Fear of Hazards in Commodity Futures Markets

Fernandez-Perez, Adrian and Fuertes, Ana-Maria and Gonzalez-Fernandez, Marcos and Miffre, Joelle

Auckland University of Technology, New Zealand, Cass Business School, City University of London, UK, University of Leon, Spain, Audencia Business School, Nantes, France

27 September 2019

Fear of Hazards in Commodity Futures Markets

ADRIAN FERNANDEZ-PEREZ[†], ANA-MARIA FUERTES[‡], MARCOS GONZALEZ-FERNANDEZ[§], JOELLE MIFFRE[¶]

ABSTRACT

We examine the commodity futures pricing role of active attention to weather, disease, geopolitical or economic threats or "hazard fear" as proxied by the volume of internet searches by 149 query terms. A long-short portfolio strategy that sorts the cross-section of commodity futures contracts according to a hazard fear signal captures a significant premium. This commodity hazard fear premium reflects compensation for extant fundamental, tail, volatility and liquidity risks factors, but it is not subsumed by them. Exposure to hazard-fear is strongly priced in the cross-section of commodity portfolios. The hazard fear premium exacerbates during periods of adverse sentiment or pessimism in financial markets.

Keywords: Commodity futures; Fear; Attention; Hazards; Internet searches; Sentiment; Long-short portfolios.

This version: May 6, 2020

---

† Research Fellow, Auckland University of Technology, Private Bag 92006, 1142 Auckland, New Zealand. Phone: +64 99219999; Fax: +64 9 9219940; E-mail: adrian.fernandez@aut.ac.nz

‡ Professor of Finance and Econometrics, Cass Business School, City University, London, ECIY 8TZ, England; Tel: +44 (0)20 7040 0186; E-mail: a.fuertes@city.ac.uk. Corresponding author.

§ Assistant Professor, University of León, León, 24071, Spain; Tel: +34 987 293498; E-mail: mgonf@unileon.es.

¶ Professor of Finance, Audencia Business School, 8 Route de la Jonelière, 44312, Nantes, France; Tel: +33(0) 240 37 34 34; E-mail: jmiffre@audencia.com.

This research was partially conducted while the third author was visiting Cass Business School supported financially by grants from Fundación Monteleón and Erasmus+ KA103. We thank the Editor, Carol Alexander, and two anonymous referees for helpful comments. The paper benefitted from discussions with Jerry Coakley, Isabel Figuerola-Ferretti, Christian Juilliard, Lola Robles, Simon Sosvilla-Rivero, Hilary Till, Adam Zaremba, and seminar participants at Aston Business School Birmingham, Eberhard-Karls-Universität Tübingen, University of Kent at Canterbury, Universidad Complutense de Madrid, and participants at Center for Empirical Finance 2019 Workshop, Brunel University, London; INFINITI 2019 conference, Adam Smith Business School, Glasgow; and Spanish Finance Forum 2019, Universidad Carlos III, Madrid; New Zealand Finance Colloquium 2019, Auckland; and Computational and Financial Econometrics Conference 2019, Senate House, London.

"Data are widely available, what is scarce is the ability to extract wisdom from them" (Hal Varian, Google Chief Economist, emeritus Professor at University of California, Berkeley.)

1. INTRODUCTION

THE COMMODITY FUTURES PRICING literature largely rests on two pillars known as the theory of storage (Kaldor, 1939; Working, 1949; Brennan, 1958) and the hedging pressure hypothesis (Cootner, 1960; Hirshleifer, 1988). The former pillar argues that the dynamics of commodity futures prices is primarily driven by inventory levels proxied by the slope of the futures curve, while the latter pillar contends that the primary determinant of commodity futures prices are hedgers' net positions. In support of these theories, a number of studies suggest that a premium can be extracted by taking long positions in backwardated futures markets and short positions in contangoed futures markets.[1] More recently, the literature has considered alternative commodity characteristics such as liquidity (Szymanowska et al., 2014), skewness (Fernandez-Perez et al., 2018), basis-momentum (Boons and Prado, 2019) or convexity (Gu et al., 2019) and has shown that they also have predictive power over commodity futures returns.

Our article hypothesizes that "fear" of rare and extreme events influences the pricing of commodity futures contracts over and beyond the factors that have been shown to price commodities. In this paper, the terminology commodity hazard fear is broadly defined as the economic agents' apprehension or concerns about potential weather, agricultural disease, geopolitical and economic events that may shift the commodity supply or demand curves. Fear can be considered as one of a set of basic or innate human emotions that is not necessarily linked to irrationality. Since fear is modulated by the process of cognition and learning, it can

[1] Rising commodity futures prices are predicted by the backwardation state as signalled by scarce inventories (Gorton et al., 2012), a downward-sloping term structure of futures prices (Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006; Szymanowska et al., 2014; Bakshi et al., 2019), net short hedging, net long speculation (Bessembinder, 1992; Basu and Miffre, 2013; Kang et al., 2019) or superior strong past performance (Erb and Harvey, 2006; Miffre and Rallis, 2007; Bakshi et al., 2019). Conversely, falling commodity futures prices are predicted by the contango state as signalled by the same characteristics at the other end of the spectrum.

thus be deemed as rational or appropriate – the fear of losing money can rationally cause agents to manage their risks actively (Lo, 2011).[2] For instance, if a storm is approaching, for as long as there is some uncertainty regarding its impact on the supply of a commodity, fear of the storm can be considered as a rational response of commodity traders to the threat. Likewise, the recollection of extreme weather that destroyed the coffee harvest in the past may trigger fear in the run-up to the current harvest season since early experiences also shape the fear system (Tottenham, 2014). While being agnostic on whether the hazard fear is purely rational or contains elements of irrationality (i.e., "excessive" fear), we hypothesize that hazard fear can affect commodity futures prices above and beyond fundamentals.

Let us first consider hazards that are supply-reducing (e.g., a frost that is likely to shift inwards the coffee supply curve) or demand-increasing (e.g., a heatwave that is likely to shift outwards the natural gas demand curve). Fear of these hazards induces expectations of a sharp rise in spot prices. We hypothesize that these expectations, in turn, influence the hedging decisions of commodity market participants; namely, producers reduce their short hedges and consumers increase their long hedges compared to the hedging strategy that they adopt in the absence of hazard-fear. The resulting increase in net long hedging ought to be matched by an increase in net short speculation, but the later may be deterred by the fact that short futures positions are seen as especially risky for speculators in a commodity market bedevilled by supply-reducing or demand-increasing hazard fears.[3] Thus, to entice short speculation the

[2] There is a large literature in psychology on whether fear is rational or irrational. A widely-held view is that an irrational fear is an emotion associated with an event or situation that an individual seeks to avoid, even though it is extremely unlikely and/or inconsequential.

[3] J.P. Morgan's Global Commodities Research (22 Sept 2017) commentary: "Non-commercial investors have been reducing their net short position across the agri commodity complex over the last fortnight amid these weather-related production risks […] We anticipate that non-commercial's will continue the wave of short covering through September, now that La Niña is a material threat, and oil prices are on the rise. This is particularly the case across markets

current price of the futures contract (relative to the expected future spot price) ought to be set higher than it would be if only fundamental forces were at play. Formally, the expected fear premium is the upward bias in the futures price as predictor of the future spot price (or mispricing) relative to what the futures price would be in the absence of any hazard fear. More explicitly, the overall commodity futures premium induced by the hazard fear can be simply formalized as $E_t\left[Premium_{t,T}\right] \equiv F_{t,T}^{CFEAR} - E_t[S_{t+T}] > 0$ with $F_{t,T}^{CFEAR} = F_{t,T} + Premium_{t,T}^{CFEAR+}$, where $F_{t,T}$ denotes the fundamental price at *t* of a futures contract with maturity T in the absence of fear and $Premium_{t,T}^{CFEAR+} > 0$ denotes the hazard-fear induced upward shift in the current futures price required to attract net short speculation. Thus, the anticipated decrease in the futures price as maturity approaches is the overall premium captured by short speculators which incorporates both a fundamental and a hazard-fear component.

Let us next consider a hazard that is either supply-increasing (e.g., a lift of an oil embargo that is likely to shift outward the oil supply curve) or demand-reducing (e.g., an economic recession that shrinks the demand for commodities). Fear of these hazards causes expectations of spot prices sharply decreasing, and producers (consumers) may then take shorter (less long) hedging positions than they would otherwise. The increase in net short hedging requires a matching increase in net long speculation. In order to induce speculators to take more long positions in this setting, the futures price ought to be lower than it would be in absence of the hazard-fear; formally, $E_t\left[Premium_{t,T}\right] \equiv F_{t,T}^{CFEAR} - E_t[S_{t+T}] < 0$ with $F_{t,T}^{CFEAR} = F_{t,T} + Premium_{t,T}^{CFEAR-}$ and $Premium_{t,T}^{CFEAR-} < 0$ is the premium induced by the supply-increasing or demand-reducing hazard fear mispricing. The rise in the futures price as maturity approaches (premium earned by long speculators) thus incorporates both a fundamental and a fear element.

---

with exposure to summer crop production in Latin America, namely CBOT Soybeans, CBOT Corn, ICE #11 Sugar and also ICE Arabica Coffee".

Building on economic psychology, we hypothesize that economic agents' fear of threats induces them to search for information (Lemieux and Peterson, 2011). This active information demand is referred to as "attention" in the recent asset pricing literature (Da et al., 2011, 2015; Han et al., 2017a, 2017b; Vozlyublennaia, 2014).[4] Motivated by this literature, we employ as proxy for attention to hazards the volume of Google search queries by keywords representing 149 hazards in the weather, agricultural disease, geopolitical and economic categories. Thus, upsurges in the search queries can signal hazard fear. We conjecture that this fear can temporarily deviate the futures price above or below its fundamental value depending on whether the underlying hazard shifts the supply and demand curves inward or outward.

Economic agents' fear can occur for many reasons. Building on the aforementioned literature on the pricing content of "attention" we are agnostic as to whether the internet searches are induced by news releases about impending hazards or simply by a phenomenon akin to the "representativeness" heuristic – when people witness a salient event their level of fear can increase independently of any economic loss they incur. For instance, a coffee producer may be anxious about the possibility of a severe frost pre-harvest because her crops were affected by such a frost in the past or because she is mindful of other extreme weather phenomena that had dramatically shifted inward the commodity supply curve.[5]

Following the above intuition, the paper contributions are threefold. Using the changes in internet search volume by 149 commodity-hazard keywords as proxy for fear surges, we adapt

[4] There is a parallel literature, largely initiated by Tetlock (2007), which establishes instead that variables related to the information supply such as the media count (number of news articles published) or the media tone (positive or negative articles) can influence asset prices.

[5] The representativeness heuristic was first described by psychologists Amos Tversky and Daniel Kahneman during the 1970s as a mental shortcut by which agents estimate the likelihood of an event by comparing it to an existing prototype that already exists in their minds (Kahneman and Tversky, 1979). When agents act on the basis of representativeness, they are more likely to make more errors by overestimating the likelihood that something will occur.

the setting of Da et al. (2015) to obtain a signal for each commodity futures (hereafter CFEAR) that reflects the nexus between past returns and hazard fear. Second, we deploy a novel CFEAR portfolio strategy that sells the commodities that appreciated the most under the influence of supply-decreasing or demand-increasing fears and buys the commodities that depreciated the most under the influence of supply-increasing or demand-decreasing fears. We formally assess the out-of-sample performance of the CFEAR portfolio and deploy time-series spanning tests to test whether the fear premium thus captured is subsumed by known systematic risk factors. Third, contributing to the commodity pricing literature, we deploy cross-sectional tests for commodity portfolios (sorted on characteristics and sectors) and individual commodities to test whether the CFEAR factor has any pricing ability beyond known systematic risk factors.

We find that the long-short CFEAR portfolio captures an economically and statistically significant mean excess return of 9.28% per annum (t = 3.35). This sizeable CFEAR premium translates into a Sharpe ratio of 0.90 that is very attractive compared to the Sharpe ratios of extant long-short commodity strategies. The CFEAR premium relates to, but is not subsumed by, fundamental risk factors (basis, momentum and convexity), tail risk factors (skewness, left- and right-tail risk), liquidity, and volatility risk factors (basis-momentum and liquidity risk). Consistent with these time-series spanning tests, cross-sectional pricing tests further suggest that the CFEAR factor has significant pricing ability beyond these factors.

Further analysis reveals a link between the CFEAR premium and overall financial market sentiment as proxied by CBOE's VIX.[6] The short leg of the portfolio, which is the main driver

[6] The Chicago Board Options Exchange (CBOE) market volatility index (VIX) measures the implied volatility of options on the S&P 500 stock index. Referred to as the "investor fear gauge" by practitioners, VIX exhibits higher levels in periods of financial market turmoil and investor fear (see e.g., Whaley, 2000). Thus, it has been employed as proxy for investors' sentiment (moods and beliefs); see e.g. Baker and Wurgler (2007) and Da et al. (2015). Gao and Süss (2015) employ it as sentiment measure in a commodity futures markets study arguing that since the equity market is still by far the most liquid, proxies from this market can be taken as representative of general financial market sentiment.

of the CFEAR premium, is made up of commodity futures contracts that are more sentiment-prone, and the CFEAR premium is significantly larger in periods of pessimism. This evidence is in line with the wisdom from human psychology (behavioural finance) that investors are more vulnerable to the fear emotion when they find themselves outside of their "comfort zones" due to market instability or large losses (Shefrin, 2002). The finding of a greater CFEAR premium in periods of overall financial market pessimism is also in line with the prediction from behavioural finance models that the higher capital constraints of informed investors and/or their lower risk absorption capacity during turmoil periods can hinder the arbitrage trades that are required to eliminate any mispricing (DeLong et al., 1990; Shleifer and Vishny, 1997; Barberis et al., 1998; Cheng et al., 2015).

This study is inspired by a nascent commodity markets literature which investigates the out-of-sample predictive linkages between investor attention (as proxied by internet searches) and commodity returns (Han et al., 2017a, 2017b; Vozlyublennaia, 2014).[7] In a broader literature, the Google search volume has been endorsed as a useful out-of-sample predictor of equity returns (Da et al., 2011, 2015; Ben-Rephael et al., 2017; Dzielinski et al., 2018), sovereign credit spreads (Dergiades et al., 2015), and macroeconomic variables such as unemployment (D'Amuri and Marcucci, 2017; Niesert et al., 2019) inter alia.

Second, our work serves to emphasize the contention by Gao and Süss (2015) that sentiment plays a role in explaining commodity futures returns. Gao and Süss (2015) show that

[7] Through the lens of purely statistical criteria such as mean squared forecast errors, Han et al., (2017a) find that the Google search volume by oil- and real economy-related keywords are good predictors of oil futures returns relative to the historical average benchmark. Han et al. (2017b) find that the predictive errors of commodity return models that include various macroeconomic variables decrease by adding as predictor the Google search volume by 13 commodity names and combinations thereof with various terms (e.g. *cost*, *price, production* and *supply*). Using Google searches by *gold price* and *oil price* as keywords, Vozlyublennaia (2014) finds that more attention decreases predictability (the ability of current/past returns to convey information about future returns) and thus argues that pricing efficiency increases.

commodity futures with low dollar open interest, high volatility, poor past performance, or low basis are more sensitive to sentiment in the sense that they perform worse when the overall financial markets are bearish or pessimistic. In a similar vein, we find that the CFEAR portfolio performance is driven by the short leg which is typically made up of commodities with these characteristics; thus, our findings also re-affirm the Gao and Süss (2015) contention that general financial market sentiment can drive the performance of commodity futures portfolios.

Finally, this study contributes to the increasing stream of literature on commodity futures pricing by showing that fear of weather, agricultural diseases, political or economic hazards affects pricing beyond exposure to known systematic risk factors relating to momentum, basis, hedging pressure, convexity, skewness, basis-momentum, market liquidity or volatility (e.g., Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006; Miffre and Rallis, 2007; Basu and Miffre, 2013; Szymanowska et al., 2014; Bakshi et al., 2019; Fernandez-Perez et al., 2018; Boons and Prado, 2019; Gu et al., 2019, amongst others). The paper not only sheds light on commodity futures pricing but also informs the design of practical investment solutions.

The remainder of the paper is organized as follows. We introduce the commodity-specific CFEAR characteristic and long-short CFEAR portfolio construction methodology in Section 2. Data and benchmarks are presented in Section 3. Section 4 presents tests of hazard fear as pricing signal through time-series spanning tests and cross-sectional tests, and examines its potential drivers. Section 5 provides extensions and robustness checks. Section 6 concludes. An online Annex provides details of further robustness checks and additional analyses.

2. COMMODITY HAZARD-FEAR

2.1. Google search volume data

Inspired by the extant literature that uses Google search volume as proxy for investor attention (or information demand) our paper introduces a commodity hazard-fear characteristic that is constructed from internet search volume data from *Google Trends*. By contrast with extant papers (e.g., Da et al., 2011; Vozlyublennaia, 2014; Han et al., 2017a,b), we are concerned with the role of attention about potential threats to the commodity supply/demand; hence, the search query terms are hazard fear keywords (see Table 1) instead of the commodity names or tickers. Google organizes the searches by their origin as region versus worldwide. We use worldwide search data in the main empirical section and U.S. data in the robustness section.

[Insert Table 1 around here]

Using various sources (Iizumi and Ramankutty, 2015; Israel and Briones, 2013; United Nations Office for Disaster Risk Reduction, 2018; and reports from Material Risk Insights[8]), we compile a list of primary keywords that reflect commodity price risks associated with weather (WE), agricultural diseases (DI), geopolitical (GP), or economic (EC) threats. Next, as in Da et al. (2015), we refine the primary keywords by examining the top ten related searches (provided by *Google Trends*) and from these we filter out the irrelevant keywords.[9] Finally, we add to the latter the *risk* and *warning* terms, e.g. we consider *tsunami*, *tsunami risk* and *tsunami warning*. We thus end up with $J = 149$ keywords as listed in Table 1 by category: 113 weather (WE), 10 agricultural diseases (DI), 14 geopolitical (GP) and 12 economic (EC) hazards.

A spell of *extreme cold* or a *frost* are examples of WE hazards that could damage the growth of cotton while simultaneously increase the demand of natural gas for heating purposes; extremely *dry weather* or *wet weather* may adversely affect the harvest of sugar and cocoa that thrive in the right mix of rain and sunshine. Among the DI hazards, an increase of *crop diseases* is likely to reduce the supply of grain commodities, and an outbreak of *La Roya* fungus is likely to reduce the supply of coffee. GE hazards such as the *Russian crisis* are threats to the supply of natural gas; likewise, a *Middle East conflict* may damage oil provision. *Recession* or *crisis*

---

[8] See www.materials-risk.com.

[9] For instance, one of the top related searches to *hail damage* is *hail storm* which we retain while we neglect searches by *flood lights* that is unrelated to the paper aim.

are EC hazards that may reduce the demand for copper or oil due to a slowdown in business activity, while the demand for gold may simultaneously rise as gold is a safe-haven.

Let j denote a search keyword and t a sample week. *Google Trends* first obtains the ratio between the volume of queries associated with keyword j during week t, denoted $V_{j,t}$, and the entire volume of queries for any keyword in the same time period, denoted $V_{k,t}$. The ratio $V_{j,t}/V_{k,t}$ is subsequently divided by its historical maximum value and multiplied by a factor of 100 to scale it between 0 and 100. The resulting variable, $s_{j,t}$, is the Google Search Volume Index (GSVI) provided by *Google Trends* which has the interpretation of a search probability: $s_{j,t}$ equals 0 if the j^{th} keyword is not searched at all on week t and equals 100 in the peak search week of the keyword. The Google searches $s_{j,t}$ are sampled at a weekly frequency with each observation capturing the search queries from Monday 00:00:00 to Sunday 23:59:59.[10]

To increase the response speed, *Google Trends* compiles the GSVI data using a random subset of the actual historical search data and therefore the GSVI time-series downloaded on two different dates d_1 and d_2 can differ, $\{s_{j,t}\}_{d_1} \neq \{s_{j,t}\}_{d_2}$; for further details, see Stephen-Davidowitz and Varian (2015). Following extant studies (see e.g. Da et al., 2011; McLaren and Shanbhogue, 2011), we download GSVI series for each of the J=149 keywords on six different dates (5th, 6th, 7th, 16th, 17th and 18th February 2019)[11] and define the search series for our analysis as their average, i.e. $s_{j,t} \equiv \frac{1}{6}\sum_{d=1}^{6}\{s_{j,t}\}_d$. Table A.1 in the online Annex summarizes the 149 raw time-series of searches thus obtained $\{s_{j,t}\}, j = 1, \ldots 149$.

[10] We download Google Search Volume Index (GSVI) data at the weekly frequency. The weekly Google search data is characterized by a better information-to-noise ratio (than monthly or daily data); namely, weekly data ought to reflect the dynamics of attention in financial markets better than the coarser monthly data while circumventing the noise that characterizes daily data (e.g. Da et al., 2011; Vozlyublennaia, 2014; Dergiades et al., 2015; D'Amuri and Marcucci, 2017; Gao et al., 2020).

[11] The average pairwise correlation between the Google search series retrieved on the above 6 dates exceeds 90% for 55 out of the 149 search terms and the average correlation is 78%.

As an illustration, Figure 1, Panel A shows the evolution of the Google search index $s_{j,t}$ for the keyword *hurricane*, and the average price of lumber futures (front-contract) in each sample month. We observe that the peaks in Google searches by *hurricane* precede the occurrence of most notorious hurricanes such as, for instance, Hurricane Irma on September 2017, and tend to coincide with, or be quickly followed by, a jump in lumber futures prices which later adjust downwards. Similar patterns are observed in Panel B for *ebola* searches versus live/feeder cattle futures prices, and in Panel C for *oil crisis* searches versus natural gas futures prices. However, the opposite is observed in Panel D where increases in Google searches by *unemployment* (a demand-reduction related fear) are associated with decreases in the price of natural gas futures contracts, which later gradually adjust upwards. We cannot and do not assert that the agents behind these searches are exclusively commodity market participants; what is key for the present purposes, as these graphical examples prima facie suggest, is that the surges in the searches convey fear. Likewise, the fear and, in turn, the attention to hazards may be triggered by current news or by intrinsic concerns driven, for instance, by memory of extreme weather phenomena that occurred in the past or by extrapolating hazards that have affected other markets (representative heuristic).

[Insert Figure 1 here]

As in Da et al. (2015), the measure of interest is the weekly log change in the Google search volume or attention to hazard j defined as $\Delta S_{j,t} \equiv \ln(s_{j,t}/s_{j,t-1}), j = 1, \dots, J$, so that sharp increases in the attention to hazards can be taken to signal a surge in hazard-specific fear. Using search changes conveniently eliminates the look-ahead bias in GSVI induced by the aforementioned division of $V_{j,t}/V_{k,t}$ by its maximum historical value; this ensures that the hazard-fear portfolio uses information that is available at the time of portfolio formation.

2.2. CFEAR portfolio construction

This section defines the so-called $CFEAR_i$ characteristic and uses it as signal for asset allocation. Note that to avoid a look-ahead bias the analysis is conducted out-of-sample; namely, the buy or sell decisions made at the end of each Monday hinge only on past data. The CFEAR portfolio formation methodology unfolds as follows.

At stage one, at each portfolio formation time *t* (Monday) we begin by standardizing the weekly histories of searches $\Delta S_{j,t}$ like Da et al. (2015) as $\Delta S^*_{j,t} \equiv \Delta S_{j,t}/\sigma^{\Delta S}_{j,t}$ for each keyword $j = 1, \dots, 149$ where $\sigma^{\Delta S}_{j,t}$ is the standard deviation of $\Delta S_{j,t}$ using past data over the preceding L weeks. This standardization ensures that the $\Delta S_{j,t}$ series are comparable across keywords.

At step two, following Da et al. (2015), we estimate for each of the commodities ($i = 1, \dots, N$) in the sample as many OLS regressions as hazard keywords ($J = 149$ regressions); each regression is aimed at measuring the strength of the relation between commodity futures returns and the surge in hazard fear over the preceding L weeks

$$r_{i,t-l} = \alpha + \beta^{CFEAR}_{i,j} \cdot \Delta S^*_{j,t-l} + \varepsilon_{i,t-l}\,, \; l = 0, \dots, L-1 \qquad (1)$$

and finally, at step 3, we obtain a CFEAR characteristic (or signal) for commodity i by aggregating the estimated $\beta^{CFEAR}_{i,j}$ coefficients across all the $J = 149$ keywords as

$$CFEAR_{i,t} \equiv \sum_{j=1}^{J} \hat{\beta}^{CFEAR}_{i,j} \qquad (2)$$

By contrast, in their analysis of the impact of attention in equity markets Da et al. (2015) retain only the keywords with the most negative slopes in Equation (1); the reason is that they are concerned with *falling* equity prices since long positions by and large predominate. In the case of assets in zero net supply, such as commodity futures, falling prices are undesirable to long traders but desirable to short traders and thus, we consider all slope coefficients regardless of their sign. What is important is that, given the prior standardization, the most (least) relevant keywords for a given commodity will be revealed through a large (small) absolute $\hat{\beta}^{CFEAR}_{i,j}$

coefficient. For instance, a large positive $CFEAR_{i,t}$ indicates that the price of commodity i co-moved positively with hazard fear and thus that the net effect of hazard fear was of a supply-reducing or demand-enhancing nature; vice versa, a large negative $CFEAR_{i,t}$ suggests that the price of commodity i co-moved negatively with hazard fear and thus that the net effect of hazard fear was of a supply-increasing or demand-reducing nature.

Next we sort the available cross-section of futures contracts at each portfolio formation time t on $CFEAR_{i,t}$; short those in the top quintile(Q5) with the largest $CFEAR_{i,t}$ and long those in the bottom quintile(Q1) with the smallest $CFEAR_{i,t}$. The constituents of the long and short portfolios are equally weighted, the positions are fully collateralized and held for a week.

The above procedure is repeated at the next portfolio formation time (next Monday end) with expanding estimation windows at steps one and two, until the sample ends. The use of increasing windows builds on Da et al. (2015) and is aimed at maximizing the accuracy of the CFEAR estimation. The intuition is that the hazards are, by definition, infrequent and therefore, a fixed-length (rolling) estimation window for Equation (1) of, say, one to five years may be too short, resulting in too noisy $\beta_{i,j}^{CFEAR}$ measures. Using longer windows reduces considerably the sample of portfolio returns. We revisit this issue in the robustness tests section of the paper.

3. DATA AND ALTERNATIVE BENCHMARKS

3.1. Data

Similar to the cross-section of extant commodities studies (e.g., Basu and Miffre, 2013; Bianchi et al., 2015; Boons and Prado, 2019) our study is based on data for 28 commodity futures contracts, as listed in Table 2, which comprise 17 agricultural (4 cereal grains, 4 oilseeds, 4 meats, 5 miscellaneous other softs), 6 energy, and 5 metals (1 base, 4 precious). The first observation is from January 2004 as dictated by the availability of weekly *Google Trends* search data. Since 52 past weeks of data are required to construct the first portfolio, the portfolios are formed over the period January 2005 to December 2018.

[Insert Table 2 around here]

We measure futures returns as $r_{i,t} = \ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right)$ where $P_{i,t}$ is the Monday settlement price of front contracts in non-maturity months or second-nearest contracts otherwise – the data source is *Thomson Reuters Datastream*. Table 2 reports summary statistics for the futures returns (mean, standard deviation, first-order autocorrelation, and Ljung-Box test statistic for the null hypothesis that the first four autocorrelations are jointly zero). Weekly returns show little evidence of predictability based on sample autocorrelations; the Ljung-Box test rejects the null hypothesis of no autocorrelation at the 5% level only for copper, gasoline RBOB, live cattle and sugar. Table 2 also reveals that the weekly CFEAR signal, as defined in Equation (2), shows variability across commodities, ranging from -0.08 (Cocoa) to 0.24 (Gasoline RBOB), with an average coefficient of variation (standard deviation per absolute mean) of 3.62.

3.2. Performance evaluation benchmarks

Throughout the paper, the performance of the CFEAR portfolio is appraised in the context of a battery of benchmarks. Following Erb and Harvey (2006), Gorton and Rouwenhorst (2006) and Bakshi et al. (2019), we first consider a long-only equally-weighted and weekly-rebalanced portfolio of all commodities (AVG) as a possible risk factor that explains CFEAR. Additional benchmarks emanate from the literature on either commodity futures pricing, in particular, or asset pricing more generally. The risk factors we use are long-short portfolios that relate to the fundamentals of backwardation and contango: backwardated commodities with high basis (Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006; Bakshi et al., 2019), good past performance (Erb and Harvey, 2006; Miffre and Rallis, 2007; Bakshi et al., 2019), net short hedging or net long speculating (Basu and Miffre, 2013; Bianchi et al., 2015; Kang et al., 2019) or a convex price curve (Gu et al., 2019) are expected to outperform contangoed commodities whose characteristics are at the other end of the spectrum. Other long-short benchmarks relate to tail risks as measured by skewness (Fernandez-Perez et al., 2018), 1% and 99% Value-at-Risk, hereafter denoted as VaR1 and VaR99 (Bali et al., 2009; Atilgan et al, 2019);[12] the goal is to test whether the CFEAR premium is a tail risk premium in disguise. Finally, we test whether the CFEAR premium relates to liquidity and volatility risks as modeled through basis-momentum (Boons and Prado, 2019) and liquidity (Amihud, 2002) portfolios.

Appendix A, Panel B, lists the *k* characteristics used in the construction of the long-short risk factors and outlines the portfolio construction method. As with the CFEAR characteristic, we sort the futures contracts at the end of each Monday by each of these *k* characteristics in turn, buy the quintile deemed to appreciate, short the quintile deemed to depreciate, assign equal weights to the constituents and hold the fully-collateralized positions for a week. The right-hand side of Appendix A presents summary statistics for the long-only and long-short characteristic-sorted portfolios; the strategies based on hedging pressure, convexity, skewness, and basis-momentum stand out with Sharpe ratios ranging from 0.45 to 0.59.

4. **IS HAZARD-FEAR PRICED?**

This section measures the CFEAR factor and accesses whether its performance reflects compensation for exposure to risk or to sentiment. The analysis is conducted using both time-series spanning tests and cross-sectional pricing tests that control for other factors.

4.1. Performance of the CFEAR portfolio

Table 3 summarizes the performance of the CFEAR-sorted quintiles and that of the long-short CFEAR portfolio over the period January 2005 to December 2018. We observe a decrease in

[12] As dictated by rational asset pricing theory, higher risk shall be compensated by higher expected returns. Thus the skewness, VaR1 and VaR99 factors are constructed as the returns of portfolios with long positions in the commodity futures with the lowest skewness, the most negative VaR1 or the least positive VaR99 and short positions in the commodity futures with the highest skewness, the least negative VaR1 or the most positive VaR99.

the excess returns of the CFEAR-sorted quintiles from 4.35% (Q1) to -14.21% (Q5). The fully-collateralized Q1-Q5 portfolio captures an economically and statistically significant premium of 9.28% p.a. (t-statistic = 3.35) which suggests that the CFEAR signal contains useful out-of-sample predictive information for commodity excess returns. The CFEAR portfolio returns translate into a Sharpe ratio of 0.9012 which is higher than that of the alternative portfolios considered in Appendix A. The CFEAR portfolio stands out as regards tail/crash risk as suggested, for instance, by a maximum drawdown of -0.1881, while the corresponding figures for the long-only and long-short commodity portfolios lie in the ranges [-0.5392, -0.1828].[13]

[Insert Table 3 around here]

Examining the excess returns of the long versus short leg of the CFEAR portfolio reveals that the premium is mostly driven by the substantial drop in price of the commodity futures contracts with the most positive $CFEAR_{i,t}$ characteristic; namely, the short leg of the portfolio yields a large negative mean excess return of -14.21% p.a. ($t = -2.59$). With an annualized mean excess return of 4.35% ($t = 0.96$), the constituents of the long portfolio contribute much less to the overall performance. We will elaborate on this finding in Section 4.3.

Are a few specific commodities driving the performance of the CFEAR portfolio? Towards addressing this question, and confirming the results of Table 2, Figure 2 shows that the frequency with which a given commodity is included in the Q1 or Q5 portfolio is often below 50% revealing that the CFEAR portfolio composition varies. The *energy* commodities

[13] We also deploy the CFEAR portfolio on second-end contracts and spreads (front- minus second-end contracts) using the same sorting signal from Equation (2). The results presented in the online Annex Table A.2 confirm the attractive predictive ability of the CFEAR signal vis-à-vis other signals. We also gauge the relative merit of the keyword groups (weather, WE; agricultural diseases, DI; geopolitical, GP; and economic, EC) by implementing the CFEAR strategy on keyword sets that exclude one group at a time. The results shown in the online Annex Table A.3 highlight the importance of the WE group which is perhaps not surprising given that the supply/demand of many commodities is fundamentally linked to the weather.

are more often in the short Q5 (than in the long Q1) portfolio which indicates that on average the fear they are exposed to is associated with supply-reducing or demand-increasing hazards.

[Insert Figure 2 around here]

Figure 3 plots the future value of $1 invested in the long-short CFEAR portfolio, long-only AVG portfolio and long-short alternative portfolios; see Appendix A. Confirming our earlier findings (c.f. Table 3), Figure 2 reveals that the CFEAR strategy is relatively attractive.[14]

[Insert Figure 3 around here]

4.2. Are the CFEAR returns compensation for extant risks?

Time-series spanning tests

The analysis thus far has revealed that the CFEAR strategy is able to capture attractive mean excess returns in commodity futures markets. We now test whether the significant CFEAR premium is merely compensation for exposure to risk factors. For this purpose, we start off with the three-factor model of Bakshi et al. (2019) that includes the AVG, basis and momentum risk factors and estimate an OLS time-series spanning regression for the excess returns of the CFEAR portfolio. We then augment this baseline specification with various factors, in turn, that emanate from the literature on the pricing of commodity futures (hedging pressure and convexity), tail-risk (skewness, VaR1 and VaR99) or for the liquidity and volatility of commodities (basis-momentum and illiquidity). For each of the specifications, we look at the sign and significance of both the betas and alpha where the latter represents the average excess return of the CFEAR portfolio that is not a compensation for the hypothesized risk factors.

[14] As Figure 2 reveals, the CFEAR strategy pulled itself apart from the alternative strategies especially from June 2014 up until February 2016, a period during which the broad commodity market was in downfall as reflected in the AVG portfolio returns. Unreported results suggest that this is because the CFEAR signal was able to "time" the decline of certain commodities (especially, crude oil) much more accurately than the alternative signals.

Table 4 reports the results and shows that the excess returns of the CFEAR portfolio are sensitive to many of the risk factors considered. For example, the CFEAR portfolio returns exhibit positive momentum, convexity, skewness, VaR99 and basis-momentum betas, and negative basis, VaR1 and liquidity betas.[15] As argued by Boons and Prado (2019), given that basis-momentum proxies for volatility and liquidity risks, the positive slope of the basis-momentum factor and the negative slope on the liquidity risk factor indicate that lack of liquidity is an important driver of the performance of the CFEAR portfolio. In fact of all the risk factors considered, lack of liquidity is the most important factor as highlighted by a highly significant slope coefficient on the liquidity risk factor and by a substantial increase in adjusted-R^2 when moving from the baseline model to a model that includes the liquidity risk premium. The last column of Table 4 reports the "kitchen-sink" model that includes all the risk factors. The only surviving factors are basis, momentum, convexity and liquidity with the liquidity risk factor still presenting the most significant slope coefficient.

[Insert Table 4 around here]

Despite the significant risk factor exposures, the CFEAR portfolio affords economically sizeable and statistically significant alphas that range from 8.23% p.a. ($t = 3.35$) to 9.47% p.a. ($t = 3.73$). Thus, compensation for risk factor exposures does not tell the whole story.

Cross-sectional pricing tests

We complement the above time-series spanning tests with cross-sectional pricing tests to establish if the CFEAR factor is priced over and above extant risk factors. Following Kan et al. (2013) and Boons and Prado (2019) inter alia, using a set of portfolios as test assets $i = 1, \dots, N$ we first estimate full-sample betas via OLS *time-series* regressions

[15] The positive betas on skewness and VaR99 are consistent with investors' preferences for lottery-type assets as predicted by cumulative prospect theory (Barberis and Huang, 2008). The negative beta on the VaR1 factor is consistent with a market's slow assimilation of bad news as argued by Altigan et al. (2019) in line with the behavioral model of Hong et al. (2000).

$$r_{i,t} = \alpha_i + \boldsymbol{\beta}_i \cdot \boldsymbol{F}_t + \varepsilon_{i,t}, t = 1, \dots, T \quad (3)$$

where $r_{i,t}$ is the time *t* excess returns of (*a*) the quintile portfolios based on CFEAR, (*b*) the quintile portfolios based on the 9 characteristics listed in Appendix A (Panel B), and (*c*) the equally-weighted and weekly-rebalanced portfolios from the 6 commodity sub-sectors reported in Table 2 (with precious and base metals as a unique sector); thus, we have $N = 56$ commodity portfolios altogether. $\boldsymbol{F}_t$ includes the CFEAR factor as well as the 10 systematic risk factors that can potentially price the cross-section of portfolio returns (Appendix A, Panels A and B) and $\varepsilon_{i,t}$ is an error term. At step two, we estimate on each week the following cross-sectional regression of average excess returns on the step-one estimated full-sample betas

$$\bar{r}_i = \lambda_0 + \boldsymbol{\lambda}\widehat{\boldsymbol{\beta}}_i + \epsilon_i, i = 1, \dots, N \quad (4)$$

where $\boldsymbol{\lambda}$ is a vector containing the prices of risk associated with each of the factors.

We consider two types of models. The baseline model entertains the three risk-factors of Bakshi et al. (2019). We subsequently expand this model by cycling through each of the additional long-short risk premia considered in the time-series spanning tests, and then all together ("kitchen-sink" model). The second set of models adds to these pricing models the CFEAR factor. We assess the added value of the CFEAR factor through the adjusted-$R^2(\%)$ and mean absolute pricing error, $MAPE(\%) = \frac{100}{N}\sum_{i=1}^{N}|\hat{\epsilon}_i|$ of each model (Equation (4)). Table 5 reports the OLS estimates $\{\hat{\lambda}_0, \hat{\boldsymbol{\lambda}}\}$, and significance *t*-tests based on the Shanken (1992) robust standard errors (t_S, to correct for error-in-variables in $\widehat{\boldsymbol{\beta}}$) and the Kan et al. (2013) standard errors (t_{KRS}, to additionally correct for model misspecification).

[Insert Table 5 around here]

Irrespective of the model considered, the CFEAR factor is positively priced (at 8.73% p.a. on average across models) with significance Shanken *t*-statistics ranging from 2.54 to 2.92 across the various specifications of the risk-return relationship. Thus, the pricing ability of CFEAR cannot be fully rationalized by the fundamentals of backwardation and contango, nor by tail risks, liquidity and volatility risks. The models that include the CFEAR factor show a notable improvement in cross-sectional fit (versus the counterpart models that exclude it) by 21.79 percentage points (pp) on average in terms of adjusted-R^2 and by 0.0084pp on average in terms of MAPE; hence, CFEAR is an important driver of commodity returns.[16]

As Daskalaki et al. (2014) inter alia argue, a bias may emerge as regards the significance of the prices of risk when the test assets are portfolios sorted by the same criterion used to construct the risk factors. We address this bias by employing as test assets the 28 individual commodities, and estimating time-varying betas over a 52-week window at the first step as in Fama and MacBeth (1973) or Boons and Prado (2019), and sequential weekly cross-sectional OLS regressions at the second step. The results gathered in the online Annex Table A.5 do not challenge the main findings (Table 5); the CFEAR factor is positively priced and statistically significant at the 5% level or better in all models, and at 7.69% a year on average across models.

4.3. Does the CFEAR effect relate to overall financial market sentiment?

The main finding hitherto is that known risk factors provide only a partial explanation for the observed CFEAR premium. This section explores the role of sentiment in financial markets.

CFEAR premium and overall financial market sentiment

We begin by summarizing in Table 6 the characteristics of the commodities allocated over time to each of the CFEAR quintiles. The CFEAR characteristic is reported in the first row, and the basis, momentum, hedging pressure, convexity, skewness, VaR1, VaR99, basis-momentum and liquidity signals, as defined in Appendix A, in subsequent rows. The last two rows of the table report the realized variance defined as the average squared daily return over the 22 days

[16] For the sake of completeness, we augment the baseline time-series pricing model of Bakshi et al. (2019) with the change in EPU index (Economic Policy Uncertainty; Baker et al., 2016) or the change in GPR index (GeoPolitical Risk; Caldara and Iacoviello, 2018). None of these variables is found significantly to explain the CFEAR premium. The cross-sectional pricing tests reaffirm this finding. Detailed results are available in Table A.4 of the online Annex.

preceding portfolio formation week t (e.g. Boons and Prado, 2019) as well as their dollar open interest defined as the product of the number of outstanding contracts (or open interest), contract size and the front-end futures settlement price (e.g. Gao and Süss, 2015). It is noticeable that the commodities in Q5 exhibit significantly greater illiquidity, variance and lottery-like payoffs, and significantly inferior past performance (momentum), smaller dollar open interest, and basis than those in Q1; the characteristics exhibited by the Q5 constituents are precisely those typical of sentiment-sensitive assets according to Baker and Wurgler (2006, 2007) and Gao and Süss (2015). Thus, the earlier finding that the CFEAR premium is driven by the commodities in the short Q5 leg (c.f., Table 3), alongside the present finding that these commodities are relatively high sentiment-sensitive represents preliminarily evidence to suggest that overall financial market sentiment plays some role in the CFEAR premium.

[Insert Table 6 around here]

Deepening our analysis of the role of sentiment in the pricing of CFEAR, we test whether there is any difference in the magnitude of the premium captured by the CFEAR strategy in periods of low financial market sentiment (or pessimism) associated with high VIX levels, and periods of high financial market sentiment (or optimism) associated with low VIX levels. For this purpose, we estimate by OLS the following weekly time-series regression

$$r_{CFEAR,t} = \alpha_0^F + \alpha_{VIX}^F \cdot D_{t-1}^{VIX} + \boldsymbol{\beta}_i \cdot \boldsymbol{F}_t + v_t,\ t = 1, \ldots, T \tag{5}$$

where $r_{CFEAR,t}$ is the excess return of the CFEAR portfolio from week $t-1$ to week t, D_{t-1}^{VIX} is a VIX dummy equal to 1 if the VIX level at $t-1$ is higher than its full sample average and 0 otherwise, and $\boldsymbol{F}_t$ are the three risk factors of Bakshi et al. (2019); namely, the AVG, basis and momentum factors. Accordingly, the parameters $\alpha_0^F + \alpha_{VIX}^F$ and α_0^F capture the CFEAR alpha in high- and low-VIX states, respectively. By setting $\boldsymbol{\beta}_i = \boldsymbol{0}$, the parameters $\alpha_0 + \alpha_{VIX}$ and α_0 capture the CFEAR premium in high- and low-VIX states, respectively.

The results in Table 7, Panel A, reveal that the CFEAR premium is larger when VIX takes on high values; namely, when the overall financial market sentiment is pessimistic.

[Insert Table 7 around here]

For example, the mean excess return of the CFEAR portfolio in high-VIX states statistically exceed that in low-VIX states by 14.19% (*t*-statistic of 2.48 for the difference in performance). Similarly, the alpha of the CFEAR portfolio relative to the Bakshi et al. (2019) model in high-VIX states (19.16%, *t*-statistic of 4.44) statistically exceeds that obtained in low-VIX states (4.26%, *t*-statistic of 1.39). Looking at the short leg of the CFEAR portfolio more specifically, we note that it performs particularly poorly in high-VIX states: the average excess return of the short leg is then statistically lower (at -30.68%, *t*-statistic of -2.82) than that obtained in low-VIX states (at -5.26%, *t*-statistic of -0.94). Thus, general pessimism in financial markets magnifies the hazard fear in commodity markets.

The finding that the CFEAR premium is greater in periods when general financial market sentiment is pessimistic can be rationalized in two related ways. Firstly, one intuitive explanation stemming from the behavioral finance literature is that investors become more vulnerable to the hazard-fear emotion when they find themselves outside of their "comfort zone" due to large market instability or sizeable losses; e.g., see discussion in Shefrin (2002).

Secondly, the aforesaid finding is also consistent with predictions from extant behavioral finance models which establish that financial markets are affected by moods-driven traders (DeLong et al., 1990; Shleifer and Vishny, 1997; Barberis et al., 1998). The common denominator to these models is that arbitrageurs (that is, informed traders who bet against the mispricing induced by moods-driven traders) may be deterred from trading away such mispricing for different reasons. One is that they fear that the mood of irrational traders could go on to become more extreme and thus prices could deviate further from their fundamental values in periods of extreme sentiment. Bearing this risk in mind, informed speculators may opt at least in the short run to not arbitrage away the mispricing of commodity futures. As a result, emotions such as hazard-fear end up impacting equilibrium futures prices in the short run and more so during periods of extreme sentiment such as pessimistic periods.

Cheng et al. (2015) provide evidence that when VIX increases, the positions of commodity futures arbitrageurs decrease as a reflection of their more constrained capital and/or their lower risk-absorption capacity during these periods; for instance, arbitrage capital was largely withdrawn from the commodity futures market over the late 2000s Global Financial Crisis. Acharya et al. (2013) formalize a model where commodity futures speculators are capital constrained during stress periods. Thus we conjecture that the CFEAR premium reflects a mispricing driven by hazard-fear and a subsequent correction; speculators fail to arbitrage away the perceived mispricing because of their binding funding constraints and/or lesser risk-absorption capacity in periods of general market pessimism or turmoil periods. To gauge this conjecture, using Equation (5) reformulated with a TED (three-month Treasury bill minus three-month LIBOR in US dollars) dummy variable as a proxy of funding liquidity risk, we obtain the mean excess return and the Bakshi et al. (2019) alpha in periods of high versus low TED for the CFEAR portfolio and for its long and short legs. As Table 7 (Panel B) shows, the absolute excess return of the short (Q5) leg is much higher in the high TED period at -21.99% p.a. than in the low TED period at -11.24% p.a.; a similar contrast is observed for the alpha. By contrast, a much smaller difference between high and low TED periods is observed in the mean excess return and alpha of the long (Q1) leg.[17]

[17] These differences in high versus low TED states are economically significant but not statistically significant which may be explained by the small number of observations on the high TED states (197 weeks or 27%) relative to the low TED states (534 weeks or 73%).

Day-of-the-week performance

The CFEAR signals are measured at the end of each Monday using past weekly returns and past Google searches data as detailed in the methodology section. The long-short CFEAR portfolio is then held for a week; namely, from a given Monday-end to the next Monday-end. While maintaining the same methodology and the same one-week holding period, we now consider other days of the week as alternative portfolio formation times. Table 7 (Panel C) presents summary statistics for the performance of the resulting portfolios. We note a monotonic decrease in the CFEAR premium throughout the week. This pattern can be explained by the wisdom from the investor psychology literature that market participants are more pessimistic on Monday which could exacerbate any hazard fear and hence, the decrease in the futures price of the Q5 quintile constituents (negative return) will be larger. At the other extreme, part of the mispricing effect of hazard fear on the Q5 futures would be counteracted by the relative more optimistic mood that characterizes Friday (e.g., Birru, 2018). Given that the CFEAR premium derives mainly from the short leg (Table 3), it is perhaps not surprising to see that the short Q5 portfolio performs worse (at -14.21%) when formed on Mondays and relatively better (at -9.96%) when formed on Fridays. The improvement in Sharpe ratio of the short portfolio (Q5) over the week is quite noticeable. To add statistical significance, we deploy the Ledoit and Wolf (2008) test for the hypothesis $H_0: SR_{Q5_{Friday}} \leq SR_{Q5_{Monday}}$ using a block size of 5. The corresponding p-value (0.0776) indicates rejection of the null at the 10% level.

Summary and discussion

Summing up, the evidence presented in this section suggests that sentiment plays a role in explaining the CFEAR premium. The commodity futures contracts in the Q5 quintile that drives most of the CFEAR portfolio performance 1) are swayed by sentiment, 2) accrue more negative weekly returns in high VIX (general financial market pessimism) than low VIX periods, and 3) accrue more negative weekly returns if the portfolio is formed on Monday when traders are typically most pessimistic. Fear of any potential hazard that shifts downward the supply (or upward the demand) will increase net long hedging and, in turn, the current futures price relative to fundamentals to attract net short speculation. The subsequent gradual downward adjustment in the futures price (negative weekly return in our analysis) represents the correction of the mispricing induced by hazard fear. Our findings are consistent with the notion that informed speculators are reluctant to engage in arbitrage trades during periods of overall pessimistic moods not only because the mispricing could in fact worsen if the moods exacerbate (DeLong et al., 1990; Shleifer and Vishny, 1997) but also because of the binding funding constraints that arbitrageurs face (Cheng et al., 2015).

5. EXTENSIONS AND ROBUSTNESS CHECKS

The purpose of this section is to appraise the CFEAR premium after transaction costs, to cycle through several aspects of the CFEAR factor construction, and to deploy a placebo test.

5.1. Turnover and transaction costs

We measure the turnover (TO) of a given portfolio as the average of all the trades incurred

$$TO = \frac{1}{T-1}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(\left|w_{i,t+1} - w_{i,t^+}\right|\right) \tag{6}$$

where $t = 1, \ldots, T$ denotes the portfolio formation times, $w_{i,t}$ is the weight assigned to the *i*th commodity as dictated by a given strategy at week *t*, $w_{i,t^+} \equiv w_{i,t} \times e^{r_{i,t+1}}$ is the actual portfolio weight right *before* the next rebalancing at $t+1$, $r_{i,t+1}$ is the weekly return of the i^{th} commodity from week t to week $t+1$. Thus the TO measure captures also the mechanical evolution of the weights due to within-week price dynamics (e.g., $w_{i,t}$ increases to w_{i,t^+} when $r_{i,t+1} > 0$). We calculate the time *t* net return of the long-short portfolio *P* as

$$r_{P,t+1} = \sum_{i=1}^{N} w_{i,t} r_{i,t+1} - TC\sum_{i=1}^{N}\left|w_{i,t} - w_{i,t-1^+}\right| \tag{7}$$

using proportional trading costs *TC*=8.6 bps (Marshall et al., 2012). Figure 4 shows the results.

[Insert Figure 4 around here]

As can be gleaned from Panel A of Figure 4, the CFEAR portfolio turnover (at 0.08) is notably inferior to that of the basis (0.38), momentum (0.27), skewness (0.21), basis-momentum (0.23) and convexity (0.56) portfolios and comparable to the turnover of the remaining portfolios. As regards performance, unreported results show that after controlling for transaction costs the CFEAR premium decreases only slightly from 9.28% p.a. (t =3.35) to 8.92% (t = 3.22), and it still represents a very attractive performance relative to alternative long-short strategies. On a risk-adjusted basis, the Sharpe ratios plotted in Panel B of Figure 4 confirm that transaction costs subsume a small part of the performance of the CFEAR portfolio.

5.2. Alternative approaches to measure the CFEAR characteristic

This section provides robustness tests related to the construction of the CFEAR signal.

First, we consider US Google searches by the users' IP address in place of the worldwide searches used thus far. Second, as in Da et al. (2015) we winsorize the Google search changes by shrinking the extreme $\Delta S_{j,t}$ towards $\overline{\Delta S}_{j,t} \pm 1.96\sigma_{j,t}^{\Delta S}$ where $\overline{\Delta S}_{j,t}$ is the mean of the time-series associated with the search term j up to time t and $\sigma_{j,t}^{\Delta S}$ its standard deviation. Third, we deseasonalize the searches $\Delta S_{j,t}$ by regressing them on month dummies and retain the residuals, also as in Da et al. (2015). The rationale for omitting these two transformations in the main analysis is that our goal is to exploit surges in Google searches and by filtering out the large hazard-search changes through winsorization we may disregard valuable information. Likewise, many weather hazards (e.g., frosts or torrential rain) are seasonal and so the fear (proxied by the search activity) may capture seasonality that has valuable predictive content.

Fourth, in order to focus on the possible distortions induced by the weeks with 0 searches ($s_{j,t} = 0$ which we replace by a very small arbitrary non-zero value 10^{-11} in the main analysis to circumvent the logarithmic transformation issue), we provide three additional robustness checks for the long-short CFEAR portfolios constructed using the same methodology except for these changes: (i) as in Han et al. (2017b) we replace the 0s by 1s so that the 0s are then turned into zero log search values, that is, $\ln(s_{j,t}) = 0$, (ii) the 0s in the search series $s_{j,t}$ are left as such and the Google search variable is instead defined as $\Delta S_{j,t} = \ln\left(\frac{1+s_{j,t}}{1+s_{j,t-1}}\right)$ instead of $\Delta S_{j,t} = \ln\left(\frac{s_{j,t}}{s_{j,t-1}}\right)$, and (iii) although we consider weeks with zero searches informative, to dispel any remaining concerns, we remove the 0 data points from the calculations. These robustness checks are labelled (4a), (4b) and (4c), respectively.

Fifth, we address the issue of noisy keywords by filtering out of the 149 original keywords those that meet any of these two criteria: (a) the time-series of weekly searches $s_{j,t}$ contains more than $\tau\%$ of 0s suggesting that the keyword is not popular, (b) the correlation among the 6 series $\{s_{j,t}\}_d$ that form $s_{j,t}$ on six different days d (Section 2.1) is less than $\kappa\%$ on average suggesting large sampling variability. We use $\{\tau, \kappa\} = \{20{,}80\}$ resulting in 72 keywords.[18]

Sixth, we address concerns related to backdating by obtaining for each of the 149 keywords new weekly search histories $\{s_{j,t}\}_d$ on the following six days d = 12th, 13th, 16th, 17th, 18th and 20th December 2019; we then define the search series per keyword $s_{j,t}$ as $\frac{1}{6}\sum_{d=1}^{6}\{s_{j,t}\}_d$.

Seventh, we measure the CFEAR signal in a manner that controls for the impact of media coverage (see e.g., Fang and Peress, 2009; Tetlock, 2015) by reformulating Equation (1) as

$$r_{i,t-l} = \alpha + \beta_{i,j}^{CFEAR} \cdot \Delta S_{j,t-l}^{*} + \gamma_i \cdot News_{i,t-l} + \varepsilon_{i,t-l}, l = 0, \dots, L-1 \quad (8)$$

$News_{i,t}$ denotes the amount of news coverage[19] of commodity i in week t with a relevance score of either 25 or 75. The rest of the portfolio formation unfolds as before.

[18] Qualitatively similar results are obtained with $\{\tau, \kappa\} = \{10{,}90\}$.

[19] We collect from WRDS-Ravenpack the weekly media coverage (or total number of news articles published from Monday to Sunday) per commodity. The WRDS-Ravenpack software assigns a score of 0 to 100 to each article to indicate how relevant the article is to the commodity at hand. For instance, a news article with relevance score of 0 for coffee means that coffee was only indirectly (passively) mentioned in the article, while an article with score 75 or higher is considered by WRDS-Ravenpack as extremely relevant to coffee (i.e., the commodity featured fairly prominently in the news story). Our main analysis is based on data extracted under the

Table 8 presents summary statistics for the performance of the resulting CFEAR portfolios. Irrespective of the approach used to measure the CFEAR characteristic, CFEAR is found to have predictive power over forthcoming futures returns. The Sharpe ratios, for example, range from 0.62 to 0.90 and thus are of a similar magnitude to that reported in Table 3 (0.90). Perhaps not surprisingly the winsorization and deseasonalization of the Google searches (columns (2) and (3)) as in Da et al. (2015) decreases the magnitude of the CFEAR premium, which serves to prove the informative content of extreme Google searches and the strong seasonality of the searches. Column (7) of Table 8 shows that taking on board media coverage does not alter the size and significance of the CFEAR premium. The rationale for this finding is twofold. On the one hand, as noted by Da et al. (2011), the response of prices to the demand of information may be different from the response to the supply of information. Second, as argued above, attention to a potential threat to a commodity supply or demand may be driven by factors unrelated to the news articles currently published such as the recollection of a hazard that shifted supply or demand in the past or the extrapolation from extreme phenomena that affected other commodities (known as the "representative heuristic" in behavioural finance).[20]

[Insert Table 8 around here]

---

conservative relevance score of 75 but, for completeness, we also report results under the rather lax relevance score of 25. The news variables are summarized in Table A.6 of the online Annex.

[20] The fact that CFEAR premium remains after controlling for the media coverage/news about each commodity can be taken as evidence of heuristic-driven bias and market inefficiency (Shefrin, 2002). Further strengthening this evidence, inspired by Vozlyublennaia (2017) we estimate predictive regressions of each commodity excess returns on past excess returns (up to four weeks) and past excess returns interacted with an aggregate attention measure ($\Delta S_t = \sum_{j=1}^{149} \Delta S_{j,t}^*$ where $\Delta S_{j,t}^* = \frac{\Delta S_{j,t}}{\sigma_{j,t}^{\Delta S}} = \frac{\ln(s_{j,t}/s_{j,t-1})}{\sigma_{j,t}^{\Delta S}}$ is the standardized "attention" variable associated with the jth hazard). The findings from these regressions (reported in Table A.7 of the online Annex) suggest that the fear-driven attention to hazards generally increases predictability of commodity futures returns which can be interpreted as a form of inefficiency.

5.3. Alternative portfolio construction methods

Further we deploy alternative CFEAR portfolios: a) considering a fixed-length rolling window of 10 years ($L = 520$ weeks) for the estimation of Equation (2), b) weighting the Q1 and Q5 constituents by the magnitude of the standardized CFEAR signal (namely, $\theta_{i,k,t} \equiv \left(x_{i,t} - \bar{x}_t\right)/\sigma^x_{k,t}$,, where $x_{i,t} = CFEAR_{i,t}$ is the hazard-fear characteristic from Equation (2) with $\bar{x}_t$ and σ^x_t its cross-sectional mean and standard deviation at time *t*), *c*) forming the long-short CFEAR portfolio with the entire cross section (*N*/2 each) of commodities weighted either by $1/N$, standardized rankings, standardized signals, or winsorized and standardized signals, and *d*) considering at each portfolio formation time the $0.8N$ commodities with the largest open interest on the prior week to further ensure that the results are not driven by illiquidity. The results, gathered in the online Annex Table A.8, suggest that the CFEAR premium remains sizeable ranging from 4.95% p.a. (*N*/2 equally-weighted commodities allocated to each leg of the portfolio) to 10.14% p.a. (only the 80% most liquid commodities are considered).

For completeness, in line with the pricing factor construction literature, we measure the premium that is captured when the long-short CFEAR portfolio is formed at each month-end and held for one month. We maintain all other aspects of the CFEAR portfolio construction as described above. We re-deploy all other portfolio strategies using the same approach. Reassuringly, the results in Table A.9 of the online Annex indicate that the CFEAR premium remains economically and statistically significant at 7.98% ($t = 3.06$) translating into a Sharpe ratio of 0.7906 that is attractive relative to the Sharpe ratio of the alternative strategies. Thus, we can assert that our findings do not hinge on the weekly portfolio formation frequency.

5.4. Placebo test

We now conduct an intuitive placebo test to ascertain whether our finding of a significant hazard-fear premium in commodity futures markets is an artefact of the CFEAR signal and factor construction methodology. For this purpose, we deploy the same methodology for cross-sections of financial futures contracts instead. The motivation is that, since it is most unlikely that fear of weather events (e.g., a frost or a tornado) or crop diseases (e.g., La Roya fungus) feed into the futures prices of equity index, currency and fixed income futures, an empirical finding of a significant hazard-fear premium also in these markets can be interpreted as suggestive that the commodity hazard-fear premium we have identified is spurious.[21]

In order to increase the power of this placebo test, we filter out the geopolitical (GP) and economic (EC) hazards that might influence pricing across asset classes and obtain the CFEAR signal using the 123 keywords/hazards in the weather (WE) and crop disease (DI) categories that are most specifically associated with commodities. We re-construct the long-short CFEAR portfolio of commodity futures using these 123 WE/DI keywords and form similar portfolios with the three cross-sections of equity index, fixed income and currency futures. Specifically, for this analysis we obtain daily settlement prices from *Thomson Reuters Datastream* for 40 equity index futures, 13 fixed income futures and 19 currency futures; see detailed composition in Table A.10 of the online Annex. The placebo test results are reported in Table 9.

[Insert Table 9 around here]

The fear premium remains sizeable and statistically significant at 8.17% p.a. ($t = 3.06$) in commodity futures markets when the keywords are restricted to the WE and DI hazards. However, in sharp contrast and consistent with the above intuition, the WE and DI hazard-fear premia are insignificant at 1.83% p.a. (t=1.62) in equity index futures, 0.19% p.a. (t=0.25) in fixed income futures and 1.16% p.a. (t=1.50) in currency futures. This plausible contrast

[21] We are mindful, however, of a literature that links rare disasters (including weather ones) and equity prices (see e.g., Barro, 2006; Hong et al., 2019, Choi et al., 2020, to name a few). Although rare events do impact the pricing of individual stocks (for example, a frost raises the valuation of producers), we expect that effect to be diversified away at the level of equity index futures (the same frost simultaneously decreases the valuation of refiners).

between commodity and financial futures suggests that the CFEAR premium uncovered in commodity futures markets is unlikely to be an artefact of the methodology.

6. CONCLUSIONS

Does the fear emotion influence commodities futures pricing? This paper addresses this question by focusing on fear about hazards such as, for instance, extreme weather, agricultural pests, geopolitical risks or a financial crisis, that represent threats to the commodity supply or demand. As in Da et al. (2011) and others, we proxy fear surges by changes in the aggregate Google search volume (or active attention) using 149 hazard-related keywords as query terms.

Through time-series spanning tests, we show that a long-short portfolio that exploits the hazard-fear as sorting signal for a cross-section of 28 commodity futures contracts earns a sizeable premium of 9.28% per annum that cannot be rationalized as compensation for exposure to a battery of known systematic risk factors. Through asset pricing tests we demonstrate that exposure to hazard-fear is a key determinant of the cross-sectional variation in the returns of commodity portfolios beyond their exposure to systematic risk factors.

The results reveal a link between the CFEAR premium and overall financial market sentiment. The short leg of the CFEAR portfolio, which drives the premium, is made up of commodity futures that are very sentiment-prone, and the CFEAR premium is significantly larger in periods of pessimism. This evidence is consistent with the wisdom from human psychology that investors are more vulnerable to the fear emotion when they find themselves outside of their "comfort zone" due to market instability or large losses. The finding of a greater CFEAR premium in periods of pessimism is also in line with the behavioural theory prediction that speculators' fear of mounting-pessimism in the short run alongside their capital and risk absorption capacity constraints deter the arbitrage needed to eliminate the mispricing.

Overall, we conclude that the presence of "animal spirits" (paraphrasing the British economist John Maynard Keynes) cannot be ruled out in commodity futures markets, namely, fear or anxiety about potential hazards, irrespective of whether they ultimately materialize, feeds into futures prices and more so during periods of general financial market pessimism.

Appendix A. Risk factors

The table focuses on the broad commodity market risk factor (long-only portfolio) in Panel A, and on alternative risk factors (long-short portfolios) in Panel B. It presents the signals used as sorting criteria for the construction of the risk factors (column 1), the criteria for allocation of commodity futures contracts to the long leg of the portfolio (col. 2), as well as the time window for signal measurement with reference to the portfolio formation time denoted t (col. 3). The right-hand section presents summary statistics for the risk factors. Mean is annualized mean excess return, StDev is annualized standard deviation, SR is Sharpe ratio, 1%VaR is the1% Cornish-Fisher Value-at-Risk, and MDD is the maximum drawdown. F_{i,t,T_1}, F_{i,t,T_2} and F_{i,t,T_3} are the time t prices of the futures contracts with respective maturities $T_1 < T_2 < T_3$. $Long_{i,t}$ and $Short_{i,t}$ are the week t long and short open interest of large speculators, respectively, as reported by the CFTC. The period is January 2005 (week 1) to December 2018 (week 4).

				Performance					
Signals		**Long positions**	**Time window to measure signals**	**Mean**		**StDev**	**SR**	**1% VaR**	**MDD**
Panel A: Equally-weighted long-only									
	$AVG_t \equiv \frac{1}{N}\sum_{i=1}^{N} r_{i,t}$	All commodities	Observations at time t	-0.0332	(-0.86)	0.1336	-0.2486	0.0562	-0.5392
Panel B: Risk-based characteristics									
Basis	$Roll_{i,t} \equiv \ln(F_{i,t,T_1}) - \ln(F_{i,t,T_2})$	Higher signal	Observations at time t	0.0346	(1.27)	0.1021	0.3387	0.0356	-0.1905
Momentum	$Mom_{i,t} \equiv \frac{1}{52}\sum_{j=0}^{51} r_{i,t-j}$	Higher signal	Observations in the 52 weeks preceeding t	0.0151	(0.51)	0.1168	0.1296	0.0421	-0.2872
Hedging pressure	$HP_{i,t} \equiv \left(\frac{1}{52}\right)\sum_{j=0}^{51} \frac{Long_{i,t-j} - Short_{i,t-j}}{Long_{i,t-j} + Short_{i,t-j}}$	Higher signal	Observations in the 52 weeks preceeding t	0.0598	(2.32)	0.1009	0.5926	0.0331	-0.1828
Convexity	$Convexity_{i,t} \equiv \frac{\ln(F_{i,t,T_1}) + \ln(F_{i,t,T_3}) - 2\ln(F_{i,t,T_2})}{T_2 - T_1}$	Higher signal	Observations at time t	0.0480	(1.85)	0.0938	0.5121	0.0301	-0.2525
Skewness	$Skewness_{it} \equiv \frac{\sum_{d=1}^{D}(r_{i,d} - \mu_i)^3 / D}{\sigma_i^3}$	Lower signal	D = Number of days in the year preceeding t	0.0444	(1.62)	0.0991	0.4481	0.0296	-0.2955
VaR1	1st quintile of the distribution of daily returns	Lower signal	Daily observations in the year preceeding t	-0.0233	(-0.77)	0.1131	-0.2058	0.0379	-0.4892
VaR99	99th quintile of the distribution of daily returns	Lower signal	Daily observations in the year preceeding t	0.0382	(1.31)	0.1141	0.3348	0.0367	-0.3429
Basis momentum	$BM_{i,t} \equiv Mom_{i,t,T_1} - Mom_{i,t,T_2}$	Higher signal	Observations in the 52 weeks preceeding t	0.0519	(1.93)	0.0967	0.5368	0.0323	-0.2376
Liquidity	$\frac{1}{D}\sum_{j=0}^{D-1} \frac{\lvert r_{i,t-j} \rvert}{\$Volume_{i,t-j}}$	Higher signal	D = Number of days in the 2 months preceeding t	-0.0019	(-0.07)	0.0963	-0.0194	0.0340	-0.5200

References

Acharya, V. V., Lochstoer, L. A., and Ramadorai, T. (2013). Limits to Arbitrage and Hedging: Evidence from Commodity Markets, *Journal of Financial Economics* 109, 441-465.

Amihud, Y. (2002). Illiquidity and Stock Returns: Cross-Section and Time Series Effects. *Journal of Financial Markets* 5, 31–56.

Baker, M., and Wurgler, J. (2006). Investor sentiment and the cross-section of returns. *Journal of Finance* 61, 1645–1680.

Baker, M., and Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives* 21, 129–152 .

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics* 131, 1593-1636.

Bakshi, G., Gao, X., and Rossi, A. G. (2019). Understanding the Sources of Risk Underlying the Cross Section of Commodity Returns. *Management Science* 65, 619-641.

Barberis, N., and Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review* 98, 2066–2100.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A Model of Investor Sentiment. *Journal of Financial Economics* 49, 307-343.

Barro, R. (2006). Rare Disasters and Asset Markets in the Twentieth Century. *Quarterly Journal of Economics* 121, 823-866.

Basu, D., and Miffre, J. (2013). Capturing the Risk Premium of Commodity Futures: The Role of Hedging Pressure. *Journal of Banking and Finance* 37, 2652–2664.

Ben-Rephael, A., Da, Z., and Israelsen, R. D. (2017). It Depends on Where You Search: Institutional Investor Attention and Underreaction to News. *Review of Financial Studies* 30, 3009–3047.

Bessembinder, H. (1992). Systematic Risk, Hedging Pressure, and Risk Premiums in Futures Markets. *Review of Financial Studies* 4, 637–667.

Bianchi, R. J., Drew, M., and Fan, J. (2015). Combining Momentum with Reversal in Commodity Futures. *Journal of Banking and Finance* 59, 423–444.

Birru, J. (2018). Day of the week and the cross-section of returns. *Journal of Financial Economics* 130, 182–214.

Brennan, M. (1958). The Supply of Storage. *American Economic Review* 48, 50-72.

Boons, M., and Prado, M. (2019). Basis-Momentum. *Journal of Finance* 74, 239-279.

Caldara, D., and Iacoviello, M. (2018). Measuring Geopolitical Risk. *Board of Governors of the Federal Reserve System International Finance Discussion Paper* No. 1222.

Cheng, I-H., Kirilenko, A., and Xiong, W (2015). Convective Risk Flows in Commodity Futures Markets. *Review of Finance* 19, 1733–1781.

Choi, D., Gao, Z., and Jiang, W. (2020). Attention to Global Warming. *Review of Financial Studies*, forthcoming.

Cootner, P. H. (1960). Returns to Speculators: Telser versus Keynes. *Journal of Political Economy* 68, 396–404.

Da, Z., Engelberg, J., and Gao, P. (2011). In Search of Attention. *Journal of Finance* 66, 1461–1499.

Da, Z., Engelberg, J., and Gao, P. (2015). The Sum of All FEARS Investor Sentiment and Asset Prices. *Review of Financial Studies* 28, 1–32.

D'Amuri, F., and Marcucci, J. (2017). The Predictive Power of Google Searches in Forecasting

US Unemployment. *International Journal of Forecasting* 33, 801-816.

Daskalaki, C., Kostakis, A., and Skiadopoulos, G. (2014). Are There Common Factors in Commodity Futures Returns? *Journal of Banking and Finance* 40, 346–363.

DeLong, J. B., Shleifer, A., Summers, L.H., and Waldmann, R. (1990). Noise trader Risk in Financial Markets. *Journal of Political Economy* 98, 703–738.

Dergiades, T., Milas, C., and Panagiotidis, T. (2015). Tweets, Google Trends, and Sovereign Spreads in the GIIPS. *Oxford Economic Papers* 67, 406–432.

Dzielinski, M., Rieger, M.O., and Talpsepp, T. (2018). Asymmetric Attention and Volatility Asymmetry. *Journal of Empirical Finance* 45, 59-67.

Erb, C. B., and Harvey, C. R. (2006). The Strategic and Tactical Value of Commodity Futures. *Financial Analysts Journal* 62, 69–97.

Fang, L. and Peress. J. (2009). Media Coverage and the Cross-Section of Stock Returns, *Journal of Finance* 64, 2023-2-25.

Fama, E. F., and MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81, 607–636.

Fernandez-Perez, A., Frijns, B., Fuertes, A.-M., and Miffre, J. (2018). The Skewness of Commodity Futures Returns. *Journal of Banking and Finance* 86, 143-158.

Gao, L., Süss, S. (2015). Market Sentiment in Commodity Futures Returns. *Journal of Empirical Finance* 33, 84-103.

Gorton, G., Hayashi, F., and Rouwenhorst, G. (2012). The Fundamentals of Commodity Futures Returns. *Review of Finance* 17, 35-105.

Gorton, G., and Rouwenhorst, G., (2006). Facts and Fantasies about Commodity Futures. *Financial Analysts Journal* 62, 47-68.

Gu, M., Kang, W., Lou, D., & Tang, K. (2019). Relative-basis. SSRN Working Paper *3404561*.

Han, L., Lv, Q., and Yin, L. (2017a). Can Investor Attention Predict Oil Prices? *Energy Economics* 66, 547–558.

Han, L., Li, Z., and Yin, L. (2017b). The Effects of Investor Attention on Commodity Futures Markets. *Journal of Futures Markets* 37, 1031–1049.

Hirshleifer, D. (1988). Residual Risk, Trading Costs and Commodity Futures Risk Premia. *Review of Financial Studies* 1, 173-193.

Hong, G.H., Li, F.W., and Xu, J., (2019). Climate Risks and Market Efficiency. *Journal of Econometrics*, forthcoming.

Hong, H., Lim, T., and Stein, J.C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance* 55, 265-295.

Iizumi, T., and Ramankutty, N. (2015). How Do Weather and Climate Influence Cropping Area and Intensity? *Global Food Security* 4, 46–50.

Israel, D. C., and Briones, R. M. (2013). Impacts of Natural Disasters on Agriculture, Food Security, and Natural Resources and Environment in the Philippines. *Economic Research Institute for ASEAN and East Asia* (ERIA) Discussion Paper 15, 1–53.

Kaldor, N. (1939). Speculation and Economic Stability. *Review of Economic Studies* 7, 1–27.

Kan, R., Robotti, C., and Shanken, J. (2013). Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology. *Journal of Finance* 68, 2617-2649.

Kahneman, D. & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47, 263–291

Kang, W., Rouwenhorst, G., and Tang, K. (2019). A Tale of Two Premiums: The Role of

Hedgers and Speculators in Commodity Futures Markets. *Journal of Finance* 75, 377-417.

Ledoit, O., and Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance* 15, 850–859.

Lemieux, J., and Peterson, R. A. (2011). Purchase Deadline as a Moderator of the Effects of Price Uncertainty on Search Duration. *Journal of Economic Psychology* 32, 33–44.

Lo, A. (2011). *Fear, Greed, and Financial Crises: A Cognitive Neurosciences Perspective* in Handbook of Systemic Risk (Ch. 23, pp 622-662) eds. J. P. Fouque and J. Langsam, Cambridge University Press.

McLaren, N., and Shanbhogue, R., (2011). Using Internet Search Data as Economic Indicators. *Bank of England Quarterly Bulletin* Q2.

Marshall, B. R., Nguyen, N. H., and Visaltanachoti, N. (2012). Commodity Liquidity Measurement and Transaction Costs. *Review of Financial Studies* 25, 599–638.

Miffre, J., and Rallis, G. (2007). Momentum Strategies in Commodity Futures Markets, *Journal of Banking and Finance* 31, 1863-1886.

Niesert, R.F., Oorschot, J.A., Veldhuisen, C.P., Brons, K., and Lange, R. J. (2019). Can Google Search Data Help Predict Macroeconomic Series? *International Journal of Forecasting*, forthcoming.

Shanken, J. (1992). On the Estimation of Beta-Pricing Models. *Review of Financial Studies* 5, 1-33.

Shefrin, H. (2002). *Beyond Greed and Fear: Understanding Behavioural Finance and the Psychology of Investing.* Oxford University Press.

Shleifer, A. and Vishny, R.W. (1997). The Limits of Arbitrage. *Journal of Finance* 52, 35–55.

Stephens-Davidowitz, S. and H. Varian (2015). A Hands-on Guide to Google Data. Google Inc. report.

Szymanowska, M., De Roon, F., Nijman, T., and Van Den Goorbergh, R. (2014). An Anatomy of commodity futures risk premia. *Journal of Financ*e 69, 453-482.

Tetlock, P.C. (2007). Giving Content to Investor Sentiment: the Role of Media in the Stock Market. *Journal of Finance* 62, 1139-1168.

Tetlock, P.C. (2015). *The Role of Media in Finance* in Handbook of Media Economics, (Vol. 1, Ch.18, 701-721) eds. S.P. Anderson, J. Waldfogel and D. Strömberg, North-Holland.

Tottenham, N. (2014). The Importance of Early Environments for Neuro-Affective Development. *Current Topics in Behavioral Neurosciences* 16, 109-129.

Vozlyublennaia, N. (2014). Investor Attention, Index Performance, and Return Predictability. *Journal of Banking and Finance* 41, 17–35.

Whaley, R.E. (2000). The Investor Fear Gauge. *Journal of Portfolio Management* 26, 12-17.

Working, H. (1949). The Theory of Price of Storage. *American Economic Review* 39, 1254-1262.

Table 1. Search query terms

This table lists all the terms or keywords (J=149) used in the Google searches grouped according to the type of hazard or vulnerability that they represent. An asterisk indicates search queries carried out specifically within the weather category of *Google Trends*. Sources: Iizumu and Ramankutty (2015), Israel and Briones (2013), United Nations Office for Disaster Reduction (2018) and Material Risk Insights (www.materials-risks.com).

Primary terms[1]	Related terms (from Google top related searches)	#terms
Weather (WE; 113 keywords)		
adverse weather	adverse weather conditions; adverse weather warning	3
blizzard*	blizzard risk;blizzard warning; weather blizzard warning	4
catastrophic weather	catastrophic events;catastrophic weather events; natural disaster: natural hazard	5
climate disturbance	climate change; cyclogenesis; global warming	4
cold*	cold spell;cold weather; freeze warning	4
cyclone	cyclone risk;cyclone warning; tropical cyclone; tropical cyclone risk; tropical cyclone warning;	6
drought	drought risk, drought warning, droughts	4
dry weather		1
El niño weather		1
extreme weather	extreme cold;extreme cold temperatures; extreme heat; extreme rain; extreme temperatures;extreme	7
forest fire	forest fires	2
flood	flood risk; flood warning; flooding; floods	5
frost*	frost risk; frost warning;frosts*	4
gust*	gusts*	2
hail	hail damage; hail risk; hail storm warning; hail storm; hail warning	6
Harmattan wind		1
heat*	heat wave; heat waves; heatwave; heatwaves	5
hot weather	high temperature; high temperatures	3
hurricane	hurricane risk; hurricane warning; hurricanes*	4
rain*	torrential rain; heavy rain*; heavy rain risk; heavy rain warning ; heavy rain fall	6
severe weather	severe heat; severe weather risk; weather risk; weather warning	5
snow*	snow risk; snow storm warning; snow warning	4
storm*	storm risk; storm warning; tropical storm; tropical storm risk; tropical storm warning	6
tornado	tornado risk ; tornado warning	3
tropical weather		1
typhoon	typhoon risk; typhoon warning	3
wet weather		1
wildfire*	wildfire risk; wildfire warning; wildfires	4
wind*	wind gust; wind gusts; wind risk; wind warning; wind speed; wind storm; strong wind; strong wind gust	9
Agricultural diseases (DI; 10 keywords)		
crop pest	crop diseases; crop pest risk; crop pests;insect pest; pest control; pest risk	7
Ebola		1
La Roya		1
rust coffee		1
Geopolitical (GP; 14 keywords)		
Middle East conflict	Middle East instability, Middle East terrorism	3
oil embargo	oil crisis, oil outage	3
Russian crisis		1
Libyan crisis		1
Syrian war		1
terrorism	Africa terrorism; Africa instability	3
terrorist attack	terrorist attacks	2
Economic (EC; 12 keywords)		
crisis	economic crisis; financial crisis	3
recession	economic recession; recession 2008; recession depression; the recession; US recession	6
unemployment	unemployment rate; US unemployment	3
Total		149

Table 2. Descriptive statistics for individual commodity futures

This table lists for the 28 commodities the sub-sector, the first and last observation dates, annualized mean excess return (Mean), annualized standard deviation (StDev), first-order autocorrelation (AC1), and Ljung-Box test statistic (LB4; H_0: first four autocorrelations are jointly zero) for the weekly excess returns, as well as the mean and standard deviation of the CFEAR characteristics. *, **, *** is significant at the 1%, 5%, and 10% level, respectively.

		First obs	Last obs	Excess return				CFEAR	
Commodity	**Sub-sector**	YYYYMMDD	YYYYMMDD	**Mean**	**StDev**	**AC1**	**LB4**	**Mean**	**StDev**
I. Agricultural sector (N=17)									
Corn	Cereal grains	20040105	20181231	-0.0671	0.2912	-0.0021	1.6121	-0.0349	0.0280
Oats	Cereal grains	20040105	20181231	0.0120	0.3475	-0.0339	7.8781 *	-0.0242	0.0295
Rough rice	Cereal grains	20040105	20181231	-0.0819	0.2488	0.0101	2.8643	-0.0208	0.0464
Wheat CBT	Cereal grains	20040105	20181231	-0.1227	0.3152	0.0129	0.6250	-0.0684	0.0329
Cotton no.2	Oilseeds	20040105	20181231	-0.0220	0.2872	0.0085	1.7628	0.0242	0.0379
Soybeans	Oilseeds	20040105	20181231	0.0525	0.2486	0.0256	0.9043	-0.0265	0.0217
Soybean meal	Oilseeds	20040105	20181231	0.1092	0.2872	0.0462	2.8353	-0.0653	0.0315
Soybean oil	Oilseeds	20040105	20181231	-0.0467	0.2460	-0.0176	2.3459	0.0498	0.0152
Feeder cattle	Meats	20040105	20181231	0.0270	0.1659	-0.0479	4.9823	0.0015	0.0220
Lean hogs	Meats	20040105	20150706	-0.0662	0.2377	0.0650	9.1910	0.0643	0.0261
Live cattle	Meats	20040105	20181231	-0.0075	0.1602	-0.0618	30.2330 ***	-0.0456	0.0119
Frozen pork bellies	Meats	20040105	20110705	-0.0228	0.2979	-0.0570	8.5047	-0.0660	0.0288
Cocoa	Misc. other softs	20040105	20181231	0.0253	0.2948	-0.0237	6.3451	-0.0797	0.0518
Coffee C	Misc. other softs	20040105	20181231	-0.0551	0.3115	0.0115	3.3936	-0.0752	0.0526
Frozen Orange juice	Misc. other softs	20040105	20181231	0.0176	0.3414	0.0344	10.0380 **	-0.0406	0.0503
Sugar no.11	Misc. other softs	20040105	20181231	-0.0417	0.3141	-0.0351	9.9182 **	-0.0284	0.0310
Lumber	Misc. other softs	20040105	20181231	-0.1229	0.3087	0.0074	3.6826	0.0209	0.0426
II. Energy sector (N=6)									
Light crude oil	Energy	20040105	20181231	-0.0753	0.3400	-0.0200	6.6687	-0.0007	0.0415
Electricity JPM	Energy	20040105	20150727	-0.1454	0.4428	0.0619	8.0159 *	0.0650	0.0732
Gasoline RBOB	Energy	20051010	20181231	-0.0305	0.3227	0.0404	14.1450 **	0.2356	0.3163
Heating oil	Energy	20040105	20181231	-0.0125	0.3095	0.0227	1.9867	-0.0179	0.0592
Natural gas	Energy	20040105	20181231	-0.3633	0.4224	-0.0102	3.7559	0.0626	0.0527
NY unleaded gas	Energy	20040105	20070102	0.1768	0.3686	-0.0146	1.9555	0.0533	0.0391
III. Metals (N=5)									
Copper (High Grade)	Base metals	20040105	20181231	0.0682	0.2720	0.0188	9.1223 *	-0.0191	0.0151
Gold 100oz (CMX)	Precious metals	20040105	20181231	0.0560	0.1785	-0.0090	3.1216	-0.0103	0.0362
Palladium	Precious metals	20040105	20181231	0.0988	0.3148	0.0220	0.7724	-0.0298	0.0553
Platinum	Precious metals	20040105	20181231	-0.0114	0.2302	0.0167	4.2287	-0.0258	0.0154
Silver 5000 oz	Precious metals	20040105	20181231	0.0421	0.3196	0.0117	2.2893	-0.0599	0.0638

Table 3. Descriptive statistics for CFEAR-sorted portfolios

The table summarizes the performance of the CFEAR quintiles and that of the long-short CFEAR portfolio. Q1 (Q5) is the quintile of commodities with the most negative (positive) CFEAR characteristic. Newey-West robust *t*-statistics are shown in parentheses for the mean. CER denotes certainty equivalent return based on power utility. The time period is January 2005 (week 1) to December 2018 (week 4).

	Long (Q1)	Q2	Q3	Q4	Short (Q5)	Q1-Q5
Mean	0.0435	-0.0210	-0.0125	-0.0391	-0.1421	0.0928
	(0.96)	(-0.46)	(-0.23)	(-0.87)	(-2.59)	(3.35)
StDev	0.1758	0.1658	0.1807	0.1615	0.1882	0.1030
Downside volatility (0%)	0.1141	0.1181	0.1267	0.1061	0.1305	0.0649
Skewness	-0.1094	-0.4580	-0.3327	-0.1203	-0.1210	-0.1307
	(-1.21)	(-5.06)	(-3.67)	(-1.33)	(-1.34)	(-1.44)
Excess Kurtosis	0.8700	1.9297	1.7840	0.7046	1.6227	0.4012
	(4.80)	(10.65)	(9.85)	(3.89)	(8.96)	(2.21)
JB normality test *p*-value	0.0010	0.0010	0.0010	0.0019	0.0010	0.0320
AC1	0.0097	0.0440	0.0311	0.0066	0.0437	0.0035
1% VaR (Cornish-Fisher)	0.0627	0.0702	0.0741	0.0584	0.0755	0.0341
% of positive months	54%	50%	50%	49%	47%	57%
Maximum drawdown	-0.4018	-0.5381	-0.6102	-0.6140	-0.8878	-0.1881
Sharpe ratio	0.2475	-0.1268	-0.0694	-0.2423	-0.7551	0.9012
Sortino ratio	0.3811	-0.1780	-0.0990	-0.3685	-1.0886	1.4299
Omega ratio	1.0926	0.9549	0.9748	0.9163	0.7579	1.3770
CER (power utility)	-0.0344	-0.0919	-0.0964	-0.1054	-0.2346	0.0660

Table 4. Time-series spanning tests

The table reports estimation results from time-series regressions of the excess returns of the long-short CFEAR portfolio onto various systematic risk factors. The base model is the commodity pricing model of Bakshi et al. (2019) which we augment with one additional risk factor at a time, and with all risk factors. Alongside the annualized alpha, we report the betas (risk exposures) with Newey West h.a.c. *t*-statistics in parentheses and the adjusted-R^2 of the regressions. The time period is January 2005 (week 1) to December 2018 (week 4).

	Base model	Base model augmented with							All risk factors
		Fundamental risk factors		Tail risk factors			Liquidity and volatility risk factors		
Annualized alpha	0.0943	0.0947	0.0881	0.0898	0.0933	0.0912	0.0891	0.0932	0.0823
	(3.69)	(3.73)	(3.48)	(3.48)	(3.63)	(3.59)	(3.35)	(3.92)	(3.35)
AVG	-0.0110	-0.0104	-0.0129	-0.0132	0.0417	0.0426	-0.0068	0.0001	0.0249
	(-0.32)	(-0.30)	(-0.37)	(-0.37)	(1.04)	(1.08)	(-0.19)	(0.00)	(0.65)
Basis	-0.1730	-0.1715	-0.2341	-0.1928	-0.1639	-0.1776	-0.1873	-0.1245	-0.2023
	(-2.89)	(-2.82)	(-3.97)	(-3.26)	(-2.82)	(-3.04)	(-3.20)	(-2.31)	(-3.82)
Momentum	0.2756	0.2778	0.2951	0.2726	0.2568	0.2674	0.2435	0.2190	0.2253
	(6.12)	(5.60)	(6.43)	(5.87)	(5.80)	(6.25)	(5.60)	(5.19)	(5.02)
Hedging pressure		-0.0094					0.1218	-0.3077	-0.0217
		(-0.17)					2.0390	-6.9259	(-0.46)
Convexity			0.1640						0.1504
			(2.99)						(3.09)
Skewness				0.1157					0.0761
				(2.24)					(1.55)
VaR1					-0.1167				-0.0175
					(-2.01)				(-0.25)
VaR99						0.1330			0.0445
						(2.58)			(0.65)
Basis-momentum							0.1218		0.0583
							(2.04)		(1.09)
Liquidity								-0.3077	-0.2741
								(-6.93)	(-6.37)
Adj.-R^2 (%)	8.48	8.37	10.28	9.56	9.50	10.04	9.48	16.23	18.54

Table 5. Cross-sectional pricing tests

The table reports the (annualized) prices of risk from cross-sectional regressions of average portfolio excess returns on full-sample betas with Shanken (1992) errors-in-variables corrected *t*-statistics in parentheses, and Kan et al. (2013) *t*-statistics additionally corrected for model misspecification in curly brackets. The base model is the commodity pricing model of Bakshi et al. (2019) which we augment with one additional risk factor at a time, and with all risk factors. The 56 test assets are the quintiles based on the CFEAR signal, alternative 9 signals listed in Appendix A, Panel B, and equally-weighted and weekly-rebalanced portfolios of commodities in all 6 sectors. The two last rows report the adjusted-R^2 and MAPE (mean absolute pricing error) of each model. The time period is January 2005 (week 1) to December 2018 (week 4).

	CFEAR	Base model		Base model augmented with														All risk factors	
				Fundamental risk factors				Tail risk factors						Liquidity and volatility risk factors					
Constant	-0.0006	-0.0002	-0.0007	-0.0001	-0.0005	-0.0001	-0.0006	-0.0004	-0.0007	-0.0024	-0.0018	-0.0020	-0.0017	-0.0006	-0.0008	-0.0004	-0.0007	-0.0014	-0.0014
	(-0.86)	(-0.24)	(-0.83)	(-0.06)	(-0.63)	(-0.15)	(-0.77)	(-0.46)	(-0.93)	(-2.13)	(-1.63)	(-1.93)	(-1.63)	(-0.69)	(-0.99)	(-0.47)	(-0.83)	(-1.26)	(-1.29)
	{-0.83}	{-0.24}	{-0.88}	{-0.06}	{-0.70}	{-0.15}	{-0.80}	{-0.49}	{-1.04}	{-1.78}	{-1.46}	{-1.60}	{-1.47}	{-0.70}	{-1.06}	{-0.52}	{-0.89}	{-1.04}	{-1.17}
CFEAR	0.0894		0.0928		0.0916		0.0913		0.0889		0.0808		0.0814		0.0868		0.0901		0.0800
	(2.56)		(2.79)		(2.75)		(2.75)		(2.69)		(2.58)		(2.54)		(2.68)		(2.92)		(2.65)
	{2.48}		{2.72}		{2.68}		{2.74}		{2.63}		{2.54}		{2.58}		{2.61}		{3.09}		{2.69}
AVG		-0.0224	0.0022	-0.0299	-0.0056	-0.0260	-0.0001	-0.0131	0.0065	0.0918	0.0613	0.0712	0.0557	-0.0040	0.0087	-0.0129	0.0023	0.0394	0.0415
		(-0.40)	(0.04)	(-0.53)	(-0.10)	(-0.47)	(-0.00)	(-0.24)	(0.12)	(1.34)	(0.90)	(1.11)	(0.86)	(-0.07)	(0.16)	(-0.23)	(0.04)	(0.59)	(0.62)
		{-0.37}	{0.04}	{-0.48}	{-0.10}	{-0.43}	{-0.00}	{-0.21}	{0.12}	{1.01}	{0.77}	{0.85}	{0.73}	{-0.07}	{0.16}	{-0.24}	{0.04}	{0.43}	{0.49}
Basis		0.0502	0.0745	0.0406	0.0643	0.0417	0.0708	0.0361	0.0621	0.0581	0.0751	0.0479	0.0691	0.0502	0.0723	0.0653	0.0753	0.0340	0.0490
		(1.58)	(2.44)	(1.32)	(2.18)	(1.28)	(2.31)	(1.15)	(2.12)	(1.84)	(2.46)	(1.50)	(2.29)	(1.58)	(2.36)	(2.12)	(2.46)	(1.17)	(1.73)
		{1.63}	{2.72}	{1.33}	{2.41}	{1.36}	{2.63}	{1.14}	{2.24}	{1.87}	{2.64}	{1.44}	{2.43}	{1.65}	{2.62}	{2.16}	{2.74}	{1.19}	{1.85}
Momentum		0.0846	0.0454	0.0703	0.0288	0.0822	0.0470	0.0668	0.0364	0.0575	0.0366	0.0562	0.0344	0.0650	0.0396	0.0586	0.0439	0.0327	0.0205
		(2.21)	(1.29)	(1.94)	(0.88)	(2.16)	(1.33)	(1.87)	(1.08)	(1.65)	(1.07)	(1.64)	(1.03)	(1.78)	(1.13)	(1.71)	(1.28)	(1.01)	(0.64)
		{2.20}	{1.37}	{1.84}	{0.93}	{2.19}	{1.41}	{1.88}	{1.19}	{1.62}	{1.16}	{1.67}	{1.14}	{1.80}	{1.22}	{1.72}	{1.37}	{1.09}	{0.71}
Hedging pressure				0.0616	0.0607													0.0485	0.0536
				(1.87)	(1.84)													(1.61)	(1.78)
				{1.71}	{1.77}													{1.56}	{1.75}
Convexity						0.0647	0.0478											0.0589	0.0503
						(2.45)	(1.87)											(2.29)	(1.97)
						{2.27}	{1.89}											{2.19}	{1.90}
Skewness								0.0707	0.0662									0.0557	0.0559
								(2.15)	(2.01)									(1.88)	(1.89)
								{1.87}	{1.88}									{1.83}	{1.90}
VaR1										-0.0517	-0.0295							-0.0382	-0.0276
										(-1.53)	(-0.90)							(-1.20)	(-0.88)
										{-1.47}	{-0.92}							{-1.26}	{-0.94}
VaR99												0.0570	0.0424					0.0321	0.0265
												(1.72)	(1.29)					(1.00)	(0.83)
												{1.69}	{1.32}					{0.97}	{0.84}
Basis-momentum														0.0866	0.0572			0.0634	0.0547
														(2.81)	(2.06)			(2.34)	(2.05)
														{2.57}	{1.85}			{2.19}	{1.96}
Liquidity																-0.0573	-0.0282	-0.0263	-0.0090
																(-1.81)	(-0.94)	(-0.91)	(-0.32)
																{-1.53}	{-0.82}	{-0.75}	{-0.28}
Adj.-R^2 (%)	41.01	32.42	62.58	37.74	69.09	40.64	63.07	44.89	68.65	48.43	66.71	49.89	68.36	42.36	64.56	44.54	62.72	65.90	77.19
MAPE (%)	0.048	0.049	0.039	0.048	0.035	0.047	0.039	0.045	0.035	0.045	0.038	0.044	0.037	0.047	0.038	0.046	0.039	0.036	0.03

Table 6. Properties of CFEAR commodity quintiles

The table summarizes the properties of CFEAR-based commodity quintiles. Q1 is the quintile of commodities with the lowest CFEAR characteristics and Q5 is the quintile of commodities with the highest CFEAR characteristics. The characteristics other than CFEAR are measured over their relevant windows as listed in Appendix A and are subsequently averaged across constituents and over time. Realized variance is the average squared daily return over the 22 days preceding portfolio formation time. Dollar open interest is the product of the number of outstanding contracts, contract size and front-end futures settlement price (/10^10). The momentum, basis-momentum and variance characteristics are annualized. The last column shows Newey-West *t*-statistics for the null hypothesis of no difference in a given characteristic across the Q1 and Q5 quintiles. The sampling period is January 2005 to December 2018.

	CFEAR					
	Long (Q1)	Q2	Q3	Q4	Short (Q5)	Q1-Q5
CFEAR	-0.0807	-0.0459	-0.0243	0.0071	0.0962	-0.1770
	(-17.80)	(-14.67)	(-12.62)	(6.76)	(16.59)	(-19.94)
Basis	-0.0076	-0.0088	-0.0082	-0.0082	-0.0131	0.0055
	(-6.06)	(-9.76)	(-7.76)	(-8.75)	(-5.21)	(2.02)
Momentum	0.0228	0.0130	-0.0040	-0.0427	-0.1160	0.1387
	(1.34)	(0.92)	(-0.18)	(-2.50)	(-4.92)	(8.10)
Hedging pressure	0.2842	0.3454	0.2952	0.2552	0.2254	0.0588
	(29.19)	(31.00)	(19.00)	(23.08)	(23.81)	(4.16)
Convexity (x1,000)	0.0482	-0.0354	-0.0343	-0.0790	-0.1154	0.1636
	(1.22)	(-1.32)	(-1.75)	(-3.38)	(-1.22)	(1.50)
Skewness	-0.0341	0.0965	0.1045	0.1509	-0.0158	-0.0183
	(-1.31)	(3.30)	(5.07)	(6.32)	(-0.78)	(-0.69)
VaR1	-0.0465	-0.0442	-0.0459	-0.0420	-0.0485	0.0020
	(-51.67)	(-47.18)	(-42.96)	(-45.19)	(-47.46)	(1.66)
VaR99	0.0456	0.0411	0.0420	0.0402	0.0474	-0.0018
	(74.07)	(55.81)	(51.66)	(43.90)	(50.64)	(-1.93)
Basis-momentum	0.0143	-0.0017	-0.0089	-0.0176	-0.0073	0.0216
	(3.50)	(-0.61)	(-2.71)	(-9.53)	(-1.85)	(4.14)
Liquidity	4.01	1.28	1.90	1.79	53.10	-49.09
	(3.78)	(3.97)	(3.23)	(4.36)	(5.04)	(-4.59)
Realized variance	0.0505	0.0245	0.0299	0.0260	0.1992	-0.1487
	(3.87)	(5.51)	(4.98)	(7.54)	(6.83)	(-5.03)
Dollar open interest	57.49	65.00	83.56	22.91	23.97	33.53
	(23.19)	(13.39)	(19.22)	(14.03)	(19.49)	(11.42)

Table 7. CFEAR effect over time

This table reports in Panel A (Panel B) the annualized mean excess return and annualized alpha from Bakshi et al. (2019) benchmark model for the long, short and long-short (LS) CFEAR portfolios in high vs. low VIX states (Panel A) and high vs. low TED states (Panel B) using the full sample average as cut-point. The last row of each panel presents *t*-statistics for the significance of differences between the high and low regimes. Panel C presents summary statistics for the long-short CFEAR portfolios formed at the end of each week day (Monday to Friday). Newey-West robust *t*-statistics are shown in parentheses. The time period is January 2005 (week 1) to December 2018 (week 4).

Panel A: CFEAR in high and low VIX states

	Mean excess return			Alpha		
	Long	Short	LS	Long	Short	LS
I. High VIX	0.0635	-0.3068	0.1852	0.1030	-0.2800	0.1916
	(0.65)	(-2.82)	(3.92)	(1.82)	(-4.60)	(4.44)
II. Low VIX	0.0341	-0.0526	0.0433	0.0662	-0.0190	0.0426
	(0.71)	(-0.94)	(1.32)	(1.80)	(-0.48)	(1.39)
t-stat (H_0: diff=0)	0.27	-2.10	2.48	0.54	-3.60	2.83

Panel B: CFEAR in high and low TED states

	Mean excess return			Alpha		
	Long	Short	LS	Long	Short	LS
I. High TED	0.0539	-0.2199	0.1369	0.1104	-0.1605	0.1355
	(0.52)	(-1.78)	(2.25)	(1.59)	(-2.21)	(2.32)
II. Low TED	0.0408	-0.1124	0.0766	0.0675	-0.0916	0.0796
	(0.86)	(-1.90)	(2.58)	(2.06)	(-2.35)	(2.97)
t-stat (H_0: diff=0)	0.12	-0.78	0.90	0.57	-0.83	0.88

Panel C: CFEAR portfolio performance and choice of portfolio formation day

	Monday-end			Tue-end	Wed-end	Thu-end	Friday-end		
	Long	Short	LS				Long	Short	LS
Mean	0.0435	-0.1421	0.0928	0.0868	0.0649	0.0623	-0.0162	-0.0996	0.0417
	(0.96)	(-2.59)	(3.35)	(3.03)	(2.35)	(2.38)	(-0.38)	(-1.66)	(1.41)
StDev	0.1758	0.1882	0.1030	0.1037	0.1115	0.1041	0.1477	0.2282	0.1160
Downside volatility (0%)	0.1141	0.1305	0.0649	0.0616	0.0686	0.0628	0.1110	0.1607	0.0788
Skewness	-0.1094	-0.1210	-0.1307	-0.0581	0.1036	-0.0066	-0.6790	0.0782	-0.3250
	(-1.21)	(-1.34)	(-1.44)	(-0.64)	(1.14)	(-0.07)	(-7.49)	(0.86)	(-3.58)
Excess Kurtosis	0.8700	1.6227	0.4012	-0.0926	0.6568	0.3547	3.2189	3.8769	3.2562
	(4.80)	(8.96)	(2.21)	(-0.51)	(3.62)	(1.96)	(17.75)	(21.38)	(17.96)
JB normality test p-value	0.0010	0.0010	0.0320	0.5000	0.0034	0.1334	0.0010	0.0010	0.0010
1% VaR (Cornish-Fisher)	0.0627	0.0755	0.0341	0.0321	0.0359	0.0336	0.0701	0.1023	0.0521
% of positive months	54%	47%	57%	54.2%	53%	54%	51%	47%	53%
Maximum drawdown	-0.4018	-0.8878	-0.1881	-0.1704	-0.1530	-0.1710	-0.5354	-0.8406	-0.2467
Sharpe ratio	0.2475	-0.7551	0.9012	0.8373	0.5816	0.5984	-0.1095	-0.4365	0.3598
Sortino ratio	0.3811	-1.0886	1.4299	1.4093	0.9453	0.9909	-0.1457	-0.6198	0.5293
Omega ratio	1.0926	0.7579	1.3770	1.3372	1.2334	1.2336	0.9599	0.8470	1.1408
CER (power utility)	-0.0344	-0.2346	0.0660	0.0598	0.0338	0.0352	-0.0729	-0.2339	0.0077

Table 8. Alternative CFEAR signal construction methods

The table summarizes the long-short CFEAR portfolio under different signal construction methods.: (1) using US searches from Google Trends; (2) winsorizing the hazard-attention variable $\Delta S^{*}_{j,t}$; (3) deseasonalizing the hazard-attention variable; (4) accounting for different treatments of the zeros in $\Delta S^{*}_{j,t}$ by replacing them by ones in (4a), by using $\Delta S_{j,t} = \ln\left(\frac{1+s_{j,t}}{1+s_{j,t-1}}\right)$ instead of $\Delta S_{j,t} = \ln\left(\frac{s_{j,t}}{s_{j,t-1}}\right)$ in (4b) or by removing the zeros in (4c); (5) excluding noisy keywords with a percentage of zeros (τ) of at least 20% or for which the average correlation amongst the six series $\{s_{j,t}\}_d$ that form $s_{j,t}$ on six different days d is less than κ=80%; (6) considering six alternative search dates (12th, 13th, 16th, 17th, 18th and 20th December 2019); (7) controling for media coverage under relevance scores 75 and 25. The time period is January 2005 (week 1) to December 2018 (week 4).

	(1)	(2)	(3)	(4)			(5)	(6)	(7)	
	US searches	Winsorized searches	Deseasonal. searches	0 searches filtering			Excluding noisy keywords (τ=20%, κ=80%)	GSVI series obtained 12th to 20th Dec 2019	Media coverage	
				(4a)	(4b)	(4c)			Relevance 75	Relevance 25
Mean	0.0738	0.0795	0.0606	0.0742	0.0813	0.0827	0.0959	0.0686	0.0835	0.0891
	(2.97)	(2.96)	(2.50)	(2.86)	(3.19)	(3.06)	(3.24)	(2.49)	(3.19)	(3.25)
StDev	0.0916	0.1034	0.0982	0.0960	0.0947	0.0993	0.1100	0.1101	0.0989	0.0993
Downside volatility (0%)	0.0570	0.0656	0.0616	0.0599	0.0580	0.0624	0.0686	0.0713	0.0596	0.0625
Skewness	0.0189	-0.0990	-0.0889	0.0262	0.0506	-0.0620	-0.1288	-0.0985	0.0164	-0.0685
	(0.21)	(-1.09)	(-0.98)	(0.29)	(0.56)	(-0.68)	(-1.42)	(-1.09)	(0.18)	(-0.76)
Excess Kurtosis	0.5735	0.6454	0.4069	0.5727	0.5459	0.7693	0.6672	0.8922	0.3767	0.3967
	(3.16)	(3.56)	(2.25)	(3.16)	(3.01)	(4.25)	(3.68)	(4.92)	(2.08)	(2.19)
JB normality test p-value	0.0110	0.0038	0.0478	0.0109	0.0136	0.0014	0.0025	0.0010	0.1027	0.0638
1% VaR (Cornish-Fisher)	0.0297	0.0350	0.0327	0.0311	0.0302	0.0335	0.0374	0.0384	0.0314	0.0323
% of positive months	56%	57%	55%	55.8%	56.2%	54%	54%	55%	55%	56%
Maximum drawdown	-0.1537	-0.1432	-0.1635	-0.1574	-0.1317	-0.1829	-0.1534	-0.1891	-0.1336	-0.1263
Sharpe ratio	0.8058	0.7695	0.6175	0.7732	0.8589	0.8329	0.8717	0.6232	0.8435	0.8973
Sortino ratio	1.2948	1.2129	0.9841	1.2387	1.4013	1.3255	1.3973	0.9617	1.4003	1.4269
Omega ratio	1.3389	1.3167	1.2450	1.3219	1.3612	1.3510	1.3606	1.2540	1.3495	1.3782
CER (power utility)	0.0528	0.0527	0.0364	0.0511	0.0588	0.0579	0.0654	0.0382	0.0589	0.0643

Table 9. Placebo test

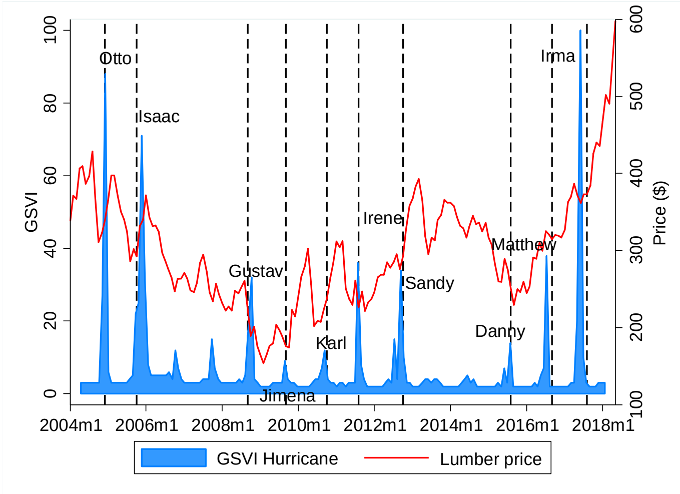
The table reports summary statistics for the long-short hazard-fear portfolios based on the 123 query terms confined to the weather (WE) and agricultural disease (DI) categories. The cross sections are as detailed in Table 2 (28 commodity futures) and in the online Annex Table A.10 (40 equity index futures, 13 fixed income futures, 19 currency futures). The weekly portfolio returns cover the time period from January 2005 (week 1) to December 2018 (week 4).

	Commodity	Equity index	Fixed income	Currency
Mean	0.0817	0.0183	0.0019	0.0116
	(3.06)	(1.62)	(0.25)	(1.50)
StDev	0.1017	0.0473	0.0277	0.0321
Downside volatility (0%)	0.0643	0.0332	0.0187	0.0212
Skewness	-0.0691	-0.0971	0.0454	0.1272
	(-0.76)	(-1.07)	(0.50)	(1.40)
Excess Kurtosis	0.6688	3.7939	2.9409	2.5147
	(3.69)	(20.94)	(16.23)	(13.88)
JB normality test *p*-value	0.0035	0.0010	0.0010	0.0010
1% VaR (Cornish-Fisher)	0.0341	0.0212	0.0114	0.0123
% of positive months	56%	53%	50%	52%
Maximum drawdown	-0.1626	-0.1151	-0.0627	-0.0613
Sharpe ratio	0.8034	0.3864	0.0674	0.3619
Sortino ratio	1.2711	0.5507	0.0998	0.5487
Omega ratio	1.3352	1.1585	1.0261	1.1446
CER (power utility)	0.0557	0.0127	0.0000	0.0090

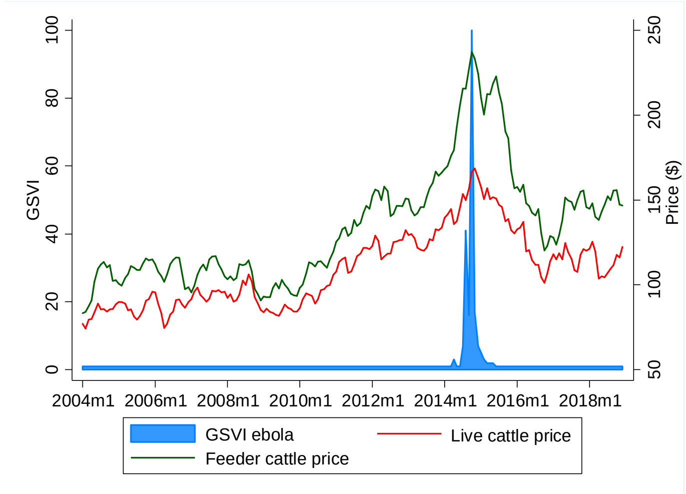
Figure 1. Google searches and commodity prices

The graphs plots the evolution of monthly search intensity or attention to *hurricane*, *ebola*, *oil crisis* and *unemployment* hazards as captured by the Google Search Volume Index (GSVI; denoted $s_{j,t}$ in the paper), alongside the monthly average of the daily commodity futures price.

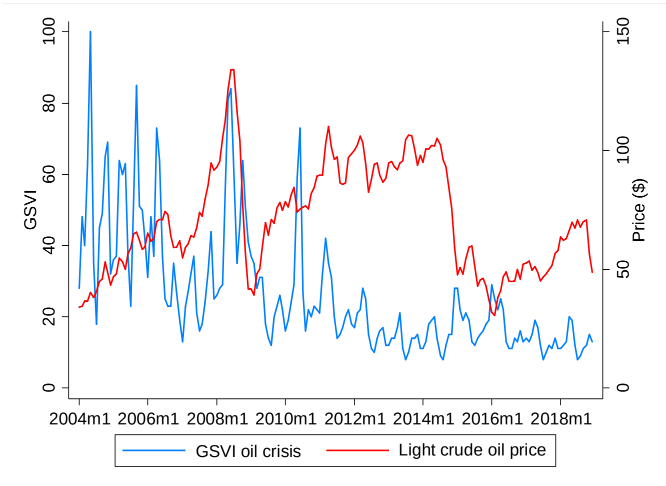
Panel A: *hurricane* (WE) searches vs lumber price



Panel B: *ebola* (DI) searches vs feeder/live cattle prices



Panel C: *oil crisis* (GP) searches vs light crude oil price



Panel D: *unemployment* (EC) searches vs natural gas price

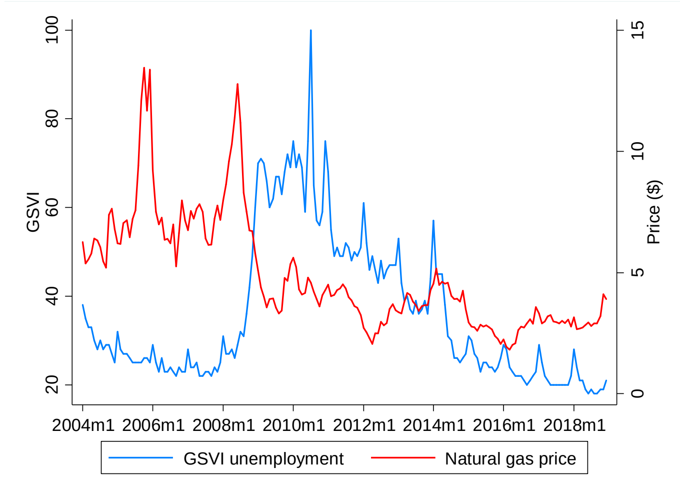


Figure 2. Constituents of long and short legs of the CFEAR portfolio

The graph plots the percentage of sample weeks from January 2005 (week 1) to December 2018 (week 4) that allocate each of the *N*=28 commodities to the top quintile (Q5) or bottom quintile (Q1) according to the CFEAR signal. The results are organized by sector.

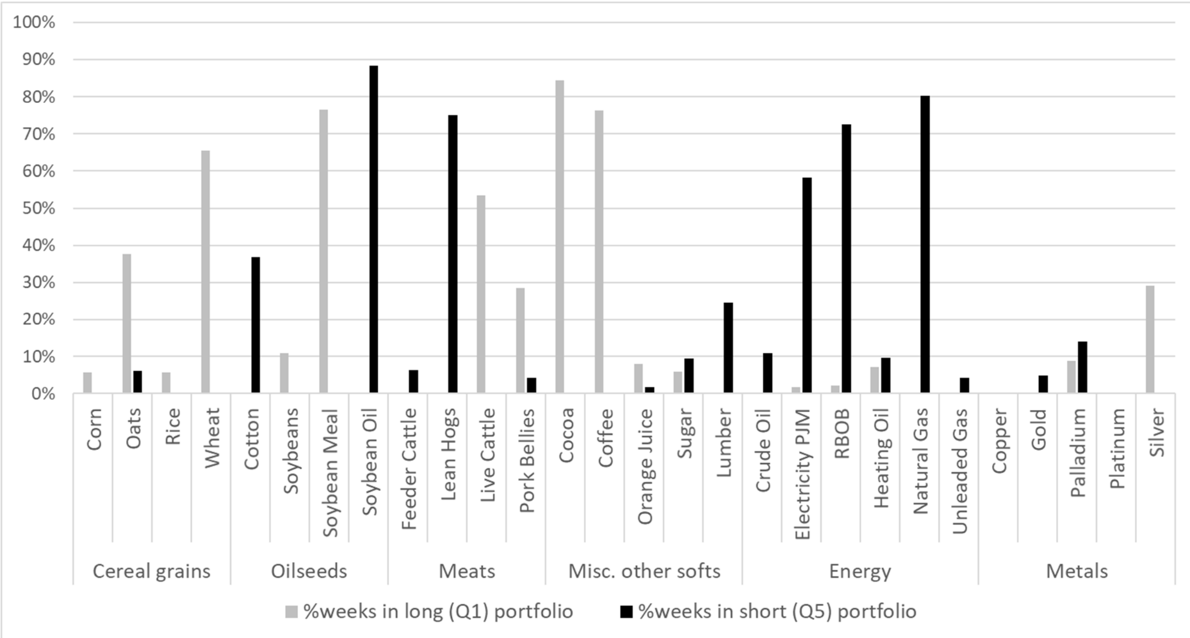


Figure 3. Future value of $1 invested in long-short and long-only commodity portfolios
The graph shows the evolution of $1 invested in the long-short portfolios based on the CFEAR signal (dark black line), on the alternative signals listed in Appendix A, alongside the evolution of $1 invested in a long-only portfolio that equally weights all commodities, AVG. The portfolio rebalancing frequency is weekly. Total returns (excess plus risk free rate) are plotted.

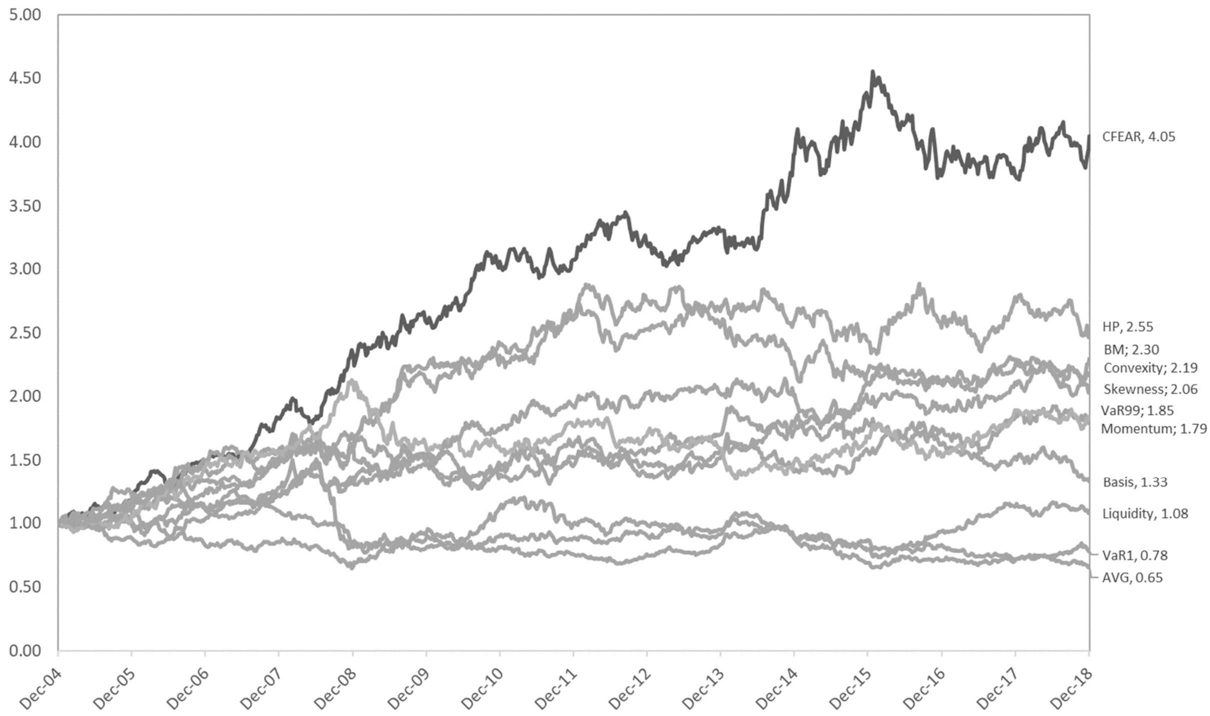
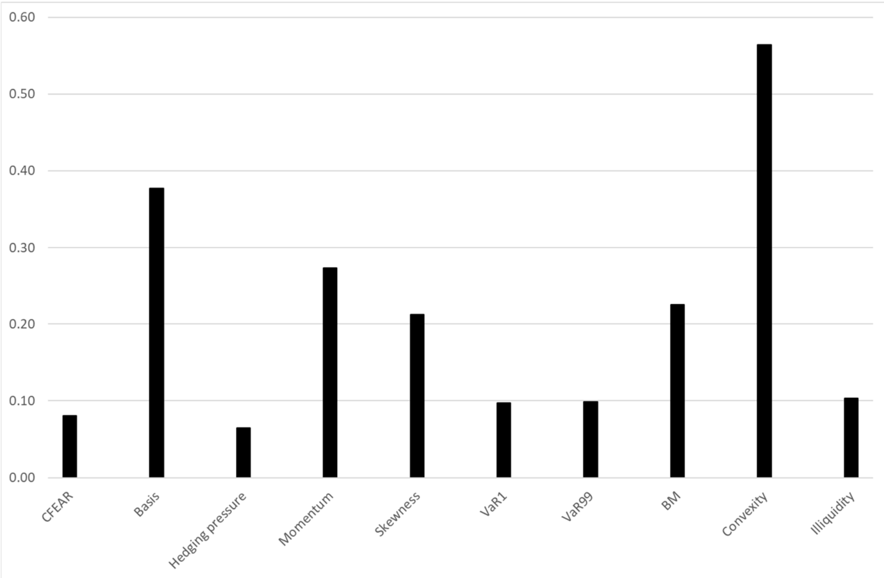


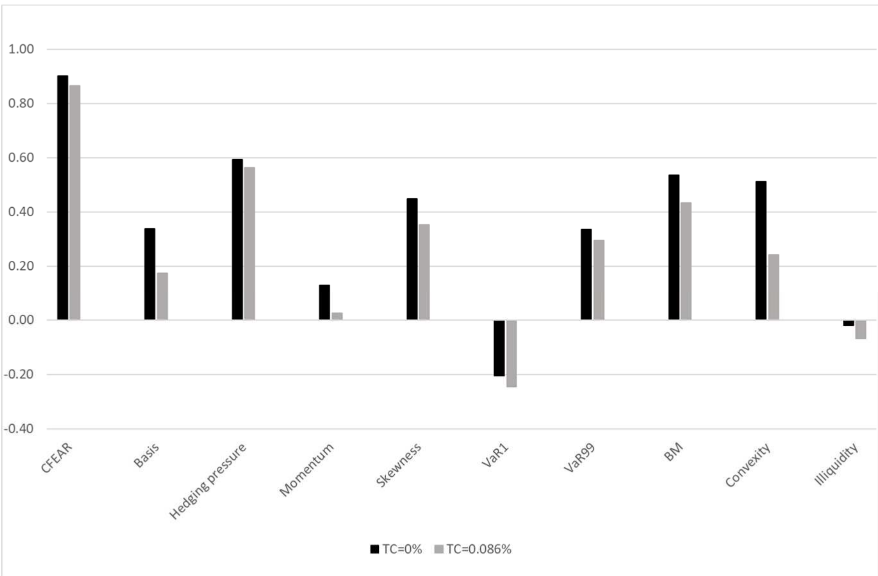
Figure 4. Turnover and net performance of commodity portfolios

Panel A plots the turnover of each of the long-short portfolios formed according to the CFEAR signals and alternative signals listed in Appendix A. Panel B plots the Sharpe ratios of each of the portfolios before and after proportional trading costs (TC) of 8.6 bps (Marshall et al., 2012).

Panel A: Turnover



Panel B: Sharpe ratio



ONLINE ANNEX

Fear of Hazards in Commodity Futures Markets

May 2020

Table A.1 Summary of Google search volume indices (GSVIs)

Each of the series summarized in this table is an average of the six weekly GSVI series per keyword $j = 1, \dots, 149$ obtained on six different dates (denoted $s_{j,t}$ in the paper) from *Google Trends*. AC1 is the first-order autocorrelation, LB4 is the Ljung-Box test statistic (H_0: first four autocorrelations are jointly zero), ADF is the Augmented Dickey-Fuller test statistic (H_0: unit root non-stationary). The observation period is January 2004 (week 1) to December 2018 (week 4).

	Search keywords	Mean	Median	St.Dev	Min	Max	% zeros	AC1	LB4	ADF
Weather (WE; 113 keywords)										
1	adverse weather	8.62	5.00	10.37	0.00	100.00	4.34%	0.346	243.14 ***	-13.040 ***
2	adverse weather conditions	7.54	5.33	10.24	0.00	100.00	20.05%	0.337	146.39 ***	-19.850 ***
3	adverse weather warning	1.88	0.00	7.79	0.00	90.33	89.14%	0.124	23.48 ***	-16.700 ***
4	blizzard	6.05	3.50	8.6	1.00	100.00	0.00%	0.434	325.75 ***	-7.500 ***
5	blizzard risk	9.88	0.00	13.56	0.00	79.67	51.09%	0.271	138.34 ***	-3.810 ***
6	blizzard warning	4.1	0.83	10.07	0.00	100.00	26.95%	0.198	54.04 ***	-15.930 ***
7	catastrophic events	29.24	25.83	18.56	0.00	99.50	3.96%	0.693	1157.22 ***	-6.170 ***
8	catastrophic weather	15.14	13.33	14.77	0.00	99.17	27.59%	0.277	189.53 ***	-10.880 ***
9	catastrophic weather events	8.6	0.00	12.55	0.00	69.05	57.09%	0.270	166.41 ***	-7.980 ***
10	climate change	36.15	32.83	14.32	9.17	100.00	0.00%	0.864	1677.29 ***	-5.490 ***
11	climate disturbance	12.76	9.00	13.97	0.00	82.83	38.95%	0.366	371.43 ***	-7.370 ***
12	cold	23.43	20.33	12.35	8.50	100.00	0.00%	0.688	978.29 ***	-7.820 ***
13	cold spell	16.08	13.50	12.12	0.00	100.00	6.64%	0.543	703.55 ***	-6.740 ***
14	cold weather	17.55	11.83	14.93	4.33	100.00	0.00%	0.792	1372.27 ***	-6.950 ***
15	cyclogenesis	12.52	8.67	12.12	0.00	100.00	9.71%	0.334	271.21 ***	-5.965 ***
16	cyclone	9.37	7.67	8.93	3.17	100.00	0.00%	0.311	142.65 ***	-14.748 ***
17	cyclone risk	12.01	10.17	12.19	0.00	94.00	24.78%	0.299	186.67 ***	-7.731 ***
18	cyclone warning	6.36	3.00	10.25	0.00	100.00	5.11%	0.277	102.55 ***	-12.072 ***
19	drought	33.64	30.50	13.42	8.33	100.00	0.00%	0.890	1912.54 ***	-6.695 ***
20	drought risk	19.82	18.17	16.2	0.00	81.33	17.24%	0.506	749.60 ***	-3.913 ***
21	drought warning	15.52	13.33	14.35	0.00	77.67	26.05%	0.264	193.68 ***	-9.021 ***
22	droughts	34.49	30.83	16.93	6.67	100.00	0.00%	0.835	1698.90 ***	-6.473 ***
23	dry weather	51.08	50.60	13.95	0.00	100.00	0.13%	0.697	1381.49 ***	-4.162 ***
24	el niño weather	14.36	12.00	14.09	0.00	99.00	23.50%	0.526	759.32 ***	-5.038 ***
25	extreme cold	15.58	12.17	12.32	0.00	100.00	0.13%	0.700	1048.59 ***	-7.636 ***
26	extreme cold temperatures	10.72	8.67	11.65	0.00	100.00	22.99%	0.434	350.43 ***	-10.477 ***
27	extreme heat	20.24	17.5	11.43	2.33	100.00	0.00%	0.556	529.87 ***	-9.870 ***
28	extreme rain	36.55	38.33	18.03	0.00	95.00	2.17%	0.618	1187.79 ***	-2.876 **
29	extreme temperatures	20.25	19.33	10.63	0.00	100.00	2.17%	0.401	345.35 ***	-8.749 ***
30	extreme weather	33.58	31.67	13.88	12.67	100.00	0.00%	0.733	1351.06 ***	-4.264 ***
31	extreme wind	29.99	28.83	14.62	0.00	100.00	1.53%	0.546	900.89 ***	-4.740 ***
32	flood	18.79	15.00	10.67	6.67	100.00	0.00%	0.647	863.55 ***	-6.681 ***
33	flood risk	18.02	17.67	8.68	0.00	100.00	0.13%	0.389	241.14 ***	-9.074 ***
34	flood warning	7.83	4.33	10.76	0.00	100.00	0.38%	0.378	199.40 ***	-9.316 ***
35	flooding	11.13	8.33	12.77	1.33	100.00	0.00%	0.386	164.14 ***	-18.565 ***
36	floods	9.16	7.83	9.81	1.00	100.00	0.00%	0.501	428.60 ***	-7.950 ***
37	forest fire	19.91	15.83	16.02	2.83	100.00	0.00%	0.790	1567.43 ***	-5.794 ***
38	forest fires	15.62	10.67	15.02	1.33	100.00	0.00%	0.634	806.28 ***	-9.430 ***
39	freeze warning	10.88	6.17	13.57	0.00	98.67	14.43%	0.229	127.47 ***	-11.446 ***
40	frost	35.62	34.00	12.78	3.00	100.00	0.00%	0.630	972.80 ***	-7.183 ***
41	frost risk	16.53	14.17	14.58	0.00	91.67	17.11%	0.469	455.58 ***	-8.924 ***
42	frost warning	10.48	7.17	12.03	0.00	100.00	13.67%	0.353	153.50 ***	-11.890 ***
43	frosts	7.49	0.00	12.25	0.00	67.00	62.45%	0.194	62.23 ***	-22.950 ***
44	global warming	43.02	39.00	18.91	10.83	100.00	0.00%	0.920	2189.03 ***	-4.227 ***
45	gust	10.68	8.83	10.21	0.00	100.00	11.24%	0.098	11,80** ***	-25.311 ***
46	gusts	8.81	7.33	8.96	0.00	100.00	16.60%	0.135	17.72 ***	-24.358 ***
47	hail	49.68	45	15.60	28.50	100.00	0.00%	0.855	2101.77 ***	-3.437 ***
48	hail damage	24.83	21.17	15.61	0.00	100.00	0.26%	0.662	1009.20 ***	-8.761 ***
49	hail risk	10.56	4.83	13.65	0.00	66.50	45.98%	0.342	332.51 ***	-3.763 ***

(*cont.*) **Table A.1 Summary of Google search volume indices (GSVIs)**

	Search keywords	Mean	Median	St.Dev	Min	Max	% zeros	AC1	LB4	ADF
Weather (WE; 113 keywords)										
50	hail storm	17.68	15.17	11.70	2.33	100.00	0.00%	0.383	275.50 ***	-10.497 ***
51	hail storm warning	4.44	0	8.87	0.00	100.00	50.83%	0.045	33.51 ***	-12.638 ***
52	hail warning	8.97	7.5	9.70	0.00	100.00	23.75%	0.113	32.19 ***	-24.845 ***
53	harmattan wind	13.02	9	14.75	0.00	91.33	39.46%	0.277	191.05 ***	-4.139 ***
54	heat	10.91	6.83	12.94	1.33	100.00	0.00%	0.743	1217.86 ***	-7.904 ***
55	heat wave	15.12	12.5	9.83	5.50	100.00	0.00%	0.667	894.82 ***	-8.037 ***
56	heat waves	18.84	17.17	11.71	0.00	96.50	0.26%	0.275	159.64 ***	-8.153 ***
57	heatwave	12.97	10.17	10.60	4.00	100.00	0.00%	0.579	483.67 ***	-14.416 ***
58	heatwaves	40.52	41.33	16.71	0.00	100.00	3.83%	0.736	1512.80 ***	-4.603 ***
59	heavy rain	11.58	9.17	10.33	0.00	100.00	5.24%	0.458	448.48 ***	-11.105 ***
60	heavy rain fall	19.16	19.00	15.53	0.00	94.00	23.63%	0.497	703.05 ***	-5.221 ***
61	heavy rain risk	7.37	0.00	13.89	0.00	72.67	69.86%	0.161	114.41 ***	-6.815 ***
62	heavy rain warning	11.44	4.83	15.01	0.00	94.00	48.28%	0.312	272.57 ***	-6.106 ***
63	high temperature	66.09	67.17	11.85	31.00	100.00	0.00%	0.840	1923.07 ***	-3.666 ***
64	high temperatures	45.55	44.00	14.33	13.00	100.00	0.00%	0.688	1011.00 ***	-5.155 ***
65	hot weather	25.77	23.17	10.88	10.17	100.00	0.00%	0.747	1212.33 ***	-7.233 ***
66	hurricane	4.44	2.00	10.97	1.00	100.00	0.00%	0.556	427.86 ***	-14.905 ***
67	hurricane risk	6.14	3.83	9.14	0.00	100.00	5.49%	0.437	316.72 ***	-12.428 ***
68	hurricane warning	4.61	2.00	9.92	0.00	100.00	0.77%	0.425	269.08 ***	-10.873 ***
69	hurricanes	10.88	7.50	12.16	1.17	100.00	0.00%	0.644	790.92 ***	-9.473 ***
70	natural disaster	25.52	28.67	14.87	3.00	100.00	0.00%	0.895	2063.86 ***	-4.452 ***
71	natural hazard	37.49	37.50	15.56	4.83	100.00	0.00%	0.651	1008.64 ***	-5.124 ***
72	rain	38.53	35.50	14.06	15.50	100.00	0.00%	0.787	1648.40 ***	-3.989 ***
73	severe heat	26.85	26.83	16.85	0.00	91.67	7.79%	0.726	1497.97 ***	-3.935 ***
74	severe weather	27.43	24.33	13.42	7.33	100.00	0.00%	0.465	547.87 ***	-6.704 ***
75	severe weather risk	12.99	9.33	14.92	0.00	86.83	34.10%	0.490	743.06 ***	-4.211 ***
76	snow	18.47	8.00	19.34	2.00	100.00	0.00%	0.846	1790.52 ***	-8.864 ***
77	snow risk	11.85	7.33	13.23	0.00	100.00	13.03%	0.458	373.43 ***	-6.639 ***
78	snow storm warning	5.71	1.00	11.75	0.00	100.00	45.08%	0.329	267.69 ***	-8.331 ***
79	snow warning	6.73	3.17	10.45	0.00	100.00	7.92%	0.442	370.42 ***	-9.642 ***
80	storm	15.64	11.83	13.35	3.00	100.00	0.00%	0.452	316.51 ***	-10.540 ***
81	storm risk	15.3	12.33	12.13	0.00	100.00	2.81%	0.462	309.64 ***	-9.107 ***
82	storm warning	20.78	16.50	14.88	0.00	100.00	0.13%	0.437	350.46 ***	-9.977 ***
83	strong wind	32.8	32.17	12.89	0.00	100.00	0.38%	0.511	663.40 ***	-6.287 ***
84	strong wind gust	11.27	3.67	14.21	0.00	78.17	49.04%	0.253	208.55 ***	-9.757 ***
85	tornado	25.45	23.67	18.33	3.00	100.00	0.00%	0.806	1886.52 ***	-3.760 ***
86	tornado risk	13.7	11.17	12.3	0.00	100.00	13.28%	0.470	578.05 ***	-5.172 ***
87	tornado warning	14.11	8.67	15.51	0.17	100.00	0.00%	0.337	249.84 ***	-7.392 ***
88	torrential rain	40.63	39.67	18.4	0.00	98.83	0.38%	0.881	2225.13 ***	-2.373
89	tropical cyclone	9.37	7.17	8.47	1.17	100.00	0.00%	0.342	150.71 ***	-11.703 ***
90	tropical cyclone risk	7.37	0.00	11.27	0.00	100.00	51.47%	0.313	185.98 ***	-10.679 ***
91	tropical cyclone warning	11.52	9.50	11.98	0.00	100.00	19.92%	0.239	105.16 ***	-9.888 ***
92	tropical storm	6.19	1.50	12.95	0.00	100.00	0.26%	0.412	215.04 ***	-18.016 ***
93	tropical storm risk	12.51	8.50	13.46	0.00	100.00	9.20%	0.483	366.20 ***	-8.709 ***
94	tropical storm warning	6.67	2.33	13.25	0.00	97.17	22.22%	0.272	152.35 ***	-9.172 ***
95	tropical weather	11.01	5.17	14.78	1.00	100.00	0.00%	0.594	573.30 ***	-8.174 ***
96	typhoon	15.90	7.00	16.94	1.83	100.00	0.00%	0.834	1780.16 ***	-2.749 *
97	typhoon risk	9.38	6.67	11.80	0.00	93.00	35.89%	0.387	236.54 ***	-7.597 ***
98	typhoon warning	10.65	5.83	13.36	0.00	100.00	0.64%	0.410	206.83 ***	-18.045 ***
99	weather blizzard warning	3.13	0.00	9.2	0.00	100.00	57.34%	0.196	40.50 ***	-22.883 ***
100	weather risk	31.82	31.00	11.67	0.00	100.00	0.38%	0.431	441.82 ***	-6.579 ***

(*cont.*) Table A.1 Summary of Google search volume indices (GSVIs)

Search keywords		Mean	Median	St.Dev	Min	Max	% zeros	AC1	LB4	ADF
Weather (WE; 113 keywords)										
101	weather warning	16.16	11.83	13.95	2.00	100.00	0.00%	0.445	428.58 ***	-6.498 ***
102	wet weather	45.1	43.50	12.44	14.00	100.00	0.00%	0.425	362.59 ***	-8.606 ***
103	wildfire	11.49	7.50	14.08	0.00	98.83	23.37%	0.531	508.92 ***	-11.054 ***
104	wildfire risk	12.48	10.67	12.31	0.00	100.00	24.52%	0.269	163.11 ***	-5.540 ***
105	wildfire warning	9.20	3.17	12.57	0.00	99.17	48.28%	0.296	145.76 ***	-5.594 ***
106	wildfires	4.25	1.83	8.95	0.00	100.00	1.66%	0.391	161.57 ***	-18.462 ***
107	wind	35.16	34.67	11.62	18.00	100.00	0.00%	0.557	674.54 ***	-7.932 ***
108	wind gust	13.71	10.83	11.38	0.00	100.00	0.89%	0.474	515.69 ***	-4.652 ***
109	wind gusts	9.75	5.67	11.71	0.00	100.00	4.73%	0.383	282.34 ***	-7.437 ***
110	wind risk	28.83	28.00	17.04	0.00	96.17	7.41%	0.519	842.55 ***	-3.451 ***
111	wind speed	23.31	23.67	9.95	9.00	100.00	0.00%	0.642	882.26 ***	-5.758 ***
112	wind storm	20.67	18.00	11.29	2.50	100.00	0.00%	0.414	221.39 ***	-17.966 ***
113	wind warning	15.19	11.83	12.96	0.00	100.00	6.00%	0.176	51.45 ***	-23.312 ***
Agricult. diseases (DI; 10 keywords)										
114	crop diseases	33.07	34.50	19.34	0.00	89.67	4.21%	0.659	1262.78 ***	-3.407 **
115	crop pest	31.95	33.17	18.74	0.00	97.33	2.43%	0.715	1509.30 ***	-2.954 **
116	crop pest risk	3.07	0.00	9.58	0.00	71.67	87.61%	0.140	35.56 ***	-24.640 ***
117	crop pests	29.69	30.33	18.57	0.00	94.50	4.98%	0.684	1316.31 ***	-4.570 ***
118	Ebola	15.36	7.33	18.17	0.00	100.00	21.20%	0.906	2198.67 ***	-3.850 ***
119	insect pest	16.02	45.67	19.18	8.33	92.50	0.00%	0.842	2146.91 ***	-3.299 **
120	La Roya	40.6	42.33	19.44	3.67	97.33	0.00%	0.808	1932.59 ***	-2.827 *
121	pest control	67.23	67.67	15.31	30.00	99.67	0.00%	0.939	2478.16 ***	-11.194 ***
122	pest risk	24.87	23.83	17.05	0.00	78.50	8.30%	0.548	890.41 ***	-3.472 ***
123	rust coffee	15.77	15.33	11.69	0.00	100.00	15.07%	0.432	522.26 ***	-5.238 ***
Geopolitical (GP; 14 keywords)										
124	Africa instability	23.02	23.17	17.16	0.00	89.67	18.52%	0.437	536.22 ***	-5.576 ***
125	Africa terrorism	24.92	22.50	12.91	0.00	100.00	0.89%	0.429	458.75 ***	-5.704 ***
126	Libyan crisis	7.30	1.50	10.84	0.00	100.00	29.89%	0.564	744.77 ***	-7.787 ***
127	Middle East conflict	16.26	14.33	11.03	0.00	100.00	0.13%	0.771	985.59 ***	-10.017 ***
128	Middle East instability	12.28	8.67	13.81	0.00	74.50	36.65%	0.342	318.81 ***	-6.707 ***
129	Middle East terrorism	21.81	19.33	12.87	0.00	100.00	0.77%	0.534	561.39 ***	-8.882 ***
130	oil crisis	33.04	29.83	15.94	6.00	98.00	0.00%	0.848	1654.46 ***	-7.968 ***
131	oil embargo	24.69	20.17	16.58	2.67	100.00	0.00%	0.831	1620.84 ***	-6.496 ***
132	oil outage	10.53	8.00	12.15	0.00	100.00	33.72%	0.438	484.36 ***	-3.076 **
133	Russian crisis	15.76	11.67	13.25	0.00	100.00	0.64%	0.762	1321.74 ***	-5.398 ***
134	Syrian war	10.75	8.83	12.64	0.00	100.00	21.07%	0.633	770.79 ***	-5.568 ***
135	terrorism	33.02	29.33	15.76	10.00	100.00	0.00%	0.870	1958.86 ***	-4.457 ***
136	terrorist attack	6.54	4.50	9.16	1.00	100.00	0.00%	0.289	81.65 ***	-20.732 ***
137	terrorist attacks	11.82	8.83	10.84	2.00	100.00	0.00%	0.543	485.03 ***	-7.016 ***
Economic (EC; 12 keywords)										
138	crisis	48.40	48.67	15.68	17.67	100.00	0.00%	0.944	2342.99 ***	-4.732 ***
139	economic crisis	24.00	18.17	18.8	1.50	100.00	0.00%	0.959	2516.64 ***	-4.084 ***
140	economic recession	26.65	20.17	19.55	0.00	100.00	0.26%	0.933	2419.65 ***	-4.365 ***
141	financial crisis	24.51	25.33	18.45	1.00	100.00	0.00%	0.952	2497.79 ***	-3.960 ***
142	recession	28.85	23.67	21.66	2.50	100.00	0.00%	0.944	2497.01 ***	-2.491
143	recession 2008	27.64	22.17	23.72	0.00	98.83	19.03%	0.936	2482.11 ***	-2.859 *
144	recession depression	23.67	13.67	23.1	0.00	100.00	2.17%	0.908	2263.30 ***	-4.789 ***
145	the recession	32.10	27.50	24.63	1.67	100.00	0.00%	0.958	2583.74 ***	-2.980 **
146	unemployment	41.28	39.17	11.2	23.17	100.00	0.00%	0.892	2111.31 ***	-5.008 ***
147	unemployment rate	48.49	47.33	16.06	11.00	100.00	0.00%	0.826	1673.84 ***	-4.472 ***
148	US recession	20.91	15.17	17.21	0.00	100.00	1.15%	0.890	2077.52 ***	-2.090
149	US unemployment	42.47	41.50	16.1	8.33	100.00	0.00%	0.770	1482.37 ***	-4.346 ***

Table A.2 Portfolios formed on second-end contracts and spreads

The table summarizes the performance of long-short strategies based on second-end contracts (Panel A) and spreads (Panel B). Spreads are defined as the return difference between front-end and second-end contracts. Newey-West robust *t*-statistics are shown in parentheses for the mean. CER denotes certainty equivalent return based on power utility. The time period covered by the portfolio returns is January 2005 (week 1) to December 2018 (week 4).

	CFEAR	Basis	Momentum	Hedging pressure	Convexity	Skewness	VaR1	VaR99	Basis momentum	Liquidity
Panel A: Second-end contracts										
Mean	0.0811	0.0599	0.0354	0.0522	0.0657	0.0469	-0.0257	0.0369	0.0429	-0.0124
	(3.26)	(2.54)	(1.30)	(2.22)	(2.83)	(1.91)	(-0.94)	(1.43)	(1.84)	(-0.47)
StDev	0.0939	0.0883	0.1086	0.0938	0.0830	0.0921	0.1057	0.1039	0.0858	0.0885
Downside volatility (0%)	0.0601	0.0573	0.0723	0.0591	0.0483	0.0525	0.0651	0.0631	0.0503	0.0570
Skewness	-0.2125	-0.1262	-0.2511	-0.0395	0.0428	0.2002	0.0224	0.0843	0.1747	-0.1007
	(-2.35)	(-1.39)	(-2.77)	(-0.44)	(0.47)	(2.21)	(0.25)	(0.93)	(1.93)	(-1.11)
Excess Kurtosis	0.7562	0.7683	0.9687	1.1404	0.1804	0.4453	0.5506	0.4431	0.6882	0.9831
	(4.17)	(4.24)	(5.35)	(6.29)	(1.00)	(2.46)	(3.04)	(2.45)	(3.80)	(5.43)
JB normality test *p*-value	0.0010	0.0010	0.0010	0.0010	0.5000	0.0085	0.0141	0.0337	0.0015	0.0010
99% VaR (Cornish-Fisher)	0.0328	0.0306	0.0402	0.0331	0.0256	0.0281	0.0362	0.0334	0.0271	0.0325
% of positive months	56%	55%	53%	54%	56%	51%	49%	52%	51.6%	49%
Maximum drawdown	-0.1890	-0.1538	-0.2101	-0.1816	-0.2261	-0.2434	-0.4943	-0.2724	-0.2383	-0.5490
Sharpe ratio	0.8633	0.6791	0.3257	0.5563	0.7918	0.5086	-0.2429	0.3551	0.4998	-0.1397
Sortino ratio	1.3481	1.0459	0.4894	0.8833	1.3612	0.8927	-0.3943	0.5846	0.8519	-0.2166
Omega ratio	1.3573	1.2757	1.1241	1.2209	1.3174	1.1975	0.9180	1.1356	1.1951	0.9513
CER (power utility)	0.0588	0.0403	0.0057	0.0301	0.0484	0.0257	-0.0537	0.0100	0.0245	-0.0320
Panel B: Spreads										
Mean	0.0145	-0.0277	-0.0182	0.0060	-0.0177	0.0036	0.0001	0.0036	0.0100	0.0076
	(1.76)	(-3.13)	(-2.76)	(1.01)	(-2.14)	(0.63)	(0.01)	(0.46)	(1.23)	(1.03)
StDev	0.0290	0.0334	0.0267	0.0215	0.0321	0.0239	0.0238	0.0269	0.0305	0.0284
Downside volatility (0%)	0.0200	0.0235	0.0189	0.0148	0.0225	0.0179	0.0173	0.0208	0.0214	0.0197
Skewness	0.3915	0.3383	-0.0457	0.0723	0.3018	0.2287	0.4533	-0.5295	-0.0119	0.1353
	(4.32)	(3.73)	(-0.50)	(0.80)	(3.33)	(2.52)	(5.00)	(-5.84)	(-0.13)	(1.49)
Excess Kurtosis	8.3974	6.0871	5.6028	2.6395	4.0826	15.1082	10.9424	6.9671	3.3884	6.8877
	(46.34)	(33.59)	(30.92)	(14.57)	(22.53)	(83.38)	(60.39)	(38.45)	(18.70)	(38.01)
JB normality test *p*-value	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010	0.0010
99% VaR (Cornish-Fisher)	0.0156	0.0166	0.0139	0.0085	0.0138	0.0187	0.0148	0.0158	0.0130	0.0149
% of positive months	51%	45%	44%	51%	47%	50%	51%	52%	52.4%	50%
Maximum drawdown	-0.0686	-0.3228	-0.2262	-0.0848	-0.2586	-0.0907	-0.1153	-0.1194	-0.1044	-0.0621
Sharpe ratio	0.5011	-0.8285	-0.6831	0.2783	-0.5500	0.1490	0.0034	0.1328	0.3291	0.2675
Sortino ratio	0.7253	-1.1797	-0.9658	0.4038	-0.7850	0.1985	0.0047	0.1723	0.4698	0.3851
Omega ratio	1.2364	0.7159	0.7623	1.1152	0.8033	1.0656	1.0014	1.0547	1.1380	1.1144
CER (power utility)	0.0124	-0.0305	-0.0200	0.0048	-0.0203	0.0021	-0.0013	0.0018	0.0077	0.0056

Table A.3 Relative importance of Google search keyword categories

The table summarizes the performance of CFEAR portfolios based on all keywords (Weather, WE; Agricultural Diseases, DI; Geopolitical, GP; and Economic, EC) and subsets thereof. Newey-West robust *t*-statistics are shown in parentheses for the mean. CER denotes certainty equivalent return based on power utility. The time period covered by the portfolio returns is January 2005 (week 1) to December 2018 (week 4).

	All keywords	Without WE	Without DI	Without GP	Without EC
Mean	0.0928	0.0475	0.0739	0.0948	0.0755
	(3.35)	(1.67)	(2.96)	(3.47)	(2.92)
StDev	0.1030	0.1033	0.0991	0.1041	0.1022
Downside volatility (0%)	0.0649	0.0658	0.0627	0.0659	0.0641
Skewness	-0.1307	-0.0350	-0.1013	-0.1776	-0.0717
	(-1.44)	(-0.39)	(-1.12)	(-1.96)	(-0.79)
Excess Kurtosis	0.4012	0.8518	0.5859	0.4217	0.5648
	(2.21)	(4.70)	(3.23)	(2.33)	(3.12)
JB normality test *p* -value	0.0320	0.0010	0.0068	0.0142	0.0101
99% VaR (Cornish-Fisher)	0.0341	0.0356	0.0334	0.0349	0.0341
% of positive months	57%	54%	56%	57%	56%
Maximum drawdown	-0.1881	-0.1784	-0.1196	-0.1464	-0.1862
Sharpe ratio	0.9012	0.4601	0.7454	0.9108	0.7381
Sortino ratio	1.4299	0.7214	1.1790	1.4384	1.1783
Omega ratio	1.3770	1.1796	1.3058	1.3796	1.3019
CER (power utility)	0.0660	0.0208	0.0492	0.0674	0.0492

Table A.4 CFEAR premium versus GPR and EPU uncertainty risks

Panel A reports the risk exposures of the long-short CFEAR portfolio in the context of the Bakshi et al. (2019) three-factor model augmented with innovations in the geopolitical risk uncertainty (GPR) index of Caldara and Iacoviello (2018) and the economic policy uncertainty (EPU) index of Baker et al. (2016). Newey-West robust *t*-statistics are shown in parentheses (Panel A). Panel B reports cross-sectional pricing tests for the same 56 commodity portfolios as in Table 5 of the manuscript. We report the (annualized) prices of risk with Shanken (1992) error-in-variables robust *t*-statistics in parentheses, and Kan et al. (2013) model misspecification and heteroscedasticity robust *t*-statistics in curly brackets. The observations for each variable cover the period from January 2005 (week 1) to December 2018 (week 4).

Panel A: Time-series tests

Constant	0.0018	0.0018
	(3.69)	(3.69)
AVG	-0.0110	-0.0113
	(-0.32)	(-0.33)
Basis	-0.1729	-0.1732
	(-2.89)	(-2.89)
Momentum	0.2758	0.2756
	(6.11)	(6.11)
ΔGPR	0.0017	
	(0.23)	
ΔEPU		-0.0322
		(-1.11)
Adj.-R^2 *(%)*	8.36	8.49

Panel B: Cross-sectional tests

Constant	-0.0002	-0.0007	-0.0001	-0.0006
	(-0.29)	(-0.86)	(-0.12)	(-0.72)
	{-0.29}	{-0.93}	{-0.12}	{-0.79}
CFEAR		0.0947		0.0915
		(2.81)		(2.74)
		{2.63}		{2.70}
AVG	-0.0200	0.0054	-0.0269	-0.0016
	(-0.35)	(0.10)	(-0.46)	(-0.03)
	{-0.33}	{0.10}	{-0.46}	{-0.03}
Basis	0.0511	0.0759	0.0552	0.0754
	(1.59)	(2.47)	(1.68)	(2.45)
	{1.62}	{2.65}	{1.70}	{2.63}
Momentum	0.0822	0.0419	0.0675	0.0393
	(2.11)	(1.17)	(1.85)	(1.15)
	{2.07}	{1.20}	{1.70}	{1.18}
ΔGPR	0.0160	0.0194		
	(1.25)	(1.50)		
	{0.59}	{1.01}		
ΔEPU			-0.0075	-0.0038
			(-1.80)	(-0.99)
			{-1.26}	{-0.81}
Adj.-R^2 (%)	34.47	65.51	39.43	64.20
MAPE (%)	0.048	0.038	0.048	0.038

Table A.5 Cross-sectional pricing tests: 28 individual commodities

The table reports the outcome of cross-sectional pricing tests using the 28 individual commodities as test assets. We report the (annualized) average prices of risk obtained in sequential (weekly) cross-sectional regressions on sequential betas with Fama-MacBeth (1973) *t*-statistics in curly brackets and Shanken (1992) corrected *t*-statistics in parentheses. The observations cover the period is January 2005 (week 1) to December 2018 (week 4).

	CFEAR	Base model		Base model augmented with															All risk factors	
				Fundamental risk factors				Tail risk factors						Liquidity and volatility risk factors					
Constant	-0.0009	0.0001	-0.0001	0.0003	0.0000	0.0001	0.0000	0.0005	0.0005	-0.0006	-0.0006	-0.0003	-0.0005	0.0000	-0.0004	-0.0003	-0.0006	-0.0005	-0.0008
	{-1.22}	{0.13}	{-0.17}	{0.34}	{0.03}	{0.09}	{-0.01}	{0.69}	{0.67}	{-0.66}	{-0.65}	{-0.36}	{-0.58}	{-0.03}	{-0.54}	{-0.38}	{-0.80}	{-0.55}	{-0.85}
	(-1.20)	(0.12)	(-0.15)	(0.32)	(0.03)	(0.09)	(-0.01)	(0.64)	(0.61)	(-0.65)	(-0.61)	(-0.35)	(-0.54)	(-0.03)	(-0.51)	(-0.37)	(-0.74)	(-0.51)	(-0.78)
CFEAR	0.0644		0.0755		0.0801		0.0742		0.0767		0.0878		0.0779		0.0711		0.0844		0.0768
	{2.17}		{2.54}		{2.73}		{2.47}		{2.60}		{3.00}		{2.66}		{2.35}		{2.84}		{2.49}
	(2.12)		(2.36)		(2.52)		(2.30)		(2.35)		(2.79)		(2.51)		(2.21)		(2.65)		(2.28)
AVG		-0.0500	-0.0402	-0.0578	-0.0477	-0.0500	-0.0466	-0.0729	-0.0739	-0.0171	-0.0183	-0.0311	-0.0222	-0.0433	-0.0231	-0.0296	-0.0144	-0.0004	-0.0051
		{-1.09}	{-0.87}	{-1.23}	{-1.01}	{-1.07}	{-1.01}	{-1.56}	{-1.57}	{-0.35}	{-0.38}	{-0.63}	{-0.45}	{-0.90}	{-0.49}	{-0.65}	{-0.31}	{-0.34}	{-0.09}
		(-1.04)	(-0.81)	(-1.17)	(-0.93)	(-1.03)	(-0.94)	(-1.43)	(-1.42)	(-0.34)	(-0.35)	(-0.61)	(-0.42)	(-0.87)	(-0.46)	(-0.64)	(-0.29)	(-0.32)	(-0.09)
Basis		0.0256	0.0507	0.0020	0.0294	0.0213	0.0454	0.0014	0.0185	0.0075	0.0455	0.0093	0.0383	0.0065	0.0401	0.0335	0.0465	0.0004	0.0242
		{0.68}	{1.39}	{0.05}	{0.78}	{0.56}	{1.21}	{0.04}	{0.50}	{0.20}	{1.22}	{0.25}	{1.04}	{0.17}	{1.05}	{0.89}	{1.24}	{0.49}	{0.54}
		(0.65)	(1.28)	(0.05)	(0.72)	(0.54)	(1.13)	(0.04)	(0.45)	(0.20)	(1.14)	(0.25)	(0.98)	(0.16)	(0.99)	(0.86)	(1.16)	(0.45)	(0.49)
Momentum		0.0691	0.0604	0.0347	0.0164	0.0607	0.0508	0.0197	0.0036	0.0544	0.0397	0.0524	0.0470	0.0570	0.0382	0.0488	0.0414	0.0002	-0.0050
		{1.86}	{1.58}	{0.94}	{0.43}	{1.57}	{1.31}	{0.53}	{0.10}	{1.43}	{1.04}	{1.40}	{1.24}	{1.47}	{0.98}	{1.27}	{1.06}	{0.27}	{-0.11}
		(1.77)	(1.46)	(0.90)	(0.40)	(1.52)	(1.22)	(0.49)	(0.09)	(1.39)	(0.97)	(1.36)	(1.17)	(1.42)	(0.92)	(1.24)	(0.99)	(0.25)	(-0.10)
Hedging pressure				0.0587	0.0540													0.0011	0.0509
				{2.00}	{1.83}													{1.83}	{1.68}
				(1.90)	(1.68)													(1.70)	(1.54)
Convexity						-0.0076	-0.0120											0.0005	0.0547
						{-0.15}	{-0.24}											{0.47}	{0.96}
						(-0.15)	(-0.22)											(0.43)	(0.88)
Skewness								0.0846	0.0780									0.0012	0.0611
								{2.50}	{2.29}									{1.70}	{1.57}
								(2.30)	(2.07)									(1.58)	(1.44)
VaR1										-0.0423	-0.0265							-0.0008	-0.0314
										{-1.29}	{-0.83}							{-1.20}	{-0.89}
										(-1.26)	(-0.77)							(-1.12)	(-0.82)
VaR99												0.0438	0.0376					0.0002	0.0148
												{1.31}	{1.14}					{0.33}	{0.41}
												(1.27)	(1.07)					(0.31)	(0.38)
Basis-momentum														0.0419	0.0539			0.0011	0.0545
														{1.28}	{1.62}			{1.49}	{1.48}
														(1.23)	(1.52)			(1.39)	(1.36)
Liquidity																-0.0240	-0.0088	0.0000	-0.0021
																{-0.81}	{-0.30}	{-0.05}	{-0.07}
																(-0.79)	(-0.28)	(-0.05)	(-0.06)
Adj.-R^2 (%)	6.95	14.93	19.76	19.25	23.35	18.37	22.79	18.90	22.83	19.60	23.07	19.47	22.65	18.64	23.25	18.40	22.81	36.25	38.55
MAPE (%)	2.543	2.322	2.193	2.208	2.092	2.225	2.101	2.212	2.093	2.201	2.097	2.199	2.101	2.221	2.099	2.221	2.099	1.656	1.569

Table A.6 Summary of commodity-specific media coverage (news) variables

The table summarizes the distribution of the total number of news articles published each week on each commodity with relevance score 25 (news mildly related to the commodity) and relevance score 75 (strongly related). AC1 is the first-order autocorrelation. LB4 is the Ljung-Box test statistic (H_0: first four autocorrelations are jointly zero) and Augmented Dickey-Fuller test statistic (H_0: unit root non-stationarity). *. ** and *** are significant at the 10%, 5% and 10% levels, respectively. The observation period is January 2004(week1) – December 2018(week4).

	News with relevance score 25/100								News relevance score 75/100							
Commodity	**Mean**	**Median**	**StDev**	**Min**	**Max**	**AC1**	**LB4**	**ADF**	**Mean**	**Median**	**StDev**	**Min**	**Max**	**AC1**	**LB4**	**ADF**
I. Agricultural sector (N=17)																
Corn	2970.7	2977.0	492.8	152.0	5331.0	0.649	892 ***	-6.71 ***	888.1	938.0	276.3	0.0	1678	0.828	1879 ***	-3.69 ***
Oats	928.4	912.0	223.6	105.0	1609.0	0.786	1636 ***	-4.46 ***	126.7	103.0	88.2	0.0	468	0.918	2403 ***	-3.77 ***
Rough rice	1264.4	1212.0	361.7	47.0	2289.0	0.604	977 ***	-6.08 ***	188.6	184.0	71.0	7.0	608	0.822	1608 ***	-6.99 ***
Wheat CBT	2832.6	2951.0	518.7	159.0	4904.0	0.712	1188 ***	-5.63 ***	1147.1	1182.0	308.6	15.0	1906	0.795	1634 ***	-4.38 ***
Cotton no.2	735.8	769.0	265.4	123.0	2193.0	0.728	1476 ***	-3.53 ***	279.6	277.0	104.7	17.0	1513	0.663	1255 ***	-1.82
Soybeans	2170.7	2663.0	502.2	145.0	4593.0	0.719	1324 ***	-5.42 ***	697.3	702.0	174.5	4.0	1204	0.753	1365 ***	-5.31 ***
Soybean meal	1190.3	1123.0	294.9	0.0	2500.0	0.862	2023 ***	-3.87 ***	293.8	291.0	79.7	0.0	1178	0.660	944.5 ***	-6.67 ***
Soybean oil	1507.0	1489.0	260.0	116.0	2327.0	0.735	1240 ***	-5.86 ***	381.5	381.0	127.4	6.0	759	0.844	1904 ***	-4.19 ***
Feeder cattle	577.4	520.0	236.5	0.0	1135.0	0.895	2373 ***	-2.51	127.4	125.0	33.2	0.0	325	0.466	267.1 ***	-16.93 ***
Lean hogs	526.3	458.0	211.8	108.0	1209.0	0.907	2337 ***	-3.05 **	260.2	233.0	94.1	0.0	706	0.835	1933 ***	-3.20 **
Live cattle	543.7	519.0	239.0	36.0	1224.0	0.912	2476 ***	-2.65 *	97.5	78.0	57.6	0.0	614	0.742	1422 ***	-5.03 ***
Frozen pork bellies	192.4	76.0	217.2	0.0	1282.0	0.959	2764 ***	-2.53	45.8	4.0	74.2	0.0	667	0.849	2085 ***	-2.88 **
Cocoa	562.6	550.0	194.1	155.0	1407.0	0.754	1455 ***	-4.88 ***	255.2	258.0	107.0	24.0	763	0.888	2295 ***	-2.50
Coffee C	668.5	704.0	241.4	133.0	1522.0	0.750	1554 ***	-3.18 **	332.9	338.0	144.7	32.0	873	0.900	2366 ***	-1.66
Frozen Orange juice	87.2	86.0	51.8	9.0	417.0	0.799	1856 ***	-3.27 **	23.8	19.0	17.1	0.0	233	0.494	674.4 ***	-4.22 ***
Sugar no.11	725.1	731.0	279.9	108.0	1461.0	0.820	1928 ***	-2.76 *	331.4	343.0	149.8	26.0	658	0.902	2387 ***	-2.72 *
Lumber	408.6	310.0	271.0	5.0	1009.0	0.954	2790 ***	-1.91	81.1	70.0	41.6	0.0	374	0.828	1675 ***	-6.66 ***
II. Energy sector (N=6)																
Light crude oil	8064.9	7789.0	2604.7	2928.0	20459.0	0.825	1607 ***	-5.28 ***	3633.5	3466.0	1257.8	1460.0	9319	0.817	1678 ***	-4.69 ***
Electricity JPM	470.8	434.0	316.9	12.0	1534.0	0.839	2140 ***	-0.41	108.9	96.0	86.6	1.0	451	0.916	2483 ***	-1.42
Gasoline RBOB	2225.9	2080.0	608.1	469.0	5371.0	0.841	1944 ***	-3.46 ***	783.4	767.0	254.7	152.0	1773	0.819	1836 ***	-3.88 ***
Heating oil	711.6	694.0	361.9	15.0	2020.0	0.934	2511 ***	-0.64	161.4	158.0	147.8	6.0	688	0.919	2565 ***	-2.20
Natural gas	3049.5	3050.0	921.8	835.0	5740.0	0.835	1584 ***	-0.95	1115.2	1113.0	427.0	321.0	3187	0.815	1514 ***	-1.15
NY unleaded gas	2225.9	2080.0	608.1	469.0	5371.0	0.841	1944 ***	-3.46 ***	783.4	767.0	254.7	152.0	1773	0.819	1836 ***	-3.88 ***
III. Metals (N=5)																
Copper (High Grade)	1857.5	1733.0	665.1	422.0	3889.0	0.889	2267 ***	-2.78 *	516.9	502.0	226.1	51.0	1310	0.849	1994 ***	-1.48
Gold 100oz (CMX)	2727.3	2514.0	911.6	741.0	6041.0	0.869	2117 ***	-2.91 **	1035.6	985.0	386.8	265.0	2713	0.820	1813 ***	-3.33 **
Palladium	456.8	426.0	218.7	105.0	1058.0	0.911	2444 ***	-2.24	161.2	197.0	156.8	25.0	991	0.706	2196 ***	-4.29 ***
Platinum	836.4	831.0	440.1	148.0	2134.0	0.951	2720 ***	-1.04	261.2	197.0	156.8	25.0	991	0.866	2196 ***	-2.87 **
Silver 5000 oz	1164.2	1115.0	338.4	256.0	2722.0	0.865	2040 ***	-2.23	323.4	307.0	134.3	93.0	1072	0.804	1730 ***	-3.58 ***

Table A.7 Coefficient estimates of predictability regressions à la Vozlyublennaia (2014)

This table reports the OLS estimation results of return predictability regressions $r_{i,t} = \varphi_i + \sum_{l=1}^{4} \alpha_{il} \cdot r_{i,t-l} + \sum_{l=1}^{4} \beta_{il} \cdot \Delta S_{t-l} + \sum_{l=1}^{4} \delta_{il} \cdot r_{i,t-1} \cdot \Delta S_{t-l} + \varepsilon_{i,t}$ per commodity ($i = 1,\ldots,28$). ΔS_{t-l} is the lagged aggregated change in searches by 149 hazards as query terms, $\Delta S_t = \sum_{j=1}^{149} \Delta S_{j,t}^{*}$ where $\Delta S_{j,t}^{*} = \frac{\Delta S_{j,t}}{\sigma_{j,t}^{\Delta S}} = \frac{\ln(s_{j,t}/s_{j,t-1})}{\sigma_{j,t}^{\Delta S}}$ is the standardized "attention" variable associated with the jth hazard. The absolute value of the coefficient of lagged return $|\alpha_{il}|$ captures the degree of predictability when there is no attention, and $|\alpha_{il} + \delta_{il}|$ with δ_{il} the coefficient of the interaction variable captures the degree of predictability when there is attention. The estimation window is January 2004 (week 1) to December 2018 (week 4). Newey-West robust *t*-statistics are shown in parentheses. Significant slopes are highlighted in bold.

	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,3}$	$\alpha_{i,4}$	$\delta_{i,1}$	$\delta_{i,2}$	$\delta_{i,3}$	$\delta_{i,4}$
I. Agricultural sector (N=17)								
Corn	-0.0032	0.0380	0.0183	0.0423	0.0004	0.0017	0.0016	0.0012
	(-0.07)	(1.02)	(0.46)	(1.05)	(0.22)	(1.07)	(1.04)	(0.69)
Oats	-0.0294	0.0002	0.0423	**-0.0756**	0.0006	0.0007	-0.0008	0.0007
	(-0.91)	(0.01)	(1.09)	(-1.92)	(0.39)	(0.45)	(-0.68)	(0.62)
Rough rice	-0.0013	-0.0337	-0.0482	-0.0210	**-0.0019**	**0.0024**	0.0000	-0.0001
	(-0.03)	(-0.84)	(-1.30)	(-0.54)	(-1.71)	(2.04)	(0.03)	(-0.05)
Wheat CBT	0.0162	-0.0186	0.0045	-0.0021	-0.0002	0.0008	**0.0020**	0.0004
	(0.41)	(-0.50)	(0.12)	(-0.06)	(-0.17)	(0.54)	(1.70)	(0.29)
Cotton no.2	0.0195	-0.0120	-0.0285	-0.0291	0.0011	**0.0025**	0.0015	**-0.0032**
	(0.50)	(-0.28)	(-0.67)	(-0.75)	(0.87)	(1.84)	(0.81)	(-2.21)
Soybeans	0.0280	-0.0074	0.0135	-0.0146	-0.0008	0.0005	-0.0014	0.0009
	(0.74)	(-0.18)	(0.37)	(-0.31)	(-0.32)	(0.34)	(-0.74)	(0.52)
Soybean meal	0.0454	-0.0011	0.0139	-0.0355	-0.0007	-0.0002	-0.0006	0.0023
	(1.13)	(-0.03)	(0.40)	(-0.75)	(-0.33)	(-0.16)	(-0.31)	(1.30)
Soybean oil	-0.0153	-0.0289	0.0348	-0.0161	-0.0006	0.0003	-0.0018	-0.0016
	(-0.39)	(-0.76)	(0.66)	(-0.31)	(-0.31)	(0.17)	(-1.31)	(-1.09)
Feeder cattle	-0.0459	-0.0166	0.0603	0.0225	-0.0014	0.0004	0.0025	-0.0020
	(-1.21)	(-0.45)	(1.56)	(0.46)	(-0.90)	(0.25)	(1.61)	(-1.39)
Lean hogs	0.0772	**0.1037**	-0.0290	-0.0642	0.0019	**0.0030**	-0.0022	-0.0023
	(1.60)	(2.26)	(-0.54)	(-1.59)	(1.16)	(2.01)	(-1.55)	(-1.31)
Live cattle	**-0.0663**	**-0.1561**	0.0386	0.0251	-0.0014	0.0002	0.0009	-0.0014
	(-1.94)	(-4.79)	(0.91)	(0.55)	(-0.97)	(0.13)	(0.59)	(-1.04)
Frozen pork bellies	-0.0485	0.0511	-0.1071	0.0347	0.0000	-0.0006	-0.0005	0.0019
	(-0.85)	(0.90)	(-1.52)	(0.68)	(0.01)	(-0.44)	(-0.26)	(0.88)
Cocoa	-0.0141	-0.0573	-0.0382	-0.0546	0.0013	0.0013	-0.0020	0.0000
	(-0.38)	(-1.35)	(-0.96)	(-1.40)	(1.04)	(0.75)	(-1.35)	(-0.01)
Coffee C	0.0092	-0.0247	-0.0365	-0.0477	0.0006	0.0002	0.0010	-0.0027
	(0.21)	(-0.61)	(-0.90)	(-1.36)	(0.38)	(0.14)	(0.71)	(-1.57)
Frozen Orange juice	0.0363	-0.0304	**-0.1007**	-0.0128	-0.0012	0.0007	0.0001	0.0016
	(0.95)	(-0.83)	(-2.86)	(-0.33)	(-0.86)	(0.50)	(0.07)	(0.95)
Sugar no.11	-0.0336	0.0228	**0.1010**	-0.0129	-0.0011	0.0002	0.0005	0.0011
	(-0.82)	(0.58)	(2.58)	(-0.36)	(-0.78)	(0.10)	(0.42)	(0.80)
Lumber	0.0198	0.0251	-0.0396	-0.0440	0.0023	-0.0002	-0.0005	0.0002
	(0.54)	(0.61)	(-1.03)	(-1.09)	(1.23)	(-0.10)	(-0.38)	(0.13)

	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,3}$	$\alpha_{i,4}$	$\delta_{i,1}$	$\delta_{i,2}$	$\delta_{i,3}$	$\delta_{i,4}$
II. Energy sector (N=6)								
Light crude oil	0.0022	0.0506	0.0488	0.0152	0.0020	0.0027	-0.0009	-0.0018
	(0.05)	(1.27)	(0.94)	(0.38)	(1.40)	(0.93)	(-0.54)	(-0.87)
Electricity JPM	0.0201	0.0301	0.0157	**0.1024**	0.0011	**-0.0038**	0.0011	-0.0002
	(0.47)	(0.71)	(0.38)	(2.13)	(0.62)	(-2.84)	(0.55)	(-0.14)
Gasoline RBOB	0.0488	0.0351	**0.1238**	-0.0109	**0.0026**	0.0034	-0.0023	-0.0024
	(1.03)	(0.75)	(2.08)	(-0.22)	(2.04)	(1.07)	(-1.49)	(-1.24)
Heating oil	0.0270	0.0003	0.0221	0.0368	0.0000	0.0016	-0.0014	-0.0016
	(0.59)	(0.01)	(0.47)	(0.97)	(0.01)	(0.71)	(-0.98)	(-0.90)
Natural gas	-0.0301	**0.0645**	-0.0182	-0.0065	-0.0001	-0.0004	0.0013	0.0012
	(-0.79)	(1.69)	(-0.41)	(-0.16)	(-0.11)	(-0.31)	(1.02)	(0.81)
NY unleaded gas	-0.0301	0.1025	0.0441	-0.0668	**0.0043**	0.0000	-0.0006	0.0028
	(-0.41)	(1.24)	(0.53)	(-0.89)	(2.14)	(-0.01)	(-0.28)	(1.56)
III. Metals (N=5)								
Copper (High Grade)	0.0136	**0.0838**	0.0324	0.0493	-0.0001	0.0017	-0.0004	0.0003
	(0.28)	(1.93)	(0.73)	(1.21)	(-0.09)	(0.76)	(-0.26)	(0.17)
Gold 100oz (CMX)	-0.0118	-0.0134	**-0.0571**	0.0146	0.0006	0.0013	-0.0015	0.0006
	(-0.30)	(-0.35)	(-1.93)	(0.32)	(0.40)	(1.30)	(-1.06)	(0.48)
Palladium	0.0201	-0.0008	-0.0082	-0.0152	-0.0016	0.0001	-0.0020	-0.0015
	(0.47)	(-0.02)	(-0.15)	(-0.40)	(-1.02)	(0.08)	(-1.29)	(-1.10)
Platinum	0.0145	0.0601	0.0327	-0.0228	0.0006	0.0022	-0.0007	0.0029
	(0.44)	(1.28)	(0.54)	(-0.48)	(0.40)	(1.61)	(-0.42)	(1.63)
Silver 5000 oz	0.0162	-0.0182	-0.0335	0.0373	-0.0003	0.0006	**-0.0035**	0.0009
	(0.37)	(-0.49)	(-0.92)	(0.93)	(-0.21)	(0.55)	(-2.07)	(0.76)

Table A.8 Robustness tests: Alternative CFEAR portfolio construction methods

The table summarizes the CFEAR factor obtained through alternative long-short portfolio construction methods where: (1) the lookback period is a fixed-length rolling window of 10 years ($L = 520$ weeks); (2) the long Q1 and short Q5 quintile constituents are weighted by the strength of the standardized signals; (3)-(6) the long Q1 and short Q5 portfolios include $N/2$ commodities each which are weighted equally, by standardized rankings, by standardized signals, and by winsorized and standardized signals; (7) at each portfolio formation time we consider only the $0.8N$ of the commodities with the largest open interest on the prior week. The portfolio returns are from January 2005 (week 1) to December 2018 (week 5), except in column (1) which are from January 2014 (week 1).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Whole cross-section				
	Quintiles, Rolling windows (L=10 years)	Quintiles, Standardized signals	Equal weights	Standardized rankings	Standardized signals	Winsorized standardized signals	80% most liquid comm
Mean	0.0740	0.0827	0.0495	0.0621	0.0593	0.0607	0.1014
	(1.66)	(2.64)	(3.08)	(3.12)	(2.40)	(2.50)	(3.47)
StDev	0.0980	0.1156	0.0601	0.0760	0.0943	0.0928	0.1114
Downside volatility (0%)	0.0575	0.0721	0.0368	0.0475	0.0595	0.0587	0.0672
Skewness	0.1889	-0.0474	-0.0591	-0.1231	-0.0839	-0.0863	-0.0217
Excess Kurtosis	0.4577	0.7776	0.3853	0.4418	0.8139	0.7258	0.3881
JB normality test p-value	0.1151	0.0013	0.0776	0.0234	0.0010	0.0019	0.0894
1% VaR (Cornish-Fisher)	0.0296	0.0392	0.0196	0.0253	0.0325	0.0317	0.0356
% of positive months	54%	54%	56%	57%	54%	54%	56%
Maximum drawdown	-0.1071	-0.1696	-0.0837	-0.1131	-0.1343	-0.1331	-0.1731
Sharpe ratio	0.7545	0.7151	0.8234	0.8176	0.6293	0.6539	0.9101
Sortino ratio	1.2856	1.1462	1.3460	1.3074	0.9968	1.0334	1.5093
Omega ratio	1.3186	1.2921	1.3372	1.3350	1.2531	1.2637	1.3801
CER (power utility)	0.0500	0.0491	0.0404	0.0476	0.0370	0.0391	0.0702

Table A.9 Long-short portfolios with monthly rebalancing

The table reports the performance of the CFEAR portfolio, alongside the set of alternative portfolios summarized in Appendix A deployed with end-of-month sorting and a one-month holding period. Newey-West robust h.a.c. *t*-statistics are shown in parentheses. CER is the certainty equivalent return based on power utility. The return observations cover the time period January 2005 (week 1) to December 2018 (week 4).

	CFEAR	**AVG**	**Basis**	**Momentum**	**Hedging pressure**	**Convexity**	**Skewness**	**VaR1**	**VaR99**	**Basis-Momentum**	**Liquidity**
Mean	0.0798	-0.0330	0.0516	0.0068	0.0516	0.0635	0.0457	-0.0174	0.0443	0.0411	0.0072
	(3.06)	(-0.72)	(1.62)	(0.24)	(1.99)	(2.33)	(1.97)	(-0.59)	(1.47)	(1.36)	(0.24)
StDev	0.1010	0.1471	0.1030	0.1108	0.0923	0.1066	0.0961	0.1141	0.1082	0.1023	0.1012
Downside volatility (0%)	0.0544	0.1170	0.0587	0.0677	0.0516	0.0561	0.0550	0.0758	0.0677	0.0712	0.0742
Skewness	0.0155	-0.8037	0.3772	0.2467	0.1402	0.4405	0.1845	-0.1963	-0.1796	-0.4079	-0.4276
	(0.08)	(-4.25)	(2.00)	(1.31)	(0.74)	(2.33)	(0.98)	(-1.04)	(-0.95)	(-2.16)	(-2.26)
Excess Kurtosis	0.2009	3.1259	1.5263	0.7557	0.3367	1.8552	0.2993	0.4661	0.4298	1.3369	1.4942
	(0.53)	(8.27)	(4.04)	(2.00)	(0.89)	(4.91)	(0.79)	(1.23)	(1.14)	(3.54)	(3.95)
JB normality test p-value	0.5000	0.0010	0.0027	0.0488	0.4535	0.0010	0.3922	0.2117	0.2733	0.0042	0.0025
1% VaR (Cornish-Fisher)	0.0622	0.1474	0.0656	0.0729	0.0568	0.0674	0.0585	0.0859	0.0759	0.0815	0.0848
% of positive months	61%	50%	55%	50%	55%	55%	54%	46%	58%	58%	55%
Maximum drawdown	-0.1664	-0.5394	-0.1977	-0.2533	-0.2070	-0.1793	-0.1975	-0.4145	-0.2511	-0.2600	-0.4237
Sharpe ratio	0.7906	-0.2244	0.5015	0.0618	0.5592	0.5956	0.4753	-0.1530	0.4089	0.4018	0.0714
Sortino ratio	1.4673	-0.2822	0.8797	0.1011	0.9996	1.1328	0.8299	-0.2303	0.6534	0.5777	0.0974
Omega ratio	1.7623	0.8357	1.4558	1.0480	1.5001	1.5744	1.4238	0.8941	1.3480	1.3517	1.0567
CER (power utility)	0.0540	-0.0943	0.0257	-0.0235	0.0304	0.0356	0.0229	-0.0510	0.0147	0.0143	-0.0191

Table A.10 Cross-sections of financial futures contracts

The table details the futures contracts employed in the placebo tests of Section 5.4 of the paper. The observation period is January 2004 (week 1) to December 2018 (week 4).

Panel A: Equity index futures (N=40)	Panel B: Fixed Income and interest rates futures (N=13)	Panel C: Currency futures (N=19)
Dow-Jones Industrial Average	1-Month Eurodollar	Australian Dollar
E-mini Dow-Jones Industrial Average	30-Day FED Funds	Brazilian Real
E-mini MSCI EAFE	3-Month Eurodollar	Canadian Dollar
E-mini MSCI Emerging Markets	2-Year U.S. Treasury Note	Chinese Renmimbi
E-mini Russell 2000	3-Year U.S. Treasury Note	Czech Koruna
E-Mini S&P500	5-Year Eurodollar Bundle	Euro
Euro Stoxx 50	5-Year U.S. Treasury Note	Hungarian Forint
MSCI Asia	10-Year U.S. Treasury Note	Israeli Shekel
MSCI Emerging Markets Latin America	30-Year U.S. Treasury Bond	Japanese Yen
MSCI India	Barclays Capital U.S. Aggregate	Korean Won
MSCI Russia	Municipal Bond Index	Mexican Peso
MSCI Taiwan	Ultra 10-Year U.S. Treasury Note	New Zealand Dollar
MSCI Thailand	Ultra Treasury Bond Index	Norwegian Krona
MSCI USA		Polish Zloty
MSCI World		Russian Rouble
Nasdaq 100		South African Rand
Nasdaq Biotechnology		Sterling
Nikkei 225		Swedish Krona
NYSE composite		Swiss Franc
Russell 1000		
Russell 1000 Growth		
Russell 1000 Value		
Russell 2000 Growth		
Russell 2000 Value		
Russell 3000		
S&P Citigroup Growth		
S&P Citigroup Value		
S&P Consumer Discretionary		
S&P Consumer Staples		
S&P Energy		
S&P Finance		
S&P Health		
S&P Industrial		
S&P Information Technology		
S&P Materials		
S&P Small Capitalization		
S&P Utilities		
S&P400 Mid Capitalization		
S&P500		
Value Line		