Is There Smart Money? How Information in the Commodity Futures Market Is Priced into the Cross-Section of Stock Returns with Delay

Steven Wei Ho* Alexandre R. Lauwers§

First draft: November 17th, 2017
This draft: October 10th, 2022
Accepted to the Journal of Financial and Quantitative Analysis (JFQA)

Abstract

We document a new empirical phenomenon in which the aggregate positions of money managers, who are sophisticated speculators in the commodity futures market, as disclosed by the Disaggregated Commitments of Traders reports, can predict the cross-section of commodity producers' stock returns in the subsequent week. We employ a number of cross-sectional methods, including calendar-time regression analysis, single-sort, double-sort, and Fama-MacBeth regressions, to confirm the predictability results. The results are more pronounced in firms with higher information asymmetry. We thus add more empirical evidence to the literature on costly information processing, which leads to gradual information diffusion across asset markets.

JEL codes: G11, G14
Keywords: Costly Information Processing, Commodity, Equity Return Predictability, Market Segmentation, Gradual Information Diffusion

We have greatly benefited from the comments by the anonymous referee and Hendrik Bessembinder (Editor). We are indebted to Brian Henderson (discussant) as well as the participants at the AFA 2019 Atlanta Annual Meeting for insights. We are also grateful for Alminas Zaldokas (discussant) and seminar participants at the Paris Winter 2018 Conference. We also thank Gang Li (discussant) and audience at the SFS Cavalcade Asia-Pacific 2019, and seminar participants at Columbia University, Queen's University, and University of Nottingham for helpful inputs. We thank Michael Johannes, Patrick Bolton, Paul Calluzzo, Daniel Chi, David Feldman, Ana-Maria Fuertes, Louis Gagnon, Lawrence Glosten, Matthieu Gomez, Lars Peter Hansen, Robert Hodrick, Harrison Hong, Wei Jiang, Harry Mamaysky, Rahul Mukherjee, Nick Pan, Neil Pearson, José Scheinkman, Cédric Tille, Neng Wang, Wei Wang, Laura Veldkamp and Andrew Zhang for comments and suggestions. All errors are our own.

*Assistant Professor, Department of Finance, University of Nevada, Las Vegas, 4505 S. Maryland Pkwy, Box 456008, Las Vegas, NV 89154-6008, email: steven.ho@unlv.edu

§PhD Student, The Graduate Institute, Geneva, Case postale 1672, 1211 Genève 1, Switzerland, email: alexandre.lauwers@graduateinstitute.ch

I. Introduction

Recent literature has documented the existence of so-called smart money in various markets. For example, Engelberg et al. (2012) find that short sellers, who are often considered to be sophisticated investors, can quickly and effectively process and respond to published news; likewise, some investors appear to be smart in the currency market (Michaelides et al., 2015, among others).[1] In this context, we seek to investigate whether money managers, who are sophisticated and specialized investors in the commodity futures market, can be deemed "smart money" with a superior information advantage on commodity fundamentals, and whether this information is passed through to the equity market in a timely manner. We answer these questions by studying the trader positions disclosed in the weekly Disaggregated Commitments of Traders (DCOT) reports published by the Commodity Futures Trading Commission (CFTC). Specifically, we focus on the MM (which stands for Managed Money, or alternatively, Money Managers, per the CFTC[2]) category of traders. The MM positions data are matched to a sample of commodity producers' equities for commodities that can be appropriately identified with an industry code, similar to the procedure proposed by Gorton and Rouwenhorst (2006), and we construct a sample from January 2007 to March 2020.

Cohen and Lou (2012) document that investors have limited resources and capacity to process information, which in turn causes the same piece of information to be impounded into firm values with differential lags. Given that the amount of rich information being produced in the market has increased, Cohen et al. (2020) find that this makes information processing more complex and investors may become inattentive to valuable information updates such

[1] Bohmann and Patel (2020) find that some investors know about upcoming energy commodity news.

[2] The CFTC uses the terms "money managers" and "managed money" interchangeably in the explanatory notes. Furthermore, per definition, the MM category of traders in the DCOT reports consists of "registered commodity trading advisers (CTAs), registered commodity pool advisers (CPOs) and unregistered funds identified by CFTC." The CFTC definition of CTAs/CPOs within MM is solely based on legal registration status under the Commodity Exchange Act (CEA) and encompasses most hedge funds (especially the sizable ones) that trade "commodity interest" (including futures) in a nontrivial manner, including many funds that are more sophisticated than simple trend followers.

as those contained in corporate filings. The theoretical work by Van Nieuwerburgh and Veldkamp (2010) concludes that market participants cannot specialize in every asset as information acquisition and processing is costly, and specialization then arises because the more an investor holds of an asset, the more valuable it is to learn about that asset; but the more an investor learns about the asset, the more valuable that asset is to hold. Other research has also documented that investors' ability to collect and process only a subset of information will lead to investor specialization, market segmentation, and gradual diffusion of information in financial markets (Menzly and Ozbas, 2010).

Based on these premises, we posit that sophisticated investors who specialize in the commodity market and trade futures would on average react to information updates related to commodity fundamentals faster than their unspecialized counterparts; also, because the fundamental of a commodity-producing firm's equity has additional firm-specific components beyond the fundamental of the commodity produced, information updates pertaining to commodity fundamentals would be gradually diffused and impounded into the equity price of these firms. Indeed, we find that the information extracted from the categorical aggregate positions of MM traders in the commodity futures market[3] can predict the cross-section of stock returns for commodity producers. In particular, if the DCOT[4] reports an increase in long position, a decrease in short position, or an increase in net position of MM, then the stock price of producers of the same commodity would increase in the following week.

As our main thesis is that return predictability arises from costly information processing, which leads to gradual information diffusion, we study the relation of our results with measures of information asymmetry and confirm that informational, rather than trading, frictions contribute to our predictability results. Specifically, we show that our results are stronger in commodity-producing firms with higher information asymmetry, as measured by *ex ante* analyst dispersion and 90-day *historical* stock volatility. However, by double-sorting

[3]The fact that trading activities are closely related to information flows is well-known in existing literature, for example, Bessembinder et al. (1996).

[4]Our study is made possible by CFTC's decision to publish the Disaggregated Commitments of Traders reports for trades after June 13, 2006, which created the MM category of traders.

our MM position signals with the illiquidity measure of Amihud (2002), we find no evidence that predictability is stronger or weaker in firms with higher trading friction. Importantly, we rule out the possibility that our predictability results would arise simply due to a mechanical link between the commodity producers' stock returns and the contemporaneous commodity futures returns. MM position changes continue to predict the residuals of stock returns after projecting these returns onto the contemporaneous futures returns. Also, we confirm that our findings are already present prior to the CFTC's Friday releases of MM positions as of Tuesdays and the results are not due to the announcement effects of DCOT reports. The pattern we document thus appears to be the result of costly information processing, as it takes time (and effort) to learn, process, and incorporate innovation in commodity fundamentals as well as firm-specific fundamentals if one were to invest in the equity of a commodity-producing firm.

The economically large and statistically significant abnormal returns attributed to this lead-lag relationship are consistent across many specifications: different factor models, different signal measures and weighting schemes, and different empirical methods, including calendar-time regression analysis, single-sort, double-sort, and Fama-MacBeth regressions. The level of abnormal returns varies across specifications and generally falls in the range of approximately 10%–13% per annum. In addition to finding alpha relative to the Carhart (1997) four-factor model and the Fama and French (2015) five-factor model, we show that abnormal return remains relative to the Stambaugh and Yuan (2016) model, which includes mispricing factors constructed from a broad set of 11 well-known equity anomalies, capturing momentum, financial distress, profitability, net stock issues, asset growth, and investment, among others.[5]

Furthermore, we decompose the MM position change signals into a momentum-driven component and a component that is orthogonal to commodity futures' past performance, and

[5] Some return anomalies generate abnormal performance mainly from short-selling overpriced stocks (Stambaugh et al., 2012), but this is not the case in our findings, where the alpha is coming from both the long and short legs of our long-short portfolios, and is not driven by a single leg.

find the non–momentum component to be driving the predictive power. Moreover, our results remain after the augmentation of the factor models with additional commodity price factors, such as futures momentum, basis (backwardation), and the recently discovered futures basis-momentum phenomenon, as well as the principal components of commodity futures returns. These results bring further support to our view that the measures of position changes of MM within the week-to-week DCOT reports would capture by and large informative reflections of "smart money" and their revised prospects on commodities' fundamentals, as opposed to merely reflecting the positions of trend followers within MM or the information content from common commodity futures strategies.

Our empirical findings are relevant in capturing some of the salient facts about today's market in that sometimes information affecting asset price movement or comovement is reflected in prices only with a lag due to a capacity constraint in information processing (Cohen and Frazzini, 2008; Cohen and Lou, 2012; Cohen et al., 2020); there is increasing specialization in investors' choice of information processing (Van Nieuwerburgh and Veldkamp, 2010); two asset markets with correlated fundamentals can be informationally segmented due to investor specialization (Menzly and Ozbas, 2010); some investors are better at private data collection and reprocessing of public information (Boehmer et al., 2020); and the presence of smart money (Engelberg et al., 2012; Huang et al., 2021, among others). Our paper is also related to the literature on the financialization in commodity markets (Henderson et al., 2014). There is also a body of literature relying on CFTC's legacy format (i.e., not disaggregated) COT reports;[6] our paper, however, studies the predictability of returns in the equity, rather

[6]Markets in which commercial hedgers—who are mostly commodity producers (as opposed to non–commercial speculators)—are net short (long) are found to have positive (negative) expected futures returns (Carter et al., 1983; Bessembinder and Seguin, 1992; De Roon et al., 2000). However, by exploring COT legacy format reports in four energy commodities, Sanders et al. (2004) find that traders' net positions, whether commercial or non–commercial, are not consistently useful in predicting weekly energy futures returns, although there is a positive (negative) contemporaneous correlation between weekly futures returns and the positions held by non–commercial (commercial) traders. Sanders et al. (2009) find the same result in the corn and live cattle futures market. In contrast, Buchanan et al. (2001) show that non–commercial positions in COT reports do provide useful and valuable information on predicting the magnitude and direction of weekly price change forecast in the natural gas futures market. In addition, by analyzing COT reports of major currency futures, Tornell and Yuan (2012) find that the peaks and troughs of commercial and non–commercial

than the futures, market. Other papers also explore the linkage between commodity markets and the bond market or equity indexes (Hong and Yogo, 2012; Fernandez-Perez et al., 2017), although they generally look at aggregate indexes (instead of firm-level data) at a long horizon, and they do not utilize the MM category in the DCOT format.

Thus, we contribute to the literature in the following respects. To the best of our knowledge, this is the first paper to investigate the cross-sectional predictability of commodity producers' stock returns that are matched with the corresponding MM positions in the commodity futures market, as recorded in the CFTC Disaggregated Commitments of Traders reports. We document that MM position changes do contain information that is conducive to the predictability of commodity producers' stock returns in the short term, which is new to the literature. This lead-lag relationship translates to large abnormal returns with respect to several asset pricing factors and is consistently confirmed through a number of empirical methods and specifications. We show that on average MM position change signals capture relevant information beyond the information already contained in past futures returns (whether past trend or 1-week-lagged futures return) or in common commodity futures strategies. Finally, we contribute by exploring potential channels related to this predictability result. We show that a mechanical link between futures return and contemporaneous equity return of commodity-producing firms is not driving our results, nor do our findings arise due to the announcement effects of DCOT reports. In addition, the lead-lag relationship is consistent with the presence of informational, not trading, friction wherein return predictability is substantially stronger for commodity-producing firms with higher information asymmetry. Our paper thus represents a contribution to the literature finding that, in settings other than ours, investors have limited information processing capacity (Cohen and Lou, 2012; Cohen et al., 2020), which leads to investor specialization, market segmentation, and gradual information diffusion.

The rest of the paper is organized as follows: Section II elaborates on the key mechanisms

traders' net positions are generally useful predictors of the evolution of spot exchange rates, and this simple trading strategy proves to be quite profitable.

underlying our analysis and provides empirically testable predictions. Section III discusses how information on traders' positions from the DCOT reports is extracted, matched to the sample of commodity producers stocks, and used as leading signals to form the long-short portfolio. Section IV presents the empirical results, including results on portfolio alpha and Fama-MacBeth cross-sectional regressions. Section V investigates potential explanations behind this documented lead-lag relationship. Section VI concludes. Further details on the procedure for computing the long-short portfolio returns are contained in the Appendix. Additional procedures and results are provided in the Supplementary Material.

II. Mechanisms and Hypotheses

Cohen and Lou (2012) find that information processing is costly and investors have limited capacity to process information. As a result, significant delay can occur in the impounding of information into the prices of complex assets relative to simple assets. Specifically, they find that the same industry shocks are incorporated into easy-to-analyze firm values before they are reflected in conglomerate firm values which require more complicated valuation analyses. Also, with limited ability to collect and process information, Cohen et al. (2020) document a finding wherein investors can be inattentive to rich information in corporate filings that is only impounded into prices with a significant delay. In addition, the theoretical literature (Van Nieuwerburgh and Veldkamp, 2010) has established that if an investor wants to form a portfolio of risky assets, she needs to first exert effort to collect information on the future value of these assets before she invests and makes a choice on the asset she learns information about and specializes in.

Thus, the mechanism of our empirical study rests on the fundamental principle of investors' limited information processing capacity, which leads to investor specialization, market segmentation, and gradual information diffusion across asset markets. In our context, we have two asset classes, commodity futures and stocks of commodity-producing firms, and they have a correlated fundamental—namely, the future prospect of the underlying commodity in question. Due to investor specialization, we believe that money managers captured in the MM category, who are sophisticated investors in the commodity futures market, would be incentivized to be proficient at gathering, analyzing, and processing public and/or private information pertaining to the commodity futures they expect to trade and thus, on average, they would be able to react to updates regarding the future prospect of the underlying commodity in question faster than unspecialized and unsophisticated traders, and their views are reflected in the positions they commit. In addition, the fundamental of a commodity producer's equity has components other than the fundamental of the commodity produced, as at a minimum it also involves the firm's capital structure, sales, local cost of production, labor relations, and management competence in decision making, among other factors. As discussed in Cohen and Lou (2012) and Cohen et al. (2020), it takes time and effort to digest, process, and incorporate firm-specific informational updates. Accordingly, in our context, comparatively speaking it would take less time (and effort) to learn, process, and incorporate revisions to the fundamental of a particular commodity, than it would to learn both the firm-specific fundamentals and the commodity fundamental. Thus, the asset prices of commodity-producing firms (especially those that are less transparent) would react slower to informational updates regarding the future prospect of commodity fundamentals.

Two main empirical return predictions follow from our discussion: i) we should observe that the positions of MM would move first in the commodity futures market, and MM's positions should on average predict the equity returns of commodity producers of the same commodity; ii) the equity returns of commodity producers consisting of firms that are relatively nontransparent (or with high information asymmetry) should be the slowest moving; for these firms we should find the strongest predictability results. These are exactly the empirical results we find. After all, if MM are voting with their money about certain commodities' prospects in the futures market that reflect their current up-to-date information regarding commodities' fundamentals, why would such information not be impounded immediately into the equity prices of commodity producers, such that there would not be any lead-lag relationship? We believe it is due to the reasons we laid out earlier in this section—that is, limited information processing capacity of investors which leads to investor specialization, market segmentation, and gradual information diffusion.

III. Data and Methods

We utilize positions data of the commodity futures market as well as firm-level returns data of the commodity producers. The stocks of commodity producers are matched to the feasible commodities based on industry classifications. We construct signal measures based on the changes in MM positions, as disclosed in the CFTC DCOT reports, and we form long-short portfolios based on the signals.

III.A. CFTC Positions Data

We use positions data and trader classification from the publicly available CFTC DCOT reports. The report provides weekly information on aggregate traders' positions for five categories of market participants who are active in the commodity futures markets (each defined in Supplementary Material Table A.1).[7] The DCOT reports display each category's open interest by long and short positions, aggregated across all contract maturities.[8] The reports are normally released every Friday at 3:30 p.m. (Eastern Time) with the positions data compiled as of the end-of-day on the Tuesday of the same week; in other words, the release dates (Fridays) are three days after the compilation dates (Tuesdays). Our analysis covers data beginning from January 2007.

We concentrate our analysis on the position changes of money managers. These traders mostly take speculative positions,[9] invest others' money in the commodity futures market on

[7]In the legacy format, however, the COT report divides reporting traders into two broad categories: the "Commercial" category, aggregating the "Producer/Merchant/Processor/User" (PM) and "Swap Dealers" (SW) categories from the DCOT reports, and the "Non–Commercial" category, comprising the "Managed Money" (MM) and "Other Reporting" (OR) DCOT categories.

[8]Spreading positions are also disclosed for three categories of traders: SW, MM, and OR. Spreading measures the extent of traders holding equal long and short positions.

[9]The London Metals Exchange (LME) also publishes its own disaggregated Commitments of Traders Reports (COTR) for certain markets in recent years, wherein for each of the trader categories, post–MiFID (Markets in Financial Instruments Directive) reports differentiate between weekly

a discretionary basis, may make use of leverage, and usually do not intend to take delivery of the underlying commodities they are trading. We focus on MM since they are the category of traders who have the most incentive to seek out and process information related to changes in commodity markets and they are believed[10] to have the expertise. Indeed, Van Nieuwerburgh and Veldkamp's (2010) model concludes that because information acquisition and processing is costly, the optimal learning strategy for investors is to concentrate on one or a small set of assets. In addition, MM may engage in a higher frequency of trading than commercial hedgers and thus are more sensitive to information related to the short term.

Money managers in the DCOT reports consist of registered commodity trading advisers (CTAs), registered commodity pool advisers (CPOs), and unregistered funds identified by the CFTC. Although a CTA or CPO does usually remind one of a passive trend follower (or momentum trader) in industry parlance, the CFTC's definition of CTAs/CPOs (which largely overlap) within the MM category is purely based on legal registration status, as opposed to self-reported fund classification or investment style. Per CFTC regulations under the CEA, any money manager that trades "commodity interest" in a nontrivial manner (unless claiming an exemption, including small pool and *de minimis* exemptions), be it a non–hedge fund or a hedge fund, generally needs to register as a CTA or CPO, or both. It follows that CTAs/CPOs within MM encompass most hedge funds (especially the sizable ones) with positions in commodity futures, among which are funds that are largely more sophisticated than simple trend followers.[11] Besides, the MM category also captures the futures-based positions of commodity exchange-traded funds (ETFs), which are essentially

positions held for hedging purposes and speculative ones, dividing them into risk-reducing and non–risk-reducing positions. As indicated on the reports, the traders under the category of Investment Firms or Investment Funds generally do not have positions for hedging purposes, thus confirming that they are indeed mostly speculating on the futures market, as expected. Furthermore, Tokic (2010) finds that the MM category potentially "behaved as a rational speculator" and that "money managers traded based on fundamentals" during the 2008 oil bubble.

[10] Krohn (2018) confirms "the existence of managerial skills among CTAs," among others.

[11] We sincerely thank CFTC Chief Economist Scott Mixon for disclosing to us the names (but not positions) of entities registered under law with the CFTC as CTAs and CPOs. The lists include entities that would have been reasonably considered to be smart hedge funds or having active skills in trading commodities, such as Bridgewater, Millennium, BlackRock, Winton, GLG Partners, Coburn Barrett, and Galtere International.

passive in nature, although they constitute only a small portion of MM's total open interest.[12]

Although we recognize that some funds in MM can be simple trend followers or more passive in nature, we posit that the metric of open interest changes in the MM category within the week-to-week DCOT reports is inherently well positioned in its ability to capture information, and there is a difference between asking whether MM position changes reflect "smart money" and contain information updates versus the question on the performance of an "average" commodity managed futures fund.[13] Positions are summed, not averaged, across all entities within the MM categories in the DCOT reports, and the data also have the advantage of being legally collected by regulators with penalties for untruthful reporting, rather than self-reported to commercial databases with associated issues such as coverage, misreporting,[14] and data quality. Furthermore, we postulate that traders who solely follow the trend signal (52-week or 26-week momentum of the commodity, which is hard to be impacted by a 1-week change) would not likely change their positions drastically in all weeks, as the trend formula prescribes. Similarly, passively managed commodity ETFs are unlikely to have large changes in positions at 100% of assets under management (AUM) in every week, notwithstanding our observation that they cannot constitute more than just a small fraction of MM positions. Hence and consistent with the empirical results we will show, we posit that MM position changes in most of the weeks are primarily contributed by the more active aspects of trading (not due to trend-following signals or passive ETFs)—that is, smart money who see fit to commit or change their positions from week to week based on their response to informational updates regarding the future prospect of commodity fundamentals. Accordingly, we will confirm empirically in Section V.C that our predictability results are mainly driven by the component in MM position changes that is orthogonal to past commodity futures returns

[12] By examining a comprehensive list of commodity-focused and futures-based ETFs that we have hand-collected, we estimate that as a whole they cannot constitute more than but a small fraction of the reported MM's total open interest (in dollars), even under generous assumptions. See Supplementary Material Section A.5 for further details.

[13] Per Du and Kane (2019), "all else equal, someone with more wealth who is *willing* and *able* to trade a contract may have more influence ... similarly, someone with more *confidence* in their estimates ... may be more willing to place their wealth at stake to back their market views."

[14] See Chen et al. (2021), among others.

(as opposed to the component that is merely picking up the positions of traders within the MM category who simply follow the trend signal), and our results survive the inclusion of commodity price factors (e.g., futures momentum and basis).

Besides MM, the CFTC DCOT reports disclose the positions of "Producers, Merchants, Processors, and Users" (PM) and "Swap Dealers" (SW) categories, which could *a priori* contain relevant information. One might expect PM traders to hold information regarding commodity fundamentals' updates by being close to local physical market conditions. From an empirical standpoint, however, the PM category is much less homogeneous than the MM category with respect to the directionality of its constituents, as it groups together the dynamics of both producers and buyers of the commodity that manifestly have different objectives. Thus, the heterogeneity of trading motives and the relative proportions of producers versus users within PM traders (which are time varying and specific to each commodity market) would blur the direction of the demand for the group as a whole.[15]

Similarly, the reported SW positions may mask important heterogeneity in the types of counterparties and their respective trading motives. Relying on internal CFTC regulatory swap counterparty data for the West Texas Intermediate (WTI) crude oil market, Mixon et al. (2018) and Mixon and Onur (2020) reveal that swap dealers take both long *and* short positions, which is consistent with their business role as intermediaries who facilitate the on average net long positioning desired by speculative traders, most of whom are passive commodity index investors, and the on average net short position desired by commercial hedgers.[16] Hence, SW positions would reflect the relative magnitudes of the exposure of the two types of counterparties, which vary across commodities and over time, weakening the inference based on SW signals. With these caveats, the usefulness of signals constructed

[15]See Ederington and Lee (2002), who find "with the assistance of officials at the Office of Policy of the US Department of Energy" that "refiners held almost twice as many short contracts as long," whereas "the end user group held six times as many long contracts as short" for heating oils.

[16]Some studies have gone as far as using the DCOT's SW category as a noisy proxy for commodity index funds positions, as argued in Cheng et al. (2015). Thanks to their customized nature, swaps have also become increasingly used by commercial firms for hedging. Acharya et al. (2013) show that 80% of oil and natural gas producers use swaps to hedge, while only 47% use futures.

from the PM and SW categories is therefore limited, and we will see that they yield marginal predictability results that are not as robust as the MM signals.

III.B. Match with Commodity Producers' Stocks

To identify and match commodity-producing firms with the commodities for which the CFTC is collecting the DCOT information, we follow a procedure similar to the industry code matching algorithm proposed by Gorton and Rouwenhorst (2006). First, for each commodity that can be appropriately identified with a four-digit U.S. Standard Industrial Classification (SIC) code, we associate all publicly traded companies with the same four-digit SIC code. Second, to expand our sample size of commodity-producing firms and to address issues related to the SIC code information provided by the Center for Research in Security Prices (CRSP) (Gandhi and Lustig, 2015), we also utilize the Bloomberg Industry Classification System (BICS) code and data on firms' breakdown of revenues (BICSRevLvlAsgn) from Bloomberg. Supplementary Material Section A.2 describes in detail the procedure used to identify commodity producers' stocks.

Ultimately, ten commodities are matched: two industrial metals—copper and steel; three precious metals—gold, silver, and miscellaneous metals (palladium and platinum); four energy commodities—biofuel, crude oil and natural gas, gasoline (refining), and coal; and one soft commodity—lumber. The SIC and BICS codes utilized for matching, as well the PERMNO of the handpicked firms, are provided in Table A.3 in the Supplementary Material, which also contains details on the futures contracts selected.

We obtain daily stock market data for the commodity producers from CRSP. Our sample includes U.S.-based ordinary common shares (SHRCD = 10 or 11) and ordinary common stocks of Canadian firms traded on the NYSE, AMEX, and NASDAQ.[17] Stocks that are

[17] We remove stocks with share code 12 (unless it is a Canadian firm) and other share codes since those corporations are incorporated outside of North America. Thus, their equity prices are subject to further country-specific risk premia. Likewise, we exclude foreign firms trading on U.S. exchanges as American Depository Receipts (ADRs) (SHRCD = 31 or 32). We include ordinary common stocks of Canadian firms to increase our sample size, as many commodity producers are located in Canada and because of the close economic integration between the United States and Canada.

primarily trading on the over-the-counter markets are excluded (i.e., we restrict our attention to EXCHCD = 1, 2, or 3). To ensure that our results are not driven by penny stocks, we retain all firms that in the previous year have an average price above \$2. Then we obtain from CRSP the daily individual stock returns with dividends, adjusted by the delisting returns. Overall, our panel sample from January 2007 to March 2020 has 116,340 observations on the firm-week level, and contains 341 firms in total, with 192 firms on average per week.

III.C. Empirical Approach and Portfolio Formation

To investigate the informativeness of MM futures positions in the DCOT reports, our empirical approach relies on three different, albeit closely related, position change measures. We denote the long, short, and spreading positions held by MM as MM_l, MM_s, and MM_{sp}, respectively. The first signal measure we use is Long Proportion Growth, which is calculated as the growth rate in MM long positions divided by MM total positions—that is, growth in $\frac{MM_l+MM_{sp}}{MM_l+MM_s+2MM_{sp}}$. Similarly, we use the signal measure Short Proportion Growth—that is, growth in $\frac{MM_s+MM_{sp}}{MM_l+MM_s+2MM_{sp}}$. We also utilize Net Change, defined as the proportional change in long MM positions minus the proportional change in short MM positions—that is, $\frac{(MM_l+MM_{sp})_t}{(MM_l+MM_{sp})_{t-1}} - \frac{(MM_s+MM_{sp})_t}{(MM_s+MM_{sp})_{t-1}}$.

Given that the DCOT reports are tabulated weekly from the beginning of trading on Wednesday to Tuesday's close (which is the compilation date), we match this time interval by computing weekly returns for each of the firms identified as a commodity producer.[18] To be precise, the report compilation date, which is usually a Tuesday unless it is a federal holiday, is considered the signal-generation date for the position change signal which we would utilize a trading day later beginning on Wednesday, in determining whether to long or short the stock of a specific commodity producer. The stock is then held until the next

[18]Figure A.II in the Supplementary Material shows the total number of stocks traded in the long-short portfolio each week from January 2007 to March 2020. In a small number of weeks, there are minor dips and spikes in the total number of stocks because the DCOT data are often missing for steel and for coal prior to August 2012. Results remain largely the same if we exclude these intermittent gaps from our sample.

compilation date. Although the CFTC does not release traders' positions as of Tuesday until Friday, we adopt this timing schedule because, as discussed in Section II, we are interested in the study of costly information processing and gradual information diffusion along the lines of Cohen and Lou (2012) and Cohen et al. (2020), among other papers.

After compounding the daily commodity producer stock returns into weekly returns, we form ten portfolios, one for each of the selected commodity. We compute the weekly return series for each commodity-equity portfolio either as the equal-weighted average of weekly stocks' returns belonging to the same commodity or as the value-weighted average, in which case the stocks' market capitalizations at the end of December of the previous year are used as weights. Thereafter, each of the commodity-equity portfolios are matched to the weekly commodity positions data, which is then further aggregated into the long and short portfolios to yield weekly returns for our long-short portfolio. For example, we form our weekly long and short portfolios starting on Wednesday, July 17, 2013 utilizing data on MM positions compiled by the CFTC on (and as of) Tuesday, July 16 compared to the value compiled on Tuesday, July 9 (using one of the three signal measures). We refer to the timing of such signals as a 1-week lag. We will also consider the case of utilizing J-week backward-looking moving averages of lagged signals to dictate our portfolio formation.

Our main empirical approach relies on two distinct procedures to compute the long-short portfolio returns. For the case of calendar-time regression analysis, after either equal-weighting or value-weighting the stocks belonging to the same commodity per above, we construct the long-short portfolio by grouping the ten commodity-equity portfolios each week into two bins (i.e., one long portfolio and one short portfolio) according to the sign of their corresponding lagged MM position signals. Specifically, we long (short) the stocks of commodity-producing firms associated with a positive (negative) signal for each of the three MM signal measures. Next, within each of the long and short portfolios, the returns of these individual commodity-level equity portfolios are weighted according to the magnitude of the commodity signals so that their signals' strength is taken into account when averaging the portfolios' returns within the long (positive signal) portfolio and within the short (negative signal) portfolio, respectively.[19] We finally construct a zero-investment long-short portfolio by taking the difference between the long and short portfolio weekly returns. Differently, for the case of the single-sort analysis, the commodity-equity portfolios are sorted weekly into three bins based on the signals' values, and the commodity-equity portfolio returns are equally weighted within each tercile. We then derive the time series of the long-short portfolios' returns by going long on the highest tercile and going short on the lowest tercile. The Appendix describes the detailed procedures and formulae employed to compute the long-short portfolio returns.

IV. Benchmark Results

Our empirical analysis exploits the joint dynamics of commodity producers' equity price changes and MM position changes in the commodity futures market. We first examine the long-short portfolios' abnormal returns by using calendar-time regressions relative to commonly used factor models and also present a number of additional analyses. We then conduct a single-sort analysis to investigate whether subsequent returns exhibit a monotonic pattern according to the strength of the signal measures. Finally, we examine return predictability with Fama-MacBeth cross-sectional regressions after controlling for a set of common determinants of stock returns and the lagged change in the commodity price itself.

IV.A. Calendar-Time Regression Analysis

For Table 1, we follow a strategy of buying the commodity producers' stocks with positive signals and selling short the stocks with negative signals for two of the three MM signal measures (Net Change and Long Proportion Growth); but for Short Proportion Growth, we long the stocks if the signal is negative and short them if the signal is positive. We then present the return moments and summary statistics of the long-short portfolios of commodity-producing firms that are constructed with the MM position changes signals, measured either

[19] For the calendar-time regression analysis, our results are generally robust to averaging with equal-weight the returns of individual commodity-level equity portfolios within the long (and short) portfolio.

TABLE 1 Long-Short Portfolio Characteristics

This table presents the return moments and summary statistics of the long-short portfolios of U.S.-listed North American commodity producers sorted by three signal measures based on the position changes of MM in the commodity futures market, as described in Section III.C. In each week, we follow a strategy of buying (selling) the producers' stocks with positive (negative) signals for the Net Change and Long Proportion Growth measures, and vice versa for the Short Proportion Growth measure. The signals measures are constructed as a 1-week lag ($J = 1$) or as a 2-week backward moving average ($J = 2$). The means and standard deviations of the weekly long-short portfolios' returns in excess of the risk-free rate, the annualized Sharpe ratios, as well as the cumulative returns of $1 invested (over the entire sample period) are presented for each of the three signal measures.

Managed Money Signal Measure	**Net Change**		**Long Proportion Growth**		**Short Proportion Growth**	
Panel A: Equal-Weight	J=1	J=2	J=1	J=2	J=1	J=2
Mean excess return (%, per week)	0.22	0.28	0.27	0.29	0.19	0.25
Standard deviation (%, per week)	2.26	2.25	2.33	2.36	2.3	2.29
t-statistics	2.61	3.27	3.01	3.24	2.18	2.88
Annualized Sharpe ratio	0.72	0.9	0.83	0.89	0.6	0.79
Cumulative return of $1 invested	4.45	6.54	5.9	6.92	3.5	5.29
Panel B: Value-Weight	J=1	J=2	J=1	J=2	J=1	J=2
Mean excess return (%, per week)	0.23	0.23	0.26	0.24	0.21	0.2
Standard deviation (%, per week)	2.14	2.06	2.19	2.11	2.14	2.09
t-statistics	2.77	2.88	3.18	2.95	2.52	2.5
Annualized Sharpe ratio	0.76	0.79	0.87	0.81	0.69	0.69
Cumulative return of $1 invested	4.56	4.6	5.92	4.94	3.97	3.82

Sample period: January 2007–March 2020

as a 1-week-lag ($J = 1$) or as a 2-week backward moving average ($J = 2$), across the three signals. We thus identify a statistically significant mean return differential between the long and short portfolios—for example, the t-statistic is 3.01 for the 1-week-lagged Long Proportion Growth measure in the equal-weight case. Potentially, exposures to risk factors could explain at least part of the return differential.

Thus, to investigate the existence of abnormal returns, we regress the weekly returns of the long-short portfolios relative to the Carhart four-factor model, as well as the Fama and French (2015) five-factor model, to control for the profitability and asset growth (i.e., investment) factors. In addition to reporting the average returns, Table 2 also presents the alphas calculated as the intercepts from the weekly calendar-time portfolio return regressions, together with their t-statistics based on White (1980) standard errors. The table compiles results for the three MM signal measures which are either constructed as a 1-week lag or as a J-week backward moving average, with J equal to 2, 6, 9, or 12 weeks. Longer look-back horizons could further alleviate concerns about high-frequency noise present in position-based signals (Hong and Yogo, 2012). For Table 2, we follow a strategy of buying the commodity producers' stocks with positive signal and selling short the stocks with negative signals for all three MM signal measures—that is, unlike in Table 1, we no longer flip the long versus short portfolio for the Short Proportion Growth measure; likewise for the rest of the paper. We thus observe in Table 2 that the alphas are all positive except for the MM Short Proportion Growth case, as expected. The estimated alphas are generally statistically significant and robust across the MM signal measures and the weighting schemes used, not only for signals with immediate look-back horizons ($J = 1, 2$) but also for signals with longer horizons (such as $J = 12$). Focusing on the $J = 2$ look-back horizon, across the three MM signals and the two weights, the alphas are economically large relative to both factor models, with magnitudes at around 25 basis points per week (corresponding to an alpha of around 13% per annum), and are highly statistically significant with t-statistics ranging from 2.73 to 3.49.

We perform additional analyses from a variety of angles and present the numeric results in the Supplementary Material. Specifically, Hou et al. (2020) find that many asset pricing anomalies disappear after dropping small-cap stocks. Commodity producers are generally large-cap stocks, and Table B.1 in the Supplementary Material shows that our results are not sensitive to the removal of stocks in the bottom 45% of market capitalization using annually updated cutoffs from the CRSP universe. In addition, Panel A of Table B.2 in the Supplementary Material reveals robust findings for the short-term signal lags when the abnormal returns in the calendar-time regression analysis are evaluated in terms of the Stambaugh and Yuan (2016) mispricing factor model, which includes, in addition to the factors of market and size, two composite factors based on a set of 11 prominent anomalies.[20]

[20]Specifically, MGMT is a composite factor constructed with six characteristics related to investment and financing, while the second cluster, PERF, is based on five characteristics, including return momentum and profitability. All our results pertaining to Stambaugh and Yuan's mispricing factor model are based on a shorter sample period ending in December 2016 due to factor data availability.

TABLE 2 Calendar-Time Portfolio Return Regression Results (%, per Week)

This table presents the average returns and alphas from weekly calendar-time regressions of the portfolio returns of U.S.-listed North American commodity producers sorted with respect to each of the three MM signal measures. The signals are constructed as a 1-week lag or as a J-week backward moving average. After equal-weighting (Panel A) or value-weighting (Panel B) the stock returns belonging to the same commodity, the commodity-equity portfolios are averaged weekly into two portfolio bins by buying the stocks with positive signal (''Pos'') and selling short the stocks with negative signal (''Neg''), as described in Section III.C. From the long-short portfolio returns (''Pos−Neg''), we then calculate the abnormal return (α) relative to the Carhart four-factor model (C4 α) and to the Fama and French five-factor model (FF5 α). The average weekly portfolio returns and α's, multiplied by 100 so they can be interpreted as percentages, are reported together with their t-statistics in parentheses (based on White standard errors). $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Managed Money Long Proportion Growth

Portfolio Rank	Panel A: Equal-Weight					Panel B: Value-Weight				
	J=1	J=2	J=6	J=9	J=12	J=1	J=2	J=6	J=9	J=12
Neg	−0.057	−0.087	−0.075	−0.083	−0.146	−0.054	−0.041	−0.006	−0.035	−0.094
Pos	0.227	0.220	0.172	0.120	0.196	0.227	0.212	0.150	0.126	0.199
(Pos−Neg)	0.284***	0.308***	0.248***	0.203**	0.343***	0.281***	0.253***	0.156**	0.161*	0.293***
	(3.21)	(3.43)	(2.89)	(2.25)	(3.41)	(3.38)	(3.16)	(2.05)	(1.86)	(3.09)
C4 α	0.299***	0.316***	0.252***	0.217**	0.371***	0.299***	0.267***	0.166**	0.178**	0.321***
	(3.36)	(3.47)	(2.88)	(2.34)	(3.40)	(3.57)	(3.28)	(2.12)	(2.01)	(3.12)
FF5 α	0.296***	0.309***	0.246***	0.210**	0.351***	0.289***	0.257***	0.158**	0.163*	0.301***
	(3.35)	(3.40)	(2.80)	(2.22)	(3.24)	(3.45)	(3.17)	(2.00)	(1.84)	(2.98)

Managed Money Net Change

Portfolio Rank	Panel A: Equal-Weight					Panel B: Value-Weight				
	J=1	J=2	J=6	J=9	J=12	J=1	J=2	J=6	J=9	J=12
Neg	−0.029	−0.096	−0.003	−0.075	−0.122	−0.046	−0.060	0.039	−0.053	−0.092
Pos	0.213	0.201	0.153	0.178	0.190	0.197	0.182	0.127	0.179	0.192
(Pos−Neg)	0.241***	0.297***	0.156*	0.253**	0.312***	0.242***	0.242***	0.088	0.232**	0.283***
	(2.81)	(3.47)	(1.76)	(2.23)	(2.61)	(2.98)	(3.10)	(1.08)	(2.28)	(2.60)
C4 α	0.264***	0.308***	0.176**	0.315***	0.364***	0.262***	0.257***	0.109	0.285***	0.330***
	(3.01)	(3.49)	(1.97)	(2.75)	(3.09)	(3.16)	(3.21)	(1.33)	(2.79)	(3.10)
FF5 α	0.264***	0.292***	0.162*	0.290**	0.323***	0.254***	0.243***	0.101	0.263***	0.298***
	(3.03)	(3.31)	(1.77)	(2.54)	(2.78)	(3.07)	(3.05)	(1.23)	(2.59)	(2.87)

Managed Money Short Proportion Growth

Portfolio Rank	Panel A: Equal-Weight					Panel B: Value-Weight				
	J=1	J=2	J=6	J=9	J=12	J=1	J=2	J=6	J=9	J=12
Neg	0.205	0.189	0.177	0.183	0.202	0.189	0.158	0.134	0.170	0.192
Pos	−0.002	−0.078	0.020	−0.048	−0.111	−0.033	−0.058	0.054	−0.030	−0.079
(Pos−Neg)	−0.208**	−0.267***	−0.157*	−0.231**	−0.313**	−0.222***	−0.216***	−0.079	−0.200*	−0.271**
	(−2.37)	(−3.07)	(−1.79)	(−1.97)	(−2.56)	(−2.73)	(−2.72)	(−0.97)	(−1.93)	(−2.48)
C4 α	−0.234***	−0.285***	−0.182**	−0.293**	−0.371***	−0.244***	−0.235***	−0.103	−0.253**	−0.323***
	(−2.62)	(−3.15)	(−2.03)	(−2.49)	(−3.10)	(−2.92)	(−2.88)	(−1.24)	(−2.44)	(−3.04)
FF5 α	−0.235***	−0.267***	−0.165*	−0.267**	−0.324***	−0.236***	−0.222***	−0.096	−0.233**	−0.290***
	(−2.65)	(−2.97)	(−1.81)	(−2.28)	(−2.75)	(−2.84)	(−2.73)	(−1.15)	(−2.27)	(−2.80)

On the other hand, it is interesting to find in Table B.3 in the Supplementary Material that total open interest growth—based on the aggregate positions of all five trader categories in the commodity futures markets—does not seem to contain information that is robustly conducive to the predictability of commodity producers' stock returns in the short term.[21] Total open interest aggregates heterogeneous groups of traders with varying incentives to trade, and unsurprisingly, this limits its usefulness as a signal. Overall, the results are in line with the notion that the MM categorical position changes reflect the traders who are sophisticated speculators, who can be and often are levered, and who have the most incentive to timely process and acquire information related to movements in the fundamentals of the commodity market.

Delving deeper, we also analyze the positions of PM and SW categories disclosed in the CFTC DCOT reports. The calendar-time regression results are presented in Tables B.4 and B.5 in the Supplementary Material. As stressed in Section III.A, however, these two trader categories, as defined by the CFTC, are less homogeneous than MM with respect to the directionality of its constituents, with diverse incentives for producers versus users of the commodity in the PM group, while the SW category masks important heterogeneity in the types of counterparties and their respective trading motives. In light of these caveats, the signals based on these two trader categories turn out to be marginally useful but not as (robustly) informative as the MM signals in the format of data that is publicly available in the DCOT reports.

[21] Hong and Yogo (2012) construct a predictor of commodity futures market returns by taking a 12-month geometric average of monthly total open interest growth (aggregated across all commodities) to study the predictability of aggregate market returns indexes. In contrast, we study at a weekly frequency the cross-sectional predictability of commodity producers' stock returns, and our signal is commodity-specific with a look-back horizon of at most 12 weeks. Overall, our insignificant result on total open interest growth does not relate to Hong and Yogo's finding at annual frequency that "commodity open interest contains information about future economic activity and inflation expectations which is impounded into the equity market (aggregate index) with delay." Instead, our results show that information extracted from total open interest growth is not useful in predicting producers' stock returns, congruent with our main thesis that predictability arises from the positions data pertaining specifically to MM due to specialization, segmentation, and gradual information diffusion.

IV.B. Single-Sort Analysis

To control for any potential nonlinear relation between the MM signal measures and subsequent stock returns, we also adopt a single-sort procedure as part of our analysis, in a way that differs in some respects from the method previously outlined. After either equal-weighting or value-weighting the stock returns belonging to the same commodity, the ten commodity-equity portfolios are sorted weekly—in this case, into three bins based on the signal's value—and averaged within each tercile with equal-weight. We then compute the long-short returns of going long on the highest tercile and going short on the lowest tercile. Henceforth, with the single-sort procedure, we are able to investigate the price impacts of different signal bins and check whether there is a monotonic relation between signal strength and subsequent portfolio returns.[22]

Table 3 shows the results for both the equal-weighted (Panel A) and value-weighted (Panel B) returns for each portfolio bin. First, for all three signal measures in both panels, the pattern of average returns for the three portfolio bins generally confirms the monotonicity of the raw returns over the signal-ranked bins. Second, the return differences ("3−1") between the highest and the lowest portfolio bins are economically large and highly statistically significant across the three MM signal measures, the two weighting schemes applied, and hold generally for any signal horizons—whether for signals utilizing the immediate past ($J = 1, 2$) or for signals with intermediate look-back horizons. The results also reveal that the abnormal returns' magnitude and high statistical significance hold not only for the Carhart four-factor model, with t-statistics ranging, for instance, from 2.42 to 3.72 for the MM Net Change signal, but also when the alphas are evaluated in terms of the Fama and French five-factor model. As shown in Panel B of Table B.2 in the Supplementary Material, qualitatively similar results can also be obtained relative to the Stambaugh and Yuan factor model.

[22] In contrast to single-sorting, the calendar-time regression analysis was performed by going long (short) on stocks according to the sign of the MM signal; the signal's strength was not a factor in determining the bins (beyond its sign) but was only used to weight the commodity-equity portfolios inside the long and short portfolios.

TABLE 3 Single-Sort Results (%, per Week)

After equal-weighting (Panel A) or value-weighting (Panel B) the producers' stock returns belonging to the same commodity, the commodity-equity portfolios are sorted weekly into three portfolio bins (bin 3 associated with the highest signals) based on the MM signals and averaged within each tercile with equal-weight, following Section III.C. The signals are constructed as a 1-week lag or as a J-week backward moving average. The average returns for each bin, and the long-short returns ("3−1") of going long on the highest tercile and going short on the lowest tercile are displayed. As a minor note, for ease of qualitative comparison, we let each of the individual portfolios start with an exposure of \$0.50 to ensure that the long-short difference has an overall exposure of \$1. From the long-short returns, we then calculate the alphas relative to the Carhart four-factor model (C4 α) and to the Fama and French five-factor model (FF5 α). The average weekly portfolio returns and α's, multiplied by 100 so they can be interpreted as percentages, are reported together with their t-statistics in parentheses (based on White standard errors). ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Managed Money Long Proportion Growth

Portfolio Rank	Panel A: Equal-Weight					Panel B: Value-Weight				
	J=1	J=2	J=6	J=9	J=12	J=1	J=2	J=6	J=9	J=12
1	−0.066	−0.113	−0.082	−0.087	−0.085	−0.063	−0.083	−0.02	−0.036	−0.074
2	0.024	0.067	0.037	0.074	−0.003	0.014	0.076	0.041	0.034	0.008
3	0.18	0.182	0.205	0.157	0.24	0.209	0.162	0.149	0.173	0.221
(3–1)	0.245***	0.294***	0.287***	0.245***	0.324***	0.272***	0.245***	0.169**	0.21***	0.295***
	(3.30)	(3.48)	(3.55)	(2.99)	(3.88)	(3.76)	(3.21)	(2.27)	(2.80)	(3.98)
C4 α	0.249***	0.285***	0.284***	0.242***	0.33***	0.277***	0.245***	0.175**	0.21***	0.304***
	(3.27)	(3.34)	(3.43)	(2.78)	(3.83)	(3.72)	(3.16)	(2.29)	(2.66)	(4.00)
FF5 α	0.247***	0.282***	0.277***	0.231***	0.312***	0.263***	0.235***	0.164**	0.196**	0.29***
	(3.27)	(3.30)	(3.34)	(2.64)	(3.62)	(3.58)	(3.05)	(2.13)	(2.50)	(3.84)

Managed Money Net Change

Portfolio Rank	Panel A: Equal-Weight					Panel B: Value-Weight				
	J=1	J=2	J=6	J=9	J=12	J=1	J=2	J=6	J=9	J=12
1	−0.057	−0.101	−0.047	−0.083	−0.033	−0.062	−0.067	0.005	−0.028	−0.024
2	−0.005	0.097	−0.02	0.047	−0.033	0.011	0.094	−0.016	0	−0.025
3	0.209	0.127	0.219	0.164	0.194	0.209	0.122	0.181	0.189	0.195
(3–1)	0.266***	0.229***	0.266***	0.247***	0.227***	0.271***	0.189***	0.177**	0.217***	0.218***
	(3.53)	(2.98)	(3.42)	(3.29)	(2.90)	(3.67)	(2.66)	(2.42)	(3.04)	(3.16)
C4 α	0.276***	0.232***	0.273***	0.249***	0.24***	0.281***	0.193***	0.182**	0.216***	0.226***
	(3.58)	(2.96)	(3.40)	(3.14)	(2.98)	(3.72)	(2.66)	(2.42)	(2.94)	(3.21)
FF5 α	0.267***	0.231***	0.256***	0.24***	0.227***	0.265***	0.187**	0.168**	0.206***	0.22***
	(3.44)	(2.92)	(3.19)	(3.03)	(2.84)	(3.51)	(2.56)	(2.23)	(2.80)	(3.13)

Managed Money Short Proportion Growth

Portfolio Rank	Panel A: Equal-Weight					Panel B: Value-Weight				
	J=1	J=2	J=6	J=9	J=12	J=1	J=2	J=6	J=9	J=12
1	0.192	0.162	0.199	0.154	0.173	0.196	0.141	0.155	0.143	0.173
2	0.01	0.053	0.001	0.013	0.035	0.034	0.068	0.01	0.009	0.031
3	−0.037	−0.069	−0.038	−0.023	−0.068	−0.057	−0.051	0.008	0.01	−0.045
(3–1)	−0.229***	−0.231***	−0.236***	−0.177**	−0.24***	−0.253***	−0.192***	−0.147**	−0.133*	−0.218***
	(−2.99)	(−2.99)	(−3.18)	(−2.29)	(−3.01)	(−3.47)	(−2.70)	(−2.08)	(−1.87)	(−3.08)
C4 α	−0.245***	−0.232***	−0.248***	−0.187**	−0.252***	−0.264***	−0.193***	−0.159**	−0.138*	−0.223***
	(−3.09)	(−2.97)	(−3.23)	(−2.30)	(−3.10)	(−3.53)	(−2.69)	(−2.20)	(−1.89)	(−3.13)
FF5 α	−0.243***	−0.223***	−0.23***	−0.18**	−0.233***	−0.257***	−0.186***	−0.149**	−0.134*	−0.213***
	(−3.05)	(−2.89)	(−2.98)	(−2.20)	(−2.89)	(−3.42)	(−2.61)	(−2.05)	(−1.81)	(−2.99)

IV.C. Fama-MacBeth Regressions

We utilize Fama-MacBeth cross-sectional regression as in Fama and French (2008) to investigate whether conditional on controls such as firm size, book-to-market ratio, short-term reversal, momentum, and past change in commodity futures prices, MM position changes can predict commodity producers' stock returns (at firm-level) in the week following the DCOT report. Specifically, we estimate the following cross-sectional regression in each week t:

$$R_{i,t} = \alpha_t + \beta_t Signal_{c,t-J} + \theta_t FR_{c,t-1} + \gamma_t' X_{i,t-1} + \epsilon_{i,t}, \qquad i = 1, 2, ..., N, \tag{1}$$

where $R_{i,t}$ is the return (subtracted by the risk-free rate) of the stock of firm i belonging to commodity c over week t (from Wednesday through the next Tuesday's compilation) following a new signal $Signal_{c,t-J}$ extracted from MM's futures market positions for commodity c. The signal $Signal_{c,t-J}$ compiles futures market positions over the previous week $t-1$, or over a $t-J$ weeks look-back horizon, where J is equal to 2, 3, 6, 9, or 12. The vector $X_{i,t-1}$ includes a set of stock-specific control variables.[23] As an additional control, we also include $FR_{c,t-1}$, which is the relative change (i.e., return) in the futures price of commodity c in the previous week until portfolio formation. If our signals remain significant in the presence of this control, it would indicate that the positions of MM do have predictive power in addition to the information already contained in commodity prices, and we indeed find an affirmative result.

The resulting parameter estimates are time series averages of weekly regression coefficient estimates. The coefficient on our signal, $\beta = \frac{1}{T}\sum_{t=1}^{T}\beta_t$, is our focus. The t-statistics are based on the time series variability of the cross-sectional slope estimates and rely on robust standard errors as the signal, and the cross-sectional regressions share the same weekly frequency. We present the results using the MM Long Proportion Growth signal in Table 4.[24] Odd-numbered

[23]These control variables are ret_{-1}, the stock return over the previous month; $ret_{-2,-12}$, the stock return over the 11 months preceding the previous month; $ln(ME)$, the log of the market value of equity at the end of the previous calendar year; and $ln(BE/ME)$, the log of the book-to-market value of the firm's equity, where the book value of equity is measured at the end of the previous fiscal year.

[24]We obtain similar results for the other two MM signal measures. See Supplementary Material Tables B.9 and B.10.

columns control for selected firm characteristics, while even-numbered columns add the prior week commodity futures returns as an additional regressor. In terms of the average adjusted R^2's, the values at around 10% are comparable in magnitude with to ones found in the literature.

Table 4 shows that our signal is robustly statistically significant across a variety of choices of lags and specifications. We find that a 1-week lag of our signal in the commodity futures market can predict future stock returns with a t-statistic of 2.93, and a t-statistic of 3.09 when $FR_{c,t-1}$ is included as a regressor, among other controls. In terms of economic magnitude, take for example the case of the 6-week moving average MM signal in column (8): the regression coefficient of 0.151 on the signal conveys that on average, *ceteris paribus*, a one-standard-deviation increase in the signal (3.85%) implies a 0.58% increase in the return of associated commodity producer in the following 1 week. Similarly in column (8), the regression coefficient of 0.104 on $FR_{c,t-1}$ conveys that, *ceteris paribus*, a one-standard-deviation increase in the 1-week-lagged commodity futures returns (3.69%) implies a 0.38% increase in the commodity producer's return in the following week.[25]

To summarize, both the single-sort and the calendar-time regression analyses show that economically large and statistically significant abnormal returns can be attributed to the lead-lag relationship between the commodity futures market and the stocks of commodity producers, utilizing signals extracted from MM's futures positions. This lead-lag effect is not captured by commonly used factors of equity returns and is robust to a variety of choices of signal measures, weighting schemes, and timing of lags. Fama-MacBeth cross-sectional regressions corroborate the portfolio results and further confirm that the positions of MM do have predictive power that is not captured by a set of commonly used determinants of stock returns and is beyond the information contained in the change in commodity price itself.

[25]Since Fama-MacBeth regressions are designed to account for a time fixed effect, we have also conducted regressions at a daily frequency if one supposes that it is best to control for daily events common to all stocks rather than weekly shocks. Specifically, regressions are performed for each of the 5 trading days following a new DCOT report on a Tuesday, even though the value of the signal is updated weekly. The estimated average slopes (based on a Newey-West correction with five lags), reported in Supplementary Material Table B.11, are qualitatively similar to their weekly counterparts.

TABLE 4 Fama-MacBeth Regressions, Managed Money Long Proportion Growth

This table shows results from Fama-MacBeth cross-sectional regressions (average slopes, and t-statistics based on White standard errors in parentheses) of firms' subsequent weekly return (subtracted by the risk-free rate) on lagged signal and other lagged controls for expected returns. The weekly return of the firm occurs within 7 calendar days (the first is always a Wednesday and the last is always a Tuesday unless they are postponed due to public holidays) following the newest DCOT report. We run the Fama-MacBeth regression at a weekly frequency. The signals from the futures market are constructed as a 1-week lag or as a J-week backward moving average. ret_{-1} is the stock return over the previous month, $ret_{-2,-12}$ is the stock return over the 11 months preceding the previous month, $ln(ME)$ is the log of the market value of equity at the end of the previous calendar year, and $ln(BE/ME)$ is the log of the book-to-market value of equity, where the book value is measured at the end of the previous fiscal year. $FR_{c,t-1}$ is the relative change in commodity price over the previous week. The row labeled Adj. R^2 displays the average of the cross-sectional adjusted R^2's. N–Companies is the number of unique firms, and N–Observations is the number of weeks utilized in the regression. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	J=1		J=2		J=3	
	1	2	3	4	5	6
Managed Money Long Proportion Growth	0.076***	0.089***	0.084***	0.067***	0.070**	0.056*
	(2.930)	(3.090)	(3.310)	(2.610)	(2.220)	(1.890)
$ln(BE/ME)$	0.000	0.000	0.000	0.000	0.000	0.000
	(−0.430)	(−0.250)	(−0.220)	(−0.190)	(−0.270)	(−0.230)
$ln(ME)$	0.000	0.000	0.000	0.000	0.000	0.000
	(0.300)	(0.530)	(0.440)	(0.590)	(0.470)	(0.660)
ret_{-1}	−0.001	−0.002	0.000	−0.002	−0.001	−0.002
	(−0.180)	(−0.560)	(−0.060)	(−0.510)	(−0.150)	(−0.550)
$ret_{-2,-12}$	0.001	0.001	0.001	0.001	0.001	0.001
	(0.520)	(0.560)	(0.520)	(0.510)	(0.510)	(0.440)
$FR_{c,t-1}$		0.037		0.088***		0.106***
		(0.960)		(2.590)		(3.200)
N–Observations	691	691	691	691	691	691
N–Companies	328	328	328	328	328	328
Adj. R^2	0.106	0.129	0.106	0.127	0.106	0.126
	J=6		J=9		J=12	
	7	8	9	10	11	12
Managed Money Long Proportion Growth	0.152***	0.151***	0.220***	0.204***	0.176**	0.201***
	(2.930)	(3.020)	(3.350)	(3.310)	(2.460)	(2.620)
$ln(BE/ME)$	0.000	0.000	0.000	0.000	0.000	0.000
	(−0.180)	(−0.140)	(0.070)	(0.020)	(−0.030)	(−0.150)
$ln(ME)$	0.000	0.000	0.000	0.000	0.000	0.000
	(0.400)	(0.470)	(0.320)	(0.510)	(0.420)	(0.520)
ret_{-1}	−0.001	−0.003	−0.002	−0.003	−0.002	−0.004
	(−0.330)	(−0.780)	(−0.420)	(−0.800)	(−0.400)	(−0.980)
$ret_{-2,-12}$	0.001	0.001	0.001	0.001	0.001	0.001
	(0.590)	(0.540)	(0.600)	(0.500)	(0.590)	(0.470)
$FR_{c,t-1}$		0.104***		0.091***		0.097***
		(3.110)		(2.800)		(2.950)
N–Observations	691	691	691	691	691	691
N–Companies	328	328	328	328	328	328
Adj. R^2	0.105	0.126	0.105	0.127	0.106	0.128

V. Contributions to Our Predictability Results

We now turn to investigating potential explanations behind this documented lead-lag relationship.[26] We begin by ruling out that our results would arise simply due to a mechanical link between the commodity producers' stock returns and the contemporaneous commodity futures returns. As laid out in Section II, our story emphasizes costly information processing, which leads to specialization, segmentation, and gradual information diffusion. Along these lines, we present evidence supporting that the return predictability likely reflects informational friction by being more pronounced for firms with higher historical stock volatility or analyst forecast dispersion, consistent with our second empirical prediction, wherein the equity returns of relatively nontransparent commodity producers are the slowest to price-adjust. Finally, we address the concern that our MM signals do not mainly capture, as we posit, informative reflections of ‘‘smart money'' and their revised prospects on commodities' fundamentals, but rather reflect the positions of trend followers within MM (or other common commodity futures strategies), a concern that finds little empirical support.[27]

V.A. A Mechanical Link with Contemporaneous Futures Returns?

One would expect the stock returns of commodity producers to be tied to the returns of the underlying commodities (Tufano, 1998, states that the relation is time varying).[28] Accordingly, to the extent to which stock returns are contemporaneously correlated to futures market returns, MM position changes may predict producers' stock returns for next week t simply because they are related to commodity futures returns.

---

[26] This section has benefited greatly from the comments and suggestions of an anonymous referee.

[27] We also confirm with further analyses that predictability is not simply the result of a self-fulfilling prophecy in which the market participants are just following MM positions after announcements of the DCOT reports on Fridays. We have conducted analyses from two angles: i) by decomposing our single-sort results separately for the Wednesday–Friday and the Monday–Tuesday intervals, and ii) by relying on the high-frequency Trade and Quote dataset to see if there is any immediate market reaction to the reports' release at 3.30 p.m. on Fridays. We find that our predictability results are already present prior to the release of MM positions and are not attributed to the announcement effects of the reports. The procedures and results are reported in Supplementary Material Section B.2.

[28] Tufano (1998) finds that a 1% change in gold prices implies contemporaneously a 2% change in the stock prices of North American gold miners, but the exposures vary across time and firms.

TABLE 5 Single-Sort Results (%, per Week), Stock Returns Orthogonalized from Contemporaneous Commodity Futures Returns

In the first stage, we project the commodity producers' weekly stocks returns in week t onto the contemporaneous (week t) commodity futures returns—that is, we run separately for each commodity c: $r_{i,c,t} = \mu_c + \beta_c FR_{c,t} + \epsilon_{i,c,t}$. We then define $\hat{r}^{residual}_{i,c,t} = \hat{\epsilon}_{i,c,t}$. In the second stage, we run the single-sort procedure, wherein after the orthogonalized stock returns ($\hat{r}^{residual}_{i,c,t}$) belonging to the same commodity are either equal-weighted (EW) or value-weighted (VW), the commodity-equity portfolios are sorted weekly into three portfolio bins based on the MM signal's value and averaged within each tercile with equal-weight. The signals are constructed as a 1-week lag or as a J-week backward moving average. From the long-short returns constructed by longing the highest tercile while shorting the lowest tercile, we then calculate the α's relative to the Carhart four-factor model (C4 α). The table presents the weekly α's (which are multiplied by 100) with their t-statistics in parentheses based on White standard errors. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Money Managers' Signal Measure:	**Net Change**		**Long Proportion Growth**		**Short Proportion Growth**	
	EW	VW	EW	VW	EW	VW
J=1	0.261***	0.268***	0.246***	0.276***	−0.207***	−0.228***
	(3.61)	(3.95)	(3.36)	(3.94)	(−2.83)	(−3.31)
J=2	0.195***	0.155**	0.296***	0.258***	−0.192***	−0.153**
	(2.71)	(2.36)	(3.95)	(3.73)	(−2.66)	(−2.32)
J=6	0.265***	0.174***	0.372***	0.263***	−0.255***	−0.166**
	(3.68)	(2.59)	(4.88)	(3.84)	(−3.63)	(−2.46)
J=9	0.159**	0.126*	0.282***	0.249***	−0.134*	−0.085
	(2.22)	(1.86)	(3.54)	(3.55)	(−1.89)	(−1.31)
J=12	0.208***	0.195***	0.306***	0.280***	−0.224***	−0.195***
	(2.86)	(2.99)	(3.91)	(3.98)	(−2.98)	(−2.92)

To explore this potential channel, we first project the equity returns onto the same-week futures returns for each commodity in our sample, and see whether our MM position change signals are still able to predict the residuals from these first-stage regressions. Table 5 presents our single-sort analysis applied to these orthogonalized stock returns. From the first-stage regressions, we confirm that indeed stock returns are positively related to contemporaneous futures returns (p-value < 0.01). Table 5 shows that the Carhart four-factor α's are sometimes reduced slightly in magnitude (e.g., for the Net Change measure) after accounting for this mechanical link with contemporaneous futures returns. Nevertheless, the MM signals still yield large and strongly statistically significant alphas. We can thus largely discount the avenue that our predictability results arise solely because of the contemporaneous link between the stock returns of commodity firms and their underlying futures returns.

V.B. Double-Sort Results and the Relation with Market Frictions

Although a mechanical link between producers' stock returns and contemporaneous futures returns fails to be the driver behind predictability, we now use the double-sort approach to focus the investigation on two sets of market frictions that could potentially contribute to our results—namely, informational friction and trading friction. As our main thesis is that return predictability arises from costly information processing, which leads to gradual information diffusion, we establish in this section that indeed, informational friction, instead of trading friction, is the cause behind the predictability patterns and return predictability is more pronounced in commodity-producing firms with higher information asymmetry.

Informational friction is at the center of a number of consistent findings in the literature on return anomalies. Sadka and Scherbina (2007) find that firms with higher analyst dispersion earn lower subsequent returns because these firms are believed to have higher information asymmetry. We proxy informational friction based on two measures, *ex ante* analyst dispersion and 90-day *historical* stock volatility. The two proxies do not necessarily coincide with each other. The first measure is forward-looking and based on market expectations of prospective performance variation, whereas the second measure is backward-looking and calculated on realized historical data. To ensure that the analyst dispersion is properly calculated from the Institutional Brokers' Estimate System (IBES) data, we require a minimum coverage of three analysts per stock. Exogenous transaction costs, demand pressure, inventory, and search friction risk are all possible sources of illiquidity, and we use the illiquidity measure proposed by Amihud (2002) as a proxy for trading friction.

We utilize double-sort as our main approach. Specifically, each week, all commodity-producing stocks are first sorted into three friction portfolios using one of the three aforementioned firm-level proxies for market friction, with the requirement that each commodity appears across those three portfolios. Within each of the friction portfolios, the producers' stock returns belonging to the same commodity are either equal-weighted or value-weighted into commodity-equity portfolios. Then, the commodity-equity portfolios are sorted dependently within each friction portfolio based on the sign of the MM signal to form two signal portfolios. The three-by-two double-sort method thus produces six portfolios.

Table 6 presents the results for the returns double-sorted with the three friction proxies and utilizing the MM Net Change signal measure as a 2-week backward moving average. We pay special attention to the difference in long-short portfolio returns (“2−1”) and whether the four- or five-factor alphas arise in the difference between the high- and low-friction bins that correspond to the (“3−1”) column at the rows (“2−1”), C4 α, FF5 α, and SY4 α, which are all marked in bold in the table.[29] By double-sorting our commodity futures market signal with Amihud’s illiquidity measure (LIQ), we find no convincing evidence that predictability is stronger (or weaker) in firms with higher trading friction. However, we do find that our results are significantly stronger in firms with higher information asymmetry, as measured by the 90-day *historical* stock volatility (VOL) and the *ex ante* analyst dispersion (AD), as compared to the firms with lower information asymmetry. For instance, if we focus on the Value-Weight column, the difference in the long-short portfolio’s C4 α between the highest and lowest terciles of *ex ante* analyst dispersion is in itself a significant difference, with a t-statistic of 2.81. Similar results are observed when we use the other two MM signal measures, which are presented in Supplementary Material Section B.3.

We thus confirm that the lead-lag relationship is due to informational friction rather than trading friction such as stock liquidity, consistent with the main thesis established in Section II—namely, that costly (in terms of time and effort) information processing leads to gradual incorporation of information, thus the equity price of a commodity-producing firm, which has high information asymmetry, is the slowest asset to price-adjust to value-relevant updates regarding commodities. This is in concordance with the finding of Cohen and Lou (2012) that the “complicated information processing channel,” rather than the “complicated trading mechanism,” is the underlying channel behind their results.

[29] As a minor note, for ease of qualitative comparison, we let each of the individual portfolios start with an exposure of \$0.25 to ensure that the values in bold have an overall exposure of \$1.

TABLE 6 Double-Sort: Managed Money Net Change (%, per Week)

This table presents results of a double-sort cross-sectional exercise. AD, VOL, and LIQ stand for the ex ante analyst dispersion, the 90-day historical stock volatility, and the Amihud illiquidity measure, respectively. In each week, all the producer stocks are first sorted into three friction portfolios using one of the three firm-level proxies of friction (AD, VOL, and LIQ) with the requirement that each commodity appears across those three portfolios. Column "3" is associated with the highest friction (column "1" with the lowest friction). Within each friction portfolio, the producers' stock returns belonging to the same commodity are either equal-weighted (Panel A) or value-weighted (Panel B) into commodity-equity portfolios. Then the commodity-equity portfolios are sorted dependently within each friction portfolio based on the sign of the MM Net Change signal to form two signal portfolios, with row "2" associated with positive signal value (row "1" with negative signal), which yields the long-short portfolio returns ("2−1"). The MM signal is constructed as a 2-week backward moving average. This three-by-two double-sort procedure produces six portfolios. We then evaluate the α's relative to the Carhart four-factor model (C4), the Fama and French five-factor model (FF5), and the Stambaugh and Yuan mispricing factor model (SY4). We pay special attention to the difference in long-short portfolio returns and whether the four- or five-factor alphas arise in the difference between the high- and low-friction bins which correspond to the ("3−1") column at the rows ("2−1"), C4 α, FF5 α, and SY4 α which are in bold. The weekly average returns and α's, multiplied by 100 so they can be interpreted as percentages, are reported together with their t-statistics in parentheses (based on White standard errors). ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

	Panel A: Equal-Weight				**Panel B: Value-Weight**			
AD / Signal	1	2	3	(3–1)	1	2	3	(3–1)
1	0	−0.055	−0.15**	−0.153***	−0.001	−0.05	−0.155**	−0.156***
	(0.01)	(−1.06)	(−2.32)	(−3.67)	(−0.03)	(−0.99)	(−2.49)	(−3.76)
2	0.043	0.036	0.047	−0.007	0.032	0.031	0.046	0.003
	(0.90)	(0.70)	(0.71)	(−0.14)	(0.71)	(0.61)	(0.72)	(0.06)
(2–1)	0.042	0.092**	0.189***	**0.147****	0.033	0.08*	0.192***	**0.159*****
	(1.02)	(2.13)	(2.97)	**(2.50)**	(0.82)	(1.84)	(3.01)	**(2.73)**
C4 α	0.05	0.096**	0.205***	**0.156****	0.042	0.089**	0.213***	**0.171*****
	(1.21)	(2.17)	(3.11)	**(2.53)**	(1.04)	(2.06)	(3.31)	**(2.81)**
FF5 α	0.046	0.083*	0.193***	**0.147****	0.04	0.078*	0.202***	**0.162*****
	(1.14)	(1.96)	(3.00)	**(2.43)**	(1.02)	(1.86)	(3.22)	**(2.72)**
SY4 α	0.037	0.093*	0.252***	**0.215*****	0.029	0.086*	0.269***	**0.239*****
	(0.72)	(1.81)	(3.01)	**(2.74)**	(0.60)	(1.65)	(3.35)	**(3.10)**
VOL / Signal	1	2	3	(3–1)	1	2	3	(3–1)
1	−0.031	−0.037	−0.145**	−0.117***	−0.033	−0.028	−0.152**	−0.121***
	(−0.70)	(−0.72)	(−2.16)	(−2.80)	(−0.77)	(−0.53)	(−2.25)	(−2.76)
2	0.002	0.05	0.045	0.032	0.006	0.045	0.051	0.032
	(0.04)	(0.97)	(0.65)	(0.68)	(0.14)	(0.88)	(0.74)	(0.69)
(2–1)	0.033	0.088**	0.183***	**0.15*****	0.04	0.07	0.196***	**0.157*****
	(0.82)	(2.06)	(2.89)	**(2.75)**	(0.95)	(1.61)	(3.08)	**(2.82)**
C4 α	0.041	0.092**	0.197***	**0.156*****	0.05	0.075*	0.212***	**0.162*****
	(1.07)	(2.12)	(2.98)	**(2.70)**	(1.27)	(1.72)	(3.22)	**(2.74)**
FF5 α	0.037	0.079*	0.182***	**0.145****	0.046	0.064	0.201***	**0.156*****
	(1.00)	(1.89)	(2.83)	**(2.53)**	(1.18)	(1.51)	(3.10)	**(2.63)**
SY4 α	0.036	0.086*	0.227***	**0.191*****	0.044	0.065	0.247***	**0.203*****
	(0.77)	(1.66)	(2.78)	**(2.59)**	(0.90)	(1.25)	(3.03)	**(2.67)**

TABLE 6 Double-Sort: Managed Money Net Change (%, per Week) (Continued)

	Panel A: Equal-Weight				Panel B: Value-Weight			
Signal \ LIQ	1	2	3	(3–1)	1	2	3	(3–1)
1	−0.063	−0.069	−0.097	−0.037	−0.067	−0.066	−0.094	−0.03
	(−1.22)	(−1.26)	(−1.60)	(−0.95)	(−1.33)	(−1.26)	(−1.56)	(−0.79)
2	−0.001	0.047	0.06	0.05	−0.002	0.044	0.064	0.052
	(−0.03)	(0.89)	(0.92)	(1.12)	(−0.04)	(0.85)	(0.98)	(1.16)
(2–1)	0.061	0.118**	0.153***	**0.092**	0.065	0.111**	0.155***	**0.09**
	(1.30)	(2.54)	(2.65)	**(1.57)**	(1.36)	(2.48)	(2.67)	**(1.54)**
C4 α	0.075	0.121**	0.165***	**0.09**	0.08*	0.114**	0.168***	**0.088**
	(1.61)	(2.58)	(2.76)	**(1.51)**	(1.75)	(2.55)	(2.82)	**(1.47)**
FF5 α	0.066	0.108**	0.153***	**0.087**	0.072	0.103**	0.157***	**0.085**
	(1.47)	(2.40)	(2.59)	**(1.45)**	(1.61)	(2.36)	(2.66)	**(1.41)**
SY4 α	0.058	0.123**	0.204***	**0.146***	0.068	0.122**	0.201***	**0.134***
	(1.02)	(2.20)	(2.67)	**(1.83)**	(1.21)	(2.27)	(2.63)	**(1.65)**

V.C. Are Results Driven by Smart MM's Positions?

To bring further support to our view that the measures of position changes of MM capture information updates pertinent to the revisions of the future prospects of commodities beyond simple commodity strategies such as commodity momentum and basis, we verify in this section that our predictability results on commodity producers' equity returns remain after controlling for a number of additional commodity price factors. In a second test, we decompose MM signal measures into a momentum-driven component and a component orthogonal to past commodity futures returns. In brief, results on both exercises favor our interpretation that MM position changes capture valuable information.

V.C.1. Inclusion of Additional Commodity Factors

Following Gorton and Rouwenhorst (2006), Gorton et al. (2013), Bhardwaj et al. (2014), Christoffersen et al. (2019), and Boons and Prado (2019), among others, we construct a number of commodity price factors, namely: i) the past 12-month futures momentum, ii) the futures basis (which captures inventory effects, that is, "backwardation," in commodities markets), iii) a benchmark commodity market index, and iv) the futures basis-momentum, which is the difference between momentum in first- and second-nearby futures contracts (the aforementioned factors are largely capable of capturing the cross-sectional variation of commodity futures returns); v) alternatively, we also capture the cross-sectional variation by conducting principal component analysis (PCA).

The construction of these commodity factors and their variants are detailed in Supplementary Material Section A.4. In particular, we construct the momentum, basis, and basis-momentum factors by calculating the difference between the long and short portfolios' returns consisting of the equity of commodity producers that are ranked based on the corresponding commodity futures signal (i.e., in the *equity space*). Alternatively, these commodity factors are constructed by calculating the long-short returns based on portfolios consisting of commodity futures (i.e., in the *futures space*). For the sake of completeness, we also identity principal components in the cross-section of futures returns, following Christoffersen et al. (2019).[30] We present the augmented single-sort results in Table 7.

We find that the abnormal returns from the Carhart four-factor (C4 α) model remain consistent once the additional four commodity factors (equity- or futures-based) are added all together to the model, or alternatively when the identified principal components of commodity futures returns are added. For example, taking the case of augmenting the Carhart four-factor model with equity-based commodity factors (i.e., eight factors in total) under the equal-weight scheme, we see on this line that the α's are slightly smaller relative to the ones previously reported, ranging from 10.14% (0.195% $\times$ 52) to 13.1% (0.252% $\times$ 52) per annum, and are generally statistically significant with t-statistics ranging from 2.53 to 3.28. Hence, if the bar or threshold is held at the level that equity investors need to outperform simple commodity strategies (in addition to the Carhart factors), we still find that MM position change signals can deliver abnormal returns beyond this bar.

[30]We identify five orthogonal principal components that explain 34.5%, 8.9%, 8.2%, 6.9%, and 5%, respectively, for a total of 63.5% of the cross-sectional variation for 21 commodity futures returns.

TABLE 7 Single-Sort Results with Additional Commodity Price Factors

After equal-weighting (Panel A) or value-weighting (Panel B) the producers' stock returns belonging to the same commodity, the commodity-equity portfolios are sorted weekly into three portfolio bins based on the signal's value and averaged within each tercile with equal-weight, following Section III.C. The MM Net Change (NET), Long Proportion Growth (LPG), or Short Proportion Growth (SPG) signals are constructed as a 1-week lag or as a 2-week backward moving average. From the return difference of the highest tercile minus the lowest tercile, we then calculate the α's relative to the Carhart four-factor model (C4 α) as well as to the augmentation of the C4 model with additional commodity price factors (basis, momentum, index, and basis-momentum, as constructed in Supplementary Material Section A.4), or to the inclusion of five commodity futures returns' principal components (PCA). The table presents the weekly α's (which are multiplied by 100) with their t-statistics in parentheses based on White standard errors. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

	J=1			J=2		
	NET	LPG	SPG	NET	LPG	SPG
Panel A: Equal-Weight						
C4 α	0.276***	0.248***	−0.245***	0.232***	0.285***	−0.232***
	(3.58)	(3.27)	(−3.09)	(2.96)	(3.34)	(−2.97)
C4 + Commodity Factors α	0.252***	0.228***	−0.213***	0.204***	0.234***	−0.195**
(equity-based)	(3.28)	(2.96)	(−2.73)	(2.62)	(2.90)	(−2.53)
C4 + Commodity Factors α	0.277***	0.247***	−0.248***	0.208***	0.269***	−0.217***
(futures-based)	(3.55)	(3.21)	(−3.13)	(2.61)	(3.02)	(−2.67)
C4 + PCA Factors α	0.272***	0.252***	−0.242***	0.231***	0.290***	−0.230***
(futures-based)	(3.53)	(3.31)	(−3.06)	(2.96)	(3.40)	(−2.94)
Panel B: Value-Weight						
C4 α	0.281***	0.276***	−0.264***	0.193***	0.245***	−0.193***
	(3.72)	(3.72)	(−3.53)	(2.66)	(3.16)	(−2.69)
C4 + Commodity Factors α	0.278***	0.274***	−0.250***	0.176**	0.217***	−0.169**
(equity-based)	(3.57)	(3.61)	(−3.30)	(2.40)	(2.87)	(−2.32)
C4 + Commodity Factors α	0.274***	0.263***	−0.262***	0.168**	0.224***	−0.176**
(futures-based)	(3.65)	(3.57)	(−3.54)	(2.29)	(2.80)	(−2.38)
C4 + PCA Factors α	0.280***	0.281***	−0.262***	0.195***	0.248***	−0.194***
(futures-based)	(3.68)	(3.77)	(−3.50)	(2.69)	(3.19)	(−2.68)

V.C.2. Non-Momentum- versus Momentum-Driven MM Signals

Although we have already shown that our results survive the inclusion of commodity factors including the momentum factor, we wish to ascertain whether our results, specifically the MM position changes signal, are mainly driven by the traders following the momentum signal within the MM category. We conjectured in Section III.A that much of the short-term position changes of MM are mainly contributed by more sophisticated funds that emphasize active management and that see fit to commit or change their positions based on their response to informational updates regarding the future prospect of commodity fundamentals, given their specialization. While, unfortunately, we do not have access to confidential individual traders' positions,[31] we can still infer from published DCOT data a component of our MM aggregated position change signals that can be predicted by past commodity futures performance.

Specifically, we decompose our MM signals, constructed either as a 1-week lag or as a J-week backward moving average, into two components by running an expanding-window recursive estimation using only the data available at the time of compilation of MM positions to avoid any look-ahead bias. With the MM Net Change signal measure (NET), for instance, we run separately for each commodity c and week t, the following first-stage regression:

$$NET_{c,t} = \mu_{c,t} + \beta_{c,t} FR_{c,t-s;t-1} + \epsilon_{c,t}, \tag{2}$$

where the futures return momentum $FR_{c,t-1;t-s}$ is computed on multiple look-back horizons from the previous $t-s$ ($s = 1$, 8, 12, 26, or 52 weeks) up to week $t-1$ in order to account for the variety of trend-following approaches that can be employed by managed futures in commodity markets. The model's fitted values $\widehat{NET}_{c,t}^{momnt.} = \hat{\beta}_{c,t} FR_{c,t-s;t-1}$ are then defined as the momentum-driven component of our MM signal, while $\widehat{NET}_{c,t}^{non\text{-}momnt.} = \hat{\mu}_{c,t} + \hat{\epsilon}_{c,t}$ isolates the component that is orthogonal to commodity futures' past performance signals. Finally, we repeat the single-sort analysis on commodity producers' stocks based on these two distinct estimated MM signals. Table 8 presents the Carhart (1997) alpha.

As reported in the right-hand column, the first-stage R^2 estimates[32] from these decomposition regressions are generally low, on average around 3.70% for the $J = 2$ MM signal (ranging from 1.20% to 11.47% across the different momentums' look-back periods), while as expected, values for the R^2 on MM signal with medium-term look-back horizon ($J = 12$ weeks) are generally higher, on average 12.38% and as high as 21.88%. Hence, the lion's share

[31] The CFTC does not disclose futures positions data at any level finer than the categories in the DCOT reports as the positions of individual reportable traders are protected information.

[32] Summary statistics on the first-stage R^2 are obtained by first taking the time series average R^2 for each commodity, then collapsing these values for all commodities. We report the range of average R^2 across the five specifications with different momentum horizons s and the overall average.

TABLE 8 Single-Sort Results on Non-Momentum- versus Momentum-Driven MM Signals

In the first stage, we start the decomposition of the MM Net Change (NET) signal into two components by running an expanding window recursive estimation separately for each commodity c and week t: $NET_{c,t} = \mu_{c,t} + \beta_{c,t} FR_{c,t-s;t-1} + \epsilon_{c,t}$, where $NET_{c,t}$ is constructed either as a 1-week lag or as a J-week backward moving average, and the futures return momentum $FR_{c,t-1;t-s}$ is computed over the previous s weeks, with $s = 1, 8, 12, 26,$ or 52. The signal is decomposed into i) a non–momentum-driven signal $\widehat{NET}^{non\text{-}momnt.}_{c,t} = \hat{\mu}_{c,t} + \hat{\epsilon}_{c,t}$ and ii) a momentum-driven signal $\widehat{NET}^{momnt.}_{c,t} = \hat{\beta}_{c,t} FR_{c,t-s;t-1}$. In the second stage, we run the single-sort procedure following Section III.C wherein after the stock returns belonging to the same commodity are either equal-weighted (EW) or value-weighted (VW), the commodity-equity portfolios are sorted weekly into three portfolio bins, based on $\widehat{NET}^{non\text{-}momnt.}_{c,t}$ in Panel A and $\widehat{NET}^{momnt.}_{c,t}$ in Panel B, and averaged within each tercile with equal-weight. From the return difference of the highest tercile minus the lowest tercile, we calculate the C4 α. The table presents the weekly alphas (which are multiplied by 100) with their t-statistics in parentheses based on White standard errors. The last column shows summary statistics on the first-stage R^2 that are obtained by first taking the time series average R^2 for each commodity, then collapsing these values for all commodities. We report the range (in brackets) of average R^2 across the five specifications with different momentum horizons s and the overall average ($O.Avg$).$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Panel A: Non-Momentum-Driven MM Sorting Signal: $\widehat{NET}^{non\text{-}momnt.}_{c,t} = \hat{\mu}_{c,t} + \hat{\epsilon}_{c,t}$

Second stage C4 α		where horizon s for momentum in the first stage is					First stage R^2 $O.Avg$ [range]
		s=1	s=8	s=12	s=26	s=52	
J=1	EW	0.190**	0.259***	0.234***	0.231***	0.213***	1.44%
		(2.46)	(3.37)	(2.99)	(2.96)	(2.68)	[0.96–2.68]
	VW	0.223***	0.273***	0.232***	0.245***	0.224***	
		(2.94)	(3.60)	(3.08)	(3.18)	(2.88)	
J=2	EW	0.239***	0.262***	0.242***	0.234***	0.232***	3.70%
		(2.88)	(2.99)	(2.98)	(2.73)	(2.77)	[1.20–11.47]
	VW	0.216***	0.191**	0.165**	0.179**	0.175**	
		(2.89)	(2.49)	(2.18)	(2.34)	(2.26)	
J=12	EW	0.253***	0.240***	0.215**	0.283***	0.217**	12.38%
		(3.09)	(2.83)	(2.55)	(3.34)	(2.47)	[4.47–21.88]
	VW	0.221***	0.224***	0.193***	0.236***	0.180**	
		(3.08)	(3.12)	(2.66)	(3.08)	(2.27)	

Panel B: Momentum-Driven MM Sorting Signal: $\widehat{NET}^{momnt.}_{c,t} = \hat{\beta}_{c,t} FR_{c,t-s;t-1}$

Second stage C4 α		where horizon s for momentum in the first stage is					First stage R^2 $O.Avg$ [range]
		s=1	s=8	s=12	s=26	s=52	
J=1	EW	0.063	−0.030	0.068	−0.054	−0.020	1.44%
		(0.79)	(−0.36)	(0.78)	(−0.63)	(−0.21)	[0.96–2.68]
	VW	−0.018	−0.023	0.031	−0.049	−0.051	
		(−0.24)	(−0.31)	(0.40)	(−0.64)	(−0.59)	
J=2	EW	0.003	0.023	0.014	−0.072	−0.097	3.70%
		(0.04)	(0.29)	(0.19)	(−0.82)	(−1.13)	[1.20–11.47]
	VW	−0.030	0.013	−0.008	−0.051	−0.133*	
		(−0.42)	(0.18)	(−0.11)	(−0.68)	(−1.74)	
J=12	EW	0.022	0.068	0.033	−0.026	0.078	12.38%
		(0.28)	(0.81)	(0.41)	(−0.29)	(0.87)	[4.47–21.88]
	VW	0.059	0.054	0.059	−0.010	0.087	
		(0.80)	(0.69)	(0.79)	(−0.13)	(1.05)	

of the variation in our MM signals, especially for short-term position changes, is orthogonal to past price momentum, and as a side note, we find that the dispersion of the non–momentum component is much larger than the momentum component, consistent with our expectation that MM position changes are mainly contributed by the more active funds. For example, taking the $J = 2$ MM signal, the standard deviation of the non–momentum component is on average seven times larger than the momentum-related component's standard deviation.

Considering that position changes may embed a price momentum component, we do sometimes observe a decrease in the α's magnitude reported in Table 8. Focusing on the MM signal with $J = 12$, from a previously reported annual α of 11.75% (0.226% $\times$ 52), abnormal value-weighted returns reduce to 9.36% (0.180% $\times$ 52) after controlling for the 52-week commodity momentum. That said, results based on non–momentum-driven signals are still large and strongly statistically significant. By contrast, the estimated α's based on the momentum-related signals are generally insignificantly different from zero. Overall, these results suggest that our signal measures are not merely picking up the straightforward trading of momentum signals within MM but rather, the skills of traders within MM who see fit to adjust positions beyond a simple trend. This further corroborates our gradual information diffusion interpretation, whereby MM position signals capture informative reflections of "smart money" that are conducive to producers' equity return predictability.

On a different but related point, one might be concerned that the abnormal returns we found in the calendar-time regression analysis and single-sort analysis (Tables 2 and 3) are mainly due to a particular subpart of the sample period. For instance, in the hypothetical case in which MM position changes would primarily capture the positions of trend followers, periods of high volatility in commodity prices could lead to autocorrelation in the residuals. Reassuringly, as shown in Supplementary Material Tables B.6 and B.7, none of the Durbin-Watson statistics indicates statistically significant evidence for autocorrelation (neither positive nor negative) in the residuals of the aforementioned regressions.[33]

[33] As shown in Supplementary Material Table B.8, we also obtain similar results for the MM Long Proportion Growth signal when applying the Cochrane and Orcutt (1949) transformation or using

VI. Conclusion

We explore in a novel empirical setting the notion that limited information processing capacity and the ensuing specialization of market participants would induce value-relevant information to diffuse gradually across segmented asset markets. By utilizing a sample of commodity producers' equities that are matched with the corresponding long and short positions that money managers (as defined in the MM category) took in the commodity futures market as disclosed in the CFTC's weekly DCOT reports, we investigate whether the money managers, who are sophisticated and specialized investors in the commodity futures market, can be deemed ‘‘smart money’’ with a superior information advantage on commodity fundamentals, and whether this commodity-relevant information would be gradually impounded into commodity producers' equity price.

We find strong evidence that the information extracted from MM's commodity futures position changes can predict the cross-section of commodity producers' stock returns during the subsequent week. Consistent with our main thesis that the predictability arises from costly information processing, our double-sort results reveal that the equity returns of commodity producers consisting of firms that are relatively nontransparent (i.e., with high information asymmetry) are indeed the slowest to price-adjust. Specifically, our predictability results are more pronounced for firms with higher *ex ante* analyst dispersion and higher *historical* stock volatility, but not so with regard to Amihud's illiquidity measure, suggesting that our findings arise from informational, rather than trading, frictions.

This lead-lag relationship is consistently confirmed through a number of empirical methods and across a range of signal measures, time lags, and weighting schemes. The relationship generates economically large and statistically significant abnormal returns, ranging from 10% to 13% a year, with respect to various factor models, such as the Carhart four-factor model, the Fama and French (2015) five-factor model, or the Stambaugh and Yuan (2016) model,

Newey-West standard errors up to five lags. We omit presentation of the other two signal measures for brevity.

and to the inclusion of additional commodity price factors capturing common commodity futures strategies such as futures momentum, basis, index, and the newly discovered basis-momentum, or the principal components of commodity returns. Furthermore, we find that on average MM position change signals capture relevant information beyond the information already contained in past commodity futures returns (whether past trend or 1-week-lagged futures return), and hence are not merely reflecting the positions of traders within MM who simply follow the trend signal. We also show that a mechanical link between the equity returns of commodity-producing firms and the contemporaneous futures returns is not driving our results, nor do our findings arise from the announcement effects of the DCOT reports, nor are results confined to small-cap stocks.

We thus present in a novel setting more empirical evidence supporting the research on complexity and return delay (Cohen and Lou, 2012; Cohen et al., 2020) that finds significant return predictability can arise as a result of investors' limited information processing capacity. In our case, as the MM are sophisticated and specialized speculators in the futures market, they would by and large react to informational updates regarding commodities relatively fast before the information is impounded into securities that are more difficult to analyze, that is, common stocks issued by commodity producers—especially the stocks of the nontransparent producers. The results are consistent with the literature that finds costly information processing can lead to investor specialization in terms of information acquisition, market segmentation, and gradual information diffusion across asset markets.

References

Acharya, Viral V, Lochstoer, Lars A and Ramadorai, Tarun. (2013). 'Limits to arbitrage and hedging: Evidence from commodity markets', *Journal of Financial Economics* 109(2), 441–465.

Amihud, Yakov. (2002). 'Illiquidity and stock returns: cross-section and time-series effects', *Journal of Financial Markets* 5(1), 31–56.

Bessembinder, Hendrik, Chan, Kalok and Seguin, Paul J. (1996). 'An empirical examination of information, differences of opinion, and trading activity', *Journal of Financial Economics* 40(1), 105–134.

Bessembinder, Hendrik and Seguin, Paul J. (1992). 'Futures-trading activity and stock price volatility', *Journal of Finance* 47(5), 2015–2034.

Bhardwaj, Geetesh, Gorton, Gary B and Rouwenhorst, K Geert. (2014). 'Fooling some of the people all of the time: The inefficient performance and persistence of commodity trading advisors', *Review of Financial Studies* 27(11), 3099–3132.

Boehmer, Ekkehart, Jones, Charles M, Wu, Juan and Zhang, Xiaoyan. (2020). 'What do short sellers know?', *Review of Finance* 24(6), 1203–1235.

Bohmann, Marc JM and Patel, Vinay. (2020). 'Information leakage in energy derivatives around news announcements', *Journal of Derivatives* 27(4), 13–29.

Boons, Martijn and Prado, Melissa Porras. (2019). 'Basis-momentum', *Journal of Finance* 74(1), 239–279.

Buchanan, William K, Hodges, Paul and Theis, John. (2001). 'Which way the natural gas price: an attempt to predict the direction of natural gas spot price movements using trader positions', *Energy Economics* 23(3), 279–293.

Carhart, Mark M. (1997). 'On persistence in mutual fund performance', *Journal of Finance* 52(1), 57–82.

Carter, Colin A, Rausser, Gordon C and Schmitz, Andrew. (1983). 'Efficient asset portfolios and the theory of normal backwardation', *Journal of Political Economy* 91(2), 319–331.

Chen, Huaizhi, Cohen, Lauren and Gurun, Umit G. (2021). 'Don't Take Their Word for It: The Misclassification of Bond Mutual Funds', *Journal of Finance* 76(4), 1699–1730.

Cheng, Ing-Haw, Kirilenko, Andrei and Xiong, Wei. (2015). 'Convective risk flows in commodity futures markets', *Review of Finance* 19(5), 1733–1781.

Christoffersen, Peter, Lunde, Asger and Olesen, Kasper V. (2019). 'Factor structure in commodity futures return and volatility', *Journal of Financial and Quantitative Analysis* 54(3), 1083–1115.

Cochrane, Donald and Orcutt, Guy H. (1949). 'Application of least squares regression to relationships containing auto-correlated error terms', *Journal of the American Statistical Association* 44(245), 32–61.

Cohen, Lauren and Frazzini, Andrea. (2008). ‘Economic links and predictable returns’, *Journal of Finance* 63(4), 1977–2011.

Cohen, Lauren and Lou, Dong. (2012). ‘Complicated firms’, *Journal of Financial Economics* 104(2), 383–400.

Cohen, Lauren, Malloy, Christopher and Nguyen, Quoc. (2020). ‘Lazy prices’, *Journal of Finance* 75(3), 1371–1415.

De Roon, Frans A, Nijman, Theo E and Veld, Chris. (2000). ‘Hedging pressure effects in futures markets’, *Journal of Finance* 55(3), 1437–1456.

Du, Xiaodong and Kane, Stephen A. (2019), Fundamental surprises, market structure, and price formation in agricultural commodity futures markets, Office of the Chief Economist, CFTC No 2020-004.

Ederington, Louis and Lee, Jae Ha. (2002). ‘Who trades futures and how: Evidence from the heating oil futures market’, *Journal of Business* 75(2), 353–373.

Engelberg, Joseph E, Reed, Adam V and Ringgenberg, Matthew C. (2012). ‘How are shorts informed?: Short sellers, news, and information processing’, *Journal of Financial Economics* 105(2), 260–278.

Fama, Eugene F and French, Kenneth R. (2008). ‘Dissecting anomalies’, *Journal of Finance* 63(4), 1653–1678.

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

Fernandez-Perez, Adrian, Fuertes, Ana-Maria and Miffre, Joelle. (2017). ‘Commodity markets, long-run predictability, and intertemporal pricing’, *Review of Finance* 21(3), 1159–1188.

Gandhi, Priyank and Lustig, Hanno N. (2015). ‘Size Anomalies in U.S. Bank Stock Returns’, *Journal of Finance* 70(2), 733–768.

Gorton, Gary B, Hayashi, Fumio and Rouwenhorst, K Geert. (2013). ‘The fundamentals of commodity futures returns’, *Review of Finance* 17(1), 35–105.

Gorton, Gary and Rouwenhorst, K Geert. (2006). ‘Facts and fantasies about commodity futures’, *Financial Analysts Journal* 62(2), 47–68.

Henderson, Brian J, Pearson, Neil D and Wang, Li. (2014). ‘New evidence on the financialization of commodity markets’, *Review of Financial Studies* 28(5), 1285–1311.

Hong, Harrison and Yogo, Motohiro. (2012). ‘What does futures market interest tell us about the macroeconomy and asset prices?’, *Journal of Financial Economics* 105(3), 473–490.

Hou, Kewei, Xue, Chen and Zhang, Lu. (2020). ‘Replicating anomalies’, *Review of Financial Studies* 33(5), 2019–2133.

Huang, Shiyang, O’Hara, Maureen and Zhong, Zhuo. (2021). ‘Innovation and informed trading: Evidence from industry ETFs’, *Review of Financial Studies* 34(3), 1280–1316.

Krohn, Ingomar. (2018), Essays in International Finance, PhD thesis, University of Warwick.

Menzly, Lior and Ozbas, Oguzhan. (2010). 'Market segmentation and cross-predictability of returns', *Journal of Finance* 65(4), 1555–1580.

Michaelides, Alexander, Milidonis, Andreas, Nishiotis, George P and Papakyriakou, Panayiotis. (2015). 'The adverse effects of systematic leakage ahead of official sovereign debt rating announcements', *Journal of Financial Economics* 116(3), 526–547.

Mixon, Scott and Onur, Esen. (2020), Risk appetite and intermediation by swap dealers, CFTC OCE Staff Papers and Reports No 2020-004.

Mixon, Scott, Onur, Esen and Riggs, Lynn. (2018). 'Integrating swaps and futures: A new direction for commodity research', *Journal of Commodity Markets* 10, 3–21.

Sadka, Ronnie and Scherbina, Anna. (2007). 'Analyst disagreement, mispricing, and liquidity', *Journal of Finance* 62(5), 2367–2403.

Sanders, Dwight R, Boris, Keith and Manfredo, Mark. (2004). 'Hedgers, funds, and small speculators in the energy futures markets: an analysis of the CFTC's Commitments of Traders reports', *Energy Economics* 26(3), 425–445.

Sanders, Dwight R, Irwin, Scott H and Merrin, Robert P. (2009). 'Smart money: The forecasting ability of CFTC large traders in agricultural futures markets', *Journal of Agricultural and Resource Economics* p. 276–296.

Stambaugh, Robert F, Yu, Jianfeng and Yuan, Yu. (2012). 'The short of it: Investor sentiment and anomalies', *Journal of Financial Economics* 104(2), 288–302.

Stambaugh, Robert F and Yuan, Yu. (2016). 'Mispricing factors', *Review of Financial Studies* 30(4), 1270–1315.

Tokic, Damir. (2010). 'The 2008 oil bubble: Causes and consequences', *Energy Policy* 38(10), 6009–6015.

Tornell, Aaron and Yuan, Chunming. (2012). 'Speculation and hedging in the currency futures markets: Are they informative to the spot exchange rates', *Journal of Futures Markets* 32(2), 122–151.

Tufano, Peter. (1998). 'The determinants of stock price exposure: Financial engineering and the gold mining industry', *Journal of Finance* 53(3), 1015–1052.

Van Nieuwerburgh, Stijn and Veldkamp, Laura. (2010). 'Information acquisition and under-diversification', *Review of Economic Studies* 77(2), 779–805.

White, Halbert. (1980). 'A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity', *Econometrica* p. 817–838.

Appendix: Procedure to Compute the Long-Short Portfolio Returns

This appendix supplements the description of the procedure to compute the long-short portfolio returns in Section III.C. First, we compound the daily returns of each stock r_{it} to obtain the weekly returns r_{iw}, where the week w runs from the beginning of day of Wednesday ($t = 1$) until the end-of-day of next Tuesday ($t = T$): $r_{iw} = \prod_{t=1}^{T} (1 + r_{it}) - 1$. Then we compute the weekly stock returns for each of the ten commodity-equity portfolios:

$$R_w^C = \frac{1}{\sum_{i \in C} W_i^V} \sum_{i \in C} W_i^V r_{iw}^C,$$

where r_{iw}^C is the stock return at week w of producer i belonging to commodity C, and

$$W_i^V = \begin{cases} marketcap_{i,year-1} & \text{if Value-Weight is applied.} \\ 1 & \text{if Equal-Weight is applied.} \end{cases}$$

For a signal with a J-week look-back horizon, $J \geq 1$, the weekly signal s of the futures market commodity C are aggregated over the look-back horizon as $S_{w,J}^C = \frac{1}{J} \sum_{k=1}^{J} s_{w-k}^C$.

In the case of the calendar-time regression analysis, if the signal $S_{w,J}^C$ is positive, then the commodity-equity portfolio C belongs to the long portfolio (L) in week w. If the signal is negative, then it belongs to the short portfolio (S). We compute the long (R_w^L), short (R_w^S), and long-short (R_w^{LS}) portfolio returns at week w, where we take into account the strength of the signal with $W_C^D = |S_{w,J}^C|$, as follows:

$$R_w^L = \frac{1}{\sum_{C \in L} W_C^D} \sum_{C \in L} W_C^D R_w^C, \qquad R_w^S = \frac{1}{\sum_{C \in S} W_C^D} \sum_{C \in S} W_C^D R_w^C, \qquad R_w^{LS} = R_w^L - R_w^S.$$

In the case of the single-sort analysis, each week w, all commodity-equity portfolios C are sorted into one of the three bins (k) based on the signal's value ($S_{w,J}^C$), with bin 1 representing the lowest tercile and bin 3 representing the highest tercile. The weekly returns for each portfolio bin k and the long-short portfolio returns (R_w^{3-1}) are computed as follows:

$$R_w^k = \frac{1}{\sum_{C \in k} W_C^D} \sum_{C \in k} W_C^D R_w^C, \qquad R_w^{3-1} = R_w^3 - R_w^1, \qquad \text{where} \quad W_C^D = 1.$$