMA representation to calculate impulse response functions. The graphs of these functions are shown in Figure IA.1. This figure plots impulse response functions for returns in response to a one standard deviation innovation in order flow and in public price news for the six commodities. The primary takeaway from these plots is that the price impacts of both order flow and public return news are mostly permanent at seven-minute horizons. For oil, gold, and copper there is essentially no reversal or continued trend in prices. For corn, wheat, and Brent there is a small reversal after a movement in prices unrelated to order flow, but for a price move corresponding to order flow we see very little reversal.

IA.4 Nonlinearities in Settlement Minute Returns

Figure IA.3 shows scatter plots of imbalance and returns in the minute prior to futures settlement for each of the seven commodities. Also presented are the linear regression line, and fitted non-parametric LOESS smoother. For all seven of the commodities, large flows generally lead to smaller impacts per dollar.

IA.5 The Accuracy of Signing Trades and Robustness with Tick-Test Trade Classification

O'Hara 2015 argues that equity markets have become a highly fragmented mix of different trading protocols and this in combination with the different speeds of various traders makes it impossible to construct a consistent picture of the state of the market, thereby making traditional imbalance calculations unreliable, especially when trying to infer the presence of informed traders. These difficulties led Easley, de Prado, and O'Hara (2016) to use futures markets as a laboratory to test alternative methods for calculating imbalances, because futures markets are centralized with a single limit book protocol. Indeed, when we compare our quote-based minute-by-minute imbalances in soybean futures with imbalances calculated using the aggressor flag for each trade, we find the correlations are well above 0.99. We recognize, however, that even in futures markets there can be issues with calculating imbalances, but given our objective to assess the impact of uninformed trading we believe that our approach is likely to overstate this impact. We provide an intuitive discussion of these ideas below, and in the next section we present a simple model of trading and returns that formalizes the intuition.

We note that trades can come from a mix of informed and uninformed traders, and the informed trades are likely to have a larger impact. We are unable to distinguish between the two types, so our estimates produce an average of the two impacts. We also note that our data only include Globex trading in the primary futures contract. To the extent that hedgers spread orders across Globex and floor trades, or use other markets, the total imbalance will be greater than the imbalance that we measure, and we are assigning the entire price impact to a subset of the imbalance. Finally, we find that a statistically significant, though economically modest, portion of imbalance is predictable at one-minute horizons, with high returns and past imbalance predicting positive future imbalance. However, we find that this again biases our estimates up, as our aggregation procedure would consider the earlier return as a result of the subsequent uninformed trading.

It is true that simple noise in the imbalance classification can result in an errors-in-variables problem that would bias our regression estimates downward.[2] One particular concern is the quote based classification method will not capture uninformed liquidity traders who place non-marketable limit orders. Easley et al. 2016 suggest that the tick test may be helpful in identifying these traders. The idea is that in order to ensure execution these traders will need to quote aggressively, so they will tend to continuously push the execution prices in the direction of their trading. For example, a large buyer will need to keep raising the bid price in order to attract sellers. Thus a series of trades at the bid at increasing prices may suggest the presence of an aggressive buyer, and these trades will be classified as buys by the tick test. Of course, a series of trades at the bid at increasing prices may also indicate a sequence of positive public signals, so using the tick test will likely produce an upwardly biased estimate of impact.

Given these issues, signing trades via the tick test provides a useful robustness check. Intuitively, considering any trade associated with a positive (negative) price move as a (buy) or sell is going to maximize the estimated impact of trading. In the next sections we report our measures of imbalance using both methods, and find that the two measures are very highly correlated. As expected the tick-test yields slightly larger impact measures, but they are very similar to those using the quote-based identification strategies.[3] Tables IA.5 and IA.6 repeat summary table for intraday analysis in the main text, as well as the univariate price impact regressions of return on imbalance but this time add-results using imbalance classified with the tick-test. As the tables show, the tick-test imbalance is highly correlated with our quote based measure, and impacts from the tick-test are

[2]We show in the next section that this will be mitigated by the fact that, measured imbalance is likely to be less volatile than true imbalance, because when there are more buys (sells) than sells (buys) in an interval, more buys (sells) are likely to be misclassified and treated as sells (buys).

[3]This is potentially due to the fact that quotes are generally only one or two ticks in our markets, making it difficult for a trader to take a position inside the best bid or offer (see Li, Wang, and Ye 2018).

very similar. We therefore, we focus on the results that use price impact estimates and imbalance from the quote-based strategy as we view them to be more accurate, but using tick-test imbalance does not materially change any of the implications of our tests.

IA.6 Scenarios for Order Flow Impact in a Simple Model

In this section we present a simple model of order flows and returns and examine how our impact estimates relate to the true impact of uninformed order flow under various scenarios. Throughout this discussion we ignore small sample issues and focus on the limiting case when the sample size is sufficiently large.

We start with a benchmark case and then consider scenarios relating to

1. Uninformed and informed order flow
2. Endogenous uninformed order flow
3. Off-exchange order flow
4. Order flow classification errors
5. Signing trades with the tick-rule
6. Using the tick-rule with microstructure effects and time-aggregation

We show that in scenarios 1-3, inclusion of these effects will bias our estimate of price impact upward, and thus lead our estimates of price impact to overstate the true price impact of uninformed trading. In contrast, in scenario 4, errors in classifying trades as buy or sell driven can lead to an underestimation of price impact, but this underestimation is bounded by the R^2 in the impact regressions. Finally, in scenario 5 we argue that using the tick-test can solve this potential issue, as signing trades in this way creates classification errors that maximize the impact estimate, and thus deliver an estimate that is the upper bound for the true impact. In scenario 6 we extend this to a more realistic setting with short-term reversals coming from microstructure dynamics of the spread, and find that for realistic distributions of imbalance, the tick-test again provides an upper bound for the true permanent impact of a trade.

In the next section we show that we obtain very similar impact estimates classifying trades via the tick-rule in our data, and thus we conclude that our price estimates are likely to overstate the actual impact of an uninformed trader.

IA.6.1 Benchmark

Consider the following simple process for returns.

$$r_t = \nu_t + \beta x_t$$

Here x_t represents order flow imbalance and ν_t represents returns coming from a public signal. For the benchmark case, we are thinking of a simple stylized model with a standard limit order book where price impacts are permanent, liquidity suppliers are competitive, buying and selling are symmetric, and there are no frictions like fixed costs of market making. In that simple case, each trade price is the new "efficient price" conditioned on the last trade sign and magnitude, so immediately after a trade the quote midpoint equals the last trade price and also equals the efficient price. Also, in this simple setup, the regression equation applies both at the trade level and aggregated across multiple trades. Later we will consider the possibility that some of the uninformed traders use limit orders at the inside of the spread, and the possibility that some of the price effects are temporary.

We assume that all right hand-side variables are drawn from symmetric distributions with mean zero. We assume that the ν_t realizations are i.i.d., and in this benchmark case that the x_t realizations are i.i.d.

Suppose that we are also able to sign trades with perfect accuracy, so that our observed order flow is $\hat{x}_t = x_t$. In this simple case, if we regress r_t on $\hat{x}_t$ as we do in our impact regressions, we will infer an estimated beta $\hat{\beta} = \beta$, and our empirical estimate of impact will be consistent.

IA.6.2 Scenario 1: Uninformed and informed order flow

Now consider a process where there are both informed and uninformed traders, and that market participants can differentiate between them, but the econometrician cannot.

$$r_t = \nu_t + \beta_{inf} x_{inf,t} + \beta_{uninf} x_{uninf,t}$$

Here assume that both types of order flow are uncorrelated with each other and the public signal, and that $\beta_{inf} > \beta_{uninf}$. Suppose again that we can sign trades again with perfect accuracy, so our estimate of order flow is $\hat{x}_t = x_{inf,t} + x_{uninf,t}$. In this case our impact estimate will be

$$\hat{\beta} = \frac{\sigma^2_{x,inf}\beta_{inf} + \sigma^2_{x,uninf}\beta_{uninf}}{\sigma^2_{x,inf} + \sigma^2_{x,uninf}}$$

Since this estimate reflects both uninformed and informed trading, it will be an upwardly biased estimate of the impact of uninformed traders.

IA.6.3 Scenario 2: Endogenous uninformed order flow

In our estimates in section IA.1, we find that a statistically significant, though economically modest, portion of order flows is predictable at one-minute horizons, with high returns and past imbalance predicting high future imbalance. This suggests that within a minute, uninformed order flows may be correlated with the realization of returns coming from either informed order flow or the public signal. We focus on the public signal. We start by again specifying a return process

$$r_t = \nu_t + \beta_{inf} x_{inf,t} + \beta_{uninf} x_{uninf,t}$$

But we now allow for positive correlations by specifying that $x_{uninf,t} = z_t + \gamma_\nu \nu_t + \gamma_{inf} x_{inf,t}$ with $\gamma_\nu > 0$ and $\gamma_{inf} > 0$. In this case, observed imbalance is the same as in the previous scenario. The estimate of impact is

$$\hat{\beta} = \frac{\sigma^2_{x,inf}(1+\gamma_{inf})(\beta_{inf} + \gamma_\nu \beta_{uninf}) + \sigma^2_z \beta_{uninf} + \sigma^2_\nu \gamma^2_\nu (1+\beta_{uninf})}{(1+\gamma_{inf})^2 \sigma^2_{x,inf} + \sigma^2_z + \sigma^2_\nu \gamma^2_\nu}$$

Consider now the case where there is no informed order flow to focus on the effect of γ_ν. Then

$$\hat{\beta} = \beta_{uninf} + \frac{\sigma^2_\nu \gamma^2_\nu}{\sigma^2_z + \sigma^2_\nu \gamma^2_\nu}$$

The endogeneity here induces extra covariance between the uninformed order flow and returns, and again biases the impact upwards.

Now consider the case where there is no public information to focus on the effect of γ_{inf}. Here we have

$$\hat{\beta} = \frac{\sigma^2_{x,inf}(1+\gamma_{inf})(\beta_{inf}+\gamma_\nu\beta_{uninf}) + \sigma_z^2\beta_{uninf}}{(1+\gamma_{inf})^2\sigma^2_{x,inf}+\sigma_z^2} = \beta_{uninf} + \frac{\sigma^2_{x,inf}(1+\gamma_{inf})(\beta_{inf}-\beta_{uninf})}{(1+\gamma_{inf})^2\sigma^2_{x,inf}+\sigma_z^2}$$

The correlation here leads to larger amounts of uninformed volume to mask informed flows, so the upward bias of the previous scenario is lessened, but the estimate will still be higher than the true uninformed impact.

IA.6.4 Scenario 3: Off-exchange volume

While the Globex is a large portion of total CME futures volume, in all of our commodities there is substantial floor trading over most of the sample period. In addition, there may be close substitutes in other futures markets. For example, we capture the COMEX copper contract, but for most of the period the LME contract has slightly higher volume. It is reasonable to assume that a trader, such as a commodity index fund, might spread orders across both the Globex and the floor to minimize impact. To see how this effects our estimates consider

$$r_t = \nu_t + \beta x_{on,t} + \beta x_{off,t}$$

We assume that the correlation of order flow ($\rho_{on,off} > 0$) is positive to represent orders being routed to both trading venues. Suppose again that we can sign trades with perfect accuracy, so our estimate of order flow is simply the imbalance on the observed exchange $\hat{x}_t = x_{on,t}$. In this case our impact estimate will be

$$\hat{\beta} = \beta\left(1 + \frac{\rho_{on,off}\sigma_{x,off}}{\sigma_{x,on}}\right)$$

Since this estimate is based on only a subset of true order flows, it will again be an upwardly biased estimate of the impact of uninformed traders.

IA.6.5 Scenario 4: Mis-classified trades

Our quote based method of classification should allow us to essentially recover the "true" buyer or seller as identified by the "aggressor" party that crossed the spread to trade. However, there is still a worry that a buy trade is executed by posting a limit-buy order which is then hit by a market-maker or HFT. We note that this is likely to be a small set of trades, as such strategies would require small tick sizes relative to the spread to execute. In our commodity futures markets spreads are usually only one, and seldom more than two, ticks wide in the liquid periods around the close for all of commodities. Nevertheless it is important to understand how this may affect our estimates, as this is a scenario which can potentially lead to erroneously small estimates of impact.

Therefore consider again the original benchmark model

$$r_t = \nu_t + \beta x_t$$

but assume that imbalance is observed with classification error $\hat{x}_t = x_t + \epsilon_t$. One of the key features of our setup is that there is likely negative correlation between the error term and x_t. This is mechanical, as any mis-classified buy will considered a sell, so a period with many buys would likely have a negative error term. In this case, we have

$$\begin{aligned}\hat{\beta} &= \beta \frac{\sigma_x^2 + \rho_{\epsilon,x}\sigma_x\sigma_\epsilon}{\sigma_x^2 + \sigma_\epsilon^2 + 2\rho_{\epsilon,x}\sigma_x\sigma_\epsilon} \\ &= \beta\left(1 + \frac{-\rho_{\epsilon,x}\sigma_x\sigma_\epsilon + \sigma_\epsilon^2}{\sigma_x^2 + \sigma_\epsilon^2 + 2\rho_{\epsilon,x}\sigma_x\sigma_\epsilon}\right)\end{aligned}$$

Therefore, while measurement errors will bias our estimates downward, the likely negative correlation between our mis-classification and true imbalance will have the opposite effect.

It can also be shown via simple algebra that the empirical R^2 in this model is

$$R^2 = \frac{\hat{\beta}}{\beta}\left(\frac{\beta^2\sigma_x^2 + \beta^2\rho_{\epsilon,x}\sigma_\epsilon\sigma_x}{\beta^2\sigma_x^2 + \sigma_\nu^2}\right)$$

Consider in the extreme case that $\rho_{\epsilon,x} = 0$ and $\sigma_\nu = 0$, then we have $R^2 = \frac{\hat{\beta}}{\beta}$, and the R^2 is a bound on the understatement of impact. Given that our R^2 are typically between 15% and 35%, this implies that this sort of mis-classification can only induce an estimated impact of 15% or 35% the true value, and this only if there is no variation from public information and there is no negative correlation between the errors and the actual imbalance.

While it seems plausible that much of the variation is coming from information not revealed from trades (either public information or trades on other venues) it is nevertheless hard to estimate precisely the degree of this bias. We therefore consider an alternate, tick-based classification, and consider its implications next.

IA.6.6 Scenario 5: Tick Classification

Consider again the benchmark model.

$$r_t = \nu_t + \beta x_t$$

In this case however, suppose that t indexes individual trades, and consider the implications of performing the tick-test classification. The econometrician observes the magnitude $|x_t|$, and the sign is then assigned as positive if $r_t > 0$ and negative if $r_t < 0$. This then gives observed imbalance

$$\begin{aligned}
\hat{x}_t &= -x_t && \text{if: } \nu_t > \beta x_t \text{ and } x_t < 0 && \textbf{MC as buy} \\
\hat{x}_t &= -x_t && \text{if: } \nu_t < \beta x_t \text{ and } x_t > 0 && \textbf{MC as sell} \\
\hat{x}_t &= x_t && \text{else} && \textbf{Correct}
\end{aligned}$$

Given this measure of imbalance, our OLS strategy yields an impact estimate of

$$\hat{\beta} = \beta + 2p_{\mathbf{MC}} \frac{E_{\mathbf{MC}}\left[|x_t r_t|\right]}{Var(x_t)}$$

where $p_{\mathbf{MC}}$ is the probability of a mis-classification, and $E_{\mathbf{MC}}$ is the expected value conditional on mis-classification. The positive values of the expectation lead to an overstated measure of price impact.

In essence, assuming that any trade that corresponded with an up (down) move in price is a buy (sell), is the most conservative approach to avoid understating impacts, and therefore yields an upper bound for the level of impact from a given level of investment. Intuitively, assuming that a given trade caused whatever movement coincided with the trade, even if that movement is the result of public information, leads to an overstatement of any impact.

However, this simple model assumes that all trade impact is immediate and permanent. In reality, when we sign trades with the tick method, we would often sign as a buy a trade near the ask, and if quotes do not react to each trade then the subsequent trade is more likely to be classified as a sell due to short-term reversal. Therefore we next consider a model to examine the implications of these effects.

IA.6.7 Scenario 6: Tick classification w/ microstructure effects and aggregation

Suppose that liquidity supply includes some combination of carrying costs and monopoly power, so that the initial trade price impact is larger than the permanent price impact. Specifically, let trade-price returns at a trade-time horizon evolve according to

$$r_t = \nu_t + \beta x_t - \Gamma \beta x_{t-1}$$

Where $\Gamma \in (0, 1)$ captures short term reversals due to liquidity and the dynamics of the spread. The permanent impact of imbalance is therefore $\beta(1 - \Gamma)$. We assume quote midpoints continue to capture only the permanent price impacts.

If we could observe x_t and estimated impact via OLS at the trade horizon, our estimate would $\hat{\beta} = \beta$, which would be an overstatement of the true impact.

However, our approach aggregates up all trades in a minute. To see the effect of this aggregation, assume that each minute m consists of N trades.

The trade-price return for a minute is

$$R_m = \sum_{s=1}^{N} r_s = \sum_{s=1}^{N} [\nu_s + \beta x_s(1 - \Gamma)] - \Gamma(\beta x_0 - \beta x_N)$$

Here the last term comes from the serial correlation generated by observations at the start and end of the aggregation period.

Since x_0 is uncorrelated with $\sum_{s=1}^{N} x_s = X_m$, but x_N is positively correlated, the last term will lead to an overstatement of impact (we miss the reversal of the last trade) in a regression using trade prices, but we use quote midpoints for the aggregate returns over the minute, so we have

$$R_m = \nu_m + \beta(1-\Gamma)X_m$$

It is clear here that if we can correctly sign trades to recover X_m, our impact estimate will be equal to the permanent impact $\beta(1-\Gamma)$. We now see how applying the tick-test impacts this estimate, and we will show once again that it biases our estimates upward above the true impact.

We use the tick test to sign trades as in the previous scenario, however a complication arises because the short-term reversal from the Γx_{t-1} term now impacts the signing of the trades.

$\hat{x}_t = -x_t$	if: $\nu_t - \beta\Gamma x_{t-1} > \beta x_t$ and $x_t < 0$	**MC as buy**
$\hat{x}_t = -x_t$	if: $\nu_t - \beta\Gamma x_{t-1} < \beta x_t$ and $x_t > 0$	**MC as sell**
$\hat{x}_t = x_t$	else	**Correct**

After signing trades, we then sum up to calculate per-minute imbalance, and we likewise calculate per-minute return over the period.

Observed imbalance with the tick rule is

$$\hat{X}_m = \sum_{s=1}^{N} (x_s + 2(\mathbb{1}_{MCBuy} x_s - \mathbb{1}_{MCSell} x_s))$$

Define

$$\pi_t = \hat{x}_t - x_t$$
$$\Pi_m = \sum_{s=1}^{N} \pi_t = \hat{X}_m - X_m$$

Here π_t is the classification error for each trade. (e.g. difference between the observed imbalance from the tick-test and the true imbalance). Note that, conditional on a misclassification, $|\pi_t| = 2|x_t|$. Π_t is the sum of these errors for each minute.

The OLS regression coefficient of our strategy is then

$$
\begin{aligned}
\hat{\beta} &= \frac{Cov(R_m, \hat{X}_m)}{Var(\hat{X}_m)} \\
&= \frac{\beta(1-\Gamma)Var(X_m) + \beta(1-\Gamma)Cov(\Pi_m, X_m) + Cov(\nu_m, \Pi_m)}{Var(X_m) + Var(\Pi_m) + 2Cov(X_m, \Pi_m)} \\
&= \beta(1-\Gamma)\left(1 + \frac{\frac{Cov(\nu_m,\Pi_m)}{\beta(1-\Gamma)} - Cov(\Pi_m, X_m) - Var(\Pi_m)}{Var(X_m) + Var(\Pi_m) + 2Cov(X_m, \Pi_m)}\right)
\end{aligned}
$$

Where the second equation follows from $R_m = \nu_m + \beta(1-\Gamma)X_m$ and $\hat{X}_m = X_m + \Pi_m$, and the final equation is obtained via some simple algebra. Since the denominator is the variance of observed imbalance, and is therefore positive, we have

$$
\hat{\beta} \geq \beta(1-\Gamma) \iff \frac{Cov(\nu_m, \Pi_m)}{\beta(1-\Gamma)} - Cov(\Pi_m, X_m) - Var(\Pi_m) \geq 0
$$

The first term captures again the fact that our impact measure using the tick-test will be an overstatement of true permanent impact if our signing errors covary positively with the arrival of public information, (which is the case by the same argument as the previous scenario). The second term comes from the fact that microstructure reversals induce a negative correlation between mis-classifications and aggregate returns. For instance, a period with many buy trades will see many negative reversals, and therefore a larger number of subsequent buys mis-classified as sells. This negative covariance will again bias our estimate up. The last term is a measurement-error term, and will bias our estimates down.

Focusing on the last two terms, and noting that all variables are mean zero, and that covariances are zero beyond one lag, we have

$$
\begin{aligned}
Var(\Pi_m) &= Np_{MC}E_{\mathbf{MC}}(4x_t^2) + NCov(\pi_t, \pi_{t-1}) \\
&= Np_{MC}E_{\mathbf{MC}}(4x_t^2) + Np_{MCMC}E_{\mathbf{MCMC}}(4x_t x_{t-1})
\end{aligned}
$$

Where $MCMC$ identifies observations where both the current trade and the previous trade are mis-classified. We also have

$$Cov(\Pi_m, X_m) = Np_{MC}E_{\mathbf{MC}}\left(-2x_t^2 - 2x_t x_{t-1}\right)$$
$$Cov(\Pi_m, X_m) = Np_{MC}E_{\mathbf{MC}}\left(-4x_t^2 - 2x_t(x_{t-1} - x_t)\right)$$

So that

$$-Cov(\Pi_m, X_m) - Var(\Pi_m) = N\left[p_{MC}E_{\mathbf{MC}}\left(2x_t(x_{t-1} - x_t)\right) - p_{MCMC}E_{\mathbf{MCMC}}(4x_t x_{t-1})\right]$$

Note that, if $\sigma_\nu = 0$, the only way there can be a mis-classification is if there are two consecutive trades in the same direction with the second trade being of smaller magnitude than the first. For instance, a large buy generates a negative subsequent return which leads to the mis-signing of a following buy (this is necessary but not sufficient). Therefore the first expectation on the right-hand size is positive. To generate two consecutive mis-classifications requires three trades in the same direction of descending magnitudes (an occurrence that happens one out of 24 sequences of three trades), and the second term on the right will typically be a second-order effect (In fact $\frac{1}{6}p_{MC} > p_{MCMC}$).

It is possible to choose probability distributions for x_t that will generate a slight understatement of permanent impact for values of Γ very close to one (i.e. highly discrete bimodal distributions so that the measurement-error term dominates). However, we find in simulations with reasonable distributions that the tick-test method will overstate impacts due to the induced negative covariance between returns and identified imbalance. Figure IA.2 demonstrates this. We simulate 10,000 minutes with 10,000 trades in each minute for imbalance drawn from the normal, logistic, and Laplace distributions (the latter two distributions are platykurtic consistent with the data). Each plot shows the ratio of our estimated imbalance method using the tick-test relative to the true permanent impact ($\frac{\hat{\beta}}{\beta(1-\Gamma)}$) as a function of Γ, and the amount of public information. For the plots, we vary σ_ν to achieve a target R^2 coming from public information. This R^2 is varied from zero ($\sigma_\nu = 0$) to 0.8. As the plot shows, the estimated impact is equal to the true permanent impact in the case where there $\Gamma = 0$ and $\sigma_\nu = 0$, but otherwise $\hat{\beta}$ is an overestimate of the true impact, even in the cases where $\sigma_\nu = 0$.

IA.7 Index Flows and Imbalance in Small and Large Trades

Table IA.7 repeats the regressions of imbalance on Index Flows shown in the main text in Panel A, and Panel's B and C repeat these regressions but decompose imbalance into small trades of a single contract (Panel B), and larger trades (Panel C). As the table shows, the relation between imbalance and index flows is primarily present in large trades. This is potentially consistent with index traders being executed by larger institutions (e.g. Kaniel, Saar, and Titman 2008).

IA.8 Index Flows, Managed Money Flows, and Momentum Trading across Subperiods

In the data we show a strong relation between daily returns and index flows that is at odds with prior literature. We also show that controlling for managed money flows largely eliminates the relation between index flows and returns over the full day, and we speculate that this result may reflect momentum trading by at least some managed money and index traders, which becomes more pronounced in the later part of the sample. Table IA.8 presents regressions that provide further evidence consistent with this hypothesis. In this table we split the sample into roughly two halves, with 2007 - 2014 shown in Panel A, and 2015 - 2020 shown in Panel B. In the first two columns, we repeat the pooled cross-sectional regressions of full day return on index flows, and find positive and significant results for both periods, though as in the paper, this effect is much stronger in the second half of the sample. In column (2), we add a control for managed money flows. As the table shows, this greatly reduces the daily relation in the second half of the sample, suggesting that in the relation between daily returns and index flows is potentially driven by traders classified as CITs following strategies similar to those of managed money traders. Columns (3) - (6) split the full day into the 30 minutes prior to settle and the remainder of the day, and show that these patterns are entirely driven by the early portion of the day, and that controlling for managed money has almost no effect on the relation between index flows and returns near the futures settlement.

These patterns are broadly consistent with a shift of some of the classified index traders into momentum strategies in the second half of the sample. Columns (7) - (9) test for this directly by regressing index flows on the previous week's return and managed money flows. As the table shows, there is little relation between index trader flows and managed money or previous weeks return in the first half of the sample. However in the second half of the sample, the previous week's return strongly predicts index flows in a manner consistent with momentum trading. This effect is also significantly reduced controlling for managed money suggesting that managed money traders are also following momentum strategies (this is confirmed in column (10) for both periods). Given that our trader position data aggregates up to weekly frequency, any momentum trading that occurs within the week will show up as a contemporaneous correlation between flows and returns. However, since managed money traders are following momentum strategies, they should provide a reasonable control for this effect. It is therefore quite interesting that after controlling for managed money, we see almost no effect on the coefficient between index flows and returns near the settlement, suggesting that this is not a mechanical relation driven by momentum trading.

IA.9 Sample Calculation of CLN Delta

Figure IA.4 uses a representative note from the sample to illustrate the calculation of the pricing date and determination date deltas. The figure illustrates the $51,437,000 Capped Market Plus Notes linked to the S&P GSCI® Crude Oil Excess Return Index that were issued on January 24, 2011 by Barclays Bank. This note is typical in that it has no payments prior to maturity and the ending return on the note is a piece-wise linear function of the return on the underlying. The note also has a path-dependent "knock-out", which is another common feature in our sample. The note "priced" based on the closing value of the index on January 14, 2011 and the 424B form was filed with the SEC on January 19, 2011. When calculating the initial delta we follow HPW and assume that the full value of the notes was committed and hedged on January 14.

The notes matured on February 1, 2012. The determination date was January 25, 2012, when the final value of the index was observed and payoff of the note was set. If the notes were hedged, then the hedge should have been removed on the determination date, so we calculate the ending delta on that date. These notes have a knock-out buffer, a contingent minimum return, and a maximum return. A knock-out occurs if the index value falls below 80% of the pricing date value on any day over the life of the notes, and if a knock-out occurs then the contingent minimum of 8% is removed.

Panel A shows the actual return path for the index and three hypothetical return paths.[4] The hypothetical return paths are shown in part to illustrate the 10,000 simulated paths that are used to value the note and calculate the delta on the pricing date, and they are also used to illustrate the possible determination date deltas in Panel B. The initial note value is the average of the risk-neutral present values of the ultimate payments to the note along each simulated path based on the specific terms of the note, including interim interest payments and early calls.[5] The initial delta is then calculated by revaluing the note with a small change in the initial value of the underlying. The pricing date delta for this particular note is 0.89, so the size of the delta hedging trade would be 89% of the face value. Because most of the notes have concave payoffs with maximum slopes less than or equal to 1.0, the average (median) delta for the notes in our sample is only 0.61 (0.63). Accordingly, the face value of a note generally overstates the amount of the hypothetical initial hedge.

As shown in Panel A, the realized path for the underlying index was below the knock-out level during the life of the note, so as shown in Panel B, the return on the note matched the realized return on the index. The final return to the notes was -1% giving a final value of $50.9 million. The payoff function had a slope of 1.0 on the determination date, so the final delta is equal to the final value of $50.9 million divided by the initial issue amount (giving a delta of 0.99). The hypothetical price path A for the underlying ends with the same return as the actual price path, but it never falls into the knock-out region so the ending return on the notes would have been the contingent minimum of 8%. The slope of the payoff is zero in this case, so if this had been the actual path, the ending delta would have been zero. The hypothetical price path B has an ending return for both the notes and the underlying of 15% which would have meant an ending value of $59.2 million for the notes, and since the slope of the payoff is 1.0 the size of the delta hedge would also be $59.2 million, or 115% of the face value, so the delta would be 1.15. The hypothetical price

[4]For all of the notes, we obtain the values of the specific index or spot price to calculate the actual path. These data are obtained either via Bloomberg or the Commodity Research Bureau.

[5]The underlying index is an excess return, so the simulated risk-neutral paths used in the valuation of this note have zero expected return.

path C has a 40% return on the underlying, which would mean the note return would have been the maximum return of 30.5%. The payoff function has a zero slope at this point, so hypothetical path C would have resulted in an ending delta of zero. In our full sample, 594 notes (out of the 597 issued through January 2014) had matured by the end of 2018. Of these, 342 of had a delta of zero on the determination date.

IA.10 Comparison of Determination Date Sample with Henderson, Pearson, and Wang (2015)

As mentioned in the main text, we do not find a significant average negative determination date return for any subset of our notes. This is contrast to the published result of Henderson, Pearson, and Wang 2015, who find a return of -42 basis points (t-stat of 2.50) on the determination dates of 42 notes with greater than $10 million outside of the Goldman Roll. They find a two-day return, that includes both the determination day and the following day, of -39 basis points (t-stat of 1.80).

We contacted HPW noting our inability to replicate this finding, and they re-examined their set of determination days. After comparing their results with our data, HPW concluded that of their original 42 days, 10 were mistakenly included, and 24 were mistakenly excluded, and therefore they produced a refined set of 56 days. They find determination day returns of -15 basis points (t-stat of 0.76) on these dates. However, when they look at the two-day window including the day after the determination date, this effect rises to -51 basis points (t-stat of 1.78).

After reviewing the refined sample of HPW, we disagree with the some of the choices made about notes in the sample. After considering the refined set of HPW, we arrived at our set of 54 notes (we had 50 in our original set, due to four where we missed the 424b filings), which excludes 3 notes in HPW's set, and includes one extra note. This is the set that is shown in column (2) of Panel D in Table 8 in the main text. Table IA.9 shows the results of Table 8 but recalculated using the two-day window, and shows that we still find no significant result. Our two-day return is -42 basis points (t-stat of 1.44) for the relevant subsample.

Table IA.10 shows the discrepancies between the two subsamples. There are two gold notes that seem to us to be clearly in the Goldman Roll period, and therefore should be excluded. These do not have much an effect on the result. Two other notes drive most the discrepancy in our return result.

The first is a Natural Gas note linked to the price of the U.S. Natural Gas ETF (ticker:UNG). We include such notes (for instance there are several linked to the SPDR Gold ETF (ticker:GLD)). Commodity linked notes typically link to various proprietary commodity indices rather than an explicit futures contract. We do not believe that being linked to a separately traded price that tracks a single commodity should disqualify a note, particularly since authorized participants can deliver futures to create shares of the underlying ETF.

The second note is a "daily liquidity note" issued by J.P. Morgan linked to corn.[6] These notes are issued for a given face value, but are not fully sold initially. Furthermore, the issuer offers to buy the note back at a market value (calculated using the underlying commodity price) at any point during the life of the note. In this sense, they behave more like an ETN than a standard note, so we do not include them in our analysis. For this particular note, the filing specifies that $2.5 million was sold, so it may have been issued, but it was certainly not a large note, and there was no guarantee that it was not sold back to the issuer prior to the determination date.

The determination date of the natural gas note coincided with a large positive return on both the determination date and the following day, and the corn note coincided with a large negative return on the day after the determination date. Therefore our inclusion of the natural gas note and exclusion of the corn note leads to a less negative estimate of average return.

IA.11 Regressions Used to Calculate Abnormal Imbalances Around CLN Pricings and Determinations

To control for any average directional imbalance occurring at a given time in the trading day we use a regression specification to calculate an abnormal imbalance. The goal is to control for the typical imbalance in each time window for each commodity. For some commodities, different CLNs are priced at different times of the day. For example, while most gold CLNs are priced based on the London 3:00 p.m. Fix (which 10:00 a.m. New York time for most of the year), some are linked to a futures index (settled at 1:30 p.m. New York time), and others are linked to the NYSE closing price of the GLD exchange-traded fund (4:00 p.m. New York time).

[6] See filing: https://www.sec.gov/Archives/edgar/data/19617/000089109210002989/e39481_424b2.htm.

To make the description more concrete, consider the regression that measures the abnormal imbalance in the five minutes leading up to a CLN pricing. The observations in this regression include the total dollar imbalance for gold futures measured over each day's five-minute interval leading up to the London Fix, the total of this imbalance for each day's five-minute interval leading up to the gold futures settlement, and the total of this imbalance for each day's five-minute interval leading up to the NYSE close (so three gold futures imbalance observations for each day in the sample). Similarly, the regression includes the daily five-minute futures imbalances for the other commodities measured over each interval where there is at least one CLN pricing for that commodity. To control for any intraday pattern in imbalances for each commodity, the regression includes fixed effects for each commodity/time interval combination. Finally, the regression includes a dummy variable set equal to one if there is a CLN priced at the end of that five-minute interval on that particular day. We report the coefficients and p-values for this dummy variable.

The regressions for the other time intervals are analogous to the regression for the five minutes leading up to the pricing that is described above. The observations include total dollar imbalances for all time intervals each day in each commodity if there is any CLN pricing for that interval in that commodity across all of the days in the sample. Each regression includes a dummy variable set equal to one if the interval is associated with a CLN pricing on that day.

References

Bessembinder, H., A. Carrion, L. Tuttle, and K. Venkataraman. 2016. Liquidity, resiliency and market quality around predictable trades: Theory and evidence. *Journal of Financial Economics* 121:142–166.

Easley, D., M. L. de Prado, and M. O'Hara. 2016. Discerning information from trade data. *Journal of Financial Economics* 120:269–285.

Hasbrouck, J. 1991. The summary informativeness of stock trades: An econometric analysis. *The Review of Financial Studies* 4:571–595.

Henderson, B. J., N. D. Pearson, and L. Wang. 2015. New evidence on the financialization of commodity markets. *The Review of Financial Studies* 28:1285–1311.

Kaniel, R., G. Saar, and S. Titman. 2008. Individual investor trading and stock returns. *The Journal of finance* 63:273–310.

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

Li, S., X. Wang, and M. Ye. 2018. Who Supplies Liquidity, and When? *Available at SSRN 3231166* .

O'Hara, M. 2015. High frequency market microstructure. *Journal of Financial Economics* 116:257–270.

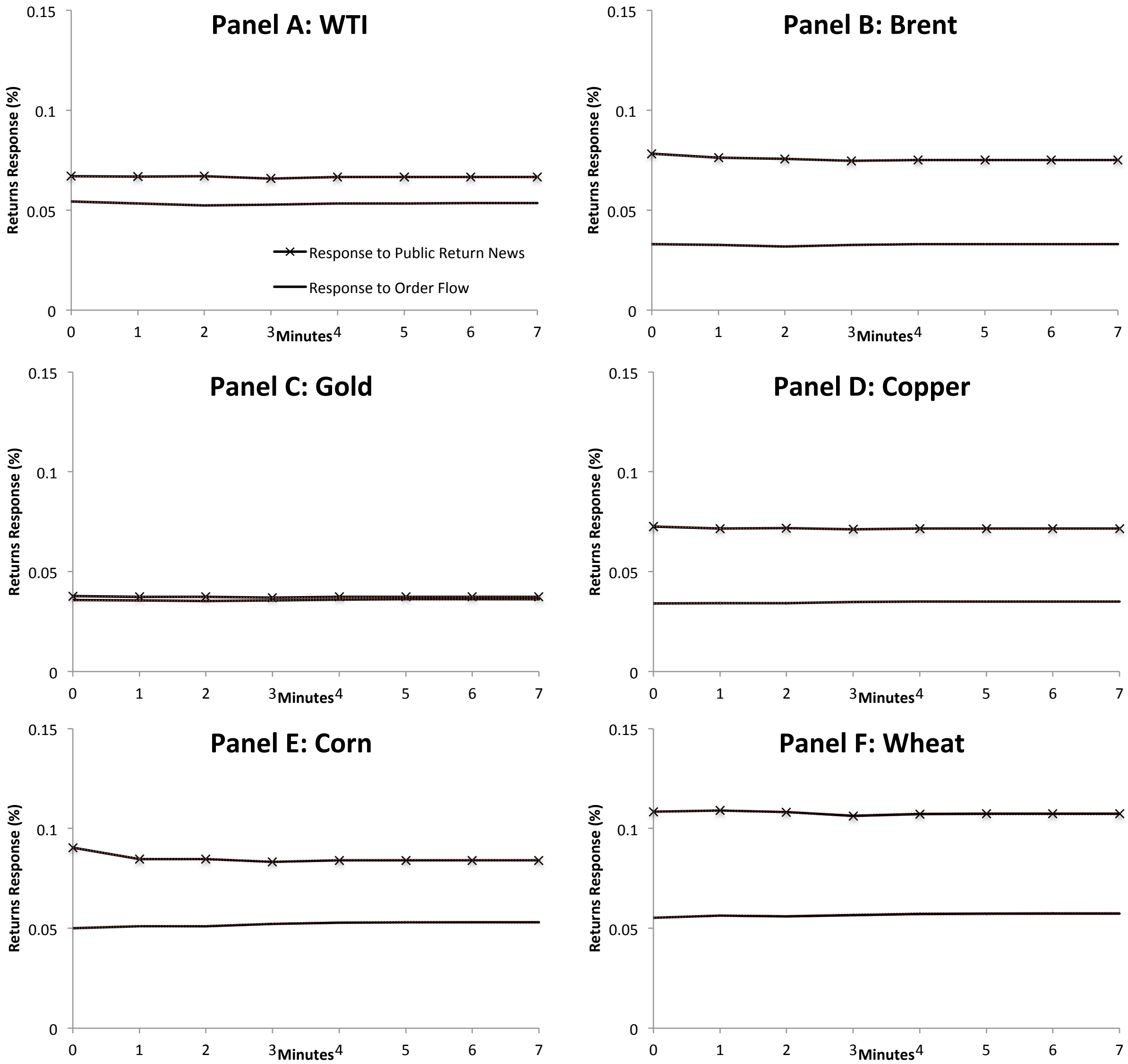


Figure IA.1. Return Impulse Response Functions for VARs
The figure shows the impulse response of returns to innovations in order flow and public news from the vector autoregression specification estimated in Table 3. Plots show return responses to one standard deviation innovations in public return news and unanticipated order flow.

Figure IA.2. Simulated tick-test impacts with microstructure effects and time-aggregation

The figure shows the ratio of estimated impacts to the true permanent impact ($\frac{\hat{\beta}}{\beta(1-\Gamma)}$) coming from simulations of the model described in section IA.6.7. The model is simulated for 10,000 minutes each with 10,000 trades, each trade is signed with tick-test and then imbalance and returns are aggregated up to the one minute level. The estimated impact of trading ($\hat{\beta}$) is then calculated via an OLS regression of minute-by-minute returns on minute-by-minute imbalance. Each plot shows the ratio of estimated impact to true permanent impact for various values of short-term reversal (Γ) and for different amounts of return variance coming from public information. Different panels show results of simulations drawing imbalance from different probability distributions.

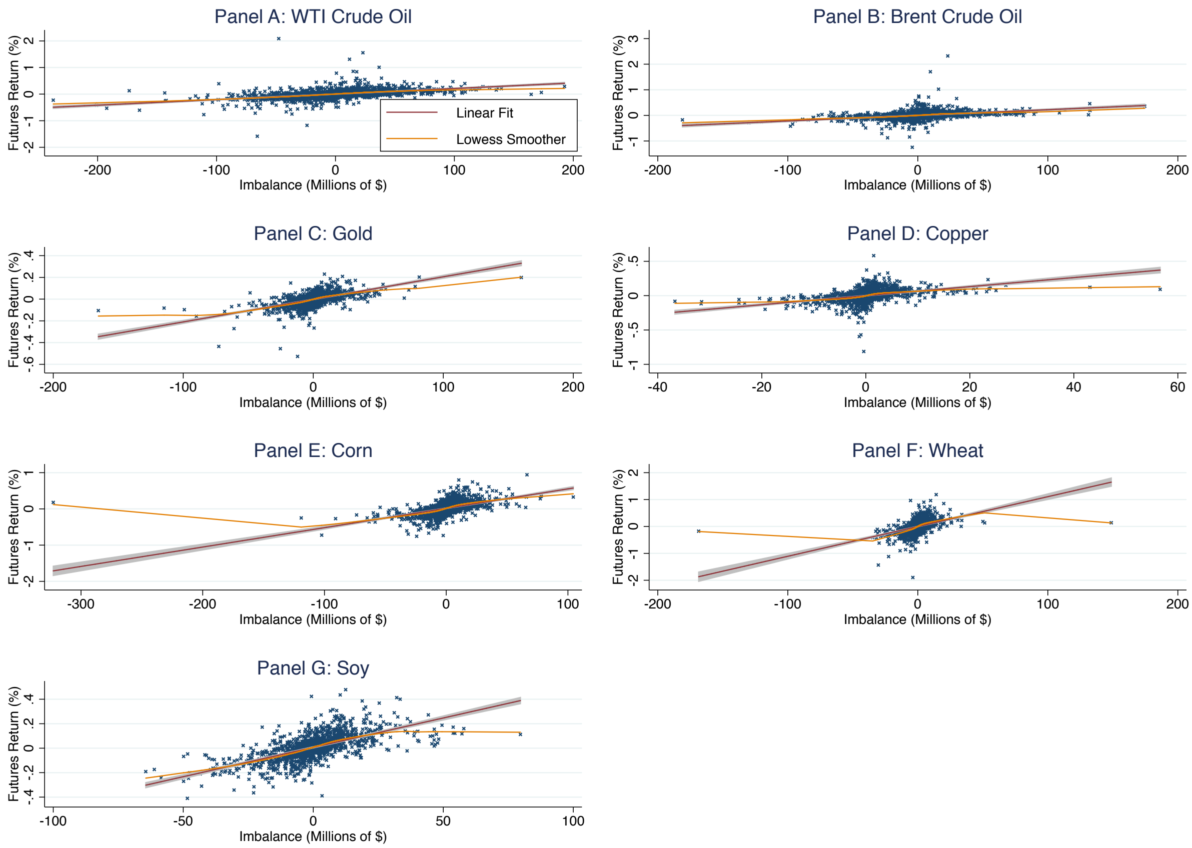
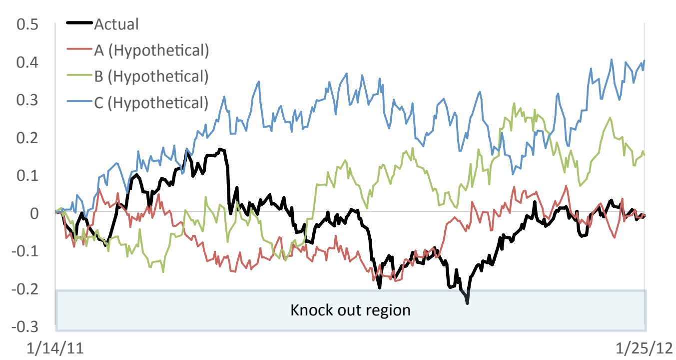


Figure IA.3. Nonlinearities in imbalance and raturns in minute prior to futures settlement

The figure shows scatter plots of order imbalance (in millions of \$) and return (in %) in the minute prior to settlement for each day across the sample. The shaded line shows linear fit and confidence interval. The single line shows a second-order LOESS smoother calculated using a tricube kernel with $\alpha = 0.8$. Data are 1/1/2007 to 4/1/2014. We exclude data prior to 1/1/2008 for Brent. Soy data begin in 2009.

Panel A: Cumulative Return Paths for Underlying Index



Panel B: CLN Return as a Function of Underlying Index

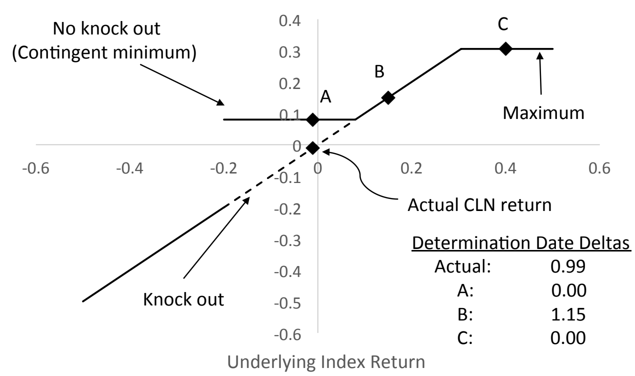


Figure IA.4. Return Paths and Determination Date Deltas for a Sample CLN
The figure illustrates the \$51,437,000 Capped Market Plus Notes linked to the S&P GSCI® Crude Oil Excess Return Index. These notes have a knock-out buffer, a contingent minimum return, and a maximum return. A knock-out occurs if the index value falls below 80% of the pricing date value on any day over the life of the notes, and if a knock-out occurs then the contingent minimum of 8% is removed. Panel A shows the actual return path for the index and three hypothetical return paths. Panel B shows the piecewise linear payoff structure across the ending cumulative returns of the underlying index along with the determination date delta for each path. The delta is calculated as (Ending Note Value × Slope of Payoff on Determination Date) / (Face Value of the Note). See section IA.9 for a detailed explanation of the four determination date delta values.

Table IA.1. Data Coverage by Commodity

The table shows the various data sources used for each commodity.

	Trade and Quote	Five Minute Futures Prices	Daily Futures Price/Volume	CFTC SCOT Report	CFTC DCOT Report	CLN SEC Filings	CFTC Index Trader Report
Frequency:	Intraday	Intraday	Daily	Weekly	Weekly	Daily	Monthly
Start:	2008	2008	2003	2008	2008	2003	2010
Stop:	2014	2020	2020	2020	2020	2014	2015
Agricultural							
Cocoa		✓	✓	✓	✓		✓
Coffee		✓	✓	✓	✓		✓
Corn	✓	✓	✓	✓	✓	✓	✓
Coffee		✓	✓	✓	✓		✓
Feeder Cattle		✓	✓	✓	✓		✓
Live Cattle		✓	✓	✓	✓		✓
Lean Hogs		✓	✓	✓	✓		✓
Soybeans	✓	✓	✓	✓	✓	✓	✓
Soybean Meal		✓	✓	✓	✓		✓
Soybean Oil		✓	✓	✓	✓		✓
Sugar		✓	✓	✓	✓		✓
Hard Wheat		✓	✓	✓	✓		✓
Soft Wheat	✓	✓	✓	✓	✓		✓
Energy							
Brent Crude	✓	✓	✓			✓	
WTI Crude	✓	✓	✓		✓	✓	✓
Heating Oil		✓	✓		✓		✓
RBOB Gasoline		✓	✓		✓	✓	✓
Natural Gas		✓	✓		✓	✓	✓
Metals							
Gold	✓	✓	✓		✓	✓	✓
CME Copper	✓	✓	✓		✓	✓	✓
LME Copper		✓	✓			✓	
Platinum		✓	✓		✓	✓	
Palladium		✓	✓		✓	✓	
Silver		✓	✓		✓	✓	✓
Nickel			✓			✓	
Lead			✓			✓	
Aluminum			✓			✓	
Tin			✓			✓	

Table IA.2. Daily WTI futures volumes for June 2013

The table shows CME WTI futures volumes in thousands of contracts for the trading days in June 2013. The table includes Globex single month trades, Globex spread trades (nearly all are calendar spread trades), and floor trades. Volumes are shown separately for the July and August 2013 contracts, and the final columns show all other contracts. The totals exclude Privately Negotiated Trades (PNT) and Globex Trade at Settlement (TAS) trades.

	July 2013 Contract			August 2013 Contract			All other contracts		
	Globex			Globex			Globex		
Trade Date	Single Month	Spread	Floor	Single Month	Spread	Floor	Single Month	Spread	Floor
20130603	214.2	55.3	2.0	13.6	61.1	0.8	17.1	234.2	0.5
20130604	226.7	56.6	0.4	13.2	58.4	0.1	21.0	265.5	0.3
20130605	189.4	56.6	0.4	11.7	40.5	0.6	13.6	218.5	0.4
20130606	178.4	68.1	1.1	15.3	71.6	0.6	24.0	275.2	0.8
20130607	219.4	74.9	0.8	19.2	76.5	0.4	34.1	363.4	0.9
20130610	124.9	67.7	2.6	14.7	69.6	0.1	13.9	212.9	0.1
20130611	174.0	59.1	2.2	23.5	57.7	0.1	19.4	186.3	0.4
20130612	170.0	52.9	2.1	26.7	71.5	0.0	16.8	174.4	0.4
20130613	144.6	57.7	1.6	38.7	61.9	1.1	18.9	184.5	0.5
20130614	161.8	50.7	2.3	48.8	66.7	0.0	45.9	303.4	1.4
20130617	150.1	71.6	2.2	54.0	78.7	0.0	30.0	182.8	0.5
20130618	81.9	50.5	0.2	65.7	75.5	0.3	18.0	188.8	1.0
20130619	31.7	45.6	0.5	144.8	92.3	0.5	28.8	268.7	0.6
20130620	7.1	13.9	0.1	282.9	81.5	0.7	51.5	337.9	1.1
20130621	-	-	-	269.6	52.9	0.3	53.4	253.8	1.3
20130624	-	-	-	223.9	75.3	0.8	41.0	334.7	1.2
20130625	-	-	-	176.4	78.9	0.3	30.4	443.7	0.5
20130626	-	-	-	221.1	59.4	0.2	34.1	254.6	1.5
20130627	-	-	-	188.4	67.2	1.3	35.4	253.2	0.4
20130628	-	-	-	177.4	52.3	0.5	37.7	255.3	2.5

Table IA.3. The Impact of Imbalance in Calendar Spreads in WTI

The table shows regressions of the high-volume month return on various measures of imbalance in WTI futures. Our benchmark imbalance used in the paper is the total net imbalance across al near contracts up through the Lead or Active month including any contract within three weeks of becoming the Lead or Active month. In this table we also include front- and next-month imbalance, as well as front- and next-month imbalance coming from explicit calendar spread trades. Panel A shows summary statistics. Panel B shows regressions of returns on various measures of imbalance for all minutes in our 2010 - 2014 sample where we have calendar spread data.

Panel A: Summary Data for Returns and Imbalance

Variable	Mean	Std. Dev	Min	Max
Hi Volume Month Return (%)	0.00	0.06	-1.68	3.10
Benchmark Imbalance (Mil $)	-0.02	15.53	-427.51	321.48
Front Month Imbalance (Mil $)	-0.04	12.62	-270.92	294.98
Next Month Imbalance (Mil $)	0.02	5.52	-247.68	152.96
Front Month Calendar Spread Imbalance (Mil $)	-0.07	20.63	-808.38	718.20
Next Month Calendar Spread Imbalance (Mil $)	0.14	20.28	-651.45	667.97

Panel B: Explaining High Volume Month Return with Imbalance

	Return (1)	Return (2)	Return (3)	Return (4)	Return (5)	Return (6)
Benchmark Imbalance	0.00277***					
	[218.940]					
Front Month Imbalance		0.00330***		0.00294***	0.00294***	0.00294***
		[224.598]		[196.003]	[195.547]	[195.241]
Next Month Imbalance			0.00477***	0.00202***	0.00202***	0.00202***
			[114.420]	[66.257]	[66.315]	[66.232]
Front Month Calendar Spread Imbalance					0.00001***	0.00002*
					[3.312]	[2.300]
Next Month Calendar Spread Imbalance						0.00000
						[0.501]
Constant	0.00012	0.00019**	-0.00005	0.00014*	0.00014*	0.00014*
	[1.916]	[3.018]	[-0.581]	[2.160]	[2.174]	[2.169]
Obs	474240	483298	474240	474240	474240	474240
R-sq	0.488	0.464	0.182	0.491	0.491	0.491

t-statistics in brackets
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table IA.4. Full Sample Price Impact VARs

The table shows the results from vector autoregressions of the form described in section IA.3. R^2_w shown in the final row is the percentage of variation in returns explained by unexpected innovations in order flow, calculated from a vector moving average representation of the VAR. Return is measured in percent, while imbalance is measured in 100s of contracts. The sample is 1/1/2007 to 4/1/2014. We exclude data prior to 2008 for Brent and minutes prior to 7:30 AM or after 4:00 PM in New York.

	WTI Crude		Brent Crude		Gold		Copper		Corn		Wheat	
	Return	Imbalance	Return	Imbalance	Return	Imbalance	Return	Imbalance	Return	Imbalance	Return	Imbalance
Imb. (t)	0.033***		0.033***		0.031***		0.097***		0.020***		0.053***	
	[778.258]		[377.948]		[730.820]		[434.507]		[407.343]		[312.961]	
Imb (t-1)	-0.003***	0.089***	-0.004***	0.127***	-0.002***	0.062***	-0.006***	0.078***	-0.000***	0.103***	-0.001***	0.085***
	[-62.650]	[66.277]	[-41.088]	[104.247]	[-36.209]	[47.071]	[-22.881]	[65.939]	[-3.181]	[63.330]	[-6.285]	[53.497]
Imb (t-2)	-0.001***	0.029***	-0.002***	0.047***	-0.001***	0.030***	-0.005***	0.039***	-0.001***	0.051***	-0.002***	0.038***
	[-25.083]	[21.670]	[-19.935]	[38.760]	[-16.365]	[22.641]	[-18.549]	[32.783]	[-15.182]	[31.108]	[-9.215]	[23.055]
Imb (t-3)	-0.001***	0.029***	-0.002***	0.042***	-0.001***	0.028***	-0.003***	0.033***	-0.001***	0.035***	-0.002***	0.027***
	[-21.857]	[21.871]	[-16.525]	[34.798]	[-18.914]	[22.021]	[-13.401]	[28.185]	[-16.457]	[22.729]	[-8.448]	[16.793]
Ret (t-1)	-0.058***	1.650***	-0.037***	0.443***	-0.061***	1.900***	-0.034***	0.175***	-0.102***	1.589***	-0.063***	0.299***
	[-55.155]	[64.120]	[-33.273]	[30.564]	[-58.543]	[75.018]	[-31.644]	[34.466]	[-74.581]	[39.908]	[-43.874]	[24.441]
Ret (t-2)	-0.014***	0.440***	-0.012***	0.144***	-0.026***	0.681***	-0.001	0.082***	-0.026***	0.486***	-0.020***	0.109***
	[-13.515]	[17.065]	[-11.029]	[9.900]	[-24.994]	[26.798]	[-0.673]	[16.211]	[-18.938]	[12.230]	[-13.035]	[8.343]
Ret (t-3)	-0.007***	0.125***	-0.009***	0.060***	-0.012***	0.288***	-0.001	0.029***	-0.011***	0.127***	-0.010***	0.027**
	[-7.101]	[4.893]	[-8.099]	[4.145]	[-11.746]	[11.419]	[-0.786]	[5.709]	[-9.606]	[3.653]	[-6.954]	[2.194]
Cons	0.001***	-0.017***	0	-0.003***	0.000***	-0.012***	0.000***	-0.001***	0.001***	-0.060***	0.001***	-0.022***
	[8.280]	[-9.939]	[0.603]	[-2.817]	[9.294]	[-12.070]	[5.297]	[-3.266]	[9.829]	[-15.818]	[5.199]	[-15.132]
Obs	919,910	919,910	792,989	792,989	915,852	915,852	874,572	874,572	493,111	493,111	475,674	475,674
R^2	0.397	0.03	0.153	0.029	0.368	0.027	0.178	0.016	0.255	0.03	0.172	0.016
R^2_w	0.382		0.145		0.354		0.171		0.248		0.165	

t-statistics in brackets

*** p<0.01, ** p<0.05, * p<0.1

Table IA.5. Summary of Trade and Quote Data for Seven Commodities with Tick-Test Imbalance

The table shows means and standard deviations for minute-by-minute returns, trading volume, and signed trading volume (imbalance). Imbalance is calculated using both a quote-based method similar to Lee and Ready 1991, and using the tick test. Statistics for volume and imbalance are reported in millions of dollars of notional value. The sample is 1/1/2007 to 4/1/2014. We exclude data prior to 2008 for Brent due to apparent data issues, and Soybeans data are only available starting in early 2009. We also exclude minutes prior to 7:30 AM or after 4:00 PM in New York.

Panel A: All Minutes

	# of Min	Return (%)		Volume (Mil $)		Quote-Based Imbalance (Mil $)		Tick-Test Imbalance (Mil $)		Corr of T-T & Q-B Imb
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Commodity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTI	920,567	0.00	0.09	33.98	42.81	-0.17	14.94	-0.08	15.74	0.95
Brent	791,420	0.00	0.08	14.72	23.58	-0.01	10.00	-0.10	10.59	0.87
Gold	919,713	0.00	0.05	21.54	35.99	-0.15	12.00	-0.17	12.65	0.93
Copper	909,542	0.00	0.07	2.79	5.43	-0.01	2.22	0.00	2.52	0.83
Corn	539,702	0.00	0.12	9.25	17.59	-0.14	7.43	0.01	8.49	0.83
Soft Wheat	509,464	0.00	0.15	4.41	9.15	-0.07	3.58	-0.03	4.03	0.78
Soybeans	467,965	0.00	0.10	10.60	18.22	-0.24	7.58	-0.11	8.40	0.94

Panel B: Settlement Minute

	# of Min	Return (%)		Volume (Mil $)		Quote-Based Imbalance (Mil $)		Tick-Test Imbalance (Mil $)		Corr of T-T & Q-B Imb
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Commodity	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
WTI	1,816	-0.01	0.11	193.25	88.46	-4.13	36.04	-2.28	36.27	0.86
Brent	1,563	0.00	0.10	81.81	68.19	0.00	22.68	-1.07	21.72	0.75
Gold	1,825	0.00	0.06	83.22	63.33	2.03	23.87	1.58	23.43	0.91
Copper	1,867	0.00	0.12	23.23	24.49	0.66	9.17	0.82	8.93	0.86
Corn	1,810	0.01	0.24	120.51	96.23	4.28	29.92	4.84	30.10	0.93
Soft Wheat	1,808	-0.04	0.36	73.17	62.36	-1.89	20.38	-1.92	19.67	0.90
Soybeans	1,271	0.00	0.17	130.89	76.94	-3.57	31.25	-1.81	32.51	0.97

Table IA.6. Price Impact Regressions with Tick-Test

The table shows the results from separat univariate regressions where the dependent variable is one-minute returns and the independent variable is the contemporaneous one-minute imbalance. Return is measured in percentage and imbalance is measured in millions of dollars, (i.e. a coefficient of 0.01 represents a return response of 0.01%, or one basis point, per million dollars of imbalance). Columns (1)-(4) show the results for all minutes using imbalance measured with a quote-based method similar to Lee and Ready 1991. Columns (5)-(8) use quote-based imbalance but only include the minute prior to daily futures settlement. Columns (9)-(11) and (12)-(14) repeat these regressions using imbalance measured with the tick test. Standard errors are shown in parentheses. The sample is 1/1/2007 to 4/1/2014. We exclude data prior to 2008 for Brent due to apparent data issues, and Soybeans data are only available starting in early 2009. We also exclude minutes prior to 7:30 AM or after 4:00 PM in New York.

Independent Variable:	Quote-Based Imbalance (Mil $)								Tick-Test Imbalance (Mil $)					
Dependent Variable:	Return (%) All Minutes				Return (%) Settlement Minute				Return (%) All Minutes			Return (%) Settlement Minute		
Commodity	Estimate (1)	Std. Err (2)	R-sq (3)	Obs (4)	Estimate (5)	Std. Err (6)	R-sq (7)	Obs (8)	Estimate (9)	Std. Err (10)	R-sq (11)	Estimate (12)	Std. Err (13)	R-sq (14)
WTI	0.0034***	(0.00001)	0.34	929,219	0.0013***	(0.00007)	0.17	1,816	0.0034***	(0.00001)	0.38	0.0015***	(0.00007)	0.26
Brent	0.0029***	(0.00002)	0.13	800,225	0.0010***	(0.00008)	0.05	1,563	0.0032***	(0.00002)	0.18	0.0016***	(0.00008)	0.12
Gold	0.0022***	(0.00002)	0.31	928,085	0.0013***	(0.00011)	0.22	1,825	0.0023***	(0.00002)	0.36	0.0015***	(0.00014)	0.30
Copper	0.0114***	(0.00010)	0.15	903,774	0.0042***	(0.00034)	0.11	1,867	0.0112***	(0.00009)	0.18	0.0054***	(0.00036)	0.17
Corn	0.0068***	(0.00019)	0.22	515,471	0.0044***	(0.00024)	0.29	1,808	0.0062***	(0.00021)	0.26	0.0047***	(0.00023)	0.34
Soft Wheet	0.0155***	(0.00050)	0.18	489,820	0.0091***	(0.00078)	0.26	1,806	0.0153***	(0.00049)	0.25	0.0105***	(0.00074)	0.32
Soybeans	0.0047***	(0.00004)	0.15	446,319	0.0032***	(0.00013)	0.37	1,271	0.0046***	(0.00003)	0.17	0.0034***	(0.00013)	0.43

*** p<0.01, ** p<0.05, * p<0.1

Table IA.7. Regressions of Futures Order Flow Imbalance on Commodity-Index Flows: Small and Large Trades

The table shows the results from regressions of weekly futures order flow imbalance on contemporaneous commodity-index flows as reported by the CFTC in their supplemental commitments of traders reports. The first three rows of Panel A show results from regressions where the independent variable is the weekly index flow (in millions of $) in a single commodity, and the dependent variable is weekly imbalance (see Table IA.5 for a description) in the associated futures market measured over the entire trading day (columns (2) - (4)), only in the 30 minutes prior to futures settlement (columns (5) - (7)), or only in the five minutes prior to futures settlement (columns (8) - (10)). Each set of three columns reports the estimate of the slope coefficient on the weekly index flows, the associated t-statistic, and the regression R-squared (regression constants are not reported). The last two rows show the slope coefficients on changes in index trader positions of pooled regressions using all three commodities. The first pooled regression is a time-series regression, and thus includes only commodity fixed effects. The second regression is cross-sectional, and thus includes both commodity and week fixed effects. For both pooled regressions standard errors are clustered by week. Panel B and C repeat the regressions of Panel A, but calculate the imbalance separately for small trades of a single contract, and trades larger than a single contract. Data are 1/1/2007 to 4/1/2014. Soybeans data begin in 2009.

Panel A: All Trades

Independent Variable:		Change in Net Position of Index Traders in Current Week								
Dependent Variable:		Weekly Imbalance in All Minutes			Weekly Imbalance in 30 Minutes Prior to Settle			Weekly Imbalance in 5 Minutes Prior to Settle		
Commodity	Obs	Estimate	T-stat	R-sq	Estimate	T-stat	R-sq	Estimate	T-stat	R-sq
Single Commodity										
Corn	382	0.383**	[2.55]	0.027	0.139***	[3.59]	0.036	0.0708***	[2.70]	0.023
Soft Wheat	382	0.527***	[4.69]	0.073	0.240***	[5.53]	0.118	0.173***	[5.11]	0.113
Soybeans	270	0.741***	[3.70]	0.094	0.125***	[2.59]	0.037	0.0733**	[2.13]	0.054
Pooled										
w/ commodity FE	1034	0.454***	[3.90]	0.056	0.152***	[4.78]	0.046	0.0865***	[3.98]	0.021
w/ com. & week FE	1034	0.320*	[1.88]	0.571	0.139***	[2.99]	0.533	0.0618**	[2.18]	0.174

Panel B: Single Contract Trades

Independent Variable:		Change in Net Position of Index Traders in Current Week								
Dependent Variable:		Weekly Imbalance in All Minutes			Weekly Imbalance in 30 Minutes Prior to Settle			Weekly Imbalance in 5 Minutes Prior to Settle		
Commodity	Obs	Estimate	T-stat	R-sq	Estimate	T-stat	R-sq	Estimate	T-stat	R-sq
Single Commodity										
Corn	382	-0.164***	[-3.11]	0.585	0.00424	[0.26]	0.501	-0.158***	[-3.42]	0.590
Soft Wheat	382	0.217***	[4.36]	0.000	0.0601***	[3.00]	0.032	0.193***	[4.00]	0.063
Soybeans	270	0.133**	[2.11]	0.046	0.0236	[1.38]	0.048	0.122**	[2.03]	0.048
Pooled										
w/ commodity FE	1034	-0.0657*	[-1.71]	0.008	0.0153	[1.27]	0.020	-0.0663*	[-1.95]	0.083
w/ com. & week FE	1034	-0.129**	[-2.22]	0.041	0.00700	[0.38]	0.166	-0.126**	[-2.42]	0.070

Panel C: Trades Larger than a Single Contract

Independent Variable:		Change in Net Position of Index Traders in Current Week								
Dependent Variable:		Weekly Imbalance in All Minutes			Weekly Imbalance in 30 Minutes Prior to Settle			Weekly Imbalance in 5 Minutes Prior to Settle		
Commodity	Obs	Estimate	T-stat	R-sq	Estimate	T-stat	R-sq	Estimate	T-stat	R-sq
Single Commodity										
Corn	382	0.547***	[3.70]	0.553	0.135***	[4.26]	0.516	0.229***	[4.29]	0.596
Soft Wheat	382	0.310***	[4.07]	0.051	0.180***	[6.38]	0.065	-0.0205	[-0.49]	0.001
Soybeans	270	0.607***	[3.39]	0.120	0.101***	[2.65]	0.001	-0.0491	[-1.00]	0.006
Pooled										
w/ commodity FE	1034	0.520***	[4.63]	0.040	0.137***	[5.38]	0.004	0.153***	[3.97]	0.248
w/ com. & week FE	1034	0.448***	[2.93]	0.068	0.132***	[3.60]	0.220	0.187***	[3.47]	0.591

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table IA.8. Index Flows, Managed Money Flows, and Momentum Trading across Subsamples

The table shows the results from weekly regressions using index flows, managed money flows, and futures returns. Columns (1) - (6) show the results of weekly returns measured across various portions of the trading day on contemporaneous index flows along with controls for managed money flows from DCOT reports. Columns (7) - (9) regress index flows on managed money flows and the previous week's return. Column (10) likewise regresses managed money flows on the previous weeks returns. Panel A shows results for the period from the start of our sample to 2014. Panel B shows results using data starting in 2015 to the end of our sample in 2020. In all regressions we pool across thirteen agricultural commodities and included fixed effects for commodity and week. Standard errors are clustered by week.

Panel A: 2007-2014

	Weekly Return									
	Full Day		Prior to 30 Min Pre Settle		30 Min Pre Settle		Index Flows			MM Flows
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Index Flows	0.254*** [3.122]	0.211*** [3.298]	0.068 [0.959]	0.028 [0.523]	0.186*** [7.212]	0.183*** [7.163]				
MM Flows		0.624*** [24.008]		0.585*** [24.910]		0.040*** [4.337]	0.010 [1.077]		0.013 [1.461]	
Lag Ret								-0.006 [-0.976]	-0.009 [-1.365]	0.172*** [13.826]
Obs	4225	4225	4225	4225	4225	4225	4225	4221	4221	4475
R-sq	0.29	0.44	0.29	0.45	0.21	0.22	0.23	0.23	0.23	0.26

Panel B: 2015-2020

	Weekly Return									
	Full Day		Prior to 30 Min Pre Settle		30 Min Pre Settle		Index Flows			MM Flows
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Index Flows	0.785*** [8.722]	0.240*** [3.287]	0.546*** [6.308]	-0.016 [-0.232]	0.239*** [8.621]	0.256*** [8.860]				
MM Flows		0.613*** [27.116]		0.632*** [31.164]		-0.019** [-2.294]	0.078*** [11.046]		0.070*** [9.990]	
Lag Ret								0.040*** [7.590]	0.021*** [4.116]	0.275*** [17.827]
Obs	3846	3846	3846	3846	3846	3846	3846	3846	3846	3855
R-sq	0.26	0.45	0.24	0.46	0.20	0.20	0.29	0.26	0.30	0.28

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table IA.9. Two-Day Futures Returns on CLN Determination Dates

The table repeats the analysis in Table 8 in the main text, but uses the average cumulative return on both the date of CLN determination and the following day. Panels A uses all days with a note that has a delta greater than zero on the a determination date prior to 1/1/2019. Panel B restricts this to days with at least $10 million of face value. Panels C and D further restricts the sample to notes with determination dates before 2/1/2014 to replicate the determination date event study of HPW.

Panel A: All Days				**Panel B: w/ $10+ Mil Face Value**		
	All Days (1)	Excluding Goldman Roll (2)	During Goldman Roll (3)	All Days (4)	Excluding Goldman Roll (5)	During Goldman Roll (6)
Realized Daily Returns						
Average	-0.06	-0.05	-0.10	-0.19	-0.31	0.15
	[-0.45]	[-0.28]	[-0.43]	[-0.91]	[-1.29]	[0.41]
Predicted Impact of Unwinding Delta Hedges						
Average	-0.10	-0.12	-0.05	-0.18	-0.22	-0.09
Obs	202	141	61	91	66	25

Panel C: Days prior to 2014/02				**Panel D: Prior to 2014/02 w/ $10+ Mil Face Value**		
	All Days (1)	Excluding Goldman Roll (2)	During Goldman Roll (3)	All Days (4)	Excluding Goldman Roll (5)	During Goldman Roll (6)
Realized Daily Returns						
Average	-0.17	-0.25	-0.01	-0.27	-0.42	0.14
	[-1.01]	[-1.12]	[-0.03]	[-1.09]	[-1.44]	[0.33]
Predicted Impact of Unwinding Delta Hedges						
Average	-0.12	-0.15	-0.05	-0.21	-0.25	-0.10
Obs	157	108	49	75	54	21

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table IA.10. Discrepancies with refined determination date sample of HPW

This table lists the discrepancies between our subset of 54 days with determination dates outside of the Goldman Roll Period, prior to February 2014, with at least $10 million of face value, and the refined set of 56 days provided to us by HPW.

Notes Included by HPW

Determination Date	. Commodity	Face Value	Day 0 Return	Day 1 Return	Explanation
10/4/13	Gold	143,249,000	-0.58	1.15	We exclude because this is the 4th trading day of the month and therefore in the Goldman Roll Period
10/7/13	Gold	35,000,000	1.15	-0.05	We exclude because this is the 5th trading day of the month and therefore in the Goldman Roll Period
7/22/13	Corn	20,000,000	-0.60	-3.37	We exclude because this is a "daily liquidity note". Only $2.5 million was sold at issue, remainder held by brokerage arm of the issuer. The issuer also offers to buy the note back at market value during the life of the note.

Notes Excluded by HPW

Determination Date	. Commodity	Face Value	Day 0 Return	Day 1 Return	Explanation
7/27/10	Natural Gas	18,319,000	1.37	2.12	Note is linked to the UNG natural gas ETF. We include notes linked to ETFs which hold only commodity futures or spot positions in a single commodity.