Journal of Financial Markets 64 (2023) 100774

Contents lists available at ScienceDirect

Journal of Financial Markets

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

Tracking speculative trading☆,☆☆

Dominik Boos [a], Linus Grob [a,b,*]

[a] ZHAW School of Management and Law, Institute of Wealth & Asset Management, Switzerland
[b] Smurfit Graduate Business School, University College Dublin, Ireland

ARTICLE INFO

JEL classification:
C14
C52
G11
Q02

Keywords:
Trend-following
Momentum
Commodity futures market
Commitments of traders
Hedge funds
Ridge regression

ABSTRACT

Managed futures funds are predominantly trend-followers. By analyzing positioning data, we provide novel evidence for this claim and estimate signals applied by these funds. We write trend-followers aggregate position as a weighted sum of past daily returns and use a generalized ridge regression for regularization and parameter estimation. This procedure prevents overfitting but remains flexible enough to capture various patterns. For the 23 commodities considered, trend-following can explain speculators' position changes with an average R^2 of more than 40%. Finally, we document that producers act as contrarians in a way that closely mirrors the behavior of momentum traders.

1. Introduction

Trend-following is the most common investment strategy in the futures markets. It is the investment style of systematic global macro funds and also the predominant input factor of many commodity funds. The popularity of trend and momentum signals is due to their consistent and well-documented long-term performance and their potential for diversification (Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006; Miffre and Rallis, 2007).

While many systematic macro funds are classified as trend-followers, details on the trading strategy and the explicit construction of the signals are kept as proprietary knowledge. Funds, for example, do not reveal the interval length used for trend estimation (which we refer to as the look-back period). Also, the way the trend signal is constructed is not made public. It is not known whether a fund uses basic momentum, moving-average crossover, or another more exotic trend signal. Finally, portfolio construction can be very different: A time series approach analyzes an asset just on its own price history (Szakmary et al., 2010; Moskowitz et al., 2012). A positive trend results in a long position, while a negative trend leads to a short position, and there is no interdependence between assets. Cross-sectional approaches focus on relative performance (Jegadeesh and Titman, 1993; Erb and Harvey, 2006). The assets with the strongest trend form a long basket and those with the weakest trend are shorted.

In this paper, we provide new evidence that commodity speculators[1] do indeed predominantly use trend signals. This is not a novelty by itself as previous studies document the strong link between trend-following and hedge funds. For example Hurst et al.

☆ We thank Jan-Alexander Posth, Valerio Poti and Peter Schwendner, the co-editor Paolo Pasquariello, and two referees for their helpful comments and suggestions, as well as seminar participants at Zurich University of Applied Sciences.

☆☆ Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

* Correspondence to: Technoparkstrasse 2, 8406 Winterthur, Switzerland.
E-mail addresses: dominik.boos@zhaw.ch (D. Boos), linus.grob@zhaw.ch (L. Grob).

[1] We use the terms (hedge) funds, non-commercial traders, and speculators interchangeably.

https://doi.org/10.1016/j.finmar.2022.100774
Received 25 October 2021; Received in revised form 25 July 2022; Accepted 26 July 2022
Available online 30 July 2022

(2013) provide a detailed analysis of managed futures and show that their returns can be explained by simple trend-following strategies. Our analysis differs, however, substantially from theirs. Firstly, we do not conjecture that certain signals are relevant for speculative trading; rather, we estimate them. More precisely, we estimate a representative trading signal used by the average hedge fund and reveal the most popular look-back periods.

Secondly, our analysis does not explain returns but rather the exposure held by hedge funds and their trading activities. Similar to Lehecka (2013), we analyze to what extent the positioning of non-commercial traders published in the Commitment of Traders Report (CoT) released by the Commodity Futures Trading Commission (CFTC) can be explained with past returns. More precisely, we show that trend and momentum signals explain more than 40% of the commodity position changes of large hedge funds. To the best of our knowledge, this is the first paper to reveal such detailed evidence about the link between the positioning of speculative traders and trend-following.

We also study other signals known to explain commodity return, specifically the basis and basis-momentum (e.g., Szymanowsky et al., 2014; Bakshi et al., 2019; Boons and Prado, 2019). However, somewhat surprisingly given their relevance in the literature, the inclusion of these variables in our model did not significantly increase its forecasting ability, in neither a time series nor in a cross-sectional specification.

Our findings on momentum are remarkably consistent across the 23 commodity futures we examine. Over the entire sample (January 1998 till the end of December 2019), every single market is predictable and only three subperiods display a negative R^2. Natural gas, the most volatile market, is the only commodity with an unstable pattern. Moreover, we find a consistent hump-shaped pattern in the signal's weighting of returns, which indicates that there are short-term and long-term traders in the market and that implementation is delayed in some way. Results are robust out-of-sample for different tuning parameters and for different maximum look-back periods.

Because there are so many different trading signals and parameters for these signals, we need an elaborate empirical design to reveal aggregate behavior. While our approach is primarily data-driven, we make the constraining assumption that the cross-sectional approach is largely irrelevant to explain the aggregate behavior. This restriction allows us to focus on the time series approach and therefore to analyze each market individually. Next, we use the representations derived in Levine and Pedersen (2016). They show that a large class of trend signals, including momentum and moving average crossing, can equivalently be represented as a weighted sum of past returns. Assuming a maximum look-back period of one year, this transformation reduces the number of trading signals to 260. We further reduce the degrees of freedom through a generalized Tikhonov regularization, also known as ridge regression, as introduced by Phillips (1962) and Tikhonov (1963). Instead of shrinking outright parameters as in the standard version, we penalize weight changes between consecutive lags of past returns. This design imposes smoothness on the weighting function of past returns. Given typical trading signals, this is a well-justified assumption. Moreover, the aim of traders to reduce turnover will inevitably lead to smoothly varying weights.

In this paper, we focus on commodities because the majority of speculative trading in these markets takes place with futures. Futures positions are therefore a very accurate proxy for the overall speculative exposure. This is not the case for the other large asset classes: Foreign exchange speculation is dominated by OTC forwards and many duration bets are implemented through swaps. While futures are the preferred instrument of macro traders, with equities there are different alternatives to implementing a trade efficiently. Another reason for the restriction on commodities is that the long history of positioning data is almost exclusively available for U.S. instruments. As the commodity market is dominated by U.S. futures, we can obtain an almost complete picture in this sector. Equity and fixed income trading is geographically more diverse and any evidence will only be fragmentary solely due to the lack of appropriate positioning data.

The focus on commodities also allows us to analyze commercial traders, who inevitably mirror the trading behavior of speculators. We investigate whether producers, users, or swap dealers take the opposite side of the speculators' trade and thus act as contrarians. In line with Cheng and Xiong (2014), we find that producers react strongly to price changes. We contribute to this strand of literature, by generalizing their results to a much broader set of commodities. In addition, we quantify the behavior of producers and show that they act in a way that closely mirrors the behavior of momentum traders.

The remainder of this paper is organized as follows: In Section 2, we describe the transformation of the trend-following strategies into the return space, the fitting algorithm, and the estimation approach. In Section 3, we document empirically that speculators are indeed trend-followers by predicting their position changes with trend signals, prove the models robustness, and point out the evolution of the empirical return signature plots. Section 4 contains the analysis of basis and basis-momentum strategies. Section 5 provides an analysis of other trading groups and reveals their trading pattern. Finally, in Section 6 we conclude the paper.

2. Methodology

2.1. Trend-following signals

Momentum, or the return over the past n days, is the most elementary trend-following signal. Using log returns, we can write the momentum $\text{mom}_t(n)$ as the sum of daily returns:

$$\text{mom}_t(n) = r_{t-n,t} = \sum_{i=0}^{n-1} r_{t-i}. \tag{1}$$

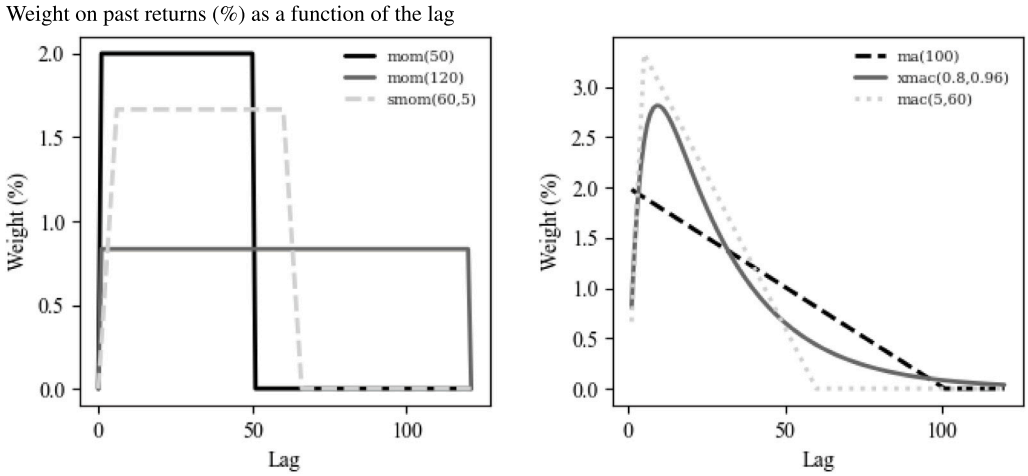

Fig. 1. Return signature plots of popular strategies.
Note: Weights on past daily returns for different momentum indicators (left) and for moving average based trend signals (right). All weights are normalized and sum up to 1.

Following Levine and Pedersen (2016), this indicator can be generalized by allowing returns with different lags to have different weights:

$$\text{trend}_t(\boldsymbol{b}, n) = \sum_{i=0}^{n-1} b_i r_{t-i}. \tag{2}$$

A straightforward example is the combination of momentum signals of different lengths, but this representation is far more general. In fact, all linear trend signals can be written in Eq. (2), as outlined by Levine and Pedersen (2016). They also derive coefficients b_i, which they call "return signature plots", for many popular strategies. Fig. 1 illustrates the most common examples.

The two solid lines in the left panel of Fig. 1 represent simple momentum signals. They give equal weight to returns on each day of the lookback period of 50 or 120 days, respectively. The dotted line in the same panel presents a stylized example of a momentum strategy with a partial rebalancing rule ($smom(60, 5)$). The implementation of the 60-day momentum signal is split over five days, which leads to a trapezoidal shape of the return signature.

Trend signals based on moving averages are shown on the right-hand side of Fig. 1. The simple moving average ($ma(100)$) is a linear function that decreases from an initial value to zero at the end of the look-back period. The moving average crossing ($mac(5, 60)$) linearly increases until the end of the shorter look-back period of five days and then linearly declines to zero at the end of the longer moving average (60 days). The exponentially weighted moving average crossing ($xmac$) has a hump-shaped form that converges to zero. Similar to the unweighted moving average crossing, it assigns the most importance to intermediate price changes and lesser weight to the most recent price changes and to very old returns.

2.2. Aggregation and portfolio construction

For the portfolio construction, we make two presumptions: Firstly, in line with Hurst et al. (2013), we assume that time series momentum is prevalent. This conjecture is underpinned by Schmid and Wirth (2021), who examine the optimal combination of cross-sectional and time series momentum in a pure trend-following strategy and find that the average optimal allocation to time series momentum is around 80% compared to a mere 20% in relative trend. Assuming a pure time series approach on one hand simplifies the model, by looking at each commodity on its own, and further allows us to capture the differences of the participants and dynamics for each market (Bosch and Smimou, 2022).

Secondly, we suppose that the average trend-follower increases the exposure proportionally to the signal.[2] Then, the aggregate position can be written in the same linear form as the signal (2). Moreover, the linear relationship between exposure and past daily returns preserves its form when aggregated across traders. Consequently, under these assumptions we can write the aggregate

[2] This is not true for individual traders. For example, going long a fixed exposure when the trend is up and short the same exposure when the trend is down is a widespread approach, which explains the returns of managed futures fairly well (Hurst et al., 2013). Also, transforming the signal with an S-shaped (or sigmoid) function, like the logistic function or arctan, to limit the risk of an instrument is common practice (e.g., Schmid and Wirth, 2021). Even interpreting very strong momentum as "overbought", therefore reducing the weight, is not uncommon. Moreover, thresholds can be applied.

position of trend-followers in a specific commodity c_t^* as[3]:

$$c_t^* = \sum_{i=0}^{n-1} \beta_i r_{t-i}. \tag{3}$$

2.3. Model-fitting

Dividing the aggregate positions of a group of traders (c_t) into trend-followers and other traders (e_t), i.e. $c_t = c_t^* + e_t$, Eq. (3) can be turned into an estimable model as follows:

$$c_t = \sum_{i=0}^{n-1} \beta_i r_{t-i} + e_t, \tag{4}$$

where e_t can also be interpreted as an error term.

Both sides of Eq. (4) show strong positive autocorrelation. For the observable traders' position on the left-hand side of the equation, this is just an empirical property stating that funds do not turn over their portfolio every week. For the right-hand side autocorrelation results because the βs are smooth, a property that we further discuss at the end of this section. Therefore, as illustrated by Granger et al. (2001), direct estimation of (4) can suffer from spurious regression. The standard solution to the problem is taking the first difference on both sides:

$$\Delta c_t = \sum_{i=0}^{n-1} \beta_i \Delta r_{t-i} + \varepsilon_t, \tag{5}$$

where $\Delta c_t = c_t - c_{t-5}$ is the weekly change of the position and $\Delta r_t = r_t - r_{t-5}$, is the difference between two daily returns that are a week apart.

While Eq. (5) results naturally from a standard statistical procedure, relating changes in positions to changes in returns is hard to interpret economically. We can bring it into a more intuitive form by rearranging coefficients:

$$\Delta c_t = \underbrace{\sum_{i=0}^{4} \beta_i r_{t-i}}_{news} + \underbrace{\sum_{i=5}^{n-6} \left(\beta_i - \beta_{i+5}\right) r_{t-i} - \sum_{i=n-5}^{n-1} \beta_{i+5} r_{t-i}}_{signal\ noise} + \varepsilon_t. \tag{6}$$

The first term on the right-hand side is a weighted sum of daily returns over the period of position changes. It reflects how contemporaneous returns – or the incoming "news" – alter the signal and hence the position. This term is related to equation (5) in Cheng et al. (2015), equation (6) of Kang et al. (2020), and the results in Table 2 of Cheng and Xiong (2014) where position changes are also explained with contemporaneous returns. Compared to these approaches, our representation is more flexible because daily returns can take on different coefficients within the period of position changes.

The other two terms on the right-hand side reflect how the date information in lagged returns is faded away. They largely determine how traders exit their position. Efficient trading must be dominated by the news term because trading on dated information, which can be seen as some sort of "signal noise", would incorporate information at a speed that is very difficult to reconcile with any reasonable level of market efficiency. Large traders trading on the signal noise term that is known in advance would also be vulnerable to front-running.

The variance of the second term in Eq. (6) is small if the coefficients of the return signature plots in Fig. 1 vary very little from one lag to the next, and the third term's variance is small if the final coefficients converge to zero. Therefore, the return signature plot must be smooth to curb trading based on the signal noise.[4] This smoothness is an important characteristic of our estimation strategy, therefore we highlight other reasons for the aggregate return signature plot to be smooth. Firstly, many of the popular signals shown in Fig. 1 are already smooth. Secondly, real-life traders will try to reduce trading and transaction costs. The split momentum strategy ($smom(60, 5)$) in Fig. 1 is an example of how partial rebalancing leads to a smoother return signature plot. Additional examples can be found in Appendix A. Thirdly, further smoothing is to be expected because funds might not rebalance daily, wait for some sort of signal confirmation or split trades over several days. Discretionary managers often include momentum signals as one input but then trade less timely than systematic funds. Finally, practitioners use a variety of signals with substantial variation in the shape and the look-back period, which is smoothing out any discontinuities in the aggregate.

[3] For simplicity of notation, we omit indicators for the group of traders and the specific commodity.

[4] For the discontinuous standard momentum, the signal noise term, which reduces to the third term of Eq. (6) in this case, is in fact as large as the news term. This unfavorable news-to-noise ratio of 1:1 is the reason why real-life traders, especially short-term traders, avoid using simple momentum. They prefer smoother signals with a much better news-to-noise ratio, such as the moving averages on the right-hand side of Fig. 1.

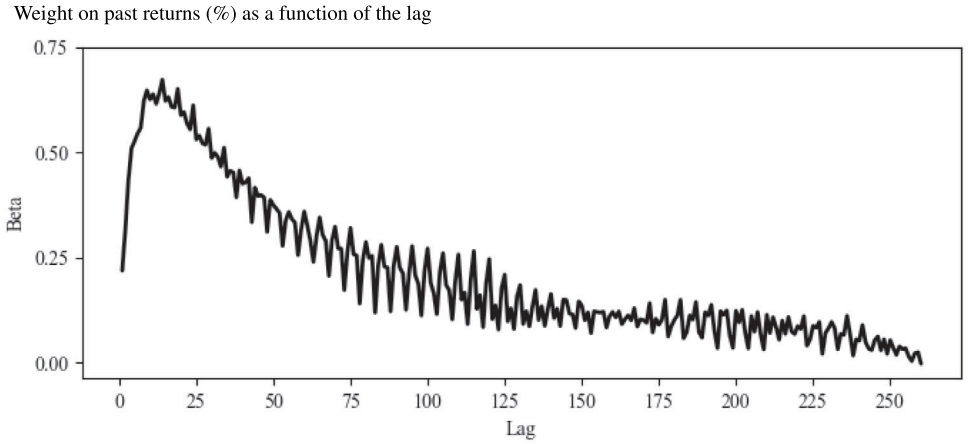

Fig. 2. Return signature plot of the panel regression.
Note: The graphs shows pooled and unregularized OLS estimates of the βs in Eq. (5) for a panel of all changes in the speculative pressure index (see Section 3.1) for net non-commercials. The sample spans the entire time period (1993–2019) and all 23 commodities (29,618 observations).

2.4. Estimation approach

As a first attempt to estimate Eq. (5) or equivalently Eq. (6), we build a panel of all position changes for the group of net non-commercials (total of 29,618 observations) stack them and run a pooled OLS regression to generate a common return signature plot. Position data are made comparable across commodities through the use of the speculative pressure index as outlined in Section 3.1. The adjusted R^2 of the regression is 24.7 and 230 out of the 260 coefficients have a $t-$statistics larger than 2. The β coefficients are plotted in Fig. 2. It shows that pooled OLS fits the lowest lags quite smoothly, but coefficients start to become slightly unstable for lags higher than 15 and then the returns signature plot turns into a sawtooth-like pattern for medium and higher lags. Apparently, even with pooled data, OLS cannot generate a stable and smooth return signature plot. On the other hand, the figure already shows a hump-shaped pattern, which we show is typical for the aggregate return signature plot.

Although this approach is not well-posed, it substantially outperforms other trend-following models that try to capture changes in positions in terms of R^2 using weekly returns. For example, our previous research follows Elaut and Erdős (2018) and finds that changes in momentum signals have a significant impact on the positioning. The results, however, are an order of magnitude weaker than the results of the pooled OLS regression. Similarly, Kang et al. (2020) fit a panel of changes in the speculative pressure index using weekly returns. They report a maximum R^2 of 6.58% and conclude that by far the largest fraction of the variation in position changes is unrelated to momentum. These comparisons show that the prediction of position changes can substantially be improved by using daily instead of weekly returns. But to use daily returns, we need some parametrization or regularization to prevent overfitting end to stabilize the return signature plot.

We guide our choice of the estimation method by two observations. Firstly, as outlined above, there are strong reasons to assume the return signature plot is smooth. Secondly, there is very little theoretical reasoning about the exact shape of the return signature plot or the optimal look-back period. Momentum is usually justified by some cognitive or behavioral bias, such as the anchoring effect of Barberis et al. (1998), the disposition effect of Barberis et al. (2001), or Hong and Stein's (1999) unified theory of underreaction and overreaction. None of the above lead to a parametric form of the return signature plot or deliver guidance about its shape or the length of the look-back period.

For these reasons, we implement a generalization of the ridge or Tikhonov (1963) regression described in Hoerl and Kennard (1970). This regularization approach imposes no parametrization, and no prior restriction on the shape of the return signature plot, but enforces its smoothness. It can be regarded as the l_2 analogous to the fused Lasso[5] estimator of Tibshirani et al. (2005). Because of this analogy, we call it "fused ridge". It imposes smoothness by penalizing the differences of the adjacent β coefficients of Eq. (5). Specifically, we minimize:

$$\min \sum_{t=1,6,\ldots,T} (\Delta c_t - \sum_{i=0}^{n-1} \beta_i \Delta r_{t-i})^2 + \sum_{i=0}^{n-2} \gamma_i^2 (\beta_i - \beta_{i+1})^2 + \gamma_{n-1}^2 \beta_{n-1}^2 \tag{7}$$

where the first term is the usual least squares loss function and the second term is an extra penalty for beta variation. As in the standard ridge, γ_{n-1} shrinks β_{n-1} (the last coefficient that drops out of the look-back period the following day) towards zero. The γs are hyperparameters determining the amount of smoothing imposed for different is. In the base case, we impose uniform smoothing by setting $\gamma_i = \bar{\gamma}$ for all $i < n - 1$.

[5] We also experimented with fused Lasso. However, this design has not proven to be numerically stable.

This approach is remarkably parsimonious and is able to regularize the problem with a minimum set of restrictions. It uses a single hyperparameter $\bar{\gamma}$ besides the look-back period T. Even more, $\bar{\gamma}$ is determined by the data as outlined in Appendix B.1 so that there remains no discretion for our base case estimator beyond the definitions in Eq. (7). Another advantage of this approach is that it is estimated using an augmented OLS regression (see Appendix B for the details) and thus additional (control) variables can be easily added as in Section 4. We therefore clearly prefer this approach to other methods, such as piece-wise polynomial restrictions.

Eq. (7) cannot be based on Eq. (6) because the rearrangement of coefficients introduces discontinuities.[6] Nevertheless, comparing the two equations further motivates our choice of regularization. The second term in (7) punishes variation in the βs, which is exactly what we need to limit the first signal noise term in (6). Similarly, the third term in (7) shrinks the final coefficients, which is the other requirement to control the signal noise term in (6).[7]

2.5. The shape of the smoothing function

In addition to our baseline model, we also look into specifications that allow the γs to vary. We conjecture that imposing more smoothing for higher-order lags relative to lower-order lags improves the forecasting results. This is because a daily return that is almost a year old should enter into the signals (2) and (3) with nearly the same weight as the return the day before or the day after. However, the weight might vary more for recent returns. To account for these considerations, we also present results for $\gamma_i = \bar{\gamma}\sqrt{i}$, which we refer to as square root smoothing. The square root function imposes only limited smoothing for the first few observations. Smoothing increases rapidly for the middle lags. At the end of the look-back period, smoothing becomes relatively homogeneous while still increasing on a small scale. In both cases, the baseline model and the square root model, $\bar{\gamma}$ is determined by the data using the leave-one-out cross-validation approach (LOOCV) as outlined in Appendix B.1. In Section 3.5, we show that imposing less smoothing for lower lags helps the model to better learn the parameters.[8]

3. Predicting position changes

3.1. Data

Our analysis covers 23 commodity futures that are either eligible for inclusion in the Bloomberg Commodity Index or constituents of the S&P GSCI (former Goldman Sachs Commodity Index) as of the end of 2019. This includes six energy futures (WTI crude oil, gasoline, heating oil, natural gas, Brent crude oil, and low sulfur gas oil), six futures from the grains complex (corn, Chicago and Kansas wheat, soybeans, soybean meal, and soybean oil), three livestock futures (lean hogs, live cattle, and feeder cattle), four softs (sugar, coffee, cotton, and cocoa), and four metals (gold, silver, platinum, and copper). LME base metal futures are not included because their positioning reports only go back to 2015.

Our positioning data are from the publicly available commitment of traders (CoT) reports. Per commodity, this report releases the number of long and short contracts held by each group of traders, as well as the open interest, which is the total amount of contracts outstanding. However, no details about specific expiry dates are disclosed.[9] The CFTC publishes positions for futures trading in the U.S. Brent crude oil and low sulfur gas oil are traded in London, and ICE Europe publishes their positioning data. The classifications from the two regulators are identical (see ICE, 2021). Positioning of the traders are usually collected weekly on Tuesday and publicly released three days later after the market closes on Friday (CFTC, 2021).

We regard the classification of commercials and non-commercials (or speculators) from the legacy report. From the disaggregated CoT reports, we use net positions for a series of subgroups: (1) managed money (also labeled as large speculators), consisting of commodity trading advisors, registered commodity pool operators, or unregistered funds; (2) swap dealers; and (3) other reportables (or small speculators). For traders in the subclass of producers/merchants/processors/users, we differ between the short position holders and long position holders because we find that the two sides react very differently to price changes. As in Cheng and Xiong (2014), somewhat simplifying, we refer to the short side as producers and the long traders as users. Further, through our decomposition, we are able to split the class of speculator into momentum traders and other speculators, which we treat as two additional subgroups. Finally, the residual group of non-reportables is also part of our analysis.

CFTC switched to weekly CoT reports in 1993 and started the disaggregated CoT on June 20, 2006. Therefore, our data spans January 1993 to December 2019 for non-commercial traders and June 2006 to December 2019 for the subgroups. Some commodities have a shorter data history. The Table 1 provides an overview of the available positioning data history, commodities' sectors, names, tickers, and exchanges.

The remainder of the data – prices, returns, and contract sizes (multipliers) – are provided by Bloomberg Finance L.P. Throughout the paper, we ignore the cash or collateral return and we only consider excess returns. They are calculated using the active futures[10] from Bloomberg.

[6] It is possible to build a regularization based on (6) but this would involve more hyperparameters to account for the discontinuity.

[7] This analogy would even be stronger if we were to observe daily position changes. In this case, Eq. (6) becomes $\Delta c_t = \beta_0 r_t + \sum_{i=1}^{n-2} \left(\beta_i - \beta_{i+1}\right) r_{t-i} - \beta_{n-1} r_{t-n-1} + \epsilon_t$.

[8] We also examine other smoothing functions, such as functions based on arctan or on the logarithm. They yield similar results as long as they impose substantially less smoothing at the front end. More general specifications have also not proven to enhance the out-of-sample predictability. Even simple combinations, such as the square root plus a constant result in unstable hyperparameters and, as so often found with more complex designs, their out-of-sample performance lags behind the simpler models.

[9] For agricultural commodities, the number of contracts held in the current (old) crop year are also released.

[10] The active futures is the contract with the highest open interest, usually the front contract.

Table 1
Data overview.

Sector	Name	BB-Ticker	Exchange	Start date	
				Legacy report	Disaggregated report
Energy	WTI	CL	NYMEX	5/01/1993	20/06/2006
	Gasoline	XB	NYMEX	14/02/2006	20/06/2006
	Heating oil	HO	NYMEX	5/01/1993	20/06/2006
	Natural gas	NG	NYMEX	5/01/1993	20/06/2006
	Brent	CO	ICE Europe	5/01/2011	11/01/2011
	Gas oil	QS	ICE Europe	5/01/2011	11/01/2011
Grains	Corn	C	CBOT	5/01/1993	20/06/2006
	Chicago wheat	W	CBOT	5/01/1993	20/06/2006
	Kansas wheat	KW	CBOT	5/01/1993	20/06/2006
	Soybeans	SB	CBOT	5/01/1993	20/06/2006
	Soybean meal	SM	CBOT	5/01/1993	20/06/2006
	Soybean oil	BO	CBOT	5/01/1993	20/06/2006
Livestock	Live cattle	LC	CME	5/01/1993	20/06/2006
	Feeder cattle	FC	CME	5/01/1993	20/06/2006
	Lean hogs	LH	CME	2/04/1996	20/06/2006
Softs	Sugar	SB	ICE U.S.	5/01/1993	20/06/2006
	Coffee	KC	ICE U.S.	5/01/1993	20/06/2006
	Cotton	CT	ICE U.S.	5/01/1993	20/06/2006
	Cocoa	CC	ICE U.S.	5/01/1993	20/06/2006
Metals	Gold	GC	COMEX	5/01/1993	20/06/2006
	Silver	SI	COMEX	5/01/1993	20/06/2006
	Platinum	PL	COMEX	5/01/1993	01/12/2009
	Copper	HG	COMEX	5/01/1993	20/06/2006

Note: The table gives an overview of the regarded futures markets including the sector, name, exchange, and the availability of the time series for both CoT Reports. The end of the interval is December 31, 2019.

Based on the these data, we construct two types of variables to characterize the positions and trading behavior of futures markets participants: (net) exposure and pressure indices. Per group of traders and commodity, the net U.S. dollar exposure at time t is defined as:

$$c_t = p_t * m_t * \text{netpos}_t, \tag{8}$$

where p_t is the price, m_t is the contract multiplier, and netpos_t is the net position (number of long contracts minus short contracts). Prices vary along the futures curve, and as only the aggregated number of all contracts is released in the CoT reports, the exact average price is not known. We take the price of the active futures as an approximation. Speculative traders size their trades in terms of U.S. dollar exposure, therefore this quantity is of particular interest when we track their trading behavior.

Pressure indices are calculated by dividing the net position of a group of trades by the total open interest oi_t:

$$c_t = \frac{\text{netpos}_t}{\text{oi}_t}. \tag{9}$$

For commercials, with opposite sign, this is a common definition of Keynes' (1930) and Hicks' (1939) hedging pressure (e.g., Dewally et al., 2013; Kang et al., 2020). Fan et al. (2020) label a similar index speculative pressure in the case of non-commercials. The major advantage of this normalization is that it is stationary by construction. This makes it comparable across commodities and over time. Stationarity is also a requirement of many test statistics.

Finally, exposure contains information about both prices and positions and is hence not appropriate to examine trading activity. Therefore, we use the raw number of contracts held by a group of traders (netpos_t) to analyze trading patterns.

3.2. Empirical design

The aim of our empirical analysis is fourfold. Firstly, we show that position changes of speculative traders can be predicted ahead of the release on Friday by using realized returns until Tuesday. Secondly, the return signature plot of aggregate speculators is revealed. Thirdly, we examine whether or not prediction can be improved through the square root smoother in Section 2.5. Finally, we identify the counterparties of speculative traders in Section 5.

A challenge for the empirical analysis is that open interest and speculative trading activity are not stable over time and both quantities show a steady upward trend over our data sample. Consequently, the parameters in Eq. (5) cannot be assumed to be stable over time, but, rather, we should assume that they adjust (slowly) to the trading activity. The upward trend, however, is quite steady and it is reasonable to presume that the most recent past is representative for the near future. Therefore, we apply a rolling-window approach for dynamic parameter learning. Specifically, we estimate the model from a five-year data window with the last observation τ and then use these parameters to forecast position changes $c_{\tau+10} - c_{\tau+5}$. Lagging the out-of-sample forecast by one-week accounts for the delayed disclosure of the CoT positions. Next, we re-estimate the model parameters by rolling the

five-year estimation period one week forward and, based on these parameters, forecast the position changes two weeks ahead. We continue until the last element of the estimation window is the last but two observations of the entire sample, which generates the parameters for the out-of-sample forecast of the last observation.[11] In line with the literature (e.g., Erb and Harvey, 2006; Fuertes et al., 2010), we use a maximum look-back period of 12 months. We are also not aware of any practical implementation using a look-back period of more than 12 months. Longer look-back periods have no material impact on the results.[12]

3.3. Evaluation measures

We primarily quantify the model's goodness-of-fit using the coefficient of determination of our out-of-sample forecast:

$$R^2 = 1 - \frac{\text{MSE}}{\text{TSS}} = 1 - \frac{\sum_{t\in W}(\Delta c_t - \widehat{\Delta c_t})^2}{\sum_{t\in W}(\Delta c_t - \overline{\Delta c_t})^2}, \tag{10}$$

where MSE is the model's mean squared error, TSS is the total sum of squares of the observed data, and W the set of weekly observations, where an out-of-sample forecast $\widehat{\Delta c_t}$ has been calculated. Note that a main objective of our analysis is to quantify the importance of trend-following. Therefore, R^2 should not just be interpreted as a measure of predictive power but rather as a measure indicating how relevant trend-following and momentum strategies are in the realm of speculative trading.

To detect biases and scaling issues, we run the Mincer and Zarnowitz (1969) regression using the same set of data points:

$$\Delta c_t = a + b * \widehat{\Delta c_t} + \eta_t. \tag{11}$$

A good forecast is unbiased and has an intercept of $a = 0$. The slope coefficient b should be close to one because otherwise a simple scaling of the forecast would improve its fit. Therefore, we test the hypothesis that the intercept equals zero and the slope is equal to one. Finally, we also test for autocorrelation in the first and second lags of the model's estimation error with:

$$\hat{\varepsilon}_t = \Delta c_t - \widehat{\Delta c_t}. \tag{12}$$

For model selection (i.e. the choice of the smoothing function), we use the Diebold and Mariano (1995) test. It tests whether two forecasts differ significantly. Moreover, we use an augmented Mincer and Zarnowitz (1969) regression that includes the forecasts of the two models:

$$\Delta c_t = a + b_1 * \widehat{\Delta c_{1,t}} + b_2 * \widehat{\Delta c_{2,t}} + \eta_t. \tag{13}$$

A superior model should squeeze out the other model in this regression. Assuming model 1 dominates, b_1 is significant and close to 1; b_2, on the other hand, is insignificantly different from zero.

3.4. Predicting position changes of non-commercials

We start the empirical analysis with non-commercials because they have the longest data series. Table 2 displays forecasting power in terms of R^2 for the three types of position changes (speculative pressure index, number of contracts, and exposure), as well as for the two different smoothing functions.

The columns (1) and (2) display the R^2 values of the position changes in the speculative pressure index. We find an observation-weighted average (shown in the last row of the table) of 26.9% for uniform smoothing and 32.1% for square root smoothing. The base model (1) works best for corn 42.9% and the square root model (2) for coffee (49, 8%). We find the weakest predictability of speculative pressure in gasoline (6.3% and 8.4% respectively) and in the livestock sector.

Predicting the number of contracts instead of the pressure index (columns (3) and (4) in Table 2) increases the average R^2 considerably to 32.8% (+5.9%) for the base model, and to 37.3% (+5.2%) for square root smoothing. In line with the results for the speculative pressure index, we find that the forecasting performance is exceptionally high for grains and softs. In these two sectors the average R^2 is 40.2% and 37.7% for uniform smoothing and even 43.9% and 43.3% for the square root model.

As mentioned in Section 3.1, speculative traders often size their trades based on the dollar amount. So we also predict exposure changes. Results of this forecast over the whole period are displayed in columns (5) and (6) in Table 2. In line with the previous results, we find that coffee and corn are the best-predicted markets with R^2s of 58.3% and 54.4% for uniform smoothing, and 61.8% and 56.4% for square root smoothing, respectively. The model works, best for commodities in the sectors of grains, softs, and metals, with a sector average R^2 of 48.5%, 50.7%, and 48.7%. For these three sectors, the predictability resulting from the square root smoothing model consistently exceeds 40%. For livestock and the less homogeneous energies, the average R^2s are markedly lower. By far the least explained market is natural gas with mere predictability of 6.0% and 9.4%, respectively. By comparing the observation-weighted average, we find that the exposure model depicts the highest overall R^2 with 38.5% and 41.8%, respectively.

[11] This rolling window approach captures the upward trend in trading activity only with some delay, and it cannot handle fast parameter changes. A complete modeling of the trend would require a regularized state–space model estimated with the Kalman filter [see Li et al. (2014) for an outline on ridge-type Kalman filtering]. With the large number of features, this approach, however, is computationally substantially more demanding, especially in an out-of-sample testing framework. In-sample hyperparameter tuning is another challenge to this setup. Therefore, it is beyond the scope of this paper to fully model the growth dynamics.

[12] Empirically, we find some marginal improvement when extending the look-back period to 270 days. For example, using square root smoothing, the average prediction R^2 for non-commercials increases by 0.3%. For 280 and 300 days, we find an average R^2 that is almost identical to the annual look-back period.

Table 2
Out-of-sample R^2 for the classification of non-commercials.

	Non-commercials									
	Position change	Pressure index		Number of contracts		Exposure					
	Start date	01/98	01/98	01/98	01/98	01/98	01/98	01/98	07/03	01/09	07/14
	End date	12/19	12/19	12/19	12/19	12/19	12/19	06/03	12/08	06/14	12/19
	Smoothing	flat	sqrt	flat	sqrt	flat	sqrt	sqrt	sqrt	sqrt	sqrt
	Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Period							(I)	(II)	(III)	(IV)
Energy	WTI	17.3	21.9	23.3	27.8	36.4	38.1	20.8	16.2	31.0	51.8
	Gasoline	6.3	8.4	8.1	10.8	27.4	28.1			30.3	25.8
	Heating oil	22.8	26.8	20.6	25.1	14.8	18.6	29.1	22.8	31.3	−29.1
	Natural gas	21.2	26.0	14.8	20.9	6.0	9.4	26.5	7.1	−0.4	17.0
	Brent	25.2	32.8	24.8	31.1	36.1	40.1				40.1
	Gas oil	28.0	36.8	27.2	34.8	17.5	24.4				24.4
Grains	Corn	42.9	47.6	46.0	49.0	54.4	56.4	58.6	55.9	52.7	61.4
	Chicago wheat	37.5	40.9	41.9	43.7	38.8	40.2	55.0	29.3	29.7	53.6
	Kansas wheat	28.4	34.0	34.5	40.3	47.4	51.5	39.1	48.7	48.5	56.0
	Soybean	34.7	37.9	38.5	41.6	45.1	47.2	45.9	45.9	44.5	50.6
	Soybean meal	33.6	37.9	39.1	42.3	49.4	51.4	45.0	49.6	44.3	56.0
	Soybean oil	30.8	37.4	41.3	46.5	39.9	44.5	32.5	28.4	32.2	64.6
Livestock	Lean hog	20.7	25.8	25.8	30.7	28.4	32.0	25.5	26.7	27.3	37.6
	Live cattle	13.5	17.9	20.6	26.4	27.5	32.1	29.0	14.9	26.2	39.1
	Feeder cattle	14.4	21.4	16.8	22.3	19.3	23.3	32.7	14.6	33.1	15.1
Softs	Sugar	30.2	37.4	32.3	38.2	46.1	50.1	26.1	36.8	48.4	56.2
	Coffee	42.3	49.8	50.2	54.7	58.3	61.8	53.9	59.0	56.2	67.8
	Cotton	34.9	42.2	37.2	42.5	45.2	48.7	41.7	41.6	33.8	62.8
	Cocoa	20.3	27.2	31.2	37.6	36.7	42.2	−36.5	40.2	41.2	45.9
Metals	Gold	20.5	22.8	35.6	38.1	49.7	51.2	2.9	53.9	41.4	56.3
	Silver	23.2	27.3	32.7	35.9	46.9	48.9	66.0	49.6	43.1	50.8
	Platinum	20.3	26.0	45.0	49.5	49.7	53.1	31.4	25.6	35.3	72.7
	Copper	35.6	40.3	39.3	43.7	37.1	41.6	61.9	17.9	19.0	56.4
Weighted average		26.9	32.1	32.8	37.3	38.5	41.8	34.6	34.2	35.8	45.2

Note: The table presents out-of-sample R^2 in percent for the classification of non-commercials. The type of the position change, the smoothing function, and the time period are indicated in the header. "flat" indicates uniform smoothing, and "sqrt" square-root-smoothing. The weighted average on the bottom of the table is weighted by the number of out-of-sample observations.

Note that we have increased the previously reported in-sample R^2 of momentum based position forecasts from 6.58% (Kang et al., 2020) to an out-of-sample R^2 of 41.8%. There are three reasons for this improvement. Firstly, and most importantly, we switch to daily returns. According to the panel regression in Section 2.4, this brings the coefficient of determination to 24.7%. Secondly, the fused ridge approach allows us to estimate each commodity on its own and thus to account for cross-sectional differences. This increases the R^2 to 32.1%. Thirdly, with an R^2 of 41.8%, the exposure can be better predicted than with the pressure index. In Section 3.6.2, we show that predictability for large speculators is 45.6%, which is even higher.

3.5. Shape of the tuning function and empirical return signature plots

A striking finding of the previous section is the higher predictive power of the square root model across all types of position changes, as well as all commodities. On average, the difference amounts to more than three percentage points. There is no forecast where uniform smoothing outperforms square root smoothing. In this subsection, we elaborate on the reasons for this improvement by looking more closely at the aggregate trading signal. We also test if the improvement is significant.

Estimated return signal plots for the exposure of speculators are depicted in Fig. 3. The shapes of these empirical β coefficients are directly comparable with the more theoretical signal plots of Figs. 1 and 7. Of course, these estimated quantities vary with the estimation window and thus over time (the evolution of the return signature plots is discussed in Section 3.7), which is why we present a snapshot at the end of our dataset.

For all commodities, the return signature plot has a hump-shaped form, similar to the shape of the *xmac*. The peak can be a bit rounder as for most of the grains and livestock or a bit sharper as, for example, for silver and bean oil. In that case, they more closely resemble the unweighted *mac*. Some futures also have plateaus — ranges where the betas are barely decline with the lag and they occur when momentum trading is dominant (e.g., for corn and soybeans). For the energy futures, the hump is narrower, more pronounced, and fairly symmetric. The curves start to decline in an almost linear fashion after roughly 30 days. Another difference is that β_0 is higher, especially for the two crude oils. Apparently, there is a good deal of short-term trend-chasing in these markets.

We cannot identify the reason for the hump-shaped pattern. Delayed trading, some kind of moving average crossing or the presence of reversal traders are all possible reasons. For the first lag, there is, however, also a technical reason: The signal resulting from today's return cannot be completely implemented today because it is only known at the end of the day.

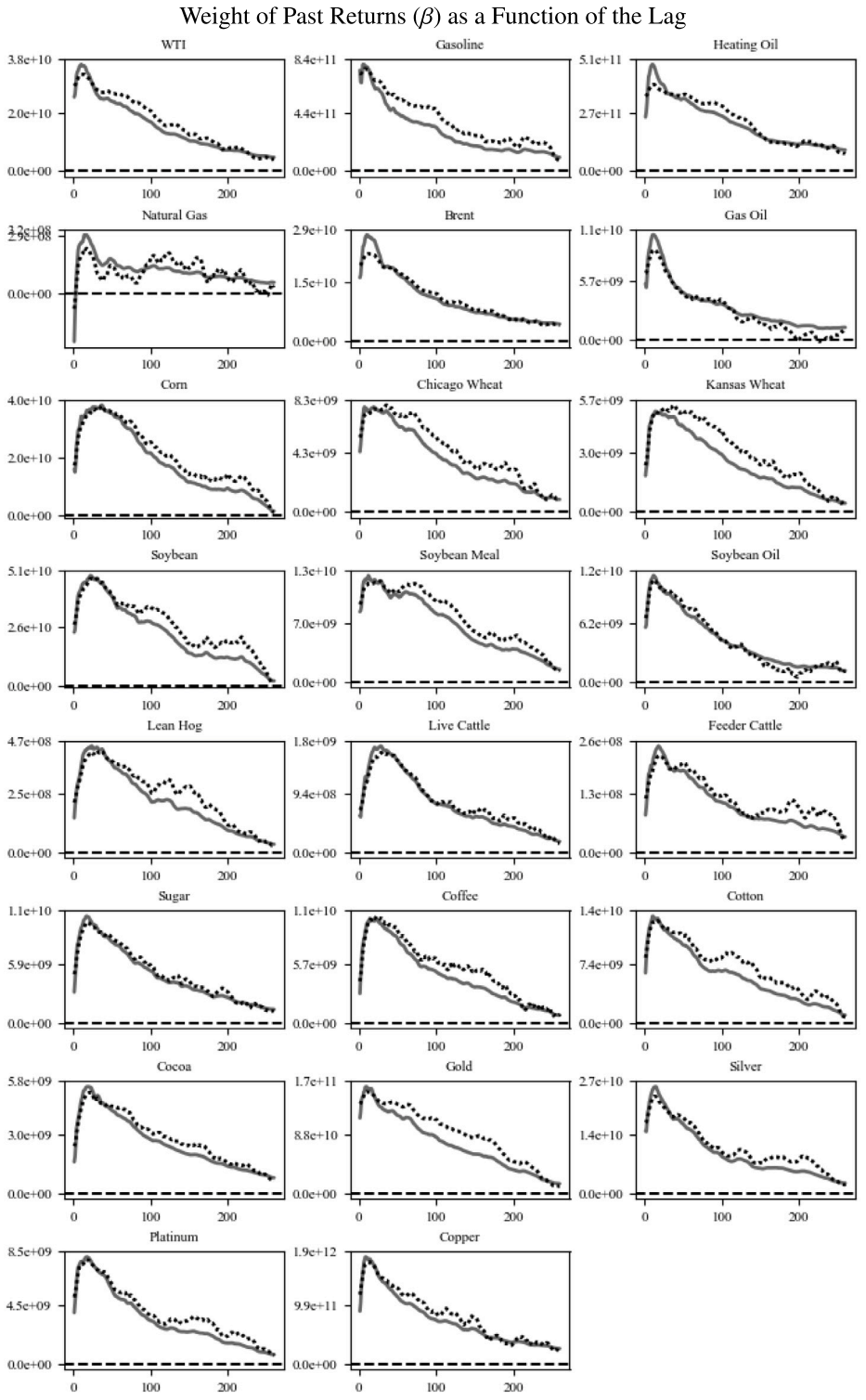

Fig. 3. Empirical return signature plots with different smoothing.
Note: The graphs show estimates of the βs in Eq. (5) for the exposure of non-commercials with uniform regularization in black (dotted line) and square root regularization in gray. The five-year subsamples end on December 31, 2019.

The loading on the first return is consistently lower for the square root model, but then the peak is higher for most of the commodities. The reduced smoothing at the near end provides the square root model with the flexibility necessary to capture this steep ascent. Contrary to this, the base model smoothes too much because otherwise the higher lags, which are smoothed with the same parameter, would become unstable. We observe that the base model produces a bumpier curve at the far end. Apparently, the difficulties the base model encounters with balancing the smoothing are the reason why the square root model outperforms it.

The variation across commodities is not easy to capture with a single model specification as we also need to reduce the variance of the βs at the same time. Clearly, the fused ridge approach does an excellent job in balancing the regularization needed to reduce the sample error and the flexibility necessary to capture all the different trading patterns. This is even more true for the square root model.

To more formally compare the forecast accuracy of the different smoothing techniques, we rely on two statistical tests. Firstly, we use the Diebold and Mariano (1995) test to determine whether the forecasts are significantly different. Secondly, we use an augmented Mincer and Zarnowitz (1969) regression, to show that the uniform weighted smoothing model gets squeezed out. The results for the exposure are displayed on the left-hand side of Table 3, the results for the speculative pressure index on the right-hand side.

For the exposure, the Diebold–Mariano test results in Table 3 strongly reject the equivalence of the uniform and the square root smoothed models for all but two commodities, namely Chicago wheat (the null hypothesis is rejected at the 10% level), and gasoline (null hypothesis cannot be rejected). The results of the augmented Mincer–Zarnowitz regression are in line with these findings. The coefficients of the square root model are all positive and significant, and for all but two even larger than 1. On the other hand, the base model often has negative loadings (in 16 cases), and only for Chicago wheat is it significantly positive at the 10% level. For soybean oil and cocoa, the loading on the base model is even significantly negative.

Table 3
Model comparison.

		Exposure					Speculative pressure index				
		Mincer–Zarnowitz regression				DM-test	Mincer–Zarnowitz regression				DM-test
	Smoothing/z	flat		sqrt		z-score	flat		sqrt		z-score
Energy	WTI	−0.20	(−0.7)	1.30	(5.0)	−3.2	−0.41	(−3.0)	1.12	(8.4)	−3.5
	Gasoline	0.20	(0.7)	0.62	(2.1)	−0.6	−0.08	(−0.3)	0.76	(2.9)	−1.2
	Heating oil	−0.41	(−1.7)	1.27	(4.9)	−3.5	−0.11	(−0.9)	0.93	(7.4)	−3.6
	Natural gas	−0.32	(−1.5)	1.04	(5.2)	−2.7	−0.38	(−2.5)	1.25	(8.9)	−3.9
	Brent	0.09	(0.2)	1.55	(4.0)	−3.2	−0.29	(−0.8)	1.68	(5.2)	−3.6
	Gas oil	−0.56	(−1.8)	1.36	(4.1)	−2.7	−0.75	(−2.4)	1.72	(6.0)	−4.0
Grains	Corn	0.15	(0.9)	1.13	(6.1)	−3.2	−0.14	(−1.3)	1.09	(10.0)	−4.8
	Chicago wheat	0.32	(1.8)	0.91	(5.0)	−1.8	−0.03	(−0.3)	0.93	(7.8)	−3.3
	Kansas wheat	−0.15	(−0.9)	1.29	(7.9)	−4.7	−0.12	(−1.1)	1.01	(9.4)	−4.1
	Soybean	0.03	(0.1)	1.22	(5.1)	−2.9	−0.18	(−1.2)	1.11	(7.7)	−3.9
	Soybean meal	0.01	(0.0)	1.34	(5.9)	−3.3	−0.10	(−0.9)	1.02	(8.7)	−3.8
	Soybean oil	−0.50	(−3.0)	1.66	(10.4)	−6.7	−0.24	(−2.3)	1.06	(10.6)	−4.9
Livestock	Lean hog	−0.07	(−0.4)	1.02	(5.0)	−2.8	−0.10	(−0.8)	1.04	(8.2)	−4.4
	Live cattle	−0.14	(−0.7)	1.19	(5.7)	−3.4	−0.44	(−2.8)	1.21	(8.1)	−4.1
	Feeder cattle	−0.08	(−0.4)	1.10	(5.8)	−3.3	−0.22	(−2.0)	1.04	(9.5)	−3.0
Softs	Sugar	−0.25	(−1.5)	1.40	(8.3)	−5.3	−0.31	(−3.2)	1.07	(11.9)	−4.0
	Coffee	−0.16	(−1.0)	1.25	(7.6)	−4.6	−0.34	(−3.4)	1.38	(13.7)	−5.3
	Cotton	−0.06	(−0.4)	1.18	(8.2)	−4.7	−0.26	(−2.7)	1.10	(12.0)	−5.2
	Cocoa	−0.48	(−2.7)	1.50	(8.4)	−5.3	−0.38	(−3.4)	1.11	(10.3)	−4.9
Metals	Gold	0.18	(0.7)	1.31	(5.3)	−2.8	−0.09	(−0.6)	0.79	(5.6)	−0.8
	Silver	−0.19	(−0.8)	1.62	(7.0)	−3.9	−0.19	(−1.4)	0.96	(7.2)	−2.8
	Platinum	−0.24	(−1.4)	1.50	(8.7)	−5.5	−0.15	(−1.3)	1.00	(9.0)	−3.1
	Copper	−0.21	(−1.0)	1.65	(7.0)	−5.1	−0.04	(−0.4)	0.94	(9.2)	−3.9

Note: The left-hand side of the table presents the predictions of exposure. First, we show the coefficients of the augmented Mincer–Zarnowitz regression "flat" shows the loadings and the t-statistics in parentheses of the base model and "sqrt" those of the square root model; the intercept is omitted. Next to it, the Diebold–Mariano z-score comparing the forecast accuracy using the two different smoothing functions is shown. t-statistics in this part are based on the jackknife estimator *HC3* of MacKinnon and White (1985). The right-hand side shows the same analysis for the speculative trading index defined in Eq. (9). For the pressure index, we report standard t-statistics. Both parts use the full sample of non-commercial traders.

The results for the speculative pressure index in Table 3 are particularly interesting because the index is stationary by construction and thus better satisfies the assumptions of the Diebold–Mariano test. The Diebold–Mariano test for the speculative pressure index is in line with its results for the exposure model specification. The null hypothesis can be rejected for all commodities but gasoline and gold. Similarly, the augmented Mincer and Zarnowitz (1969) regression consistently displays significantly positive coefficients, close to one. The coefficients for the base model are all negative, in most cases not significantly negative.

The results of this subsection unanimously show that square root smoothing significantly outperforms uniform smoothing. Therefore, for the remainder of the paper, we rely on this specification.

3.6. Robustness

In this subsection, we analyze the robustness of the model. Firstly, by analyzing the performance over four non-overlapping subperiods, which are: (I) January 20, 1998, till June 30, 2003, (II) July 1, 2003, till December 31, 2008, (III) January 1, 2009, till June 30, 2014, and (IV) July 1, 2014, till December 31, 2019. Secondly, we apply the model to the narrower classification of managed money. Thirdly, we run Mincer–Zarnowitz regressions and take a closer look at the autocorrelation of the residuals.

3.6.1. Time consistency

The columns (7)-(10) of Table 2, present the results of the square root model across the four subperiods (I)-(IV) of equal length. The average R^2 is consistently around 35% in the periods (I)-(III) followed by a substantial and fairly consistent increase to an average of 45.2% in the last period. While these findings show that the predictability is robust across different time periods, they also indicate that the full sample R^2 is somewhat overstated. As there is an upward trend in the magnitude of speculative trading and open interest, the later observations with stronger predictability affect the R^2 more than the somewhat weaker forecasts at the beginning of the sample period.

Across all commodities and subperiods, only three display a negative R^2. Besides the poorly fitted natural gas in period (III), it is the first subperiod of cocoa and the last subperiod of heating oil (columns (7) and (10) in Table 2). In the latter two cases, a relatively rapid decline in momentum trading and the associated decline in the βs is the reason for the poor performance of our model. As outlined in Section 3.2, our empirical design cannot handle rapid parameter shifts.

3.6.2. Predicting position changes for large speculators

To further underline the robustness of the model, we apply it to the classification of managed money (2) and also compare its results to non-commercials (1) in Table 4.

The R^2 for the large speculators is on average 2.4% higher than the R^2 of non-commercials (estimations are over the same sample period). This increase confirms the conventional wisdom that trend-following funds, especially the large ones, are usually classified as managed money (CFTC, 2021). The results are in line with the findings for non-commercials in Section 3.4. The model works best for commodities in the sectors of grains, softs, and metals and is less accurate for livestock commodities and the energy sector, although the results for the energies are quite heterogeneous. This model depicts the highest R^2 for platinum (68.9%), but continues to work very well for coffee and corn.

Table 4
Out-of-sample R^2 for large speculators.

	Classification	Non-commercials	Managed money		
	Position change	Exposure	Exposure	Pressure index	Number of contracts
	Column	(1)	(2)	(3)	(4)
Energy	WTI	46.2	42.2	37.7	38.6
	Gasoline	28.1	28.1	23.6	29.7
	Heating oil	11.1	17.7	26.0	32.2
	Natural gas	15.5	33.0	34.5	35.4
	Brent	40.1	51.3	33.0	35.5
	Gas oil	24.4	44.5	47.2	45.5
Grains	Corn	57.0	53.5	48.8	49.2
	Chicago wheat	43.0	50.9	52.0	52.2
	Kansas wheat	51.8	49.6	43.1	44.7
	Soybean	46.6	47.9	43.1	45.3
	Soybean meal	51.7	53.1	44.0	47.4
	Soybean oil	52.7	52.5	53.3	56.1
Livestock	Lean hog	35.3	39.7	34.9	36.1
	Live cattle	33.8	35.7	24.3	30.7
	Feeder cattle	21.0	29.3	22.0	26.6
Softs	Sugar	51.4	51.0	43.9	46.9
	Coffee	63.2	61.3	55.6	53.1
	Cotton	53.4	49.9	30.2	41.8
	Cocoa	44.4	53.4	47.5	53.5
Metals	Gold	52.2	50.0	37.2	44.7
	Silver	49.3	46.3	36.7	36.2
	Platinum	74.8	68.9	59.9	59.0
	Copper	46.1	49.5	43.4	50.0
Weighted average		43.2	45.6	39.7	42.9

Note: The table presents out-of-sample R^2 in percent for the three different type of position changes (exposure, speculative pressure index, and the number of contracts). The estimation period is based on the sample period of the disaggregated CoT report, also for the class of non-commercials. All forecasts are based on square root smoothing.

Table 5
Model biases.

		Mincer Zarnowitz regression				Autocorrelation			
		Intercept		Slope		$corr(\hat{\varepsilon}_t, \hat{\varepsilon}_{t-1})$		$corr(\hat{\varepsilon}_t, \hat{\varepsilon}_{t-2})$	
	Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Energy	WTI	3.3E+07	(0.9)	1.11	(1.4)	0.16	(5.6)	−0.02	(−0.6)
	Gasoline	5.3E+08	(0.2)	0.82	(−2.5)	0.24	(5.1)	0.09	(1.9)
	Heating oil	1.7E+08	(0.1)	0.88	(−1.2)	0.14	(4.9)	−0.04	(−1.4)
	Natural gas	−2.9E+05	(−0.2)	0.78	(−2.8)	−0.05	(−1.6)	0.06	(1.9)
	Brent	9.5E+04	(0.0)	1.63	(3.9)	0.17	(2.6)	0.13	(1.8)
	Gas oil	−3.5E+07	(−0.8)	0.85	(−1.1)	0.10	(1.3)	0.05	(0.7)
Grains	Corn	2.4E+06	(0.2)	1.28	(5.1)	0.28	(10.0)	0.05	(1.8)
	Chicago wheat	−3.8E+04	(−0.0)	1.21	(3.1)	0.00	(0.0)	−0.08	(−2.6)
	Kansas wheat	1.2E+05	(0.0)	1.15	(3.1)	0.17	(6.0)	0.08	(2.7)
	Soybean	2.9E+06	(0.2)	1.25	(4.3)	0.18	(6.1)	0.13	(4.5)
	Soybean meal	1.0E+06	(0.2)	1.35	(6.2)	0.19	(6.7)	0.05	(1.8)
	Soybean oil	−1.1E+05	(0.0)	1.19	(3.8)	0.15	(5.2)	−0.01	(−0.3)
Livestock	Lean hog	1.2E+05	(0.3)	0.96	(−0.6)	0.33	(10.9)	0.13	(3.9)
	Live cattle	3.5E+05	(0.5)	1.07	(0.8)	0.20	(7.0)	0.00	(0.0)
	Feeder cattle	3.1E+03	(0.0)	1.03	(0.3)	0.22	(7.6)	0.09	(3.0)
Softs	Sugar	−1.6E+04	(0.0)	1.16	(3.3)	0.16	(5.4)	−0.03	(−0.8)
	Coffee	−4.0E+05	(−0.1)	1.15	(3.2)	0.13	(4.3)	−0.02	(−0.6)
	Cotton	1.7E+06	(0.3)	1.12	(2.6)	0.24	(8.4)	0.04	(1.5)
	Cocoa	8.5E+05	(0.3)	1.06	(0.9)	0.26	(9.1)	0.09	(3.2)
Metals	Gold	2.5E+07	(0.7)	1.48	(5.7)	0.10	(3.4)	−0.05	(−1.6)
	Silver	2.5E+06	(0.3)	1.43	(5.7)	0.24	(8.5)	0.02	(0.8)
	Platinum	1.7E+06	(0.6)	1.27	(5.3)	0.26	(9.3)	0.05	(1.7)
	Copper	−9.5E+07	(−0.1)	1.45	(5.4)	0.17	(5.9)	−0.07	(−2.4)

Note: The table provides results for the classification of non-commercial traders (entire data samples) with square root smoothing. The left panel provides the results of the Mincer Zarnowitz regressions. The regression coefficients are followed by corresponding t-statistics in parentheses. We are testing if the intercept is different from zero and if the slope coefficient is different from one. The right panel "Autocorrelation" shows first and second-order autocorrelations of the models prediction error in Eq. (12) with the corresponding t-statistics in brackets. All t-statistics are based on the jackknife estimator *HC3* of MacKinnon and White (1985).

3.6.3. Model bias and autocorrelation

The results of the Mincer–Zarnowitz regression are reported in the first part of Table 5. Column (1) shows the value of the intercept a of Eq. (11) and the corresponding t-statistics in column (2). While the intercepts are large in absolute terms, they are all insignificant, and, therefore, the bias in the forecasts is negligible for all commodities. The column (3) displays the slope coefficient b and the t-statistics for b being different from 1 in (4). For the energy sector, the results are rather heterogeneous. For the two varieties of crude oil, b is significantly larger than 1, but for the other energy commodities, it is below 1. This mirrors the fact that the model is less accurate in the energy sector and, more specifically, that it is unstable for natural gas and has some problems fitting heating oil in the last quarter of the sample. For the other sectors, the results are consistent. Beta is either insignificantly different from 1 or significantly higher than 1. As we document above, there is an upward trend in open interest and speculative activity. If trading activity increases, estimates from the most recent past will most likely underestimate the magnitude of trading. This is exactly what we find in the data.

Finally, we also report tests for autocorrelation of the forecasting error $\hat{\varepsilon}_t = \Delta c_t - \widehat{\Delta c_t}$ in Table 5. First- and second-order autocorrelations are depicted in columns (5) and (7) with the corresponding t-statistics in (6) and (8). First-order correlations are significantly positive, usually highly positive, for all but three commodities. Only natural gas displays an (insignificantly) negative first-order autocorrelation. Lean hog (0.33) and corn (0.28) show the largest coefficients and also the highest significance.

Recalling the definition of the residuals in Eq. (5) – position changes of speculators that are not trend-followers – it is not surprising that we see some autocorrelation: Like the trend-followers, other speculators might also implement trades slowly, potentially over many days. This generates a certain level of autocorrelation in the position changes. It largely vanishes after a week, as shown in the last two columns of Table 5. More specifically, the second-order coefficient remains significantly positive only for five markets. It even turns negative for eight commodities and significantly negative for Chicago wheat and copper.

3.7. Evolution of the return signature plots

Fig. 4 shows the estimated return signature plots at different estimation periods for non-commercials. All commodities show a distinct upward trend in the magnitude of trading, and in particular, a large jump from the data point at the end of 2009.

Another remarkably stable characteristic of these plots is the hump-shaped pattern we observe over time and across commodities. Given the relatively large changes in the magnitude of the signal, the shape remains surprisingly consistent. In unreported results, we also inspect other characteristics of the shape: For instance, the average look-back period defined as the β-weighted average of the lag ($\sum i \times \beta_i / \sum \beta_i$), the location of the peak of the hump defined as the lag i with the highest loading, and the range of the lags

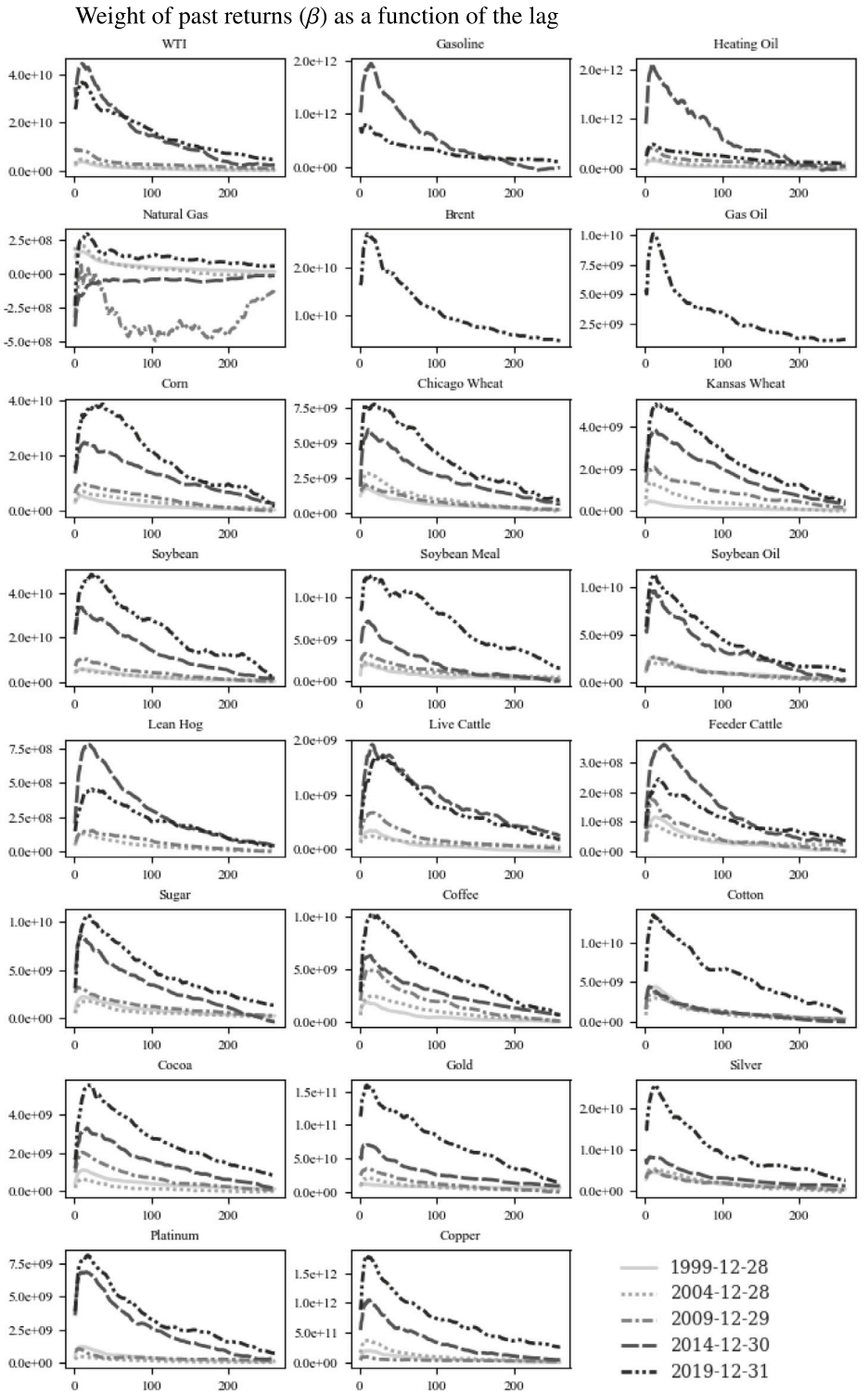

Fig. 4. Evolution of return signature plots for the exposure.
Note: The graphs show estimates of the βs in Eq. (5) for the exposure of non-commercial traders using square root smoothing. The five five-year subsamples are evenly spread over the entire data sample.

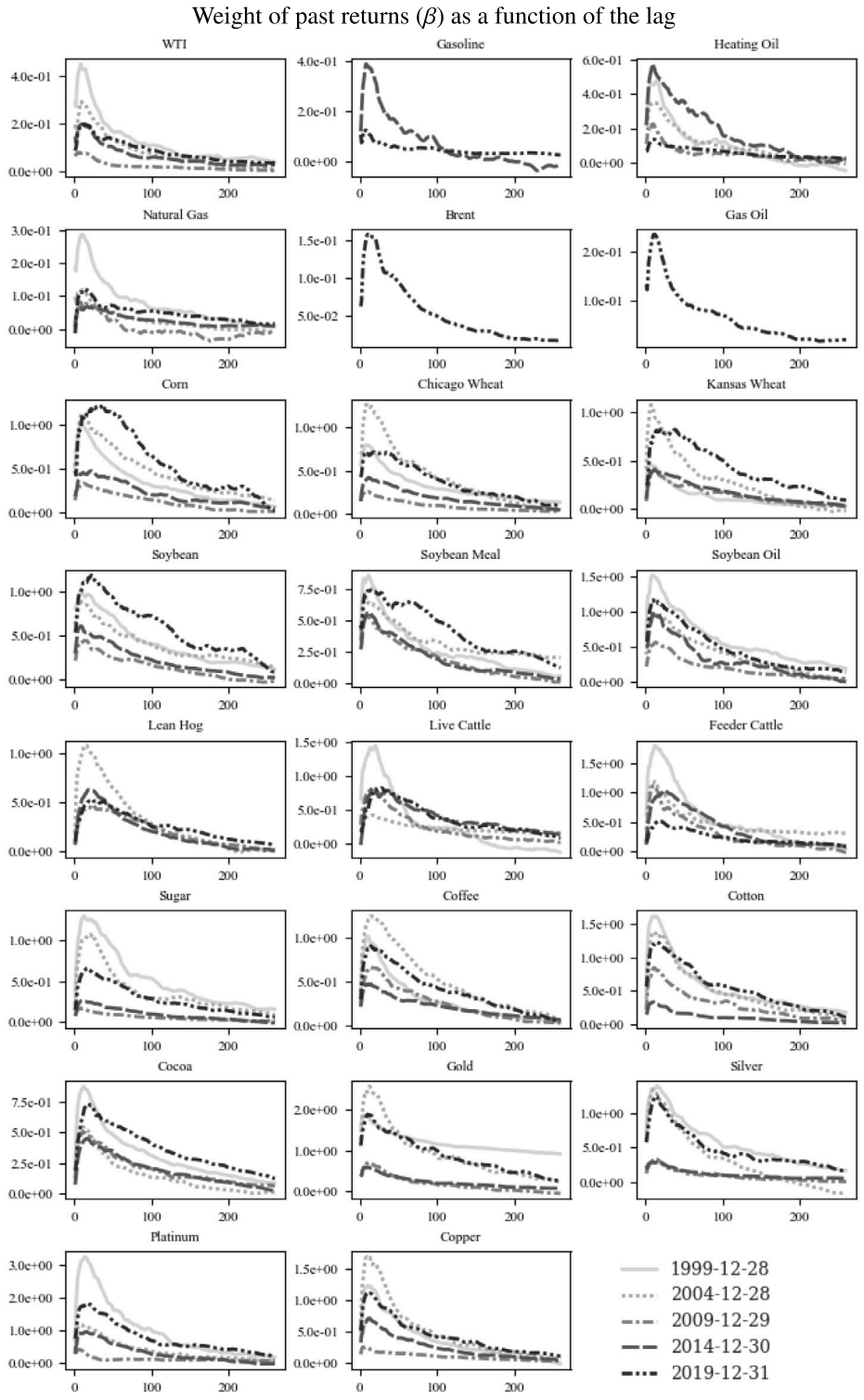

Fig. 5. Evolution of return signature plots for the speculative pressure index.
Note: The graphs show estimates of the βs in Eq. (5) for the speculative pressure index for the classification of non-commercial traders using square root smoothing. The five five-year subsamples are evenly spread over the entire data sample.

containing the 10 largest values of β. While of course fluctuating, all these figures remain, in general, stable over time. There was no striking tendency, for example, for a longer or shorter average look-back period or for a drift of the peak of the hump. The only exception is natural gas, which has an unstable return signature plot caused by periods of strong reversal trading. This instability is in line with the lower predictability of position changes reported in Tables 2 and 4.

Finally, we also take a closer look at the return signature plots of the speculative pressure index, displayed in Fig. 5. The speculative pressure index is stationary by design, therefore we can better analyze trading activity relative to open interest. The graphs show that the hump-shaped pattern is consistent over time and across commodities, also for this measure. However, the upward trend in speculative trading has vanished in relative terms, so we see a very heterogeneous picture where for many commodities trading activity is higher at the beginning of the sample. Moreover, in most markets, the trading activity is very low around the financial crisis. This finding might be related to the convective risk flow observed by Cheng et al. (2015), who show that speculators reduced their net long positions during the crisis.

4. Importance of other factors

The results in Section 3 show that the well-documented commodity momentum premium is indeed exploited by speculators. There are, however, other factors known to explain commodity returns. For example, Erb and Harvey (2006) implement a simple long–short strategy that buys the six most backwardated commodities and shorts the six most contangoed commodities. The Sharpe ratio of this portfolio is almost twice as high as the Sharpe ratio of the long-only index. Fuertes et al. (2010) document that double-sort strategies that exploit both momentum and the basis clearly outperform the single-sort strategies. Among others, Gorton et al. (2012), Szymanowsky et al. (2014), and Koijen et al. (2018) also document the performance potential of basis or carry strategies. Boons and Prado (2019) show that changes in the basis or basis-momentum strongly outperform benchmark characteristics in predicting commodity spot and term premiums in both the time series and the cross-section. Davis et al. (2022) extend these findings to other asset classes.

We experimented with incorporating basis and basis-momentum into our out-of-sample position forecast; however, this did not lead to a better forecasting power of the model. Therefore, we switched to an in-sample framework, which confirmed our findings: changes in the basis do not significantly improve a momentum-based position forecast.

In our in-sample analysis, we use the definition of Fuertes et al. (2010) and for the basis signal, we use the following:

$$\text{basis}_t = \ln(F_{t,1}) - \ln(F_{t,2}), \tag{14}$$

where $F_{t,1}$ is the price of the first generic futures contract, and $F_{t,2}$ is the price of the second generic futures contract. Hong and Yogo (2012) and Gorton et al. (2012) also use the first and the second contract to define the basis. With this definition, the signal change for a basis strategy becomes the difference of the first ($r^w_{t,1}$) and the second weekly futures return ($r^w_{t,1}$):

$$\Delta\,\text{basis}_t = r^w_{t,1} - r^w_{t,2}. \tag{15}$$

This resembles the signal in Boons and Prado (2019), who define basis-momentum as the difference between momentum in a first and second nearby futures strategy. In fact, if we apply decomposition (6) to the basis-momentum signal, $r^w_{t,1} - r^w_{t,2}$ is the news term with constant coefficients. Therefore, adding the basis change in (15) to our momentum-based forecast (5) is informative about position changes related to both basis and basis-momentum strategies.

We forecast changes in the pressure index (because it is stationary) with the basis change signal (15) in addition to the lagged returns. The momentum signal is regularized as described in Section 2 using square root smoothing. Due to the inevitable bias, in-sample testing with regularization is not straightforward and usually not precise. We conduct a non-exact t-test that can be interpreted in exactly the same way as the standard t-test. Details about the procedure and its pitfalls are outlined in Appendix C. The t-statistics of the carry signal per commodity are reported in Table 6. The results are highly inconclusive: 8 commodities have negative t-statistics, and WTI and natural gas even a significantly negative one. Out of the 15 futures with the correct sign, only 5 are significant. Silver, the most significant one, is a commodity with normally minimal variation in the basis. This confirms our (unreported) findings from the out-of-sample analysis that incorporating basis changes do not add predictive power to the position forecast.

A reason for this finding could be that carry strategies are usually described (cf. the references at the beginning of this section) and implemented in a cross-sectional framework. Therefore, relative rather than absolute carry changes might drive speculative trading. To test this possibility, we deduce the cross-sectional average among the 23 commodities from the basis change in Eq. (15) and rerun the augmented regression using this signal instead of the unadjusted basis change. This regression also does not indicate substantial trading on basis changes. Specifically, only 4 commodities are significant with the right sign, as can be seen in the fourth column of Table 6.[13] We conclude that basis and basis-momentum strategies play, if at all, only a minor role in practice.[14]

In addition to these carry-based signals, we also looked into risk-based signals and position adjustments (unreported) but we could not find robust results. The only signal we have found to improve time series momentum is cross-sectional momentum. Just

[13] In unreported analyses, we also calculated the cross-sectional average of the basis using weighted averages accounting for the varying number of contracts per year and hence the varying lag between the expiry of the first and the second futures. In addition, we examined lagged basis changes, and lagged relative basis changes. All these settings failed to predict position changes.

[14] Basu and Miffre (2013) find a link between the term structure and hedging pressure in a univariate framework. Our basis change signal is also significant in a univariate regression, confirming these findings.

Table 6
Other strategies.

		Base	Basis changes		Relative basis changes		Cross-sectional momentum	
		R^2	t-value	R^2	t-value	R^2	t-value	R^2
Energy	WTI	34.2	−4.00	36.9	−3.90	36.8	0.34	34.2
	Gasoline	28.9	2.03	29.2	1.97	29.2	3.14	29.2
	Heating oil	32.0	0.13	31.9	0.09	31.9	3.58	32.3
	Natural gas	31.8	−2.18	32.1	−1.79	31.9	2.79	31.9
	Brent	29.6	0.80	29.6	1.16	29.7	−0.38	29.6
	Gas oil	38.7	−0.63	38.6	−0.40	38.6	2.19	38.6
Grains	Corn	38.9	0.12	38.8	0.18	38.8	2.91	39.2
	Chicago wheat	41.0	−0.88	41.1	−0.06	40.9	5.66	42.7
	Kansas wheat	31.0	−0.78	30.9	0.35	30.9	2.79	31.1
	Soybean	36.0	0.39	35.9	0.06	35.9	4.18	36.8
	Soybean meal	46.2	0.77	46.2	0.62	46.2	4.23	47.0
	Soybean oil	41.6	1.45	41.5	0.46	41.5	3.00	41.8
Livestock	Lean hog	37.1	−0.61	37.1	−0.73	37.1	1.87	37.1
	Live cattle	27.9	−0.07	27.8	0.21	27.8	0.93	27.8
	Feeder cattle	27.0	1.73	27.1	2.63	27.5	3.26	27.6
Softs	Sugar	35.6	−0.53	35.5	−0.27	35.5	2.72	35.8
	Coffee	58.5	1.86	58.6	2.90	58.9	4.13	58.9
	Cotton	32.1	0.98	32.0	0.78	32.0	1.22	32.0
	Cocoa	40.1	0.75	40.0	0.78	40.0	3.12	40.4
Metals	Gold	32.0	1.61	32.2	0.83	32.0	2.15	32.2
	Silver	27.9	3.91	29.4	1.53	28.0	3.07	28.1
	Platinum	49.7	1.31	49.7	0.11	49.6	4.23	50.4
	Copper	33.1	2.88	33.8	2.57	33.6	6.35	35.6
Weighted average		36.1		36.3		36.2		36.5

Note: The table presents the adjusted in-sample R^2 in percent as well as the non-exact t-values (as described in Appendix C) for the additional factors. The additional factors are basis changes, relative basis changes and cross-sectional momentum. The results are based on the estimation of the speculative pressure index, where we regard the full sample of the large speculators.

as with the relative carry, we demeaned weekly return by the cross-sectional average and run another augmented regression using this signal, in addition to the time series momentum part. The last two columns of Table 6 report t-statistics of the relative momentum signal and R^2s. Nineteen display significant t-statistics with the right sign, many of them are highly significant. Brent crude oil is the only contract with the wrong sign. The increase in R^2, however, is moderate and less than 0.5% on average. We conclude that cross-sectional momentum plays a subordinate but significant role in practice.

5. The role of producers

5.1. Momentum trading of other groups

Not only speculators but also hedgers respond strongly to price changes as outlined in Cheng and Xiong (2014). They short more contracts when futures price rises and reduce their short position as futures price fall, or to put it differently, they pursue a reversal strategy. This observation is in fact just the other side of the coin from our finding that speculators are predominantly trend-followers and thus increase their long position when prices rise. By the adding up constraints and because non-reportables make up only a small fraction of traders, hedgers must follow strategies that mirror those of speculators. Using the disaggregated CoT report, we can further analyze which subgroup of traders is taking the other side of momentum traders. For this aim, we split commercial traders into the three subgroups of producers, users, and swap traders as described in Section 3.1. We also examine other reportables and non-reportables.

For each of these five subgroups, we predict position changes (number of contracts) and estimate the return signature plots in the same way as we did for speculators. The left-hand side of Table 7 shows the correlation between the momentum-based trading of speculators and each of the subgroups. The right-hand side of the table shows the out-of-sample predictability of our momentum-based forecast.

We find the highest overall R^2 among producers with an observation-weighted average of 24.7%. The energy sector is pulling this figure down significantly, but the results of the other sectors are consistently high and only in two cases below the overall average. The correlation is above 0.6 for all commodities, so producers are collectively reversal traders in all markets. For agricultural products and metals, the correlation is even above 0.9 without exception. This observation extends the findings of Cheng and Xiong (2014), that as the price rises, producers increase their short positions while as the price falls, they reduce their short position to a substantially larger set of commodities.

Fig. 6 further illustrates this relationship. It shows the return signature plots for producers and speculators at the last data point of our sample. Leaving aside the energy sector, the curves of the two groups have a very similar shape; in particular, the characteristic

Table 7
Correlation and R^2 for position changes in the number of contracts.

		Correlation					Number of contracts				
		Prod	User	Swap	Orep	Nrep	Prod	User	Swap	Orep	Nrep
Energy	WTI	0.60	−0.31	−0.88	−0.64	0.11	−0.5	0.5	27.7	6.6	21.4
	Gasoline	0.81	−0.53	−0.59	−0.67	0.90	1.8	0.8	0.7	22.6	28.3
	Heating oil	0.89	−0.80	−0.86	−0.70	0.88	9.4	1.2	1.1	11.4	29.7
	Natural gas	0.62	−0.27	−0.97	−0.59	0.67	0.3	−0.5	23.3	27.3	6.7
	Brent	0.79	−0.27	−0.25	−0.40	0.74	8.8	−2.8	3.6	−0.6	−3.0
	Gas oil	0.92	−0.48	−0.90	−0.64	0.91	3.3	−0.6	7.1	4.8	15.9
Grains	Corn	0.97	−0.77	0.38	−0.34	0.07	40.0	8.5	−1.4	−0.1	6.7
	Chicago wheat	0.95	−0.95	−0.19	−0.89	0.53	37.2	15.9	−2.2	18.6	0.1
	Kansas wheat	0.94	−0.91	0.61	0.31	0.47	33.8	8.8	6.6	−1.9	2.0
	Soybean	0.98	−0.90	0.68	−0.19	0.68	28.1	9.1	1.4	1.3	12.0
	Soybean meal	0.98	−0.92	0.20	0.22	0.90	33.6	13.5	−0.8	1.6	37.1
	Soybean oil	0.98	−0.95	0.38	−0.27	0.87	35.9	28.8	−1.2	0.0	33.3
Livestock	Lean hog	0.97	−0.73	0.61	−0.45	−0.04	42.0	0.8	8.1	−0.9	1.9
	Live cattle	0.95	−0.69	0.59	0.02	−0.78	28.5	2.3	4.2	−2.7	13.1
	Feeder cattle	0.92	−0.59	0.79	−0.52	−0.92	8.5	−1.2	0.4	1.0	19.2
Softs	Sugar	0.95	−0.87	−0.60	−0.53	0.86	23.8	8.0	2.3	5.1	36.9
	Coffee	0.95	−0.93	0.57	−0.54	0.40	29.4	18.6	0.0	3.1	1.5
	Cotton	0.98	−0.90	0.17	0.44	0.88	39.0	2.8	1.1	0.2	28.3
	Cocoa	0.94	−0.92	0.56	−0.30	0.63	25.9	17.0	0.4	1.7	17.6
Metals	Gold	0.95	−0.75	−0.98	−0.53	0.87	26.1	2.2	35.4	−0.1	16.8
	Silver	0.97	−0.77	−0.97	−0.88	0.64	28.5	3.2	32.8	12.5	8.5
	Platinum	0.99	−0.87	−0.97	−0.21	0.53	53.3	6.3	53.0	9.8	0.9
	Copper	0.96	−0.89	−0.86	−0.91	0.67	22.3	11.5	4.4	24.0	5.0
Weighted average		0.91	−0.75	−0.11	−0.40	0.48	24.7	7.2	8.4	6.4	15.4

Note: The table depicts on the left-hand side the correlations between the momentum-based trading of speculators and each of the subgroups (form left to right: Producers, user, swap dealers, other reportables, and non-reportables). The right-hand side of the table shows the out-of-sample predictability (R^2 in percent) of our momentum-based forecast with square root smoothing.

hump is also present for producers. For most commodities, the magnitude of the signal is stronger for speculators, but for lean hog, live cattle, and cotton they are virtually identical.

For users, the pattern is similar but less consistent and less pronounced overall. Their correlation is negative for all commodities, and the average absolute correlation remains quite high at −0.75. Like the producers, they are therefore reversal traders. But the R^2 values are much lower and average a mere 7.2%. Predictability is non-existent in the energy sector and negligible in the livestock sector.

Swap dealers show a more diverse behavior. For WTI, natural gas, and the three precious metals, they clearly take a counter-position to momentum traders. The other commodities show low predictability with often negative R^2s and negative correlation. This heterogeneity corresponds well with the fact that swap dealers trade on futures markets solely to hedge and manage risks arising from swap transactions. Their clients can be traders who trade for speculative or hedging purposes; therefore, identifying the clients' motivation for entering a swap contract is difficult (Bosch and Smimou, 2022). This heterogeneity of customers becomes apparent in our aggregate analysis.

Other reportables (i.e. small speculators) are only marginally predicted with an average R^2 of 6.4%. They tend to be reversal traders, which also supports the findings from Section 3, that the sub-groups of large speculators are more pronounced momentum traders than the parent class of non-commercials. Finally, the majority of non-reportables are moderate momentum traders with an average R^2 of 16.4% and an average correlation with speculative momentum of 0.49.

5.2. Counterparties of trend-followers

The analysis in the previous subsection shows that producers, and to a lesser extent consumers, are on the other side of the speculative momentum traders. Fig. 6 indicates an important role for producers but lacks quantification and attribution to the subgroups. To obtain a breakdown, using the adding up constraint, we decompose our position change forecast for non-commercials (mom_t) into five components

$$\Delta \text{mom}_t = \Delta \text{prod}_t - \Delta \text{user}_t - \Delta \text{swap}_t - \Delta \text{ospec}_t - \Delta \text{nrep}_t, \tag{16}$$

where prod, user, swap, ospec, and nrep represent the number of contracts held by producers, users, swap dealers, and other speculators (as the residual of our position change forecast) and non-reportables respectively. The producers enter with a positive sign because they hold short contracts. The variance of momentum trading can now be decomposed into five covariance terms:

$$\text{var}(\Delta \text{mom}) = \text{cov}(\Delta \text{mom}, \Delta \text{prod}) - \text{cov}(\Delta \text{mom}, \Delta \text{user}) - \text{cov}(\Delta \text{mom}, \Delta \text{swap}) - \text{cov}(\Delta \text{mom}, \Delta \text{ospec}) - \text{cov}(\Delta \text{mom}, \Delta \text{nrep}). \tag{17}$$

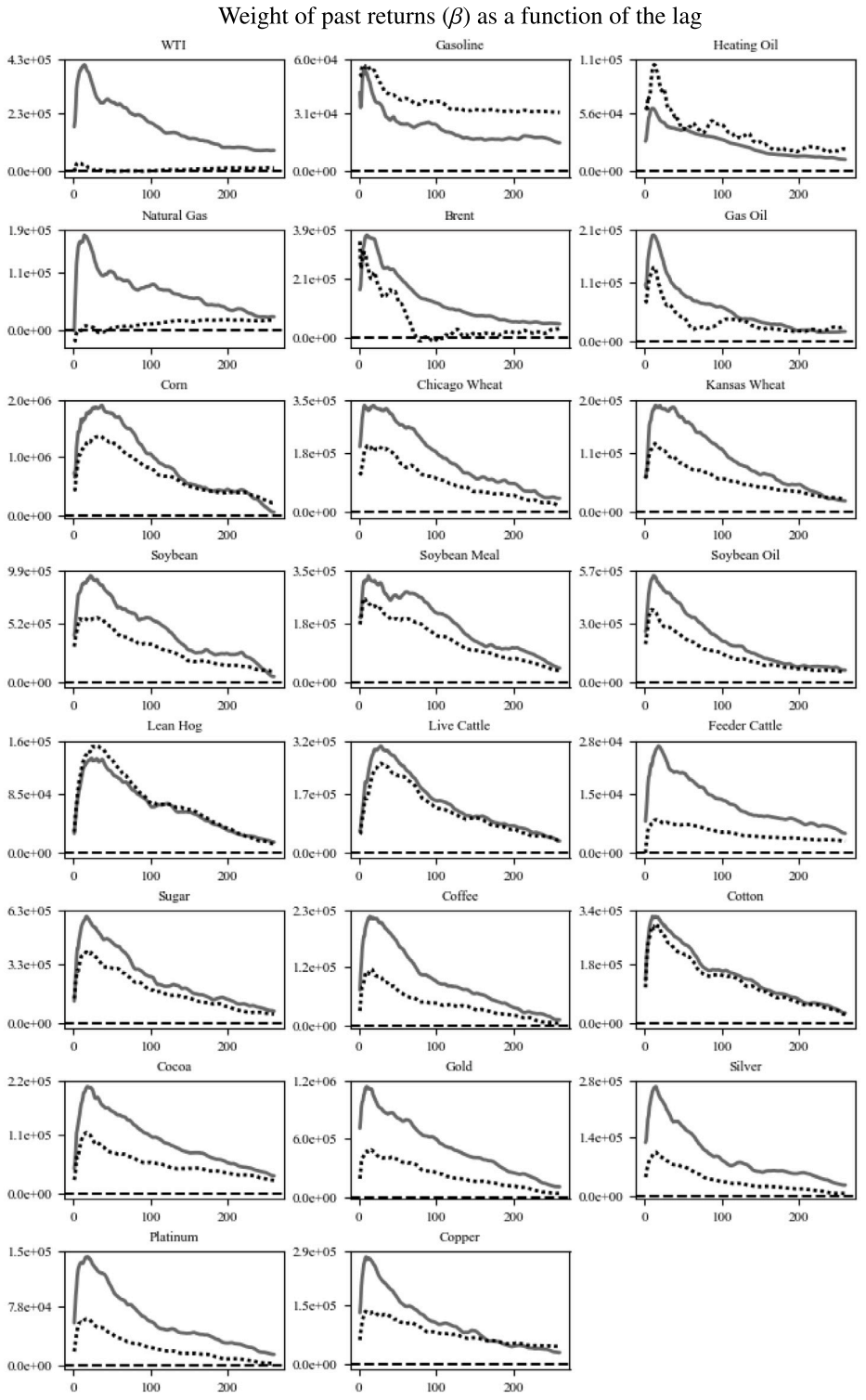

Fig. 6. Empirical return signature plots for non-commercials and producers.
Note: The graphs show the estimation of the βs in Eq. (8) for the number of contracts for producers in black (dotted) and non-commercials in gray. The plots are based on the last data point of our sample.

Table 8
Variance decomposition.

		On actual CFTC position changes					on fitted momentum trades				
		Prod	User	Swap	Ospec	Nrep	Prod	User	Swap	Nrep	ε
Energy	WTI	0.17	0.09	0.77	−0.17	0.15	0.13	0.05	0.76	0.09	−0.03
	Gasoline	0.70	0.45	0.11	0.22	−0.47	0.88	0.52	0.07	−0.49	0.02
	Heating oil	0.89	0.42	0.20	0.07	−0.58	0.78	0.35	0.21	−0.58	0.23
	Natural gas	0.27	0.16	1.03	−0.30	−0.16	0.22	0.05	0.84	−0.19	0.08
	Brent	1.55	0.23	0.20	−0.89	−0.09	0.90	0.09	0.08	−0.10	0.02
	Gas oil	1.15	0.06	0.41	−0.40	−0.22	0.68	0.14	0.18	−0.22	0.23
Grains	Corn	0.99	0.25	−0.08	−0.24	0.07	0.73	0.18	−0.02	0.04	0.07
	Chicago wheat	0.79	0.42	−0.02	−0.17	−0.02	0.66	0.36	0.02	−0.06	0.02
	Kansas wheat	0.85	0.34	−0.16	−0.13	0.09	0.75	0.31	−0.10	−0.02	0.06
	Soybean	1.09	0.49	−0.15	−0.34	−0.08	0.83	0.35	−0.08	−0.11	0.02
	Soybean meal	1.17	0.51	−0.03	−0.29	−0.37	0.95	0.35	−0.01	−0.34	0.05
	Soybean oil	0.97	0.58	−0.08	−0.23	−0.25	0.77	0.48	−0.02	−0.25	0.01
Livestock	Lean hog	1.09	0.10	−0.19	0.00	0.00	0.98	0.07	−0.08	0.01	0.02
	Live cattle	0.79	0.17	−0.11	−0.08	0.22	0.67	0.18	−0.04	0.15	0.04
	Feeder cattle	0.48	0.09	−0.13	−0.03	0.60	0.43	0.08	−0.09	0.52	0.06
Softs	Sugar	1.02	0.50	0.11	−0.24	−0.39	0.82	0.30	0.07	−0.32	0.13
	Coffee	0.83	0.47	−0.04	−0.25	−0.01	0.62	0.38	−0.03	−0.01	0.03
	Cotton	1.07	0.17	−0.04	−0.02	−0.19	0.99	0.18	−0.01	−0.20	0.04
	Cocoa	0.79	0.43	−0.07	−0.04	−0.11	0.72	0.38	−0.05	−0.14	0.10
Metals	Gold	0.62	0.12	0.68	−0.25	−0.17	0.50	0.07	0.53	−0.16	0.05
	Silver	0.57	0.16	0.63	−0.25	−0.11	0.47	0.06	0.57	−0.13	0.03
	Platinum	0.71	0.11	0.85	−0.63	−0.04	0.45	0.07	0.53	−0.08	0.02
	Copper	0.70	0.50	0.15	−0.33	−0.02	0.59	0.26	0.14	−0.02	0.02
Weighted average		0.81	0.31	0.16	−0.19	−0.08	0.67	0.24	0.15	−0.11	0.06

Note: The table displays the results of the variance decomposition, outlined in Section 5.2. The left-hand side depicts the results of the variance decomposition based on the actual CFTC data using Eqs. (16) and (17). The right-hand side presents the results based on the fitted momentum trades from Eq. (18). The analysis is based on the position changes in the number of contracts over the entire out-of-sample period.

Dividing both sides by $\mathrm{var}(\Delta\,\mathrm{mom})$ yields a variance decomposition. The results of this decomposition are shown in Table 8. For all but six commodities, the covariance with producers is the largest contributor. The exceptions are WTI, natural gas, and precious metals, where swap traders deliver the highest contribution, as well as for feeder cattle where non-reportables play the most important role. These exceptions are the same as in the forecast and correlation analysis in Section 5.1. Overall, producers contribute 81% of the variance of speculative momentum trading. The contribution of users is also consistently positive and averages 31%.

Another way to decompose speculative momentum trades is to use the momentum (reversal trades) of other subgroups (marked with a leading m for each subgroup):

$$\Delta\,\mathrm{mom}_t = \Delta\,\mathrm{mprod}_t - \Delta\,\mathrm{muser}_t - \Delta\,\mathrm{mswap}_t - \Delta\,\mathrm{mnrep}_t - \rho_t. \tag{18}$$

We use our out-of-sample forecast, which implies that the first four terms do not add up to the left-hand side as they would with an unregularized in-sample OLS. The last term ρ is a residual, which is chosen in such a way that the equation holds. This decomposition delivers very similar results as the previous decomposition: Producers dominate with an average contribution of 67%, and the contribution of users remains consistent but smaller (24%). Overall, both variance decompositions in this section and the analysis of the previous section clearly indicate that producers are the dominant counterparty of speculative momentum traders.

5.3. Trading motives of commercials

The aim of this paper is to reveal trading patterns and describe them quantitatively. The methods we use are very well suited for this purpose but cannot provide information about the underlying reasons for the trading activity. While the speculators' rationale is quite clear, they replicate a known anomaly to make a profit, the motivation of the producers is not as evident. Many motives are conceivable and are also described elsewhere [see Cheng and Xiong (2014) and the references therein]. Reversal trading, however, is rarely justified by simple profit-taking, although this would fit well with the behavioral explanations for the momentum effect. For instance, producers could suffer from a displacement effect as in Barberis et al. (2001) and thus realize profits too early or they could anchor prices to previous levels as in Barberis et al. (1998) or to production cost.

There are also risk-based explanations of profit-taking. If prices soar, producers can end up with a very large profit in dollar terms resulting from unhedged exposure to the commodity they produce. It is hardly rational to hold this concentrated position solely in a single commodity, rather it is advisable from a risk point of view to reduce or diversify it. Moreover, according to the theory of storage, higher prices are associated with lower inventories and thus higher volatility (e.g., Symeonidis et al., 2012; Geman and Smith, 2013; Basu and Miffre, 2013). Then, if producers increase their hedging with the risk, they also increase it with prices and thus act as reversal traders.

Finally, one can reflect on the subordinate role of users relative to producers. This difference can be explained by the fact that producers usually run a business that is concentrated on a single or a few similar raw materials. With users, on the other hand, a single commodity is usually just one input factor among many. As a result, the commodity risk of users is better diversified per se and therefore less risky. Accordingly, the need for hedging is also lower. This echoes Keynes' (1930) and Hicks' (1939) theory of hedging pressure and normal backwardation, which also assigns a dominant role to producers.

6. Conclusion

In this paper, we demonstrate that trend-following and momentum trading are indeed the predominant investment strategies of speculative traders in the commodity futures market. We fit typical signals and deduce strong out-of-sample forecasts of position changes and hence the flow generated by speculators. While these results are interesting per se, our flow analysis can also be used for further research. It can, for example, be used to assess whether the popularity of trend-following leads to capacity constraints and thus to potential inefficiencies (Baltas and Kosowski, 2013). A related issue is the concern that excessive and unidirectional speculative trading in food commodities can force prices up or have destabilizing effects with adverse effects on the real economy [see Haase et al. (2016) for a literature review on the topic]. A better understanding of the flow pattern could add another perspective to this strand of literature.

If momentum trading is indeed at the capacity limit or has destabilizing effects, our results can also be useful to optimize the execution of trend strategies. An even more detailed understanding of speculative trading patterns can be used to refine the understanding of the liquidity needs of speculators. This knowledge could then be used to derive a liquidity providing strategy along the lines of Kang et al. (2020) but with more accurate and real-time tracking of trend-followers.

We also point out an interesting interplay between momentum traders and producers. However, this analysis is purely empirical and lacks an economic rationale. A utility-based framework that rationalizes producer behavior would be a natural continuation of our research.

Our model might be extended to other futures markets traded by CTAs. With a complete universe, it can also be used to replicate the time series and return characteristics of CTAs, yielding another empirical examination of the return signature plots derived in this paper.

Finally, the paper also offers an innovative regularization method through the generalized ridge regression and the in-sample tuning of the hyperparameters using LOOCV, and it demonstrates the usefulness of uneven smoothing along the ordered features. Our choice of the weighting function, however, is rather ad hoc and a more stringent choice of this function might be a topic of further research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Jan-Alexander Posth for pointing us to the Levine and Pedersen (2016) article, Valerio Poti, Peter Schwendner, the co-editor Paolo Pasquariello, and two referees for their helpful comments and suggestions, as well as seminar participants at Zurich University of Applied Sciences.

References:

– Levine, A., Pedersen, L.H., 2016. Which trend is your friend? Financ. Anal. J. 72 (3), 51-66.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Return signal plots

A.1. Moving average

The moving average signal is defined as the current price minus the average price over the look-back period:

$$ma_t = p_t - \frac{1}{n}\sum_{i=1}^{n} p_{t-i} = \frac{1}{n}\sum_{i=1}^{n} p_t - p_{t-i} = \frac{1}{n}\sum_{i=1}^{n} r_{t,t-i} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=0}^{i-1} r_{t-j} = \frac{1}{n}\sum_{i=0}^{n}(n-i)r_{t-i},$$

so, the triangle plotted in the right graph of Fig. 1.

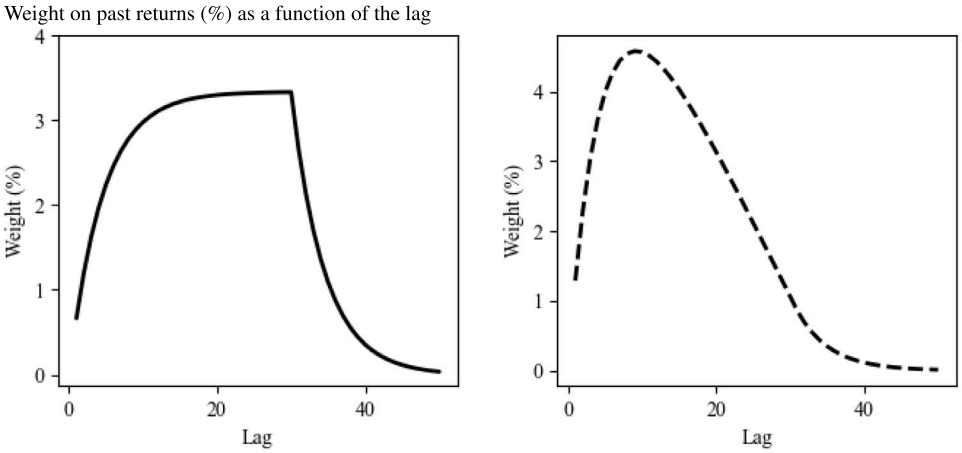

Fig. 7. Return signature plots for partially rebalanced signals.
Note: Weights of the partially rebalanced momentum signal (left) and the partially rebalanced moving average signals (right). Weights are normalized to sum to 1.

A.2. Partial rebalancing

Fig. 1 depicts a very simple rebalancing rule that averages the momentum signal over five days. Another rebalancing rule is that a fixed proportion α of the difference between the model-implied weight (so the signal) and the current weight is traded. These rules are path-dependent but reach a steady state when traded over a longer period. For a momentum signal with a look-back period of n $\beta_i = 1/n$ for all $i < n$, so for the partial rebalancing rule the loading on the new return is $\beta_0^\alpha = \alpha/n$. In the steady state, the weight of the one-day lagged return must be equal to α times $(1-\alpha)$ times the previous β_0^α and so on:

$$\begin{aligned}\beta_1^\alpha &= \alpha\beta_1 + (1-\alpha)\beta_0^\alpha\\ \beta_2^\alpha &= \alpha\beta_2 + (1-\alpha)\beta_1^\alpha\\ &\vdots\end{aligned}$$

This can be solved recursively. The same can be done for any linear trend signal. Fig. 7 shows the resulting return signature plots for mom_{30} and a ma_{30} as a basis signal and $alpha = 0.2$. The momentum plot increases until the end of the look-back period but still has a sharp brink at the end of the look-back period. The *ma* shows a nice hump shape form similar to the *xmac* and to what can be observed in the data.

Appendix B. Details on the estimation procedure

As in the ridge case, the parameters can be estimated with OLS by augmenting the data with pseudo-observation reflecting the regularization terms. The augmented model can be visualized as follows:

$$\begin{bmatrix}\Delta c_1\\ \Delta c_6\\ \vdots\\ \Delta c_T\\ 0\\ 0\\ \vdots\\ 0\\ 0\end{bmatrix} = \begin{bmatrix}\Delta r_1 & \Delta r_0 & \Delta r_{-1} & \dots & \Delta r_{2-n}\\ \Delta r_6 & \Delta r_5 & \Delta r_4 & \dots & \Delta r_{7-n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ \Delta r_T & \Delta r_{T-1} & \Delta r_{T-2} & \dots & \Delta r_{T+2-n}\\ \gamma_1 & -\gamma_1 & 0 & \dots & 0\\ 0 & \gamma_2 & -\gamma_2 & \dots & 0\\ \dots & \dots & \ddots & \ddots & \vdots\\ 0 & 0 & \dots & \gamma_{n-2} & -\gamma_{n-2}\\ 0 & 0 & \dots & 0 & \gamma_{n-1}\end{bmatrix}\begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_{n-1}\end{bmatrix} + \begin{bmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_T\\ \eta_1\\ \vdots\\ \eta_{n-1}\end{bmatrix} \tag{B.1}$$

If $\gamma_n = 0$, the last line can be skipped.[15] Matrix notation is useful and we write:

$$\begin{bmatrix}\Delta C\\ 0\end{bmatrix} = \begin{bmatrix}X\\ \Gamma\end{bmatrix} B + \begin{bmatrix}E\\ H\end{bmatrix} \tag{B.2}$$

with the self-explanatory definition of the matrices. Finally, we define the hat matrix as:

$$A(\Gamma) = X(X^T X + \Gamma^T \Gamma)^{-1} X^T. \tag{B.3}$$

[15] We experimented with shrinking the last observation to zero, i.e., using $\gamma_{n-1} > 0$. This specification did not improve the forecasting power.

Eqs. (B.1) and (B.2) remain valid if we enlarge the set of predictive variables as in Section 4.

B.1. Hyperparameter tuning

Following Allen (1974), a set of hyperparameters Γ can be efficiently determined using leave-one-out cross-validation (LOOCV). This is done by estimating Eq. (5) with all but the τth observation. This yields estimated β_is denoted as $\hat{\beta}_{i,-\tau}(\Gamma)$ and an out-of-sample forecast of the position changes:

$$\widetilde{\Delta c}_\tau(\Gamma) = \sum_{i=0}^{n-1} \hat{\beta}_{i,-\tau}(\Gamma)\Delta r_{\tau-i}. \tag{B.4}$$

Allen's PRESS (prediction sum of squares) statistics is the average mean squared error over all observations:

$$P(\Gamma) = \frac{1}{T}\sum_{\tau=1}^{T}(\widetilde{\Delta c}_\tau(\Gamma) - \Delta c_\tau)^2. \tag{B.5}$$

Using the Sherman–Morrison–Woodbury formula, it can be shown that:

$$P(\Gamma) = \frac{1}{T}\|B(\Gamma)(I - A(\Gamma))\Delta C\| \tag{B.6}$$

where $B(\Gamma)$ is a diagonal matrix with jjth entry $1/(1 - a_{jj}(\Gamma))$, a_{jj} being the jjth entry of $A(\Gamma)$ and $\|\cdot\|$ is the Euclidean norm. See van Wieringen (2020) and Golub et al. (1979) for further details. The optimal set of hyperparameters is then the minimizer of the PRESS statistics.

In our base model with all γ_i being equal, there is only one hyperparameter to optimize, and no specification mentioned in this paper has more than two free hyperparameters. Using (B.6), the optimal parameters can efficiently be estimated numerically. Note that the hyperparameters are tuned by the data. Therefore, beyond the shape of the γ-curve, there is no discretion in their choice.

Appendix C. In-sample t -test

Our in-sample test follows Cule et al. (2011). Using notation from Eq. (B.2), the covariance matrix of the regression coefficients resulting from a ridge regression is given by:

$$\mathrm{var}(B) = \sigma^2(X^TX + \Gamma^T\Gamma)^{-1}X^TX(X^TX + \Gamma^T\Gamma)^{-1}, \tag{C.1}$$

e.g., van Wieringen (2020). In practice, σ^2 is replaced by its estimate, given by the residual mean square of the ridge model:

$$\sigma^2 = \frac{(\Delta C - XB)^T(\Delta C - XB)}{\nu}, \tag{C.2}$$

where ν is the effective degree of freedom defined as:

$$\nu = T - \mathrm{tr}\,(2A - AA^T). \tag{C.3}$$

Standard errors of B ($\mathrm{se}(\beta_i)$) are the square root of the diagonal of the covariance matrix in Eq. (C.1) and the test statistics are given as $\beta_i/\mathrm{se}(\beta_i)$. Halawa and Bassiouni (2000) refer to this testing procedure as a "non-exact" t-type test because it ignores the bias of the ridge regression. Their simulation study, however, shows that the procedure works well when ridge is a desired estimator. Then, it outperforms the standard OLS-based t-test in most cases.

References

Allen, D.M., 1974. The relationship between variable selection and data augmentation and a method for prediction. Technometrics 16 (1).

Bakshi, G., Gao, X., Rossi, A.G., 2019. Understanding the sources of risk underlying the cross section of commodity returns. Manage. Sci. 65 (2), 619–641.

Baltas, A.-N., Kosowski, R., 2013. Momentum strategies in futures markets and trend-following funds. SSRN Available at: https://ssrn.com/abstract=1968996.

Barberis, N., Huang, M., Santos, T., 2001. Prospect theory and asset prices. Q. J. Econ. 116 (1), 1–53.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment1. J. Financ. Econ. 49 (3), 307–343.

Basu, D., Miffre, J., 2013. Capturing the risk premium of commodity futures: The role of hedging pressure. J. Bank. Financ. 37 (7), 2652–2664.

Boons, M., Prado, M.P., 2019. Basis-momentum. J. Finance 74 (1), 239–279.

Bosch, D., Smimou, K., 2022. Traders' motivation and hedging pressure in commodity futures markets. Res. Int. Bus. Finance 59, 101529.

CFTC, 2021. Disaggregated Commitments of Traders Report. Explanatory Notes, Commodity Futures Trading Commission, n.d. https://cftc.gov/idc/groups/public/@commitmentsoftraders/documents/file/disaggregatedcotexplanatorynot.pdf, Last accessed on 2021-04-22.

Cheng, I.-H., Kirilenko, A., Xiong, W., 2015. Convective risk flows in commodity futures markets. Rev. Financ. 19 (5), 1733–1781.

Cheng, I.-H., Xiong, W., 2014. Why do hedgers trade so much. J. Legal Stud. 43 (2S), 183–207.

Cule, E., Vineis, P., De Iorio, M., 2011. Significance testing in ridge regression for genetic data. BMC Bioinformatics 12, 372.

Davis, J., Dorsten, M., Gillmann, N., Tsai, J., 2022. Carry momentum. Financ. Anal. J. 78 (1), 5–38.

Dewally, M., Ederington, L.H., Fernando, C.S., 2013. Determinants of trader profits in commodity futures markets. Rev. Financ. Stud. 26 (10), 2648–2683.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. J. Bus. Econom. Statist. 13, 253–263.

Elaut, G., Erdős, P., 2018. Trends' signal strength and the performance of CTAs. Financ. Anal. J. 7.

Erb, C.B., Harvey, C.R., 2006. The strategic and tactical value of commodity futures. Financ. Anal. J. 62 (2), 69–97.

Fan, J.H., Fernandez-Perez, A., Fuertes, A.-M., Miffre, J., 2020. Speculative pressure. J. Futures Mark. 40 (4), 575–597.

Fuertes, A.-M., Miffre, J., Rallis, G., 2010. Tactical allocation in commodity futures markets: Combining momentum and term structure signals. J. Bank. Financ. 34 (10), 2530–2548.
Geman, H., Smith, W.O., 2013. Theory of storage, inventory and volatility in the LME base metals. Resour. Policy 38 (1), 18–28.
Golub, G.H., Heath, M., Wahba, G., 1979. Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics 21 (2).
Gorton, G.B., Hayashi, F., Rouwenhorst, K.G., 2012. The fundamentals of commodity futures returns. Rev. Financ. 17 (1), 35–105.
Gorton, G., Rouwenhorst, K.G., 2006. Facts and fantasies about commodity futures. Financ. Anal. J. 62 (2), 47–68.
Granger, C.W.J., Hyung, N., Jeon, Y., 2001. Spurious regressions with stationary series. Appl. Econ. 33.
Haase, M., Seiler Zimmermann, Y., Zimmermann, H., 2016. The impact of speculation on commodity futures markets – A review of the findings of 100 empirical studies. J. Commod. Mark. 3 (1), 1–15.
Halawa, A., Bassiouni, M.E., 2000. Tests of regression coefficients under ridge regression models. J. Stat. Comput. Simul. 65 (1–4), 341–356.
Hicks, J., 1939. Value and Capital, second ed. Clarendon Press, Oxford.
Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 42 (1).
Hong, H., Stein, J.C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. J. Finance 54 (6), 2143–2184.
Hong, H., Yogo, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices? J. Financ. Econ. 105 (3), 473–490.
Hurst, B., Ooi, Y.H., Pedersen, L.H., 2013. Demystifiying managed futures. J. Invest. Manage. 11 (3), 42–58.
ICE, 2021. ICE Futures Europe Commitments of Traders Report. Explanatory Notes, Intercontinental Exchange, Inc., n.d. https://www.theice.com/publicdocs/futures/CoT_Notes.pdf, Last accessed on 2021-04-22.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. J. Finance 48 (1), 65–91.
Kang, W., Rouwenhorst, K.G., Tang, K., 2020. A tale of two premiums: The role of hedgers and speculators in commodity futures markets. J. Finance 75 (1), 377–417.
Keynes, J.M., 1930. A Treatise on Money, Vol. II. Macmillan, London.
Koijen, R.S., Moskowitz, T.J., Pedersen, L.H., Vrugt, E.B., 2018. Carry. J. Financ. Econ. 127 (2), 197–225.
Lehecka, G.V., 2013. Hedging and speculative pressures: An investigation of the relationships among trading positions and prices in commodity futures markets. In: Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. St. Louis, MO.
Levine, A., Pedersen, L.H., 2016. Which trend is your friend? Financ. Anal. J. 72 (3), 51–66.
Li, Y., Gui, Q., Gu, Y., Han, S., Du, K., 2014. Ridge-type Kalman filter and its algorithm. WSEAS Trans. Math. 13, 852–862.
MacKinnon, J.G., White, H., 1985. Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties. J. Econometrics 29, 305–325.
Miffre, J., Rallis, G., 2007. Momentum strategies in commodity futures markets. J. Bank. Financ. 31 (6), 1863–1886.
Mincer, J., Zarnowitz, V., 1969. The evaluation of economic forecasts. In: Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance. National Bureau of Economic Research, Inc, pp. 3–46, URL: https://EconPapers.repec.org/RePEc:nbr:nberch:1214.
Moskowitz, T.J., Ooi, Y.H., Pedersen, L.H., 2012. Time series momentum. J. Financ. Econ. 104 (2), 228–250.
Phillips, D.L., 1962. A technique for the numerical solution of certain integral equations of the first kind. J. ACM 9 (1), 84–97.
Schmid, O., Wirth, P., 2021. Optimal allocation on time series and cross sectional momentum. J. Portf. Manage. Multi Asset Spec. Issue.
Symeonidis, L., Prokopczuk, M., Brooks, C., Lazar, E., 2012. Futures basis, inventory and commodity price volatility: An empirical analysis. Econ. Model. 29 (6), 2651–2663.
Szakmary, A.C., Shen, Q., Sharma, S.C., 2010. Trend-following trading strategies in commodity futures: A re-examination. J. Bank. Financ. 34 (2), 409–426.
Szymanowsky, M., De Roon, F., Nijman, T., Van Den Goorbergh, R., 2014. An anatomy of commodity futures risk premia. J. Finance 69 (1), 453–482.
Tibshirani, R., Michael, S., Saharon, R., Ji, Z., Keith, K., 2005. Sparsity and smoothness via the fused lasso. J. R. Stat. Soc. Ser. B Stat. Methodol. 67, 91–108.
Tikhonov, A.N., 1963. Solution of incorrectly formulated problems and the regularization method. Sov. Math. 1035—1038.
van Wieringen, W.N., 2020. Lecture notes on ridge regression.